



HAL
open science

Méthodes de traitement et d'analyse d'image par réseaux de neurones pour la microscopie de fluorescence

Emmanuel Bouilhol

► **To cite this version:**

Emmanuel Bouilhol. Méthodes de traitement et d'analyse d'image par réseaux de neurones pour la microscopie de fluorescence. Imagerie médicale. Université de Bordeaux, 2022. Français. NNT : 2022BORD0418 . tel-04086957

HAL Id: tel-04086957

<https://theses.hal.science/tel-04086957>

Submitted on 2 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR
DE L'UNIVERSITÉ DE BORDEAUX

ECOLE DOCTORALE MATHÉMATIQUES ET
INFORMATIQUE

SPECIALITÉ : INFORMATIQUE

Par **Emmanuel BOUILHOL**

Méthodes de traitement et d'analyse d'image par réseaux de
neurones pour la microscopie de fluorescence

Sous la direction de : **Macha NIKOLSKI**

Soutenue le 15 décembre 2022

Membres du jury :

Laure BLANC-FÉRAUD	Directrice de recherche	Université Côte d'Azur	Examinatrice
Pascal DESBARATS	Professeur	Université de Bordeaux	Président
Christine FERNANDEZ-MALOIGNE	Professeur	Université de Poitiers	Rapportrice
Macha NIKOLSKI	Directrice de Recherche	Université de Bordeaux	Directrice de thèse
Jean-Christophe OLIVO-MARIN	Directeur de Recherche	Institut Pasteur	Examinateur
Emeline RIBOT	Chargée de Recherche	Université de Bordeaux	Examinatrice
Thomas WALTER	Professeur	Ecole des Mines de Paris	Rapporteur

« *Le hasard ne favorise que les esprits préparés.* »

Louis PASTEUR, Douai, le 7 décembre 1854, Discours d'inauguration de la Faculté des sciences de Lille

Méthodes de traitement et d'analyse d'image par réseaux de neurones pour la microscopie de fluorescence

Résumé :

Le développement des méthodes de microscopie de fluorescence permet aujourd'hui l'acquisition de données d'imagerie cellulaire à haut débit et à une résolution moléculaire. Ces techniques sont désormais un élément essentiel pour faire la lumière sur les processus cellulaires. Ainsi, la microscopie rend possible l'étude de phénotypes cellulaires couvrant les phénomènes globaux tels que les modifications morphologiques de la cellule entière et de ses composants, ou encore la détection de l'altération des processus cellulaires grâce à la quantification et localisation des molécules. Dans cette thèse, nous nous intéresserons particulièrement à l'emploi de méthodes informatiques et mathématiques pour analyser de manière quantitative les données d'imagerie de microscopie de fluorescence.

Dans le premier chapitre, il s'agit d'étudier comment l'utilisation de réseaux de neurones à convolution dédiés au rehaussement de l'intensité du signal de fluorescence des spots d'ARNm dans les images de microscopie permet d'automatiser leur détection en aval par des outils conventionnels largement utilisés par la communauté d'analyse d'images de microscopie. Nous présentons dans ce chapitre la méthode DeepSpot, et nous montrons qu'en renforçant le signal de tous les spots de manière à ce qu'ils aient la même intensité sur toutes les images, indépendamment du contraste, du bruit ou de la taille des spots, il devient possible d'automatiser leur détection en aval.

Dans le second chapitre, nous proposons une approche de deep learning non supervisée de rehaussement du signal des noyaux cellulaires. Plus précisément, nous appliquons le transfert de style entre des images d'organoïdes acquises avec une faible intensité lumineuse vers des images sur lesquelles les noyaux sont visiblement saillants, sans qu'une annotation préalable des images ne soit requise. Pour cela, nous avons implémenté un CycleGAN avec une fonction de coût sur mesure dédiée à l'augmentation du signal des noyaux. Nous avons entraîné ce réseau de neurones et évalué l'impact du rehaussement du signal des noyaux sur leur segmentation en aval, à l'aide de méthodes conventionnelles ainsi qu'avec des méthodes basées sur l'apprentissage.

Dans le troisième chapitre, nous présentons DypFISH, une librairie d'analyse de données spatiales, qui propose des outils d'analyse quantitative pour étudier les localisations subcellulaires des ARNm et des protéines de manière statistiquement robuste. Entre autres, DypFISH introduit des techniques qui permettent l'évaluation conjointe des données ponctuelles d'ARNm dans les images smFISH et des données continues de protéines par immunofluorescence (IF).

Mots-clés : Deep Learning, Analyse d'image, Microscopie de fluorescence, Analyses spatiales

Neural network-based methods for image processing and analysis of fluorescence microscopy images

Abstract:

The advent of fluorescence microscopy methods make it currently possible to acquire cellular imaging data on a molecular scale at that in high-throughput fashion. These techniques now constitute an essential tool for unraveling cellular processes. Thus, microscopy makes it possible to study cellular phenotypes covering global phenomena such as morphological modifications of the whole cell and of its components, or the detection of alterations of cellular processes thanks to the quantification and localization of molecules. In this thesis, we will focus on the use of computational and mathematical methods to quantitatively analyse fluorescence microscopy imaging data.

In the first chapter, we study how the use of convolutional neural networks dedicated to the enhancement of the intensity of fluorescence signal of mRNA spots in microscopy images allows to automate their downstream detection by conventional tools widely used by the microscopy image analysis community. In this chapter, we present the DeepSpot deep learning method, and show that by enhancing the signal of all spots up to the same intensity in all images, regardless of the contrast, noise or size of the spots, it enables seamless automation of their downstream detection.

In the second chapter, we propose an unsupervised deep learning approach for cell nuclei signal enhancement. More precisely, we apply style transfer between images of organoids acquired with low light intensity to images where the nuclei are visibly prominent, without requiring prior annotation of the images. For this purpose, we implemented a CycleGAN with a custom cost function dedicated to the signal enhancement of nuclei. We trained this neural network and evaluated the impact of nuclei signal enhancement on their downstream segmentation, using conventional methods as well as learning-based methods.

In the third chapter, we present DypFISH, a spatial data analysis library, which provides quantitative analysis tools to study subcellular localizations of mRNAs and proteins in a statistically robust manner. In particular, DypFISH introduces techniques that allow joint evaluation of point data of mRNA in smFISH images and continuous immunofluorescence (IF) protein data.

Keywords: Deep Learning, Image analysis, Fluorescent microscopy, Spatial analysis

Unité de recherche

UMR 5095 Université de Bordeaux, 33000 Bordeaux, France.

Remerciements

S'il y a une seule conviction que j'ai pu acquérir au cours de ces années, c'est bien la certitude qu'on ne fait jamais rien tout seul, et encore moins une thèse. J'ai, à ce titre, de nombreux remerciements à adresser.

Je tiens à remercier en premier lieu les membres du jury d'avoir accepté de lire et de juger ce travail, Laure Blanc-Féraud, Pascal Desbarat, Christine Fernandez-Maloigne, Jean-Christophe Olivo-Marin, Emeline Ribot, Thomas Walter. Que ce soit par vos travaux, par des cours d'analyse d'image, par votre présence à différents moments de cette thèse, vous avez chacun participé à édifier ce projet de près ou de loin et je suis ravi que vous soyez là, à la fin, pour en juger. Merci également à Olivier Thoumine, Laurent Reveillère et Jean-Christophe Olivo-Marin pour leurs conseils et leur bienveillance lors des comités de suivi de thèse.

Enfin, je suis profondément persuadé que le principal facteur qui détermine si une thèse a une chance de succès ou non, reste la qualité de l'encadrement. À ce titre, Macha, j'aurais bien du mal à écrire en quelques lignes ce que tu as pu m'apporter durant ces quatre années de thèse. Je ne saurais donc que te remercier d'avoir réussi le difficile équilibre, de me laisser toute latitude et en même temps d'avoir été là dans les moments où c'était nécessaire. J'ai beaucoup appris grâce à toi, scientifiquement et humainement. J'ai pu apprécier tes nombreuses qualités durant tout ce temps, aussi bien en tant que cheffe d'équipe, que directrice de recherche, directrice de thèse(s), mais aussi en tant que scientifique et surtout en tant que personne. Si j'arrive à n'avoir un jour que le quart de ces qualités, je considérerais déjà cela comme un succès. Merci.

Merci à l'équipe de "début de thèse" et même un peu plus tôt, Ben et les histoires dont toi seul a le secret, merci également pour les découvertes musicales et les soirées de guitare, Aurélien et Alexis, notre bureau des légendes locale, merci pour vos blagues et votre disponibilité, Hayssam, ta bienveillante faculté a poussé les gens à être meilleurs restera un mystère et en même temps une inspiration. La liste est longue et j'en oublie, Émeric, David, Jeff, Tristan, Marie, Galadriel, Raluca, Katarzyna, Emrah, June, Nicolas, Peter, Yvan, Damien, chacun a votre manière avez réussi l'exploit de rendre votre passage enrichissant. Antoine, je te décerne la médaille du meilleur stagiaire, mais ça tu le sais déjà. Justine, tu ne soupçonnes même pas à quel point tu as pu avoir un impact sur cette thèse, mais plus généralement, sur moi. Ta bienveillance, tes questionnements et tes remises en question ont toujours été pour moi une source profonde d'enrichissement et pour rien au monde je n'aurais changé de bureau malgré les coups de pied. Claire, ton énergie débordante aura été un bon coup de boost en milieu de thèse. Cédric, ta culture et ta jovialité auront été une bouffée d'air, promis je me suis (un peu) amélioré à LoL.

Et enfin un grand merci à l'équipe de "fin de thèse", Edgar ta capacité à creuser des sujets aussi variés que les réseaux de neurones ou les systèmes audios en passant par la cuisson d'une entrecôte aura été très enrichissante. Grigorii, merci pour ta façon d'avoir de nouvelles idées et de les mettre en pratique, merci pour tes questionnements, tes réponses, pour l'initiation au yoga et a bien d'autres sujets. Christer merci pour ta bonne humeur. Domitille, la force tranquille, j'admire tes capacités scientifiques autant que tes capacités sportives, une source d'inspiration là encore. Johanna, merci de ne pas m'avoir laissé seul au labo les soirs de cette fin de thèse, pense quand même à rentrer de temps en temps. Merci pour la gaieté et la vitalité que tu apportes. Ulysse, je ne saurais trop exprimer à quel point tu auras été d'un soutien incroyable sur cette fin de thèse. Merci donc pour ta patience, ton écoute, m'avoir redonné l'envie de reprendre la guitare, ta capacité à t'impliquer dans des sujets totalement variés comme la muscu, le basket, les pédales d'effets, les échecs ou encore les cryptomonnaies, c'est pour moi une grande inspiration à en faire autant.

Merci à mes amis, les Aurillacois, les Poitevins, merci aux glandus (ils se reconnaîtront.), quelque part votre capacité à tryhard sur des sujets variés aura été une motivation à entamer cette aventure. Ivan, je ne peux pas ne pas mentionner la personne qui a probablement eu le plus d'influence sur ma vie, ta curiosité naturelle, la découverte des sciences et vie juniors, de StarGate et de tellement d'autres choses, le tout avec la bienveillance qui te caractérise, restent pour moi un modèle que j'espère atteindre un jour.

Un grand merci à mes parents pour avoir très tôt éveillé ma curiosité scientifique en m'emmenant voir les "connaissance du monde", les mercredis à la bibliothèque, les livres de sciences et techniques, démonter, souvent sans retour, des appareils électroniques pour regarder dedans et enfin, à me montrer patiemment que la connaissance peut prendre des visages très variés, mais qu'elle est toujours utile, même en lisant un Astérix. Papa, merci pour tout ce que tu m'auras appris, ton infatigable capacité de travail est une inspiration dans les moments difficiles. Maman, merci également pour tout ce que tu m'as appris, ton sixième sens pour flairer un danger avant tout le monde, ou pour flairer quand j'ai besoin d'encouragements sont tellement précieux. Merci à vous de m'avoir poussé les 25 premières années et de m'avoir soutenu les 10 suivantes. Un grand merci à Caro, Steph et Armelle, je vous dois cette curiosité en grande partie, bien entendu uniquement pour essayer d'avoir toujours raison.

Bélinda, quelques lignes ne suffiront pas. Tu auras dû me supporter bien plus que je ne l'aurais souhaité et tu l'as fait, avec tant de bienveillance, merci pour ton soutien infaillible toutes ces années et dans bien des circonstances. Ta curiosité de chaque instant et ta créativité me sont indispensables pour continuer à apprendre. Ce fut précieux d'avoir quelqu'un qui croit en vous, même lorsque vous-même n'y croyez plus.

Table des matières

Remerciements	v
1 Introduction	1
1.1 Contexte scientifique	1
1.2 Microscopie à fluorescence	3
1.2.1 Principes généraux	4
1.2.2 Contraintes fondamentales	5
1.2.3 Modalités de microscopie de fluorescence	6
1.3 Acquisition et analyse de données de microscopie à fluorescence	12
1.3.1 smFISH : Marquage par fluorescence pour mesurer l'expression des gènes	12
1.3.2 Immunofluorescence	18
1.3.3 Marquage par fluorescence des noyaux cellulaire	20
1.4 Analyses spatiales et statistiques pour la microscopie de fluorescence	23
1.4.1 Fondements de l'analyse spatiale	24
1.4.2 Analyses spatiales du premier ordre	26
1.4.3 Analyses spatiales du second ordre	28
1.4.4 Colocalisation	32
1.5 Deep Learning	35
1.5.1 Fondamentaux du Deep Learning	35
1.5.2 Réseau de neurone à convolution (CNN)	39
1.5.3 Réseaux de neurones génératifs	44
1.6 Objectif global de la thèse	50
2 Rehaussement du signal dans les images smFISH pour une détection plus précise des ARNm	51
2.1 Contexte scientifique	51
2.2 Contribution scientifique	52
2.2.1 Réseau de neurones pour le rehaussement du signal de petits objets	52
2.2.2 Rehaussement du signal des ARNm dans les images smFISH	55

2.2.3	Amélioration de la précision de détection des spots d'ARNm	57
2.3	Discussion	59
2.4	Article	61
3	Rehaussement du signal des noyaux cellulaires dans l'imagerie à faible intensité	87
3.1	Contexte scientifique	87
3.2	Contribution scientifique	89
3.2.1	Segmentation des noyaux	89
3.2.2	Méthode d'évaluation	90
3.2.3	Modèle d'apprentissage non supervisé pour le rehaussement de signal	90
3.2.4	Jeux de données	94
3.2.5	Rehaussement de la saillance des noyaux	96
3.2.6	Amélioration de la segmentation des noyaux	97
3.3	Discussion	98
3.4	Article	100
4	Analyses de la distribution spatiale subcellulaire de l'ARNm et des protéines dans les données smFISH	113
4.1	Contexte scientifique	113
4.2	Contribution scientifique	116
4.2.1	Enrichissement localisé	116
4.2.2	Degré de clustering des ARNm et des protéines	119
4.2.3	Corrélation de localisation des ARNm et des protéines avec la position du centrosome	120
4.2.4	Colocalisation des distributions des ARNm et des protéines	123
4.3	Discussion	125
4.4	Article	128
5	Discussion et perspectives	159
5.1	Synthèse	159
5.2	Limites	161
5.3	Perspectives	164
5.4	Conclusion générale	167
A	DeepMeta : un réseau de neurone pour la détection de métastases pulmonaires dans les images de poumons de souris	169

B FluoSIM : Simulateur de diffusion de molécules fluorescente de type SMLM	183
C Prédiction de la localisation subcellulaire des ARNm : Apprentissage à partir de sources de données hétérogènes	199
Bibliographie	211

Table des figures

1.1	Contexte scientifique	2
1.2	Déplacement de Stokes	4
1.3	PSF	6
1.4	Microscopie de fluorescence	7
1.5	Comparaison d'images de tubuline	9
1.6	SMLM	10
1.7	Hybridation	13
1.8	Autofluorescence	14
1.9	Pipeline traitement d'image	16
1.10	Laplacien de Gaussien	17
1.11	Immunofluorescence	19
1.12	Immunofluorescence exemples	20
1.13	Grille de lecture spatiale	25
1.14	Densité	28
1.15	Fonction K et G	31
1.16	Ripley K	32
1.17	Colocalisation vecteur	33
1.18	Neurone formel	36
1.19	Convolution	40
1.20	Convolution à trous	41
1.21	Pooling	42
1.22	Convolution blocks	43
1.23	GAN	45
1.24	CycleGAN	48
2.1	Détection de petits objets	53
2.2	Architecture DeepSpot	54
2.3	DeepSpot évaluation	56
2.4	DeepSpot exemples	57

3.1	Styles d'images pour CycleGAN	88
3.2	Évaluation SaliencyNet	91
3.3	Architecture SaliencyNet	92
3.4	SNR noyaux	96
3.5	Résultats SaliencyNet	98
4.1	DypFISH	115
4.2	Enrichissement périphérique	117
4.3	Densité locale muscle	118
4.4	Degré de clustering	120
4.5	Degré de clustering par ARNm	121
4.6	Enrichissement par rapport au centrosome	122
4.7	MTOC polarity index	122
4.8	Colocalisation méthode	123
4.9	Colocalisation score	124
4.10	Colocalisation timepoints	125

Liste des tableaux

2.1	Performance DeepSpot vs deepBlink	58
3.1	Jeux de données SaliencyNet	96
3.2	IoU SaliencyNet	97
4.1	Analyses spatiales DypFISH	127

Chapitre 1

Introduction

1.1 Contexte scientifique

L'analyse quantifiée des données d'imagerie microscopique par des méthodes informatiques, mathématiques et d'intelligence artificielle est désormais un outil essentiel pour la compréhension précise des processus cellulaires. Les techniques de microscopie moderne offrent aujourd'hui aux biologistes un moyen robuste et quantitatif d'acquisition de données de l'imagerie cellulaire à l'échelle moléculaire, et ce, à haut-débit. Ainsi, il est possible d'obtenir des informations allant de la morphologie des cellules et des noyaux à la quantification des molécules qui y sont présentes, telles que les ARN messagers et les protéines de façon localisée. Ces informations sont de plus en plus utilisées pour déchiffrer les mécanismes moléculaires impliqués dans le fonctionnement cellulaire, mais aussi dans les dysfonctionnements pathologiques.

Grâce aux progrès apportés par l'automatisation des microscopes et le développement des sondes fluorescentes, la microscopie optique est désormais devenue une approche complémentaire des technologies "omiques", telles que la transcriptomique et la protéomique, en rendant possible la quantification spatio-temporelle des processus biologiques à l'échelle de la cellule. Ainsi, le screening à haut contenu (HCS) est une approche d'imagerie qui permet l'étude de phénotypes cellulaires couvrant les phénomènes globaux tels que les modifications morphologiques de la cellule entière et de ses composants, ou encore la détection de l'altération des processus cellulaires grâce à la quantification et la localisation de molécules (voir le schéma de la figure 1.1). L'analyse automatisée de ces données du phénotype morphologique et du phénotype de localisation et leur quantification rend possible les stratégies de criblage cellulaire à haut débit pour des applications telles que par exemple le développement de nouveaux médicaments.

L'émergence de techniques d'acquisition de données de phénotypage cellulaire à haut-débit en imagerie microscopique rend cruciale l'automatisation des analyses. En effet, les

tâches classiques telles que la segmentation de cellules et de noyaux, la détection, la quantification et la localisation d'objets ou encore leur tracking sont consommatrices de temps et souvent source d'erreur si elles sont réalisées manuellement. Naturellement, un grand nombre de méthodes et d'outils logiciels ont été développés pour répondre à ce besoin et proposent des solutions automatisées ou semi-automatisées, couvrant des applications très variées, allant du simple seuillage aux modèles de live-tracking de molécules. Un nouveau challenge dans l'application de ces méthodes se présente désormais aux utilisateurs (chercheurs biologistes), celui du paramétrage optimal de ces algorithmes pour obtenir des résultats escomptés.

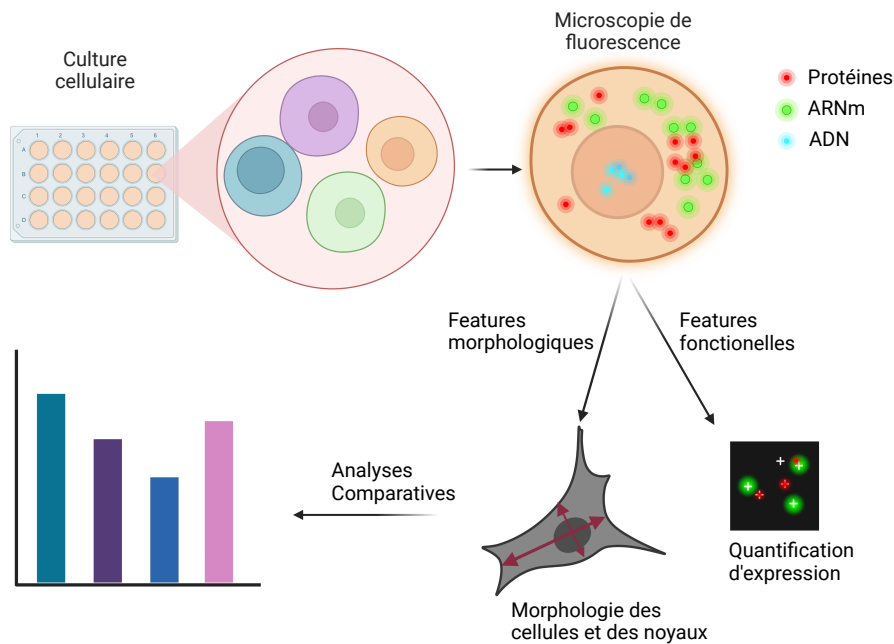


FIGURE 1.1 – Pipeline de microscopie cellulaire de phénotypage à haut débit. Les cellules sont cultivées dans des plaques contenant de multiples puits permettant de créer des conditions différentes, avant d'être marquées par fluorescence, cette fluorescence est ensuite imagée par plaque par le microscope, permettant d'extraire des informations morphologiques et des informations sur la localisation des molécules. Ces informations serviront enfin à l'analyse comparative et quantifiée entre les conditions d'intérêt.

Le développement de l'analyse d'images par apprentissage machine permet de créer de nouvelles méthodes capables d'apprendre une partie de la diversité des données et nécessitant donc moins de paramétrage. De ce fait, l'apprentissage profond fournit une boîte à outils puissante pour répondre au besoin de l'automatisation. Les grandes tâches que l'on cherche à automatiser grâce aux méthodes d'intelligence artificielle peuvent être classées en trois catégories principales : (1) la restauration d'images, où l'image en entrée

est transformée dans une image améliorée en sortie, (2) la segmentation, où l'on cherche à déterminer les régions ou bien les objets d'intérêt et (3) la quantification, où les objets sont localisés, classifiés, comptés ou tracés.

Dans cette thèse, nous introduirons les méthodes utilisées au cours des différents chapitres, portant sur les principes généraux, l'acquisition et l'analyse des données de microscopie de fluorescence ainsi que sur les analyses spatiales de ces données et enfin sur les méthodes de Deep Learning. Nous nous intéresserons dans un premier temps à ce comment l'apprentissage par les réseaux de neurones peut être utilisé pour la restauration d'images en microscopie cellulaire à haut débit et nous étudierons l'impact de cette restauration sur l'automatisation de la segmentation de noyaux et de la localisation des spots d'ARN messagers. Nous couvrirons ainsi les deux types d'informations issues de l'acquisition d'images cellulaires à haut débit, celui de type morphologique et celui de type fonctionnel. Dans un deuxième temps, nous présenterons comment les informations ainsi obtenues peuvent être analysées de façon statistiquement rigoureuse pour l'analyse et la comparaison de phénotypes d'intérêt.

1.2 Microscopie à fluorescence

La microscopie de fluorescence s'est imposée comme une technique fondamentale pour l'imagerie à haut débit et à fort contraste. Elle utilise des molécules fluorescentes pour mettre en évidence des régions cellulaires clés. Les images obtenues par microscopie cellulaire à haut débit peuvent provenir d'un large éventail de techniques de microscope, chacune exploitant à sa manière les principes de l'optique.

Dans cette thèse, nous nous concentrerons sur l'utilisation de l'imagerie de fluorescence pour mettre en évidence la localisation des ARNm, des protéines, mais également des structures cellulaires comme les noyaux. Dans cette section, nous détaillerons les principes des techniques utilisées pour ce faire, telles que l'immunofluorescence, l'hybridation in situ ou encore la microscopie de molécule unique. Nous nous intéresserons dans un premier temps aux fondamentaux de la microscopie de fluorescence avant d'aborder ces diverses techniques d'acquisition des données.

Dans la suite de la section, après avoir couvert les principes généraux de la microscopie de fluorescence, nous présenterons les méthodes d'acquisition suivantes : (i) la microscopie confocale pour imager les ARNm et les protéines utilisées dans les chapitres 2 et 4, (ii) les différentes variantes de microscopie à molécule unique utilisées dans les chapitres 3 et 5 pour imager des noyaux et simuler la diffusion des protéines dans les cellules.

1.2.1 Principes généraux

Le terme de fluorescence est introduit pour la première fois en 1852 par Sir George Gabriel Stokes. S'appuyant sur les travaux préliminaires de John Frederick William Herschel, il établit que les rayons lumineux produits par fluorescence sont systématiquement moins réfringibles que les rayons qui les induisent et il publie cette découverte en 1852. Ce décalage entre la longueur d'onde d'absorption et la longueur d'onde de fluorescence sera appelé "déplacement de Stokes" (voir figure 1.2). Il se calcule en faisant la différence entre le pic d'émission et le pic d'absorption d'un fluorophore. Plus le décalage est important, plus il est aisé de différencier les longueurs d'ondes associées à l'excitation et celles associées à l'émission. Ce principe sera la fondation du développement de la microscopie de fluorescence dont le principe de base repose sur le marquage d'échantillons par des molécules fluorescentes, les fluorophores.

Lorsqu'un échantillon est soumis à un laser dans une bande spécifique de longueur d'onde, ces longueurs d'onde sont absorbées par les fluorophores qui entrent dans un état d'excitation électronique. En quelques nanosecondes, les fluorophores vont retourner vers leur état électronique fondamental et cette transition peut alors s'accompagner d'une émission lumineuse, la fluorescence. L'énergie du photon émis est alors inférieure à celle du photon absorbé, ce qui se traduit en un décalage de longueur d'onde, le déplacement de Stokes. Grâce à des filtres optiques tels que les miroirs dichroïques, il est alors possible de trier les longueurs d'onde spécifiques de la lumière émise par ces fluorophores [93].

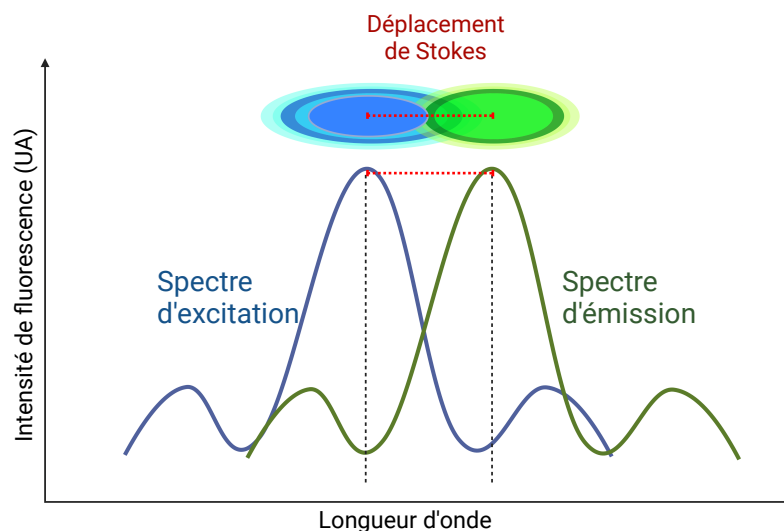


FIGURE 1.2 – Déplacement de Stokes, le spectre d'émission présente une longueur d'onde plus élevée que le spectre d'absorption.

Contrairement à la microscopie classique, où l'échantillon est observé avec la lumière incidente ou transmise, la différence entre les longueurs d'onde d'excitation et d'émission permet de bloquer la lumière incidente et d'obtenir une image présentant un contraste élevé entre l'échantillon et le fond. L'un des principaux avantages de la microscopie à fluorescence est l'augmentation du signal des structures et des molécules marquées, par rapport au fond sombre.

1.2.2 Contraintes fondamentales

La microscopie de fluorescence présente cependant des contraintes fondamentales imposées par la diffraction et dues au caractère ondulatoire de la lumière, qui sont susceptibles d'impacter la qualité des images acquises.

Limite de résolution

Premièrement, la résolution optique, qui se définit par la capacité d'un système à distinguer l'image de deux fluorophores séparés par une distance d . Le critère de Rayleigh permet de définir la distance minimale entre deux points afin qu'ils soient différentiables, en d'autres termes, de prédire le pouvoir séparateur théorique d'un système optique : $d_R = \lambda/2NA$ où d_R est la distance de Rayleigh, λ est la longueur d'onde de la lumière d'excitation, généralement un laser, tandis que NA (*Numerical Aperture*) correspond à l'ouverture numérique de l'objectif, avec $NA = n * \sin \alpha$ où n est l'indice de réfraction du milieu entre l'objet et l'objectif ($n = 1$ pour l'air et $n = 1,51$ pour l'huile d'immersion), et α est le demi-angle d'ouverture de l'objectif [93] comme détaillé sur la figure 1.3.

Deux fluorophores sont alors résolus si $d > d_R$, non résolus si $d < d_R$ et ils sont dits à la limite de résolution lorsque $d = d_R$.

Fonction d'étalement du point

Deuxièmement, un autre élément perturbateur peut impacter la qualité des images. Il s'agit de la fonction d'étalement du point, plus communément appelée PSF pour *Point Spread Function*. Il s'agit de la réponse impulsionnelle à une source lumineuse ponctuelle qui, dans un système optique parfait, correspond au profil de diffraction de l'image. En microscopie, la PSF définit la relation entre le fluorophore et son image obtenue par le microscope, l'image consiste alors en une somme de toutes les PSF des sources lumineuses ponctuelles observées [197]. La PSF peut être modélisée par une tâche d'Airy, plus la tâche d'Airy est petite, meilleure est la résolution et meilleure sera la qualité de l'image obtenue. Plus deux objets sont proches et leurs PSF sont étalées, plus leurs PSF vont se

superposer et empêcher la résolution des sources ponctuelles.

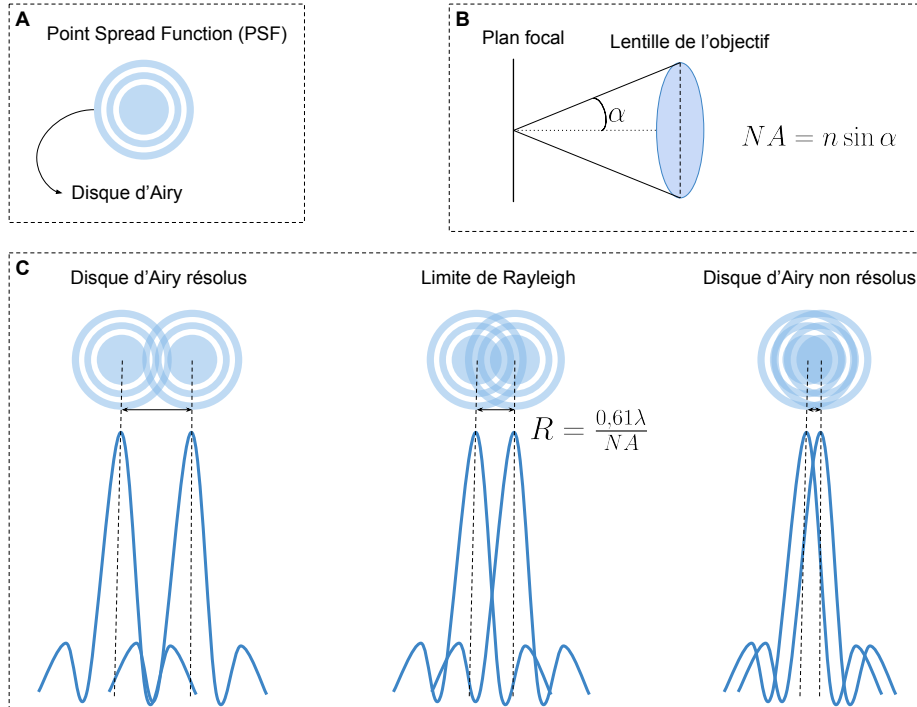


FIGURE 1.3 – (A) La fonction d'étalement du point (PSF) se présente sous forme d'un disque d'Airy dans le cas d'un système optique parfait. (B) Mesure de l'ouverture numérique d'un système optique (objectif). (C) Le critère de Rayleigh permet de définir si deux sources ponctuelles lumineuses seront résolues par le système optique utilisé pour imager. Adapté de [38].

De ces deux propriétés optiques (limite de résolution et PSF), il découle que la distance de résolution minimale ne peut être inférieure à la moitié de la longueur d'onde de la lumière d'excitation. Par conséquent, pour une longueur d'onde de 400 nm, qui est proche de la limite de ce qui est utilisable en imagerie biologique, la distance de résolution maximale approximative est de 200 nm, bien qu'en pratique, les aberrations optiques limitent la résolution du champ large à plus de 250 nm [27].

1.2.3 Modalités de microscopie de fluorescence

Suite aux travaux de Stokes sur la fluorescence, il a fallu attendre 1904 pour la création du premier microscope à fluorescence [38, 73]. Ce n'est que bien plus tard avec la découverte de la GFP isolée à partir de la méduse *Aequorea victoria* en 1962, puis son séquençage en 1992, le développement de variant en 1994 et leur utilisation en tant que

marqueurs, que la microscopie de fluorescence permettra la visualisation simultanée de plusieurs composants cellulaires dont les ARNm et les protéines.

Diverses modalités de microscopie à fluorescence, que nous allons présenter dans la suite de ce chapitre, ont fait leur apparition au cours des dernières décennies à la faveur d'avancées scientifiques et technologiques à la fois sur les marqueurs fluorescents et sur les systèmes d'acquisition.

Microscope confocal

On peut citer en premier lieu la microscopie confocale, développée et brevetée par Marvin Minsky en 1957, [40] qui utilise un éclairage ponctuel via un sténopé (pinhole), placé dans un plan focal conjugué à celui de l'échantillon et du sténopé du détecteur (d'où l'appellation confocal) et qui ne laisse passer que la lumière émise par le point focal souhaité, toute lumière en dehors de ce plan est bloquée, ce qui évite de recueillir la lumière réfléchie ou fluorescente provenant de l'extérieur du point focal, donnant un avantage en termes de contraste et de bruit de fond. Le microscope confocal permet de réaliser des sections optiques dans la profondeur du tissu ou de la cellule. La technique est cependant plus coûteuse et plus complexe à utiliser qu'un microscope classique, aussi appelé microscope à épifluorescence [27, 123].

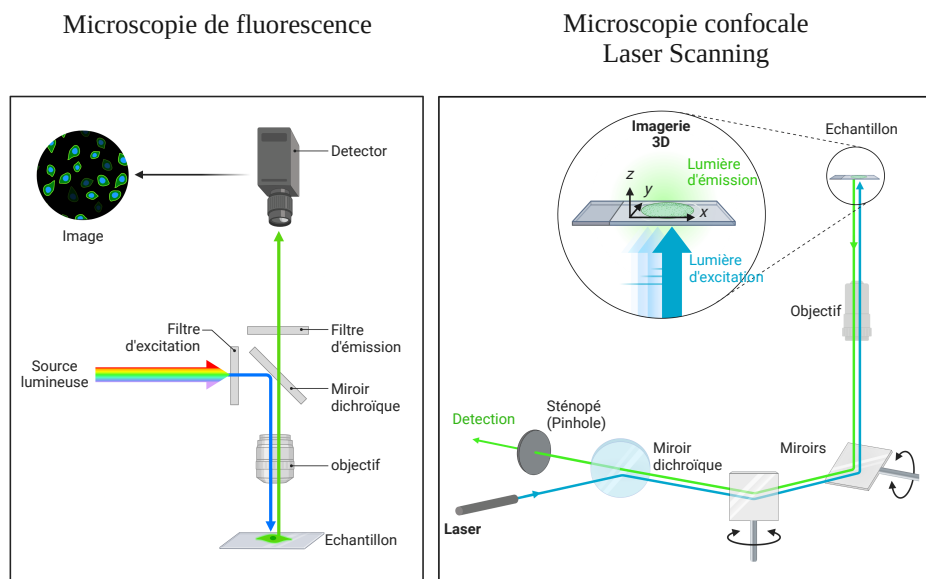


FIGURE 1.4 – Comparaison des systèmes d'acquisition d'image de fluorescence. Le panneau de gauche présente la microscopie plein champ, le panneau de droite présente la microscopie confocale de type Laser Scanning.

Light-sheet

D'autres modalités, comme la microscopie de fluorescence à feuille de lumière (LSFM, pour Light-Sheet Fluorescence Microscopy) [163], utilisent une fine feuille de lumière pour exciter uniquement les fluorophores situés dans le volume focal désiré. Les microscopes à feuille de lumière (LSM) offrent une bonne résolution axiale, limitent le photoblanchiment et la phototoxicité à la fraction de l'échantillon illuminée et peuvent imager à très haut débit [164]. Les LSM sont utilisés pour des analyses dans la profondeur de grands échantillons et pour les observations tridimensionnelles (3D). La microscopie à feuille de lumière se décline sous plusieurs formes, nous citerons ici deux techniques SPIM et soSPIM, cette dernière étant utilisée au cours de ces travaux pour l'imagerie de noyaux cellulaire.

La microscopie à illumination sélective SPIM, pour Single Plane Illumination Microscope, utilise une feuille de lumière pour éclairer l'échantillon par le côté [69]. Ce type de microscopie présente un certain nombre d'avantages, comme des acquisitions rapides et une faible phototoxicité en éclairant sélectivement un seul plan de l'échantillon tout en capturant le signal de fluorescence à travers un second objectif positionné perpendiculairement à la feuille de lumière. Une variante développée par les auteurs de [47] ne nécessite, elle, qu'un seul objectif, appelée soSPIM, pour Single Objective SPIM. Cette technique, plus polyvalente, permet l'imagerie 3D de structures biologiques allant de la cellule unique à l'embryon complet. Elle permet également un plus grand champ de vision et d'imager des molécules uniques sur une plus grande surface.

SMLM

Pour circonvenir aux contraintes fondamentales et obtenir une résolution suffisante pour localiser avec précision les molécules, un certain nombre de modalités de microscopie se concentrent sur la localisation de molécules uniques et sont regroupées sous le terme générique de Single Molecule Localization Microscopy, SMLM. Il s'agit d'une famille de techniques d'imagerie qui améliorent considérablement la résolution spatiale par rapport aux techniques de microscopie de type hybridation in situ qui sont, elles, limitées par la diffraction (voir l'exemple de la figure 1.5).

La SMLM est usuellement rangée sous la catégorie des techniques de microscopie de super-résolution aux côtés de la Stimulated Emission Depletion (STED) [61] et de la structured illumination microscopy (SIM) [150] (voir figure 1.6), qui toutes, permettent d'obtenir des images de structures biologiques à l'échelle nanométrique. La microscopie SMLM ayant un ordre de résolution allant de 10 à 20 nm, cela permet d'imager des molécules distinctes. Contrairement à l'imagerie microscopique standard qui produit des images sous forme de matrices de pixels (2D) ou de voxels (3D), la SMLM génère un film d'émissions de fluorophores individuels à partir duquel on peut extraire une liste de

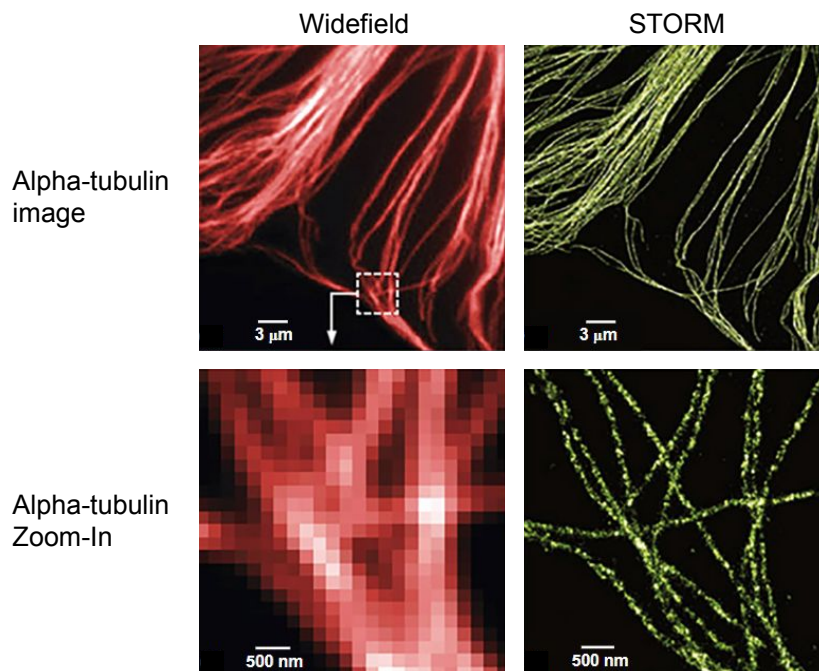


FIGURE 1.5 – Illustration de la différence de résolution entre la microscopie de plein champ à gauche, et la microscopie SMLM de type STORM à droite. Adapté de [173].

coordonnées, contenant au moins les positions temporelles (x , y) des émetteurs individuels, souvent complétée par des informations sur l'incertitude de localisation, l'intensité de l'émetteur ainsi que sur la position axiale (z) [29]. Ceci permet deux utilisations principales, l'imagerie à super-résolution (SR) et le suivi de particules uniques (SPT) [110].

La SMLM est basée sur le fait que les coordonnées spatiales des fluorophores peuvent être déterminées avec une grande précision si leurs PSF ne se chevauchent pas. Pour ce faire, l'imagerie séquentielle de sous-ensembles aléatoires de fluorophores suffisamment distants pour que leurs PSF soient bien différenciés est commune à la plupart des techniques de SMLM. Une fois les images acquises, le calcul de la position des fluorophores se fait par fitting gaussien. Ces techniques sont possibles en grande partie grâce au développement de fluorophores ayant des propriétés permettant de moduler la fluorescence par photoblanchiment ou par photoactivation, puis de suivre l'évolution de leur intensité au cours du temps. En induisant la fluorescence de façon stochastique (blinking), il est possible de localiser individuellement et de façon successive les molécules émettrices. Plusieurs milliers d'images partielles successives sont ainsi acquises avant de procéder à la reconstruction de l'image globale [79].

Au sein des SMLM, on peut mentionner le Single Particle Tracking (SPT) qui vise

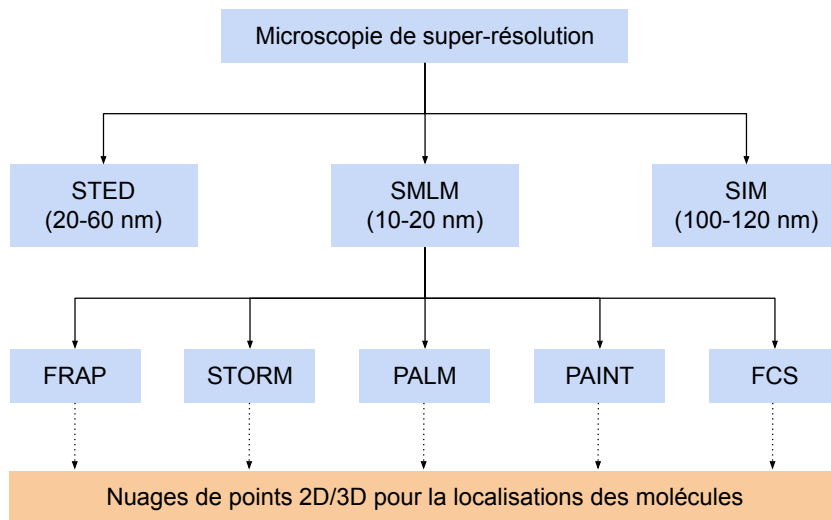


FIGURE 1.6 – Classification des méthodes de microscopie de super-résolution, est indiqué en nanomètre l’ordre de grandeur de la résolution atteignable par chaque catégorie de systèmes. Adapté de [79].

à observer le mouvement des molécules individuelles puis à déterminer leur trajectoire, leur vitesse ainsi que leur coefficient de diffusion. Existente également les méthodes par photoactivation et photoblanchiment, à savoir la PhotoActivation of Fluorescence (PAF) et la Fluorescence Recovery After Photobleaching (FRAP) qui consiste à photoblanchir une zone d’intérêt puis à suivre la récupération de la fluorescence dans cette même zone au cours du temps ; la Fluorescence Correlation Spectroscopy (FCS), qui analyse la dynamique moléculaire en corrélant les fluctuations d’intensité. Des approches plus récentes basées sur la localisation de molécules uniques, telles que la PhotoActivated Localization Microscopy (PALM) [14] la Stochastic Optical Reconstruction Microscopy (STORM) [143] et la Point Accumulation In Nanoscopic Topography (PAINT) [34]. Ces techniques diffèrent essentiellement par la manière dont l’activation aléatoire d’un sous-ensemble de molécules est induite [173].

Un assez grand nombre de méthodes d’analyse d’image peut être appliqué aux images acquises par SMLM. Ces dernières années les méthodes basées sur le deep-learning ont pris le pas, on peut citer Deep-STORM [120] et ANNA-PALM [126], DeepLoco [131] ou encore DECODE [160].

Imagerie "Live-cell"

Pour mieux comprendre les processus biologiques, capturer les mouvements et les comportements des cellules vivantes, l'imagerie des cellules vivantes (live-cell) s'est imposé comme une méthode de choix pour comprendre les processus cellulaires complexes, étudier les réponses cellulaires à la perturbation de leur environnement et à la modulation des voies cellulaires ainsi que pour comprendre les comportements des cellules individuelles ou des populations de cellules.

Pour acquérir des images de cellules vivantes, le principal défi est de minimiser les dommages causés par la lumière tout en conservant un rapport signal/bruit suffisant. [43], car une exposition prolongée à la lumière sur des cellules vivantes peut entraîner une phototoxicité, un photoblanchiment et, par conséquent, la mort cellulaire [165]. Par ailleurs, il est essentiel de maintenir les cellules dans leurs conditions de croissance optimales et d'éviter les éventuelles réactions de stress. À cette fin, de nombreux paramètres doivent être optimisés afin de trouver un compromis entre la quantité de lumière suffisante pour imager et la quantité de lumière évitable pour limiter la dégradation des échantillons vivants.

Pour limiter la phototoxicité, les méthodes de microscopie à faible luminosité (low-light) font l'objet d'une attention accrue, car elles permettent l'imagerie de cellules sensibles à la lumière sans phototoxicité et sans photoblanchiment des fluorophores. Cependant, cela se fait au détriment du contraste des images acquises [80].

Les images résultant de ces diverses méthodes d'acquisition sont des matrices discrètes reflétant l'intensité lumineuse de l'échantillon capturée par le microscope. Elles peuvent être bidimensionnelles (x, y) ou tridimensionnelles (x, y, z) , z étant alors la dimension dans la profondeur de l'échantillon étudié. Dans ce dernier cas, on parle couramment de "stack" d'images, c'est-à-dire une pile d'images bidimensionnelles. Les images sont généralement encodées sur 8, 10 ou 12 bits, soit respectivement des intensités de pixels de 256, 1024 et 4096 niveaux de gris possibles [159]. Si plusieurs signaux provenant de fluorophores de couleurs différentes sont acquis en même temps, par exemple en 8 bits, l'image totale contiendra un canal pour chaque couleur.

1.3 Acquisition et analyse de données de microscopie à fluorescence

Dans cette section, nous introduirons les techniques qui ont servi à produire les données utilisées dans cette thèse.

Premièrement, nous présentons l'hybridation in situ à molécule unique (smFISH) servant à marquer par fluorescence les ARN messagers pour quantifier l'expression de gènes dans une cellule. En effet, dans le chapitre 2 nous nous intéresserons à la façon de dépasser les limitations de cette technique en mettant en place une méthode de rehaussement du signal de fluorescence lié aux ARNm pour permettre une meilleure détection. Dans le chapitre 4 nous en tirerons profit pour l'analyse spatiale de l'expression des gènes.

Deuxièmement, nous introduirons l'immunofluorescence (IF), qui permet de marquer par fluorescence des protéines cibles afin de les quantifier. Ainsi la conjonction de données smFISH et IF nous permettra dans le chapitre 4 d'appliquer des analyses de colocalisation spatiales ARN-protéine en vue de caractériser des processus cellulaires.

Pour ces deux modalités, nous expliquerons comment sont acquises les images, quelles sont leurs particularités et quels sont les algorithmes couramment utilisés pour en extraire de l'information.

Enfin, nous aborderons le marquage par fluorescence des noyaux cellulaires ainsi que les méthodes utilisées pour segmenter les images résultantes.

1.3.1 smFISH : Marquage par fluorescence pour mesurer l'expression des gènes

L'hybridation in situ fluorescente à molécule unique (smFISH) s'est imposé comme la méthode préférentielle pour mesurer l'expression des gènes grâce à sa capacité à détecter des transcrits individuels au sein de cellules uniques et donc permettre des analyses quantitatives [177]. Elle repose sur le principe d'hybridation des acides nucléiques, selon lequel deux brins complémentaires d'acides nucléiques s'hybrident facilement l'un à l'autre dans les conditions appropriées pour former un duplex ARN-ARN ou ADN-ADN, appelé hybride. La smFISH utilise de nombreuses sondes oligonucléotidiques courtes portant un fluorophore, pour cibler différentes régions du même transcrit ARNm (voir figure 1.7) et ainsi atteindre la résolution suffisante pour imager des molécules uniques [135].

Différentes modalités de smFISH ont été développées, toutes néanmoins se basent sur un processus commun [44, 136, 194]. Pour certains types d'échantillons, la première étape est la décontraction, un signal smFISH peut être obscurci ou rendu inutilisable s'il est masqué par une morphologie contractée. Suite à cela, viennent les étapes suivantes.

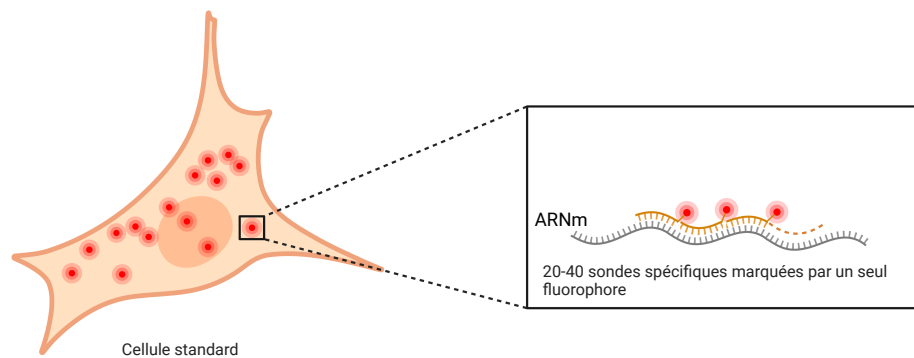


FIGURE 1.7 – smFISH : l’hybridation de multiples sondes fluorescente sur un ARNm cible permet de visualiser la répartition d’un type d’ARNm dans la cellule.

- **La fixation de l’échantillon** Cette étape va permettre de figer la cellule ou le tissu dans l’état souhaité.
- **La perméabilisation** Les parois membranaires des échantillons sont généralement perméabilisées à l’aide d’un détergent pour permettre une bonne pénétration des réactifs d’hybridation. C’est une étape cruciale ou l’intégrité de l’échantillon peut être dégradée. Dans le cas de DNA-FISH cette étape est suivie par une étape de dénaturation de l’ADN afin de permettre l’hybridation.
- **L’hybridation entre les sondes et leurs cibles** Deux méthodes de marquage existent, le marquage direct et le marquage indirect. Pour le marquage indirect, les sondes sont marquées avec des nucléotides modifiés qui contiennent un haptène, alors que le marquage direct utilise des nucléotides qui ont été directement modifiés pour contenir un fluorophore [194]. Les sondes fluorescentes ont usuellement une taille comprise entre 20 et 50 nucléotides et embarquent un nombre variable de fluorophores [135]. Afin d’optimiser la sensibilité et la spécificité des marqueurs, les plus récentes variations utilisent de courtes séquences avec peu de fluorophores, mais en plus grand nombre, ce qui permet une détection et une quantification plus robuste des transcrits [55, 44, 25].
- **Le lavage post-hybridation** L’objectif est de séparer les hybrides non spécifiques et d’éliminer les sondes non liées au tissu afin de minimiser le bruit de fond [176].

Une étape finale de clarification des tissus peut être réalisée avant d’imager, pour éviter la diffusion de la lumière dans le tissu [138].

La technique FISH présente plusieurs limites. D’une part, comme les cellules sont fixées, smFISH ne peut pas être utilisée pour l’analyse temporelle de l’expression des gènes dans la même cellule (live-imaging) [55]. D’autre part, le faible nombre de marqueurs

fluorescents séparables (généralement environ cinq canaux dans le spectre visible) limite considérablement le nombre d'espèces d'ARNm qui peuvent être sondées simultanément, généralement de 1 à 4 gènes par expérience [24]. Par ailleurs, comme tous les ARNm que l'on souhaite imager doivent être illuminés en même temps, cela constitue une limite dans la densité possible d'ARNm, car chaque molécule doit être distincte de ses molécules voisines pour être détectée de manière fiable.

Images issues de smFISH

Les images acquises en smFISH sont des images en niveaux de gris, avec un signal blanc sur fond noir. Le signal est constitué par les fluorophores qui se sont hybridés avec un ARNm cible (voir figure 1.7) et ils se visualisent sous forme d'un point, car la plupart des ARNm sont plus petits que la limite de diffraction. Dans l'image, chaque ARNm est représenté par sa PSF, occupant quelques pixels sur l'image. L'arrière-plan, lui, se présente sous la forme d'un fond sombre, plus ou moins bruité, d'une part, par la diffraction de la lumière émise par les fluorophores hybridés dans les diverses structures cellulaires ou par l'autofluorescence de certains composants cellulaires tels que les mitochondries [118], d'autre part par l'hybridation non spécifique des fluorophores sur d'autres ARNm ou protéines non ciblés [117] tel que présenté sur la figure 1.8.

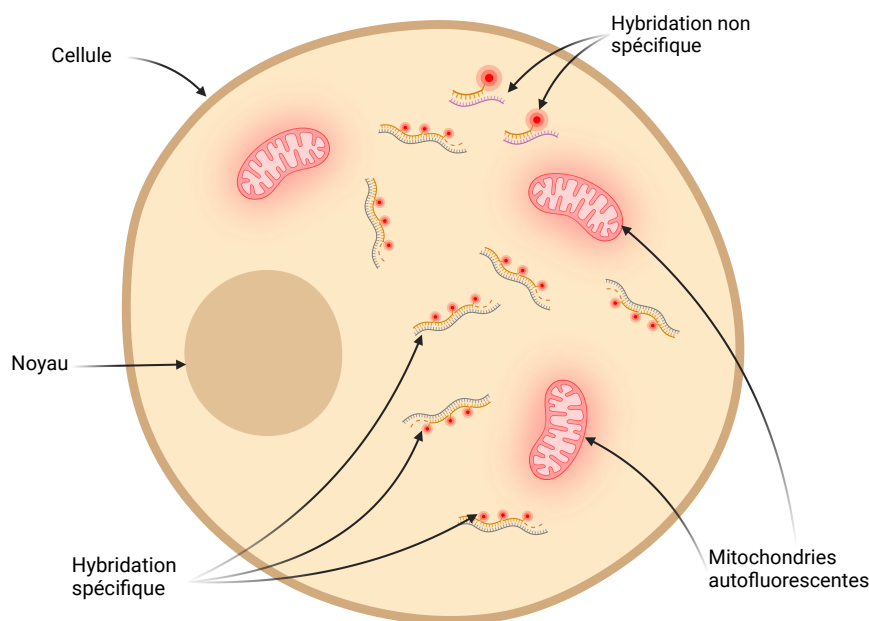


FIGURE 1.8 – Lors de l'acquisition des images, des photons sont émis par les organites autofluorescent tels que les mitochondries, mais également par des sondes fluorescentes hybridées de manière non spécifique sur des ARNm non ciblés.

La définition des images résultantes dépend du microscope, plein champ ou confocal, ainsi que du capteur qu'il embarque. Les images de microscopie confocale issues de photomultiplicateurs peuvent avoir des pixels de très petite taille, environ 10 fois plus petits que ceux des caméras CCD [20]. La plupart des caméras CCD de qualité scientifique ont des pixels qui ont généralement 6 à $7\mu\text{m}$ de côté. Ceci peut impacter la qualité finale de l'image, car plus les pixels sont fins, meilleure sera la définition. Cependant, cela implique que moins de photons émis sont captés par pixel et donc un rapport signal sur bruit plus bas [159]. Dans le cas des images 3D, la dimension z des voxels est généralement bien plus élevée, car elle est définie par la capacité du microscope à faire des coupes optiques dans la profondeur de l'échantillon [15].

Challenges pour l'analyse d'images issues de microscopie smFISH

Nous présentons ici les principales caractéristiques des images acquises par smFISH ainsi que les difficultés qu'elles présentent lors de leur analyse.

Dans la plupart des cas, la fluorescence résultant des sondes liées à chaque copie d'un ARNm dépasse largement le bruit de fond. Cependant, en fonction de l'échantillon et des conditions expérimentales, le rapport signal sur bruit peut être bas, ce qui rend difficile un certain nombre d'applications telles que le multiplexage ou la détection en aval de petits ARNm trop courts pour accueillir suffisamment de fluorophores. Enfin, le bruit de fond est généralement plus prononcé dans les tissus complexes [117].

Outre la résolution spatiale limitée par la diffraction, une autre source importante d'aberrations introduites dans le processus d'imagerie est le bruit photonique intrinsèque, qui résulte de la nature aléatoire de l'émission des photons. Le bruit des photons (bruit de Poisson), qui dépend de l'électronique du détecteur, peut être réduit uniquement en augmentant l'intensité lumineuse ou le temps d'exposition. Cependant, l'augmentation de l'intensité lumineuse dans le but d'améliorer la qualité de l'image entraîne un affaiblissement permanent du signal fluorescent en raison du photoblanchiment [191].

Enfin, pour la plupart des microscopes, l'illumination ainsi que l'efficacité de la réception du signal n'est pas uniforme sur tout le champ de vision, c'est le cas par exemple du vignettage qui induit une perte de signal fluorescent sur les bords de l'image. Ceci implique une efficacité de la détection également non uniforme sur les différentes parties de l'image acquise [17].

La difficulté dans l'analyse de grands volumes d'images smFISH réside alors dans leur variabilité en termes de bruit de fond, d'intensité et de qualité de l'image en général et des spots en particulier [39].

Méthodes de détection de spots dans les images smFISH

L'analyse quantitative des données smFISH implique la détection de nombreuses sources ponctuelles de lumière, usuellement dénommées comme “spots”. En vision par ordinateur (computer vision), les spots se définissent assez simplement, comme des régions de l'image nettement plus claires ou plus sombres que les pixels alentour [82] et comportant un extremum local [95]. Dans le cas de la fluorescence, son étendue spatiale peut se limiter à la figure de diffraction de la source lumineuse.

Le processus courant de la détection de spots dans les images de microscopie consiste à débiter par une étape de débruitage, puis une étape de rehaussement du signal pour terminer par l'étape de seuillage qui permet à proprement parler de détecter les spots [158], comme illustré sur la figure 1.9.

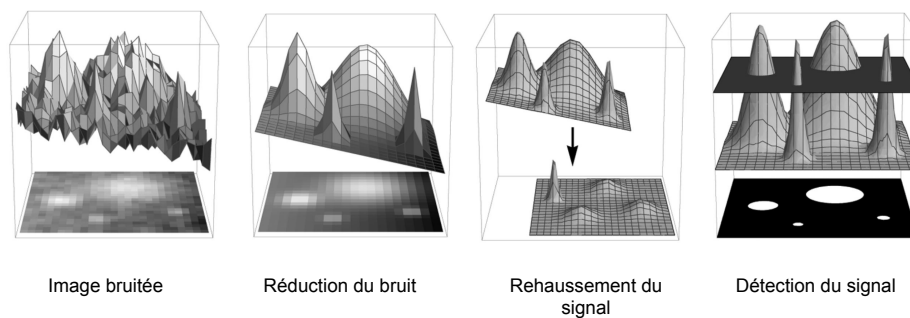


FIGURE 1.9 – Processus classique de détection de spots dans des images de microscopie. L'image est d'abord débruitée pour enlever le bruit de photon, l'image résultante est filtrée pour rehausser le signal d'intérêt avant d'être seuillée pour extraire les informations de localisation du signal. Adapté de [158].

L'étape de rehaussement du signal peut faire appel à des algorithmes classiques (filtres, morphologie, soustraction des bruits de fond, ondelettes), ou bien des algorithmes de machine learning, supervisés ou non.

La majorité des outils de détection de spots en smFISH utilise l'opérateur mathématique qu'est le Laplacien de Gaussien (LoG), aussi appelé filtre de Marr ou chapeau mexicain [135, 52, 121, 145, 5, 82, 72]. Par rapport à d'autres méthodes d'acquisition d'images, la méthode de déconvolution du signal par la PSF théorique ou réelle du microscope est peu utilisée pour le traitement des images smFISH, car elles ont un fort bruit de fond prépondérant par rapport à la PSF. Le filtre Laplacien est un filtre détecteur de bord utilisé pour calculer les dérivées secondes d'une image, mesurant la vitesse à laquelle les dérivées premières changent. Cela détermine si un changement dans les valeurs de pixels adjacents provient d'un bord ou d'une progression continue [82]. Le filtre Laplacien étant

très sensible au bruit, on applique au préalable un filtre Gaussien pour lisser le bruit. Plus formellement, étant donné un pixel (x, y) , le Laplacien, dans un système de coordonnées cartésiennes bidimensionnelles, se définit par la somme des dérivées partielles secondes d'une divergence de gradient a deux dimensions :

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

$$\frac{\partial^2 f}{\partial x^2} = f(x + 1) + f(x - 1) - 2f(x) \rightarrow [1; -2; 1]$$

$$\frac{\partial^2 f}{\partial y^2} = f(y + 1) + f(y - 1) - 2f(y) \rightarrow \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Dans le cas d'une image, il n'est pas possible de calculer directement la dérivée seconde, car les pixels d'une image sont discrets, il faut alors l'approcher en utilisant l'opérateur de convolution. Dans ce cas, l'opérateur laplacien discret est construit en combinant deux filtres de dérivée seconde à une dimension, en un seul à deux dimensions :

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Cet opérateur peut être combiné avec un filtrage Gaussien pour obtenir le Laplacien de Gaussien (LoG). La combinaison des deux opérateurs de convolution donne alors le kernel LoG tel que présenté sur les matrices 3×3 ci-dessus. Ce filtre appartient à la grande famille des ondelettes ou wavelets.

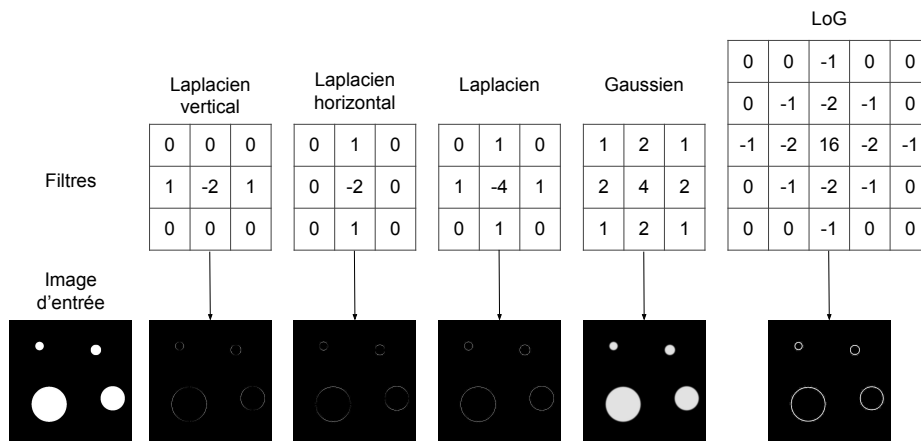


FIGURE 1.10 – Exemple de détection de spot par le filtre LoG. Les filtres Laplacien sont combinés avec un filtre Gaussien pour donner un filtre Laplacien de Gaussien.

Dans les méthodes hors apprentissage, on peut également citer l'utilisation des ondelettes multi-échelles [30, 134, 124]. Ces méthodes se basent sur l'idée que les spots seront présents à chaque échelle de décomposition en ondelettes et apparaîtront donc dans le produit multi-échelle. De nombreuses autres méthodes hors machine learning s'attaquent à la détection de spots en microscopie, les revues suivantes dressent un tableau assez complet de la situation [10, 101, 158].

Il existe assez peu de méthodes utilisant des algorithmes de Deep Learning pour localiser les spots dans les images smFISH, on peut néanmoins citer deepBlink [115].

Méthodes de détection et quantification des spots

La détection des spots se fait majoritairement par segmentation, le plus souvent par seuillage binaire. Les méthodes les plus utilisées sont le seuillage d'Otsu ou la ligne de partage des eaux (watershed). Otsu est un seuillage de type automatique, qui calcule la variance intraclasse pour tous les niveaux de gris possibles dans l'image. Le seuil optimal est alors celui qui obtient la plus haute variance. Dans la segmentation par watershed, une image est considérée comme un paysage topographique avec des crêtes et des vallées. Les valeurs d'élévation du paysage sont généralement définies par les valeurs de gris des pixels respectifs ou par l'amplitude de leur gradient. Le watershed transforme alors l'image en bassins versants, et pour chaque minimum local, un bassin versant comprend tous les pixels dont le chemin de descente le plus accentué se termine à ce minimum. La transformation en bassin versant décompose complètement une image et attribue donc chaque pixel un bassin versant, c'est-à-dire une région [133].

La quantification peut se présenter sous deux formes : (i) la localisation précise de chaque spot, dans ce cas, le centre de masse est calculé pour chaque région obtenue, ce qui permet d'obtenir une coordonnée (x, y) que l'on utilise généralement comme coordonnées de localisation du spot, (ii) les spots sont comptés sans pour autant que l'on connaisse leurs coordonnées, permettant néanmoins de quantifier l'expression des gènes.

1.3.2 Immunofluorescence

L'immunofluorescence (IF) est une technique qui permet de détecter de très nombreux composants dans les cellules ou les tissus en se basant sur les antigènes de la cible. Il peut s'agir de protéines, d'hormones, de composés médicamenteux, etc. Cette capacité est obtenue par des combinaisons d'anticorps spécifiques marqués par des fluorophores. Elle peut être réalisée sur des cellules cultivées ou des suspensions cellulaires, ainsi que sur des cibles spécifiques, dans des échantillons de tissus ou encore sur des organismes entiers.

Tout comme pour la technique smFISH, plusieurs étapes sont nécessaires avant de pouvoir imager les échantillons, notamment la fixation [45], le paraffinage [41] et le blocage, qui prévient le marquage non spécifique. Vient ensuite l'étape de marquage proprement dite, qui peut être directe si les fluorophores sont directement fixés aux anticorps primaires, ou indirecte si le fluorophore est fixé sur un anticorps secondaire [156], comme présenté sur la figure 1.11. Cette seconde méthode permet en théorie d'avoir un meilleur rapport signal sur bruit, car plusieurs anticorps secondaires peuvent reconnaître un anticorps primaire, donnant ainsi lieu à une amplification du signal de fluorescence. On peut également mentionner les méthodes d'amplification du signal par multiplexage, qui utilisent des anticorps secondaires capables de conjuguer plusieurs fluorophores [71]. Enfin, une étape de lavage permet de supprimer la majorité des liaisons non spécifiques susceptibles d'ajouter du bruit de fond à l'image.

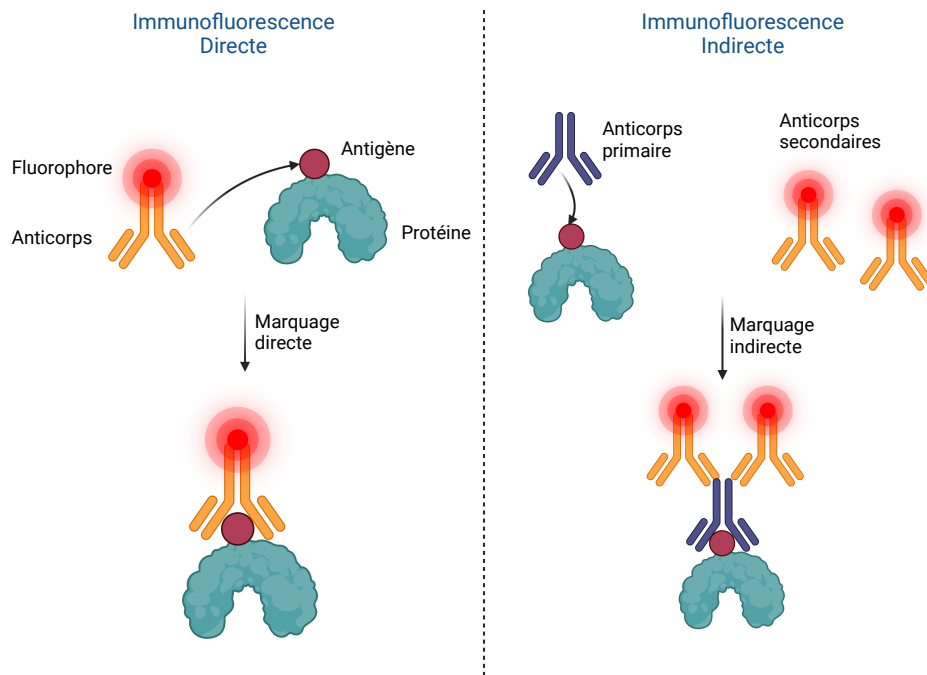


FIGURE 1.11 – Illustration des deux techniques principales d'immunofluorescence, à droite l'immunofluorescence directe ou l'anticorps portant le fluorophore cible directement un antigène, à gauche l'immunofluorescence indirecte ou un anticorps primaire sert d'attache à l'anticorps secondaire portant la fluorescence. Adapté de [71]

Les images sont généralement acquises via des microscopes confocaux et les images obtenues par IF peuvent présenter une plus ou moins grande densité de signal, homogène ou moucheté [156] (voir figure 1.12).

Méthodes de quantification du signal dans les images d'immunofluorescence

Contrairement à l'imagerie smFISH, la détection du signal de chaque molécule est rarement possible du fait de haut taux de fluorescence (sauf en SMLM), la mesure du taux d'expression se fait alors en quantifiant l'intensité lumineuse [21]. Cette tâche est généralement réalisée manuellement ou à l'aide de scripts maison et il existe assez peu d'outils spécifiquement développés pour quantifier cette tâche, on peut cependant citer EzColocalization [162] et AIABS [21]. Les features que l'on cherche à extraire sont des intensités, des régions de haute densité formant des blobs (région plus claire que le voisinage avec, au moins, un maximum local et dont la taille peut être assez variable). Sur la base de ces informations, il est possible de faire des analyses de distance ou de clustering entre régions plus ou moins denses ou encore entre images entières afin de caractériser des morphotypes [62].

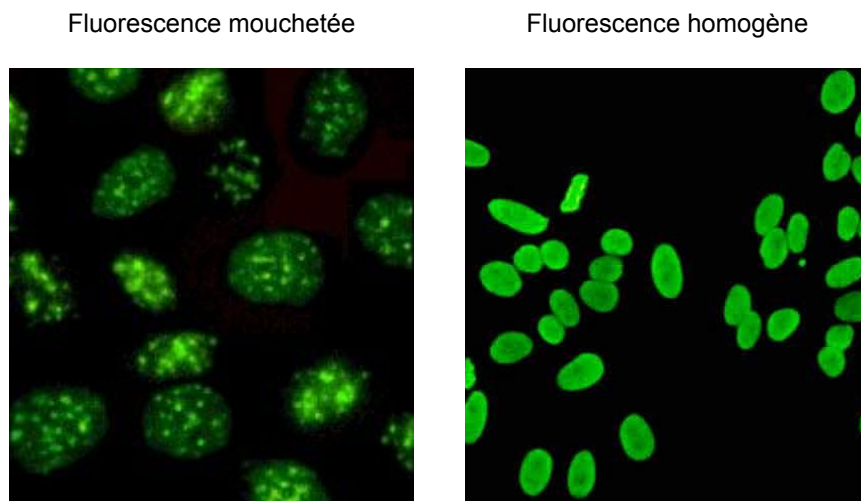


FIGURE 1.12 – Exemple d'acquisition d'image d'immunofluorescence, à droite une fluorescence mouchetée, car la densité de la protéine cible est faible, à gauche une fluorescence homogène, car la densité de protéine cible est élevée. Photos réalisées au Centre Hospitalier Lyon-Sud <https://www.memobio.fr/>

1.3.3 Marquage par fluorescence des noyaux cellulaire

Pour rendre les noyaux visibles par fluorescence, la méthode préférentielle est d'utiliser le DAPI, un fluorophore qui se lie fortement à l'adénine (A) et la thymine (T), présente en abondance dans le noyau. Le DAPI absorbe la lumière dans l'ultraviolet (358 nm) et à un maximum d'émission dans le bleu à 461 nm [169].

Segmentation des noyaux

La détection et la segmentation des noyaux cellulaires dans les images de microscopie à fluorescence constitue souvent une des premières étapes de l'étude des profils d'expression cellulaire. Il s'agit de segmenter non plus des spots, mais plutôt des régions d'une image, plus larges, qui partagent des caractéristiques communes.

Le niveau de difficulté de la détection d'un noyau unique dans une image varie considérablement en fonction des cas. Dans les cas les plus simples, les noyaux présentent un contraste élevé et sont spatialement bien séparés. Dans d'autres cas, la segmentation est moins évidente, par exemple dans les images où les noyaux se superposent dans la profondeur de l'échantillon, ou si le contraste est faible et les contours flous [64].

La segmentation des noyaux peut être divisée en deux types d'approches : celles utilisant de l'algorithmie classique et celles utilisant du Deep Learning.

Approches classiques

Dans les méthodes classiques, viennent en premier lieu les seuillages [92], qui supposent que les intensités respectives des noyaux et du bruit de fond soient suffisamment uniformes pour qu'il existe un seuil optimal permettant de séparer les noyaux entre eux. Des seuillages multiples sont également utilisés [16], cela permet, avec différentes valeurs de seuillage, de définir un jeu de régions redondantes à différents niveaux, correspondant aux noyaux, supposant là encore une certaine uniformité dans l'intensité des noyaux et du bruit de fond. On retrouve également des approches par détection de bords [181] utilisant des dérivées du premier ou second ordre (LoG, filtre de Canny), supposant un bruit de fond faible et que les bords des noyaux soient nets, ce qui, en fonction de l'expérimentation, n'est pas toujours le cas. Viennent ensuite les approches par segmentation en régions où le principe consiste à associer les pixels voisins partageant des caractéristiques communes. Les plus utilisées sont, ici encore, le watershed et la croissance en région [103, 175, 53].

Approches par Deep Learning

La majeure partie de la littérature actuelle, au sujet de la segmentation des noyaux, repose sur l'utilisation d'algorithmes d'apprentissage de type Deep Learning. Pionnier en la matière, le U-Net [140] utilise un réseau de type d'encodeur-décodeur doublé de "skip connections" qui lui permettent de propager une information spatiale dans le réseau, en plus de l'information sémantique obtenue par l'encodeur. Cette architecture, nécessitant peu de données d'entraînement, facile à mettre en œuvre et généralement assez robuste, a donné d'excellents résultats pour les cas les moins complexes et a été adaptée avec succès à beaucoup d'autres domaines de l'imagerie biomédicale. À ce stade, on peut considérer

deux types d'architectures majoritairement utilisées pour la détection de noyaux, celles se basant sur le U-Net, et celles se basant sur l'architecture Mask R-CNN, ce dernier utilisant un système de bounding-box pour localiser puis segmenter les noyaux.

Dans la première catégorie se trouvent des méthodes se voulant génériques à la détection de cellules et de noyaux, tels que StarDist [153], qui utilise un U-Net pour apprendre à détecter les polygones convexes de type partie étoilée, ainsi que CellPose [166] qui pour sa part apprend à un U-Net à prédire des gradients de cartes typologiques de type watershed, des images de noyaux.

Dans la seconde catégorie, se trouve une méthode dédiée à la détection de noyaux, NucleAIzer [63], qui entraîne un Mask R-CNN à créer des bounding-box autour des noyaux puis à les segmenter. Ce réseau nécessite que les images soient au préalable redimensionnées afin d'obtenir des noyaux ayant tous une taille équivalente.

Enfin, les auteurs de [180] ont entraîné deux réseaux de neurones, un U-Net et un Mask R-CNN, afin de retenir comme segmentation, l'intersection des deux segmentations prédites.

Enfin, une approche qu'il convient de mentionner ici, qui ne fait cependant pas directement parti de la famille des outils de segmentation, mais plutôt des outils de restauration d'image, est CARE [188]. En utilisant des paires d'images de noyaux cellulaires acquises en même temps, mais avec deux illuminations différentes, une de faible intensité et une de haute intensité, CARE, avec un réseau de neurone basé sur le U-Net, permet de restaurer le signal des noyaux acquis avec une faible luminosité. Cette technique a été ensuite généralisée aux tâches de débruitage, de projection de surface et de super-résolution. Très performante et proposant aujourd'hui, sous le nom de CSBDeep, une boîte à outils comportant des modèles pré-entraînés pour les tâches mentionnées précédemment, cette méthode souffre de quelques inconvénients. Premièrement, elle n'est pas généralisable aux très petites structures comme les spots [188]. Deuxièmement, elle nécessite au préalable l'acquisition de jeux de données pairées, et donc d'acquérir des images à la fois dans la condition de départ et dans celle que l'on souhaite obtenir (basse et haute résolution, par exemple). Ceci peut constituer une forte limitation, car l'acquisition de tels jeux de données n'est pas souvent possible.

1.4 Analyses spatiales et statistiques pour la microscopie de fluorescence

La compréhension de types et d'états cellulaires est essentielle pour le progrès dans la connaissance du rôle des cellules dans l'émergence de phénotypes d'intérêt. Traditionnellement, des techniques telles que l'immunohistochimie, la cytométrie de flux et les méthodes morphologiques sont utilisées pour identifier les types de cellules [137]. Plus récemment, les technologies de séquençage de cellules individuelles (*single cell*), permettent d'obtenir des informations sur les profils d'expression de cellules individuelles, caractérisant ainsi leurs types et leurs fonctions [200].

Avec le développement très récent de la transcriptomique à résolution spatiale [112] puis des autres "omics" à résolution spatiale à l'échelle de tissus, il est aujourd'hui possible d'étudier la distribution de différents types cellulaires en observant la localisation des ARNm et des protéines sur les coupes de tissus [48]. De façon encore plus fine, la position des molécules dans une cellule donnée, peut fournir des informations utiles pour définir l'état cellulaire, la fonction de la cellule [189] et, finalement, la fonction du gène observé [65] au sein d'un type cellulaire particulier.

Un certain nombre de découvertes biologiques ont été rendus possibles grâce aux analyses de localisation. Nous citons par exemple le clustering de certaines protéines du noyau telles que les protéines de liaison CCCTC et les complexes protéiques du groupe Polycomb, à propos desquelles il a été établi l'importance de cet état cellulaire dans l'organisation du génome et dans la régulation de la transcription et dont la perturbation peut provoquer des troubles neurologiques du développement [90]. Un autre exemple est la localisation clusterisée mais distincte des récepteurs T-cell (TCR) et des linker d'activation des T-cell (LAT) qui deviennent colocalisées lors de l'activation des T-cell [90]. Enfin, la traduction locale d'un ARNm en la protéine correspondante est connu être un phénomène important pour un grand nombre de fonctions cellulaires dans les cellules neuronales [66].

Des mécanismes régulant la localisation des ARNm ont été observés, le plus évident concerne le transport des ARNm le long du cytosquelette, qui peut être soumis à des contraintes pour favoriser ou empêcher le déplacement de transcrits vers certaines régions [147]. Un second mécanisme observé, appelé diffusion-trapping permet de faire diffuser les molécules dans la cellule jusqu'à une région spécifique où elles seront alors ancrées [86]. Enfin, des exemples de localisation sélective ont également été observés, le transcrit est alors soumis à des mécanismes de restriction qui l'adressent à une région spécifique tout en lui empêchant l'accès à d'autres régions de la cellule [70]. Ces différents mécanismes ne sont pas mutuellement exclusifs et peuvent agir simultanément pour définir une localisation

spécifique d'un transcrit donné [96, 171].

L'acquisition de données à haut débit par les techniques de phénotypage cellulaire en imagerie de microscopie (souvent par SMLM), donne la possibilité d'étudier les différences entre les états cellulaires définis par la localisation de molécules, de façon systématique. Il devient alors possible d'établir si les distributions de molécules suivent différents motifs entre deux conditions (par exemple avec ou sans administration d'une drogue), ou si deux molécules différentes se trouvent colocalisées, et ce, de façon statistiquement robuste [90].

De manière générale, toute analyse spatiale peut être divisée en trois étapes principales :

1. L'identification des régions à analyser,
2. La mesure du signal que l'on cherche à analyser ainsi que son affectation aux régions,
3. Application sur ces mesures des analyses statistiques et comparatives.

Dans cette thèse, nous nous intéresserons à deux types de signaux, (i) les signaux discrets, comme c'est le cas du signal des ARNm dans les images smFISH ; (ii) les signaux continus, comme le signal protéique issu des images d'immunofluorescence. Dans cette section, nous présentons les bases des analyses spatiales et leurs applications aux statistiques spatiales pour la corrélation de signaux discrets et continus, que nous utiliserons plus en détail dans le chapitre 4 de cette thèse.

1.4.1 Fondements de l'analyse spatiale

Les premières méthodes d'analyses spatiales ont été développées pour d'autres domaines tels que la géographie. Les premiers travaux remontent aux années 1820 et 1830 en France où des mathématiciens (Charles Dupin), agronomes (Adolphe d'Angeville), criminologues (André-Michel Guerry), statisticiens (Adolphe Quetelet) établirent ce qu'on dénommera plus tard comme les "statistiques spatiales de la morale de France", portant sur l'instruction, la criminalité, la mortalité, le recensement de la population. Les statistiques spatiales seront également appliquées et développées pour l'analyse des cartes de la ville de Londres lors de l'épidémie de choléra de 1854 par John Snow [155].

Grille de lecture

La première étape, lors de l'établissement de statistiques spatiales, est de considérer la région qui va être étudiée ainsi que son potentiel découpage en sous-régions. Dans les travaux cités précédemment, il pouvait s'agir de couper des cartes par région, département, arrondissement, quartier, etc.

Une zone étudiée peut être considérée dans son ensemble (une carte entière, une image entière) ou bien être morcelée pour établir des statistiques plus fines. Le découpage ne suit pas nécessairement de règle précise, il dépend essentiellement de la donnée observée et de l'information que l'on cherche à obtenir. Le découpage peut se faire en régions disjointes (des îles dans un océan, des cellules dans une images) ou en tessellation, c'est-à-dire un découpage intégral, sans lacune et sans superposition. Le découpage s'effectue en régions qui peuvent prendre de nombreuses formes, cadrans, carrés, polygones, iso-lignes, diagramme de Voronoï, etc (voir figure 1.13). et qui n'ont pas nécessairement besoin d'être uniformes. Une région peut également être encore découpée en régions plus petites pour créer différentes mailles d'analyse (système particulièrement utilisé dans la construction des modèles de météorologie).

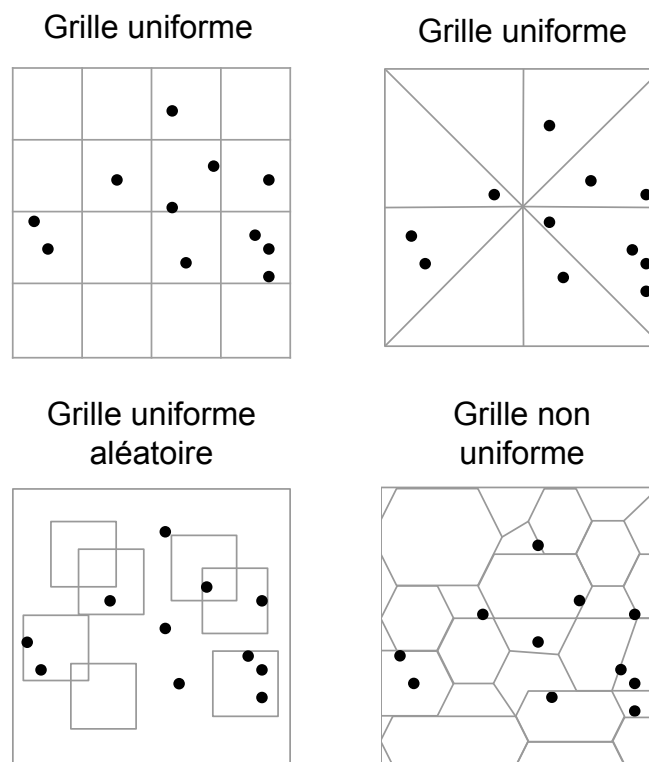


FIGURE 1.13 – Le cadre de lecture d'une analyse spatiale peut être choisi de façon très diverse en fonction de ce que l'on cherche à étudier. Il peut s'agir d'un découpage intégral uniforme (cadres du haut) ou non uniforme (cadre en bas à droite), placé aléatoirement (cadre en bas à gauche) ou non aléatoirement.

La difficulté du découpage en régions réside alors essentiellement dans le choix de la forme et du nombre de cadrans afin de pouvoir capturer les nuances dans la distribution

spatiale entre les différentes images.

Une contrainte supplémentaire s'ajoute lorsque l'on souhaite analyser des images de cellules, celle de l'alignement des images. En effet, pour comparer deux cartes d'une ville par exemple, il est assez aisé de les superposer avec l'aide des points cardinaux et de repères physiques. Les cellules n'ont pas de points cardinaux, et pas systématiquement de repères morphologiques uniformes d'une cellule à l'autre qui permettraient de les aligner aussi aisément qu'une carte géographique. Dans ce cas, soit le découpage doit se concentrer sur des régions comparables (cellules entières, noyaux, organites particuliers), soit il est nécessaire de définir des points de comparaisons possibles au regard d'une question biologique. Les repères les plus utilisés sont le noyau, la membrane plasmique et s'y ajoutent parfois d'autres biomarqueurs ainsi qu'une direction de polarisation [25].

1.4.2 Analyses spatiales du premier ordre

Quantification

La statistique la plus évidente est la quantification du signal, souvent désignée comme la mesure d'abondance. Dans le cas des données basées sur les coordonnées des points (comme c'est le cas des spots d'ARNm), il s'agit d'une simple somme de tous les points contenus par région ou par cadran. Cette mesure, bien que basique, est souvent la première que l'on cherche à extraire, car elle donne une idée très générale du taux d'expression d'un gène. Pour les signaux continus (comme c'est le cas pour les protéines), la mesure correspondante sera la somme de toutes les intensités, donnant également une idée très générale du taux de protéines.

L'analyse de motifs ponctuels comporte d'autres mesures élémentaires, comme le point moyen (*mean center*), correspondant à la moyenne des valeurs des coordonnées (x, y) des points, donnant une idée de la tendance centrale de la distribution des points. La distance standard, elle, permet de mesurer la dispersion des points autour du point moyen, elle se définit comme suit :

$$d = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu_x)^2 + \sum_{i=1}^n (y_i - \mu_y)^2}{n}} \quad (1.1)$$

où μ_x et μ_y sont les coordonnées du point moyen, x_i, y_i sont les coordonnées d'un point i donné et n le nombre total de points. Ceci permet de mesurer le degré de dispersion, mais ne permet pas de montrer un quelconque effet directionnel dans la distribution des points. Pour pallier ce problème, une variante de la distance standard est la distribution directionnelle ou l'ellipse de l'écart standard. Dans le cas de données bidimensionnelles,

l'écart-type est calculé séparément pour chaque axe et permet de définir un axe long et un axe court tout en montrant la dispersion des points le long de ces deux directions.

Densité

Outre les mesures de quantification, des mesures basées sur la densité permettent d'étudier les variations de densité de points dans l'espace. Les mesures de densité peuvent être divisées en deux catégories : la densité globale et la densité locale. La densité globale fait référence au rapport entre le nombre de points observés et la taille de la zone d'étude entière et n'est calculée qu'une fois, à l'aide de l'équation suivante :

$$\lambda = \frac{n}{a} \quad (1.2)$$

où n est le nombre de points observés et a , l'aire de la région étudiée.

La densité globale fournit une mesure unique de la densité de points dans l'ensemble de la zone d'étude, mais elle ne permet pas de saisir les variations locales. La densité locale, en revanche, montre des densités de points variables à différents endroits de la zone d'étude, en utilisant également l'équation 1.2 avec a correspondant à l'aire de la sous-région où les points ont été comptés. La mesure est appliquée alors sur tout ou partie des sous-régions (voir figure 1.14 pour exemple).

Dans le cas du signal continu, la mesure équivalente est la moyenne des intensités sur la région étudiée.

Une autre approche pour mesurer la densité est l'estimation par noyau (*kernel density*), qui est une mesure de densité par fenêtre glissante. Là encore, la mesure se fait par cadran, mais les cadrans sont recouvrants. Le résultat est une grille de valeurs de densité où chaque cellule se voit attribuer la valeur de densité calculée pour la fenêtre du noyau centrée sur cette cellule, à la manière d'une convolution.

La densité est utilisée en bioimagerie, cependant au sein d'une population de cellules, le volume peut considérablement varier, ce qui implique que la densité doit fluctuer également d'une cellule à l'autre. Toutefois, il a été montré que de nombreuses espèces différentes d'ARNm au sein d'une lignée cellulaire présentent une densité constante pour des écarts de volume allant jusqu'à 6 fois [127]. De même, il a été montré que les cellules quiescentes et sénescents conservent également la même densité, malgré les changements dans le nombre moyen d'ARNm [127]. Ceci indique que pour de nombreux gènes, l'abondance des ARNm est contrôlée pour produire la quantité appropriée à la taille de la cellule, impliquant un mécanisme de conservation de la densité [127]. Ainsi, la mesure de densité permet tout de même, dans certains cas, de comparer l'expression des gènes entre des cellules de même lignée, mais ayant un volume totalement différent.

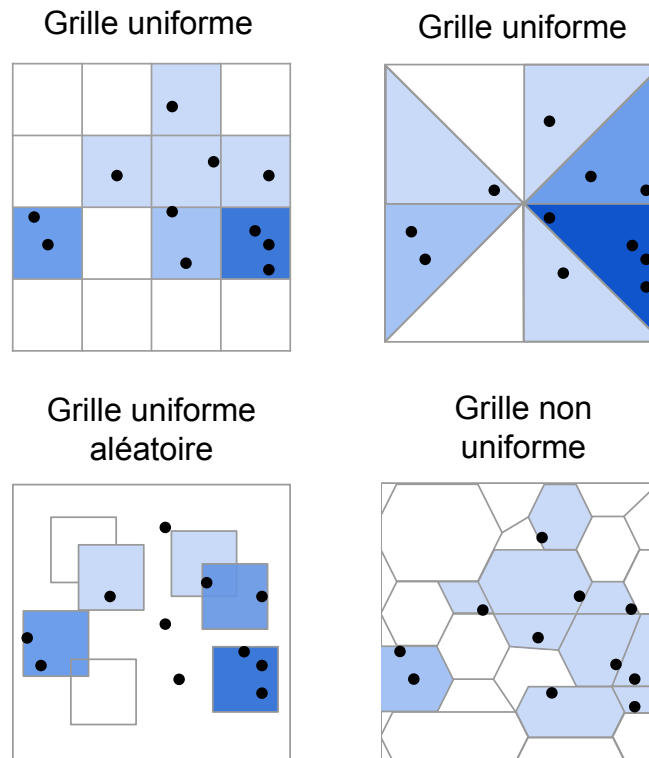


FIGURE 1.14 – Le choix du cadre de lecture d’une image va impacter le résultat du calcul des densités. Sur les trois grilles uniformes, la région bleu foncé indique l’existence de régions de forte densité, tandis que sur la grille non uniforme, cette densité n’est pas visible.

1.4.3 Analyses spatiales du second ordre

Des alternatives à la densité sont les méthodes basées sur la distance entre les points, pour étudier la façon dont ils sont distribués les uns par rapport aux autres.

Clustering

La mesure la plus commune pour distinguer les cas où les ARNm se clusterisent dans l’espace des cas où ils sont dispersés, est la distance moyenne du plus proche voisin (*average nearest neighbor*, ANN), qui est la moyenne de la distance du plus proche voisin entre toutes les paires de points (voir équation 1.3). Elle est utilisée comme une mesure globale de la configuration spatiale des points.

$$ANN = \frac{\sum_{i=1}^n d_i}{n} \quad (1.3)$$

où d_i est la distance du plus proche voisin de i et n est le nombre de points.

Cet indicateur est souvent comparé à la valeur d'ANN pour une distribution aléatoire de Poisson (*Complete Spatial Randomness, CSR*), afin d'évaluer la significativité d'un clustering :

$$CSR = \frac{0.5}{\sqrt{\frac{n}{A}}} \quad (1.4)$$

où A est l'aire du rectangle minimal permettant de délimiter l'ensemble des points.

Le score z , qui est le ratio de la différence entre l'ANN et la distribution aléatoire et la variance, est calculée selon l'équation 1.5.

$$z = \frac{ANN - CSR}{\frac{0.261}{\sqrt{n^2/A}}} \quad (1.5)$$

Si la distance moyenne est inférieure à la moyenne calculée pour une distribution aléatoire, la distribution des points est considérée comme agrégée, si elle est supérieure alors les points sont considérés comme dispersés. Le score z indique la significativité de cette distribution.

Cette idée peut s'étendre au k -ième plus proches voisins (KNN) et permet d'évaluer la configuration spatiale pour différents ordres de voisinage.

Cette métrique permet de mesurer le degré de regroupement d'un ensemble de points, mais ne donne cependant pas d'information particulière sur la complexité des configurations.

L'algorithme K-means est un algorithme itératif qui tente de partitionner les données en K sous-groupes distincts, où chaque point n'appartient qu'à un seul groupe. Il essaie de rendre les points intra-groupes aussi similaires que possible tout en maintenant la distance inter-groupes aussi grande que possible. Il affecte les points à un cluster de telle sorte que la somme des carrés de la distance entre les points et le centroïde du cluster soit minimale. Moins il y a de variation au sein des clusters, plus les points sont homogènes au sein d'un même cluster [1]. L'algorithme suit les étapes suivantes :

1. Initialisation aléatoire des k centroïdes,
2. Assigne les points au cluster le plus proche en calculant la distance euclidienne point-centroïde,
3. Calcul des nouveau centroïdes en calculant la moyenne de tous les points du cluster,
4. Répéter les étapes 2 et 3 jusqu'à ce que les points ne changent plus de cluster.

Un autre algorithme, DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*) se base, lui, sur la densité. Il peut découvrir des clusters de différentes formes et

tailles à partir d'une grande quantité de données, il est robuste au bruit et aux outliers [42]. Il utilise deux paramètres :

1. ϵ : détermine la distance des points entre eux pour être considérés comme faisant partie d'un cluster,
2. m : le nombre de points nécessaire a la formation d'un cluster.

L'algorithme procède en choisissant arbitrairement un point dans l'ensemble des données. S'il y a au moins m points dans un rayon de ϵ autour du point, alors il considère que tous ces points font partie du même cluster. Les clusters sont ensuite étendus en répétant récursivement le calcul du voisinage pour chaque point voisin, jusqu'à ce que tous les points aient été visités.

Fonction K de Ripley et fonctions associées

Des fonctions ont été développées pour obtenir plus de détails quant à la configuration spatiale étudiée. La première est la fonction $G(r)$ qui calcule la fréquence cumulée de la distribution des distances des plus proches voisins (équation 1.6).

$$G(r) = \frac{\{\#d_i : d_i \leq r, \forall i\}}{n} \quad (1.6)$$

avec $\#$ le nombre de points qui ont au moins un voisin entre le point i et la distance r .

Il s'agit d'une fonction cumulative qui augmente avec la distance r . Cette mesure est comparée à une distribution de Poisson homogène. G augmente lorsque les points sont regroupés et diminue lorsque les points se dispersent.

La fonction G ne prend en compte que le plus proche voisin et ne permet donc pas d'analyser la distribution à différentes distances. Pour répondre à ce problème, les statistiques spatiales de Ripley ont été introduites. C'est une famille de méthodes utilisées pour décrire si des points discrets dans un espace ont une distribution aléatoire, dispersée ou agrégée. Une fonction très utilisée dans l'analyse des motifs de points, est la fonction K, ou fonction K de Ripley (*Ripley's K function*) [139], qui mesure la distance entre les points pour toutes les distances possibles (voir figure 1.15).

Cette fonction définit un cercle de rayon d autour de chaque point i puis compte le nombre total n de points qui tombent à l'intérieur de ce cercle. Ensuite, la distance d est incrémentée pas à pas et le nombre de points à l'intérieur du cercle de rayon d est recalculé à chaque pas.

Pour une distance donnée d , $K(d)$ se définit selon l'équation 1.7.

$$K(d) = \frac{1}{\lambda^2 A} \sum_{i=1}^n \sum_{j \neq i}^n I_d(d_{ij}) \quad (1.7)$$

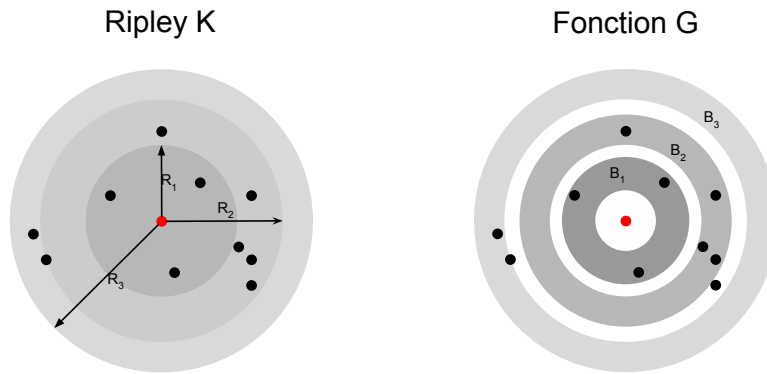


FIGURE 1.15 – La fonction K de Ripley est calculée pour des disques successifs de rayon r croissant et est donc cumulative par nature. La fonction G utilise des bandes de distance B croissantes au lieu de disques et permet l'analyse des concentrations de points à différentes échelles de distance. Adapté de [174]

où A est l'aire étudiée, λ est l'intensité (le nombre de points divisé par l'aire), d est la distance par rapport à i , j est l'ensemble des points en excluant i , et I_d est une fonction qui prend 0 si j n'est pas à une distance d de i , et qui prend 1 sinon (équation 1.8).

$$I_d(d_{ij} < d) = \begin{cases} 1 & \text{si } d_{ij} < d \\ 0 & \text{sinon} \end{cases} \quad (1.8)$$

Cette mesure se compare là encore à une distribution de Poisson $K_{expected} = \pi d^2$ pour estimer à quel point $K(d)$ dévie de cette distribution ponctuelle régulière. Lorsque $K(d) < K_{expected}$ cela indique que la distribution suit un processus régulier dans la bande de distance d , tandis que $K(d) > K_{expected}$ indique que la distribution suit un processus de regroupement (*clustering*), comme présenté sur la figure 1.16. Cette fonction, cumulative, a tendance à rendre peu évidente la lecture des différences entre $K(d)$ et $K_{expected}$ pour les valeurs de d faibles.

La fonction K peut être normalisée pour donner la fonction L (équation 1.9) de Besag [12], donnant une ligne théorique et non plus une courbe et rendant plus aisée la lecture du graphique pour les petites distances.

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d \quad (1.9)$$

Cette fonction peut encore être normalisée de manière que $k_{expected} = 0$, donnant alors la fonction H (équation 1.10).

$$H(d) = L(d) - d \quad (1.10)$$

Ces fonctions sont généralement utilisées pour comparer une distribution de points

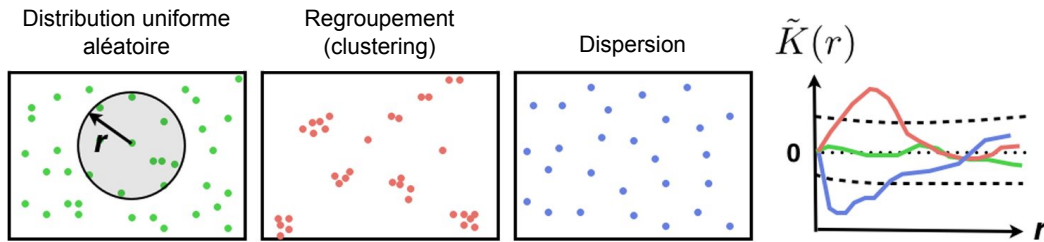


FIGURE 1.16 – Analyse de la distribution spatiale de points avec la fonction de K de Ripley. $\tilde{K}(r)$ est proportionnelle au nombre de paires de points à l'intérieur d'un cercle de rayon r autour d'un point central donné. La déviation $\tilde{K}(r)$ de la distribution uniforme indique un clustering si $\tilde{K}(r) > 0$ ou une dispersion si $\tilde{K}(r) < 0$. Adapté de [84].

donnée, à une distribution aléatoire, pour établir si le clustering détecté est statistiquement significatif ou non. Plus précisément, la distribution de points étudiée est testée par rapport à l'hypothèse nulle selon laquelle les points sont distribués de façon aléatoire et indépendante. Une valeur positive de $H(d)$ est alors dite cohérente avec un clustering sur l'échelle spatiale considérée, tandis qu'une valeur négative est cohérente avec une dispersion.

1.4.4 Colocalisation

Les analyses de colocalisation cherchent à étudier les relations spatiales entre différents jeux de données. Dans le cas de la bioimagerie, il s'agit la plupart du temps d'étudier les liens entre les molécules différentes marquées par fluorescence, en quantifiant leur cooccurrence et leur corrélation [183], par exemple entre un ARNm et la protéine correspondante. Ceci implique donc des analyses comparées entre deux jeux d'images différents. Ces approches reposent sur le paradigme selon lequel la proximité spatiale est une indication d'interactions entre les molécules à l'intérieur de la cellule.

Il s'agit donc de pouvoir établir des motifs de distribution du signal dans deux jeux de données puis de les comparer. Comme mentionné dans l'introduction de cette section, afin d'établir des motifs de distribution du signal et de les comparer, l'étape préalable essentielle est de rendre les images comparables entre elles en termes de grille de lecture. Il s'agit donc d'aligner les images en utilisant les repères spatiaux biologiques.

La comparaison va alors se faire sur des vecteurs contenant les caractéristiques spatiales préalablement mesurées. De manière la plus simple, il est donc possible de comparer deux images $I_1 = \vec{x}_i$ et $I_2 = \vec{y}_i$ où \vec{x}_i et \vec{y}_i sont les vecteurs de valeurs représentant l'image,

tel que par exemple tous les pixels ou alors les densités des régions (voir figure 1.17). S'il s'agit de deux ensembles d'images, $\{I_1\}$ et $\{I_2\}$, il devient alors nécessaire de construire des vecteurs les représentant, par exemple avec les moyennes de densité par région (ce qui est possible uniquement si les images sont "alignées").

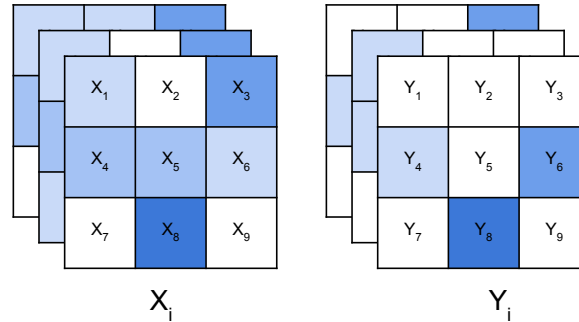


FIGURE 1.17 – Exemples de comparaison par corrélation calculée pour des jeux d'images sur une grille de lecture uniforme découpant chaque image en 9 régions dans lesquelles la densité a été mesurée. La colocalisation c sur les données $\{I_1\}$ et $\{I_2\}$ se calcule alors sur les vecteurs correspondants, $c(\mu(\vec{x}_i), \mu(\vec{y}_i))$, où $\mu(\vec{x}_i) = \langle \mu(x_1) \dots \mu(x_9) \rangle$ et $\mu(\vec{y}_i) = \langle \mu(y_1) \dots \mu(y_9) \rangle$.

La comparaison utilisée le plus souvent est la corrélation. Pour calculer des coefficients de corrélation, les tests statistiques préférentiellement utilisés sont le coefficient de corrélation de Pearson et les coefficients de recouvrement de Manders [18].

Le coefficient de corrélation de Pearson, initialement formulé par Karl Pearson en 1896, a été appliqué à la microscopie de fluorescence prêt de 100 ans plus tard par Erik Manders [105]. Il permet de connaître l'intensité et la direction d'une relation d'interdépendance entre deux variables, x et y , et le coefficient r est le ratio entre la covariance et le produit des déviations standards (voir équation 1.11). Dans l'article de Manders [105], il s'agit de comparer deux images et x et y sont alors des pixels.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1.11)$$

où n est la taille de la population étudiée, x_i et y_i sont les point de l'échantillon i , \bar{x} et \bar{y} sont respectivement les moyennes des variables x et y .

Le coefficient de corrélation de Pearson mesure la covariance des niveaux de signal de deux images. Comme l'intensité moyenne est soustraite à chaque intensité de pixel, ce coefficient est indépendant du niveau de signal et de bruit, ce qui en fait une mesure assez robuste, sous réserve de ne l'appliquer qu'au sein d'une cellule et de ne pas prendre en compte le fond de l'image qui peut fausser la mesure en ajoutant de fort taux de vrais négatifs [37]. Le coefficient r est compris entre -1 et 1 , une valeur de r proche de 0

indique l'absence de corrélation, une valeur proche de 1 indique une corrélation positive et une valeur proche de -1 indique une corrélation négative. Si les valeurs extrêmes sont aisément interprétables, les valeurs intermédiaires peuvent être plus difficile à interpréter.

Les coefficients de Manders [106] ($M1$ et $M2$, voir équations 1.12 et 1.13) sont des mesures de cooccurrence qui calculent le pourcentage du signal total d'une images A qui chevauche le signal de l'autre image B . Dans le cas de la microscopie de fluorescence, A et B sont généralement deux canaux d'acquisition de fluorescence, par exemple le vert pour les ARNm et le rouge pour les protéines.

$$M1 = \sum_i \frac{A_{i, coloc}}{A_i} \quad (1.12)$$

$$M2 = \sum_i \frac{B_{i, coloc}}{B_i} \quad (1.13)$$

ou A_i et B_i sont les intensités des pixels i sur les images A et B , $A_{i, coloc}$ prend la valeur A_i si $B_i > 0$ et $B_{i, coloc}$ prend la valeur B_i si $A_i > 0$.

$M1$ et $M2$ expriment le pourcentage de fluorescence ayant une contrepartie dans l'autre image et permettent de quantifier la direction de la colocalisation.

Outils d'analyses spatiales pour les images de microscopie

Plusieurs outils qui intègrent des analyses spatiales ont été développés pour la microscopie de fluorescence. On peut citer Giotto [35], une boîte à outils d'analyses spatiales et de visualisation; SquidPy [128] qui permet de faire de la visualisation, de la détection/segmentation et des statistiques spatiales; Big-fish [72] un outil permettant de mettre en place des pipelines allant de l'analyse d'images jusqu'au calcul des statistiques spatiales; SODA [85], un outil qui se base sur la fonction K de Ripley pour caractériser le positionnement spatial relatif de plusieurs distributions de molécules (colocalisation) ainsi que nos propres travaux sur DypFISH [148] une librairie de fonctions d'analyses de motifs de distribution et d'interdépendance spatiale entre la localisation d'ARNm et de protéines.

1.5 Deep Learning

1.5.1 Fondamentaux du Deep Learning

Les récentes avancées en matière de vision par ordinateur et d'apprentissage automatique apportent une collection d'algorithmes dotés d'une très bonne aptitude à déchiffrer le contenu des images. Ces algorithmes d'apprentissage, rassemblés sous la dénomination de *Machine Learning* constituent un ensemble de techniques donnant la capacité aux machines d'apprendre automatiquement un ensemble de modèles à partir d'exemples : les données fournies en entrée. Une partie de ces algorithmes utilisent des réseaux de neurones dont la dimension peut varier en fonction du nombre de couches de neurones utilisées pour les construire. Cet empilage de dizaines, parfois de centaines, de couches de neurones donne une certaine profondeur à cette méthode que l'on dénomme alors l'apprentissage profond (*Deep Learning*). Cette profondeur apporte une plus grande complexité et permet l'apprentissage de modèles plus complexes.

Les techniques d'apprentissage profond peuvent être distinguées en deux grandes catégories : non supervisées et supervisées. Une troisième distinction est généralement appliquée pour les techniques dites semi-supervisées, mais ne faisant pas l'objet de cette thèse, elles ne seront pas traitées ici. L'apprentissage supervisé consiste à apprendre une correspondance entre un ensemble de variables d'entrée et une variable de sortie sur la base d'exemples de paires entrées-sortie et à appliquer cette correspondance pour prédire les sorties pour des nouvelles données d'entrée. L'apprentissage supervisé est la méthodologie la plus utilisée [31] lorsque des données possédant une annotation définissant la caractéristique à apprendre sont disponibles.

L'apprentissage non supervisé permet, lui, de mettre en œuvre le processus d'apprentissage en l'absence de données annotées. Dans ce cas, le réseau de neurones apprend des caractéristiques significatives et une représentation interne nécessaire pour découvrir la structure ou les relations cachées dans les données d'entrée.

Ces dix dernières années, les méthodes d'apprentissage profond ont été développées pour l'analyse des images de microscopie et ont modifié le paysage du traitement et de l'interprétation des données de bioimagerie. Ces avancées sont en mesure de rendre les analyses difficiles plus routinières et permettent d'automatiser les analyses quantitatives sur de grands volumes de données, tant pour la classification d'images ou d'objets que pour la segmentation ou la détection d'objets ou encore pour la microscopie augmentée telle que le débruitage ou la super-résolution.

Cette thèse faisant largement appel aux réseaux de neurones et à l'apprentissage profond, cette section explicite les propriétés et les éléments de base de l'apprentissage profond qui font le lit d'une grande partie des récents développements en analyse d'images de fluorescence ainsi que des méthodes que nous présentons dans les chapitres 2 et 3.

Neurones

À l'ère moderne, l'histoire des réseaux de neurones débute avec la description du neurone formel en 1943 par Warren McCulloch et Walter Pitts [113]. Ce neurone formel se veut être un modèle simplifié du neurone biologique tel que décrit dans les travaux du neuroscientifique Santiago Ramón Y Cajal qui s'attacha à décrire la morphologie et le fonctionnement des neurones. Il s'agit ni plus ni moins d'une fonction mathématique, qui prend une ou plusieurs valeurs en entrée et renvoie en sortie un signal binaire (0 ou 1) comme présenté sur la figure 1.18.

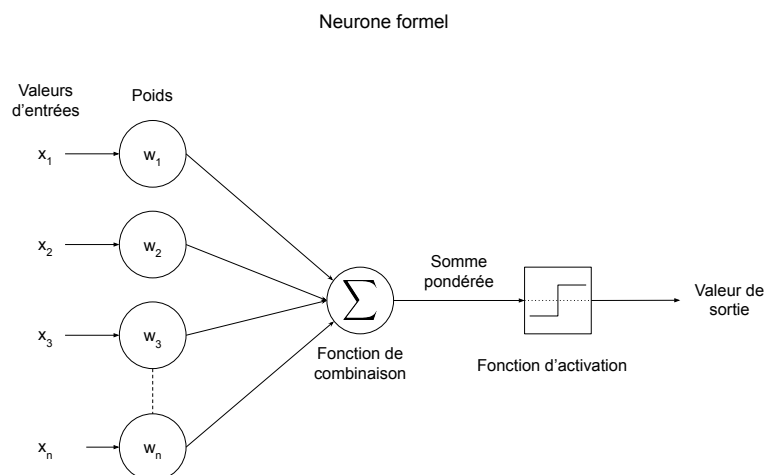


FIGURE 1.18 – Le neurone formel présenté Warren McCulloch et Walter Pitts, les valeurs d'entrée sont pondérées avant d'être combinées. Le résultat est seuillé par une fonction d'activation à seuil dur et produit une valeur de sortie 1 ou 0. Adapté de [144].

Ce neurone se comporte comme un opérateur logique à partir duquel il est possible de construire un ensemble d'opérateurs, capable de résoudre des opérations logiques. Ainsi prennent naissance les réseaux de neurones artificiels. Dans la première version du neurone formel, les valeurs d'entrées sont pondérées par un poids (ω) avant d'en calculer la somme et d'appliquer une fonction non linéaire, appelée fonction d'activation. Dans le cas [113] il

s'agit d'une fonction de type seuil dur, le neurone est alors "activé", c'est-à-dire non-nul, que si la somme de ses entrées pondérées dépasse une valeur seuil [88].

Ce neurone formel deviendra par la suite la base de ce que Marvin Minsky, encore lui, désignera en 1956 comme l'intelligence artificielle. Sur ces prémisses, Rosenblatt deux ans plus tard créera le perceptron, permettant ainsi d'entraîner un premier neurone, plus précisément, d'apprendre à modifier les poids des différentes entrées.

Couche de neurones (layer)

L'idée d'associer des neurones viendra encore plus tard avec le perceptron multicouche, composé de trois couches de neurones, une couche (layer) d'entrée, une couche de sortie et au milieu une couche dite "cachée". La sortie d'une première couche de neurones est alors l'entrée de la couche suivante. L'idée reste la même dans les réseaux de neurones actuels, cependant du fait de l'amélioration des capacités de calculs, ces réseaux contiennent davantage de couches cachées, alors dites "profondes", on parle alors d'apprentissage profond (*Deep Learning*). En ajoutant plus de couches et plus de neurones dans une couche, un réseau profond peut représenter des fonctions d'une plus grande complexité.

Rétropropagation du gradient

L'introduction de fonction d'activation différentiable telle que la fonction sigmoïde à la place des fonctions de seuil permet également d'introduire la rétropropagation (*back-propagation*) du gradient [142, 141] à travers les différentes couches de neurones. Il s'agit d'une approche itérative permettant d'optimiser les poids ω de chaque neurone du réseau en calculant les dérivées partielles d'un critère appelé fonction de coût ou fonction de perte que l'on cherche à minimiser lors de l'apprentissage. Cette optimisation est généralement réalisée par descente de gradient stochastique.

Fonction de coût (loss)

La fonction de coût, qui est un sujet de recherche important dans le domaine de l'apprentissage automatique [182], joue un rôle central dans la construction des algorithmes d'apprentissage automatique et dans l'amélioration de leurs performances. On peut considérer qu'un algorithme d'apprentissage automatique se résume à la résolution d'un problème d'optimisation principalement représenté par la fonction de coût. La fonction de coût permet de calculer l'erreur entre la sortie réelle attendue et la sortie prédite par le réseau de neurones.

Au début de l'entraînement, les paramètres du réseau, les poids, sont initialisés de manière aléatoire et sont ensuite modifiés de façon itérative afin de réduire la fonction de coût. En pratique, cette optimisation se fait par descente de gradient, une méthode

d'optimisation convexe qui permet de trouver le minimum de la fonction en se déplaçant dans la direction opposée au gradient. Plusieurs types de fonctions de perte sont utilisés pour répondre à des problèmes bien spécifiques.

Pour les problèmes de régression, consistant à prédire une valeur continue, deux fonctions de pertes sont très majoritairement utilisées. Il s'agit de la Mean Absolute Error (MAE) aussi appelée *l1 loss* (voir équation 1.14), et de la Mean Squared Error (MSE) aussi appelée *l2 loss* (voir équation 1.15) qui calculent une erreur entre les données réelles X et les données prédites Y .

$$MAE = \sum_{i=1}^D |x_i - y_i| \quad (1.14)$$

$$MSE = \sum_{i=1}^D (x_i - y_i)^2 \quad (1.15)$$

Ces deux fonctions de coût ont toutes deux pour propriété d'éviter les valeurs négatives et donc l'annulation réciproque des erreurs. Dans la MSE, l'erreur augmente de manière quadratique et de ce fait les erreurs importantes sont fortement pénalisées. La contrepartie est que l'effet des erreurs relativement petites en sera d'autant plus affaibli. Le choix de l'utilisation de l'une ou l'autre des fonctions dans un réseau de neurone dépend alors des données et de l'application que l'on souhaite en faire.

Dans le cas de la classification, où l'on cherche à prédire des classes discrètes, les fonctions préférentielles, les fonctions les plus utilisées sont la Binary Cross Entropy (BCE), la Categorical Cross Entropy (CCE) et la Kullback Leibler Divergence (KLD) (voir respectivement les équations 1.16, 1.17 et 1.18).

$$BCE(x, y) = -(y \log(p) + (1 - y) \log(1 - p)) \quad (1.16)$$

$$CCE = - \sum_{c=1}^M y_{o,c} \log(p_{o,c}) \quad (1.17)$$

$$KLD(\hat{y}||y) = \sum_{c=1}^M \hat{y}_c \log \frac{\hat{y}_c}{y_c} \quad (1.18)$$

où M est le nombre de classes, y est un indicateur binaire (0 ou 1), si la classification c est correcte pour l'observation o , p est la probabilité prédite que l'observation o appartienne à la classe c et \hat{y} la classe prédite par le réseau de neurones. La BCE correspond à la CCE appliquée au cas binaire. Ces trois fonctions visent toutes à minimiser la dissimilarité entre deux distributions de probabilités [99, 100].

1.5.2 Réseau de neurone à convolution (CNN)

Convolution

Dans le cas de la vision par ordinateur, le type de réseau le plus utilisé fait appel à l'opérateur de convolution pour apprendre la correspondance entrées-sortie. Il s'agit des Convolutional Neural Network (CNN), réseaux qui ont permis d'excellentes performances dans le traitement vidéo, la reconnaissance d'objets, la classification et la segmentation d'images, entre autres applications devenues maintenant courantes [97, 78].

Les premiers travaux sur les CNN modernes ont été proposés par Yann LeCun [87]. En démontrant qu'un CNN qui agrège des caractéristiques plus simples en caractéristiques progressivement plus compliquées peut être utilisé avec succès pour la reconnaissance de caractères manuscrits, il ouvre la voie au développement des CNN dans d'autres domaines. Bien que l'intégration de la rétropropagation soit introduite dans les CNN en 1989 [89], leur véritable essor viendra en 2012 à la suite d'une compétition d'analyse d'image (ILSVRC) où le CNN baptisé AlexNET dominera la compétition, démontrant ainsi tout le potentiel de ces réseaux de convolution.

Un produit de convolution est une opération de multiplication d'une matrice, par exemple une image, par une autre matrice, généralement plus petite, appelée matrice de convolution, masque de convolution, filtre ou encore noyau (kernel), afin de produire une nouvelle matrice. Dans le cas de l'image, la convolution permet de modifier le spectre spatial, par exemple atténuer ou accentuer les hautes fréquences ou bien réduire le bruit. Dans le cas des CNN, la taille la plus utilisée pour le masque de convolution est de 3×3 [157], cependant on trouve également des filtres 5×5 ou 7×7 [83].

Pour une image f et un masque de convolution g , l'image filtrée s'obtient en effectuant le produit de convolution entre f et g tel que défini par l'équation 1.19, où les indices x et y sont les coordonnées dans l'image f et les indices i et j , les coordonnées du noyau g .

$$\forall (x, y) \in Z^2, (f * g)(x, y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i, j)g(x - i, y - j) \quad (1.19)$$

Le noyau de convolution g contient les coefficients de pondération. Il comporte une opération de translation par le vecteur (x, y) et possède une symétrie centrale du fait de l'inversion des coordonnées i et j , il joue alors le rôle de fenêtre glissante sur l'image f . La symétrie centrale est avant tout due à un détail technique qui implique que la convolution ne peut s'appliquer sur les rangées de pixels bordant l'image. En Deep Learning ce problème est contourné en ajoutant une ligne de pixels sur chaque bord, opération appelée *padding*.

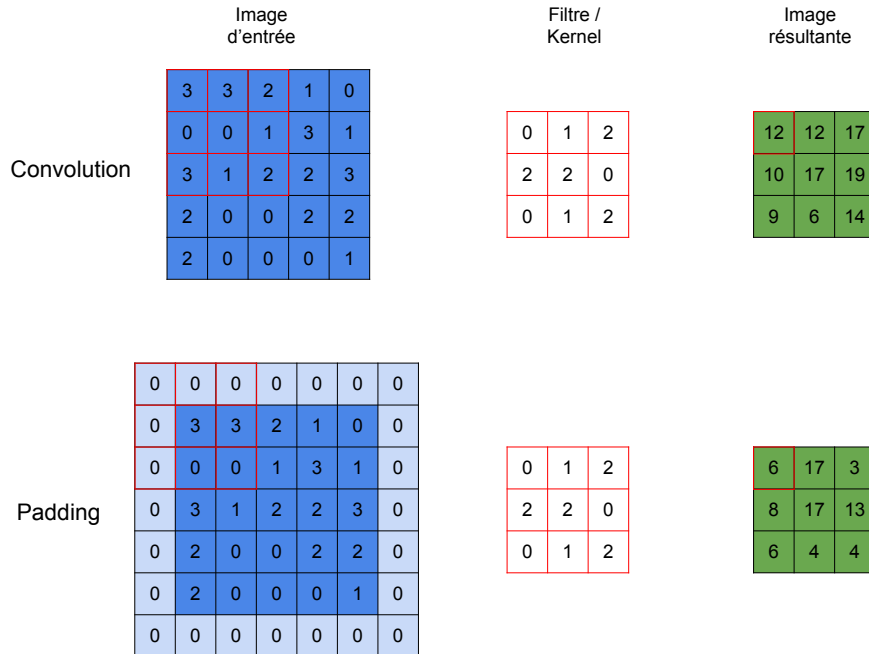


FIGURE 1.19 – En haut une convolution classique, en bas une convolution avec bordure (padding), les valeurs de l’image résultante s’en trouvent modifiées sur les bords.

Pour les images 2D, l’opération de convolution fonctionne comme indiqué dans la figure 1.19. Dans un CNN, l’image d’entrée est convoluée avec un filtre qui correspond au poids du réseau et en résulte une image de sortie qui est une carte des caractéristiques (*feature map*) mises en exergues par le filtre. En ajustant le poids des filtres pendant l’apprentissage, les CNN peuvent extraire des caractéristiques différentes et leur donner plus ou moins d’importance.

Convolution à trous

Les convolutions à trous, ou convolutions dilatées, reposent sur la même formule que les convolutions discrètes, mais comportent en plus un facteur de dilatation l , comme indiqué dans l’équation 1.20 et où P est le champ réceptif.

$$s(P) = (f *_l g)(P) = \sum_{s+lt=P} f(s)g(t) \quad (1.20)$$

Les convolutions dilatées ou à trous, sont utilisées pour augmenter ce que l’on appelle le champ réceptif des unités de sortie sans augmenter la taille du filtre de départ. Le champ réceptif est un carré dont la taille augmente de façon exponentielle avec la dilatation,

comme illustré dans la figure 1.20. Il est défini comme étant la taille de la région sur l'image d'entrée sur laquelle va s'appliquer le filtre de convolution.

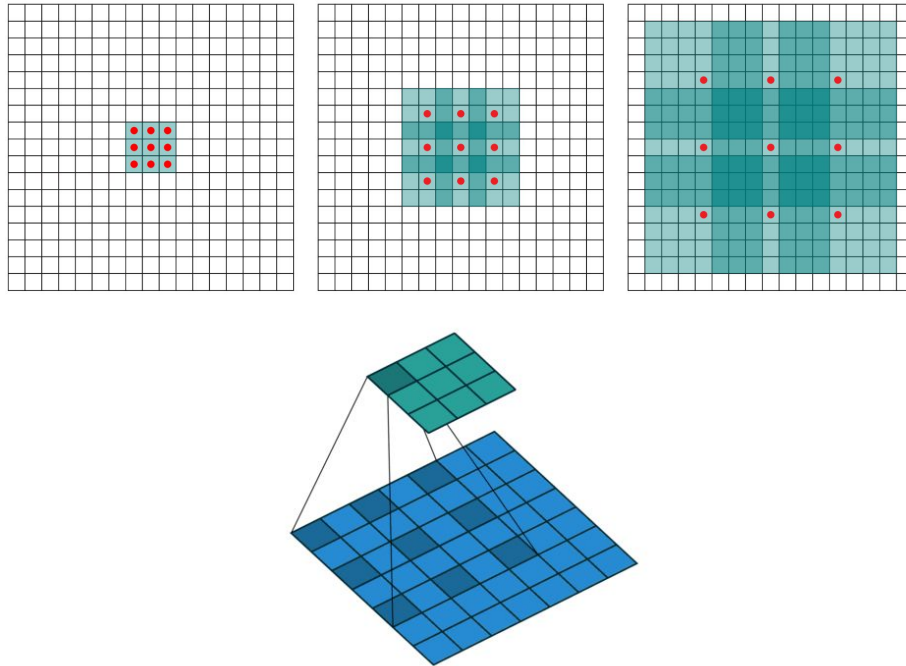


FIGURE 1.20 – Champ réceptif de la convolution à trous. À gauche, si l'espacement est égal à 1, cela correspond à une convolution classique. Au centre, si l'espacement est égal à deux, le champ réceptif augmente. À droite, l'espacement est de 3, le champ réceptif augmente de manière exponentielle. Adapté de [36]

Les convolutions à trous permettent donc intégrer un contexte plus global sans pour autant perdre en résolution [195]. Par ailleurs, si les convolutions dilatées avec de grands taux de dilatation permettent de capturer des contextes plus larges sans augmenter les coûts de calcul, elles introduisent également plus de zéros entre les valeurs du filtre et, de ce fait, perdent plus d'informations de contexte local dans les petites régions de l'image [179]. Pour les convolutions dilatées avec de faibles taux de dilatation, elles peuvent être utilisées pour améliorer les performances de localisation des CNN, mais elles produisent également des cartes de caractéristiques avec des champs réceptifs plus étroits et donc une moins bonne extraction du contexte. Une façon de contourner ces problèmes est d'utiliser une architecture en cascade de couches de convolution dilatée pour extraire des informations contextuelles à plusieurs échelles. Cette architecture en cascade est composée de couches de convolution consécutives, chacune d'entre elles utilisant des noyaux de convolution dilatés avec le même taux de dilatation pour produire des cartes de caractéristiques denses ayant le même champ réceptif [190].

Composition des blocs de convolution

Les CNN sont habituellement composés de différents types de couches, ils ne contiennent pas que des couches de convolution.

On retrouve au sein des réseaux de neurones, des couches dites d'*activation*, qui appliquent une fonction différentiable et non linéaire qui détermine si le neurone doit être activé ou non en fonction de la donnée d'entrée. Les fonctions les plus utilisées sont la fonction *sigmoïde* $f(x) = \frac{1}{1+\exp -x}$ qui fait correspondre à toute entrée, une sortie comprise entre 0 et 1 majoritairement utilisée pour la classification binaire ; ainsi que la fonction *ReLU* (Rectified Linear Unit) qui fait correspondre toute entrée à une valeur positive $f(x) = \max(0, x)$. Cette fonction est utilisée dans les couches intermédiaires, mais pas dans la couche de sortie. Sa variante, *Leaky ReLU* permet elle de prendre en compte les valeurs négatives tout en réduisant fortement leur valeur. Finalement, la fonction *Softmax* $\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$ pour $i = 1, 2, \dots, K$ permet de transformer un vecteur réel en vecteur de probabilité, elle est le plus souvent utilisée dans la couche finale d'un modèle de classification, notamment pour les problèmes multiclasse [2].

Un autre couche très utilisée dans les CNN est la couche dite de *pooling* (figure 1.21). Cette couche permet d'agréger des valeurs en sortie de convolution, pour condenser l'information et rendre le CNN invariant aux petites déformations locales, appelée invariance locale à la translation. La couche la plus commune est le *max-pooling*, généralement par un carré de 2 pixels de côté, qui vient appliquer l'opérateur maximum sur les 4 pixels de son champ réceptif. Il existe aussi un opérateur *average-pooling* qui vient alors moyenner le champ réceptif. Ainsi, alors que le *max-pooling* donne la caractéristique la plus importante dans un champ réceptif particulier d'une *feature map*, l'utilisation du *average-pooling* donne, lui, la moyenne des caractéristiques présentes.

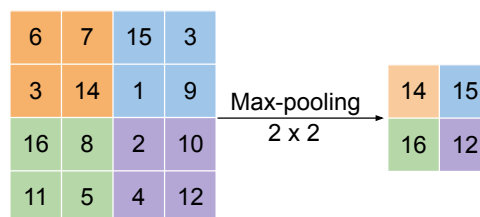


FIGURE 1.21 – Illustration de la couche de pooling, ici le max-pooling.

L'organisation des CNN se fait en répétant plusieurs fois de suite les différentes couches mentionnées précédemment. Cela constitue ce que l'on dénomme des "blocs de convolution", comme montré sur la figure 1.22.

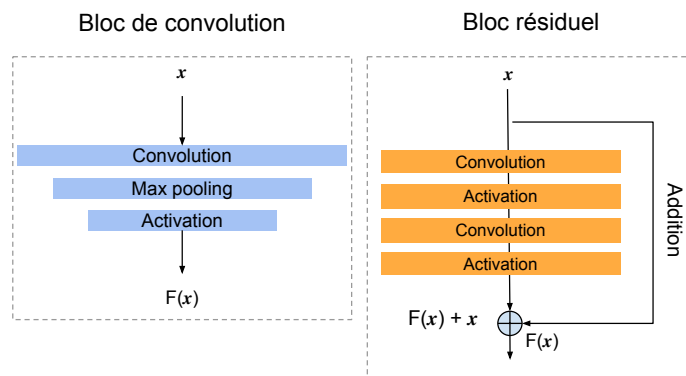


FIGURE 1.22 – Exemples de différents types de blocs de convolution. À gauche, un bloc de convolution classique, utilisé pour la réduction spatiale (downsampling). À droite, un bloc de convolution résiduel utilisé pour propager l'information dans la profondeur du réseau de neurones.

La composition usuelle d'un bloc de convolution contient une première couche de convolution, suivie d'une couche d'activation puis d'une couche de pooling. Le CNN complet est alors composé de plusieurs de ces blocs qui se suivent. L'opération de pooling divisant par 2 la dimension des images, en appliquant plusieurs blocs de manière consécutive, on obtient rapidement de très petites *feature map*. Ceci ne pose pas de problème dans le cas de la classification d'image, en revanche pour des tâches telles que la segmentation, la reconstruction ou la super résolution, cela implique d'avoir une image en sortie du réseau de neurone dont la taille est au moins équivalente à la taille de l'image d'entrée. Pour y parvenir, la méthode usuelle consiste à utiliser un autre type de couches qui permettent d'augmenter (up-sampling) les dimensions spatiales des *feature map* intermédiaires. Il s'agit de la convolution transposée, qui permet d'inverser les opérations de sous échantillonnage induites par les blocs de convolution.

Bloc résiduel

À mesure que les réseaux deviennent de plus en plus profonds, un problème de dégradation a été mis en évidence : plus la profondeur du réseau augmente, plus la précision se dégrade rapidement [59]. Cette dégradation n'est pas causée par un sur-apprentissage (overfitting) et l'ajout de couches supplémentaires à un modèle suffisamment profond conduit à une erreur d'apprentissage plus élevée [161]. La solution couramment apportée à ce problème est l'utilisation de bloc de convolution dit "résiduels". Si on considère une entrée x et une fonction $F(x)$ qui transforme cette entrée x en *feature map*, si la transformation ne change pas la dimension de x alors, il est possible d'associer x et $F(x)$ dans

une fonction résiduelle ou la sortie y ou $y = F(x) + x$ (voir figure 1.22).

Dans le réseau de convolution, l'opération $F(x) + x$ est réalisée en faisant un court-circuit puis une somme entre l'image d'entrée x et l'image convoluée $F(x)$, ce qui n'introduit pas davantage de paramètres ni de complexité dans le réseau [60] et permet de propager l'information spatiale dans la profondeur du réseau. Ces types de blocs peuvent également être utilisés pour l'étape de up-sampling et dans certaines conditions, l'utilisation de blocs résiduels pour la convolution transposée améliore les performances des réseaux génératifs en préservant les caractéristiques utiles, depuis l'espace des caractéristiques de faible dimension vers l'espace des caractéristiques de haute dimension [185].

1.5.3 Réseaux de neurones génératifs

Au cours de ces travaux, nous utiliserons pour certaines tâches un type de CNN bien spécifique, à savoir des réseaux de neurones génératifs et plus spécifiquement des réseaux de neurones génératifs adverses, appelés GAN pour Generative Adversarial Networks. Il existe un autre grand type de réseau génératif, l'auto-encodeur variatif (VAE), qui ne sera pas traité ici.

GAN

Les GAN appartiennent à la famille de modèle dit génératif, c'est-à-dire qu'ils tentent d'apprendre une fonction de densité de probabilité à partir d'un ensemble d'apprentissage, puis de générer de nouveaux échantillons tirés de la même distribution. Le but est alors de générer de nouvelles données synthétiques qui ressemblent à des données réelles.

De par leur nature versatile, les GAN ont eu un succès rapide dans des domaines aussi variés que l'augmentation de données, la super résolution d'image, la translation d'image à image et pour des applications allant de la médecine à l'astronomie (voir [32] pour une revue plus complète de l'étendue de l'application des GAN).

Les GAN ont été originellement proposés par [51] et fonctionnent en introduisant un entraînement adversarial entre deux réseaux de neurones.

Un GAN se compose d'un réseau générateur G et d'un réseau discriminant D qui sont en concurrence l'un avec l'autre dans un jeu minimax. Le générateur G applique une transformation aux données d'entrée, généralement un vecteur de bruit de Poisson, tandis que le discriminateur D est essentiellement un classificateur qui tente de déterminer si les données qu'il reçoit en entrée sont réelles (X) ou fausses (Z) c'est-à-dire créées par le générateur (voir figure 1.23).

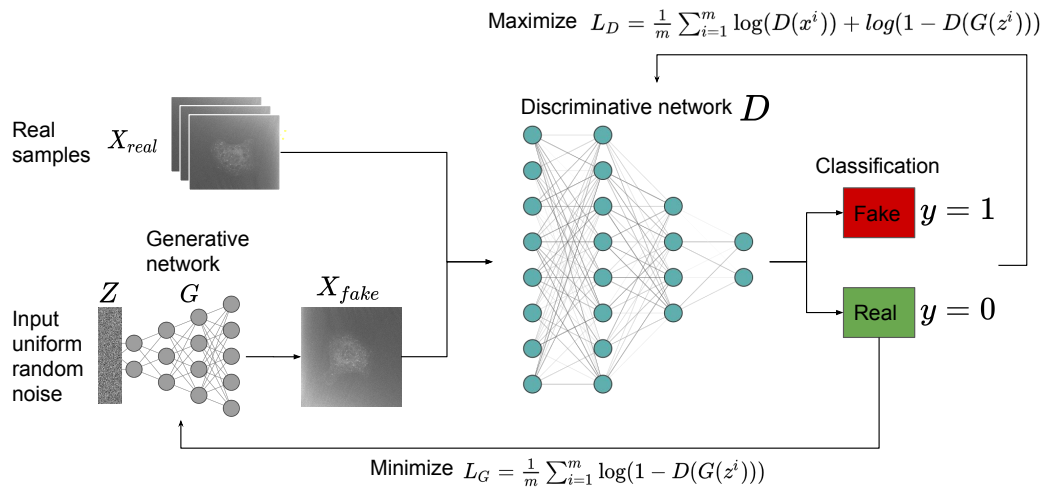


FIGURE 1.23 – Les GAN sont composés de deux réseaux de neurones, un générateur G générant des images et un discriminateur D différenciant les images réelles des fausses. La fonction de perte est calculée séparément pour chaque réseau et rétro-propagée à travers G et D individuellement.

Pour deux ensembles de données : X contenant des données réelles et Z contenant les données créées par G , le but de D est de maximiser les différences entre les deux distributions de données de X et Z afin de faciliter la discrimination entre les données fausses et réelles. Le Générateur, quant à lui, cherche à minimiser ces différences. L'intuition derrière ce jeu entre D et G , est que G doit produire des sorties X qui sont proches de Z afin de gagner le jeu.

Plus précisément, la fonction de perte de D est de maximiser la probabilité logarithmique de prédire 1 pour une donnée réelle $\log(D(x))$ et de maximiser la probabilité logarithmique de prédire 0 lorsqu'il s'agit d'une image générée $\log(1 - D(G(z)))$. Dans le même temps, G essaie de minimiser la probabilité logarithmique que D prédise correctement. Les deux fonctions de perte qui jouent l'une contre l'autre sont définies comme suit :

$$\begin{aligned} \mathcal{L}_D &= \max E[\log(D(x))] + E[\log(1 - D(G(z)))] \\ \mathcal{L}_G &= \min E[\log(1 - D(G(z)))] \end{aligned} \tag{1.21}$$

En utilisant les équations de 1.21, la fonction objective de l'ensemble du GAN peut être décrite comme un jeu minimax entre G et D . Soit $p_{data}(x)$ la distribution des données X et p_z la distribution du vecteur de bruit entré dans G , on obtient la fonction objective suivante :

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)}[\log D(x)] + E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (1.22)$$

Le problème d'optimisation consiste alors à trouver une solution qui maximise $D(x)$ et minimise $G(z)$. Dans ce contexte, pour un G donné, le D optimal, noté D^* est :

$$D^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \quad (1.23)$$

où $p_g(x)$ est la distribution des données générées par G . Si l'on remplace alors l'équation 1.23 dans l'équation 1.22, la fonction obtenue est une somme de divergence de Kullback-Leibler (KLD) :

$$\min_G \max_D V(D, G) = KL[p_{data}(x) \parallel \frac{p_{data}(x) + p_g(x)}{2}] + KL[p_g(x) \parallel \frac{p_{data}(x) + p_g(x)}{2}] - \log(4) \quad (1.24)$$

Le minimum de cette fonction est $-\log(4)$, ce qui est possible uniquement si les deux termes de l'équation 1.24 sont égaux à 0, soit uniquement possible si les deux probabilités $p_{data}(x) = p_g(x)$.

L'optimum quand $p_{data}(x) = p_g(x)$ est donc atteint quand $\frac{p_{data}(x) + p_g(x)}{2} = 0.5$, ce qui est possible quand $p_{data}(x) = 0.5$ and $p_g(x) = 0.5$

Concrètement, D prédit une valeur proche de 1 pour les images réelles et une valeur proche de 0 pour les images générées. Le but de cet entraînement adversarial est d'atteindre un point où D prédit 0.5 pour chaque classe (indécision), ce point est appelé l'équilibre de Nash. Lorsque cet équilibre est atteint, la fonction objective peut être exprimée sous la forme d'une KLD symétrique qui correspond à la divergence de Jensen-Shannon (JSD) entre deux distributions P et Q :

$$JSD(P \parallel Q) = \frac{1}{2} D_{KL}(P \parallel M) + \frac{1}{2} D_{KL}(Q \parallel M) \quad (1.25)$$

avec $M = \frac{1}{2}(P + Q)$

En fin de compte, la fonction objectif des GAN est donc une divergence de Jensen-Shannon entre la distribution des données réelles et celle des données factices créées par G .

Dans la pratique, cela impliquerait de mettre à jour les poids de D en utilisant l'ascension stochastique plutôt que la descente stochastique. Pour des raisons d'efficacité informatique, cette fonction est généralement mise en œuvre sous la forme d'une classification binaire utilisant la perte d'entropie croisée binaire (BCE), avec des étiquettes de 0 pour les images générées et de 1 pour les images réelles, ce qui permet d'utiliser la descente de gradient stochastique.

Ce jeu minimax entre deux fonctions de coût concurrentes présente quelques inconvénients durant l'entraînement. D'une part le jeu minimax peut rester instable et donc la descente du gradient peut ne jamais converger. D'autre part, si l'apprentissage favorise le générateur, la perte deviendra trop faible trop rapidement et le gradient sera trop petit pour mettre à jour correctement les poids du GAN. Ce problème, très courant, est connu sous le terme de disparition ou fuite de gradient (*vanishing gradient*). Enfin, un problème récurrent sur les GAN est la perte de mode (*mode collapse*), qui survient quand le générateur ne produit plus qu'un petit sous ensemble des sorties possibles. Le générateur concentre ses prédictions sur l'un des modes de la distribution cible et produit des points hautement réalistes, mais très peu diversifiés.

De nombreuses solutions ont été proposées pour circonvier à ces limitations [32], par exemple le LSGAN [107] ou le WGAN [3] qui utilisent respectivement la MSE et la distance de Wasserstein comme fonction de perte, qui permettent d'améliorer la stabilité de ce réseau [81].

À la suite des améliorations successives de la stabilité des GAN et de la qualité des images générées, de nombreux développements se sont tournés vers la génération d'images réalistes, à partir de vecteur de bruits dans un premier temps, et dès 2018, en apprenant à générer des images réalistes à partir d'autres images. Pionniers en la matière, les GAN Pix2Pix [74] et CycleGAN [201] se sont attachés à résoudre le problème de la translation image à image et du transfert de style.

CycleGAN

La notion de style peut être assez complexe à définir, nous choisirons ici une définition globale, c'est-à-dire qu'un style correspond à l'ensemble des caractères esthétiques d'une chose ou d'un ensemble. Dans le cas des images, il peut s'agir de textures particulières, de teintes, d'intensités, mais il peut s'agir aussi de caractéristiques implicites, bas niveau, qu'il peut être difficile de définir avec exactitude, par exemple la différence entre deux styles de deux peintres différents. C'est d'ailleurs avec cet exemple que le CycleGAN a été introduit, en convertissant des images du style de Monet vers celui de Van Gogh ou Cézanne [201]. CycleGAN introduit d'autres applications possibles comme le changement de textures et de couleurs, le remplacement d'une peau de cheval par une peau de zèbre

et il a été utilisé pour des tâches nécessitant un décodage de caractéristiques bas niveau comme changer la saison dans des photographies de paysages ou coloriser des photos en noir et blanc [201].

L'intérêt du CycleGAN par rapport aux méthodes précédentes est de proposer la translation d'images entre deux domaines de manière totalement non supervisée sans avoir recours aux paires "image x " - "image x annotée" comme le requiert le mode d'apprentissage supervisé. Le CycleGAN permet alors, à partir de deux collections d'images A et B de différents styles (parfois appelés domaines), de réaliser un traducteur de l'un à l'autre en utilisant des réseaux de neurones de type U-Net ou ResNet.

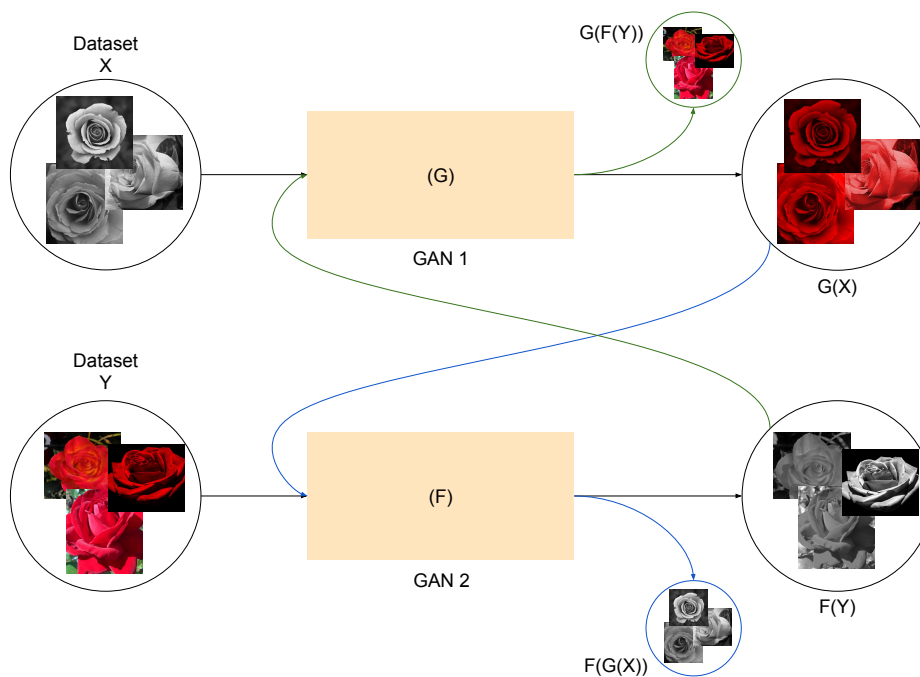


FIGURE 1.24 – Les CycleGAN sont composés de deux GAN distinct, G transformant les données X qui sont ici des images en niveau de gris en images RGB $G(X) \approx Y$ et F transformant les données Y en données $F(Y) \approx X$. Pour assurer une consistance cyclique, les images RGB $G(x)$ sont à nouveau transformées en images en niveaux de gris par le GAN F , de même, les images $F(y)$ sont re-colorisées par le GAN G .

Pour ce faire, le CycleGAN introduit une paire supplémentaire de discriminateur et de générateur par rapport au GAN. Un générateur apprend la conversion $G : X \rightarrow Y$ tandis que l'autre apprend $F : Y \rightarrow X$, comme présenté sur la figure 1.24. Les discriminateurs respectifs entraînent ces générateurs de manière adversariale. Dans ce cadre, un cycle est formé à partir de conversion inversée, c'est-à-dire que chaque image x passant dans G et donnant une image transformée y , suivra également le chemin inverse ou y passera

dans F pour donner une image x' reconstruite. Cela permet d'éviter la perte de cohérence entre les deux GAN et les deux domaines d'images, en assurant que $F(G(X)) \approx X$ et que $G(F(Y)) \approx Y$, avec une fonction de perte cyclique présentée dans l'équation 1.26.

$$\mathcal{L}_{cyclique}(G, F) = E_{x \sim p_{data}(x)}[\|F(G(x)) - x\|] + E_{y \sim p_{data}(y)}[\|F(G(y)) - y\|] \quad (1.26)$$

La fonction objective totale du CycleGAN est alors composée de la fonction de perte de chaque GAN, et de la fonction de perte cyclique (équation 1.27).

$$\mathcal{L}(G, F, D_x, D_x) = \mathcal{L}_{GAN}(G, D_Y) + \mathcal{L}_{GAN}(F, D_X) + \mathcal{L}_{cyclique}(G, F) \quad (1.27)$$

De cette façon, les mêmes images d'entrée sont utilisées pour contraindre la cohérence du cycle, ce qui libère l'exigence d'images appariées. Les générateurs sont conditionnés par leur image d'entrée au lieu du vecteur de bruit d'un GAN classique.

Les GANs en microscopie

L'utilisation des GAN et CycleGAN en analyse d'imagerie biologique reste pour l'heure relativement limitée, on peut citer néanmoins CytoGAN [49], qui utilise des GAN pour effectuer du profilage morphologique de cellules humaines imagées par microscopie à fluorescence. Les auteurs de [58] ont également utilisé un GAN pour produire des images réalistes de tissus imagés par microscopie électronique, tandis que les auteurs de [125] ont utilisé un GAN pour produire des image de cellules en microscopie de fluorescence. Le CycleGAN a été utilisé par [184] pour augmenter la résolution d'images de microscopie et [94] utilise également un CycleGAN pour réaliser de la déconvolution sur images de microscopie de type SMLM.

1.6 Objectif global de la thèse

L'analyse des images de microscopie de fluorescence a bénéficié d'avancées majeures au cours des dernières années, notamment dû au regain d'intérêt suscité par l'introduction de méthodes d'apprentissage profond. Ces avancées ont permis un grand nombre d'études basées sur la quantification du signal dans les données d'imagerie cellulaire. Néanmoins, un certain nombre de questions restent non résolues et rendent la détection du signal et la segmentation des objets d'intérêt complexe.

Dans cette thèse, nous proposons d'apporter des réponses à certains de ces challenges non pas en s'attaquant directement à la détection du signal et à la segmentation des objets dans les images de microscopie de fluorescence, mais plutôt en cherchant à mettre en place des méthodes de Deep Learning permettant d'améliorer le signal des images de microscopie en vue de faciliter les analyses en aval. Les développements principaux s'appuient sur le Deep Learning et concernent le rehaussement du signal des spots d'ARNm à l'intérieur des images smFISH, ainsi que le rehaussement du signal des noyaux cellulaires dans les images ayant un faible ratio signal sur bruit. Enfin, nous proposons un cadre statistique rigoureux pour étudier les distributions spatiales dans les cellules pour permettre l'observation des états cellulaires en aval de la quantification du signal, rendue possible par nos méthodes de rehaussement de signal.

Chapitre 2

Rehaussement du signal dans les images smFISH pour une détection plus précise des ARNm

2.1 Contexte scientifique

La technique smFISH permet d'établir le profil d'expression des gènes avec une précision subcellulaire pour déterminer les états moléculaires de divers types de cellules [77] et ce, à haut débit [7]. Le répertoire des méthodes de quantification de l'expression de l'ARNm est vaste et comprend la microscopie smFISH, clamp-FISH, amp-FISH et des versions multiplexées (MerFISH), toutes permettent la localisation de l'ARNm au niveau subcellulaire. S'il existe des différences technologiques entre ces méthodes en termes de nombre d'ARNm détectés et de nombre de cellules traitées, toutes produisent des données d'imagerie avec des spots d'ARNm qui peuvent être associés à des coordonnées (x, y) .

L'automatisation de l'acquisition d'images a entraîné une augmentation du nombre de projets à haut débit ayant une composante de localisation de transcrits. Ainsi, le besoin de méthodes de détection précises et automatisées des spots d'ARNm dans les images de microscopie à fluorescence est devenu important. Malgré les progrès réalisés ces dernières années, il s'avère encore difficile de localiser les spots correspondant à différents ARNm de manière entièrement automatisée, pour des raisons principalement liées à l'intensité et l'irrégularité du bruit de fond induit par l'autofluorescence et la liaison non spécifique des sondes. De plus, les images smFISH peuvent avoir un faible rapport signal/bruit (SNR) impactant la capacité des algorithmes à distinguer les signaux faibles des bruits forts, la frontière entre le fond (bruit) et le signal (spots) n'étant alors pas toujours évidente [199].

Dans ce contexte, le principal inconvénient des méthodes classiques de détection des spots d'ARNm est la nécessité d'une intervention humaine importante pour déterminer les meilleurs paramètres des algorithmes de détection (voir chapitre 1, section 1.3). Souvent,

de petites différences dans les caractéristiques des images conduisent à la nécessité d'un réglage des paramètres [23]. La qualité de la détection dépend alors largement de la capacité de l'utilisateur à choisir correctement les paramètres de la méthode de détection, en fonction des propriétés de chaque image (contraste, spots, artefacts, bruit).

Dans ce chapitre, nous étudierons l'utilisation de réseaux de neurones à convolution dédiés au rehaussement de l'intensité du signal de fluorescence (correspondant à la troisième étape des pipelines classiques de détection du signal, voir figure 1.9), dans les images de microscopie et permettant ainsi leur détection en aval par des outils conventionnels largement utilisés par la communauté d'analyse d'images de microscopie, sans qu'il soit nécessaire de régler finement les paramètres. Nous présentons dans ce chapitre la méthode DeepSpot, et nous montrons qu'il est possible d'éviter les étapes de réglage manuel des paramètres en renforçant le signal de tous les spots de manière à ce qu'ils aient la même intensité sur toutes les images, indépendamment du contraste, du bruit ou de la forme des spots.

2.2 Contribution scientifique

2.2.1 Réseau de neurones pour le rehaussement du signal de petits objets

La problématique de la détection de spots d'ARNm peut être reliée à celle de la détection de petits objets en analyse d'images. Cette dernière reste une partie difficile du problème général de détection d'objets en raison des informations limitées contenues dans les petites régions d'intérêt. Par exemple, il a été démontré que la taille de l'objet a un impact majeur sur la précision des réseaux de détection d'objets de Deep Learning tels que VGG, ResNet ou Inception V2 (voir figure 2.1) [68].

L'hypothèse souvent avancée est que les petits objets ne contiennent pas suffisamment d'informations sémantiques du fait de leur taille [98], le défi devient alors de capturer les caractéristiques sémantiques de l'objet tout en gardant suffisamment d'informations spatiales [46]. En outre, il a été démontré que l'ajout d'un contexte supplémentaire améliore la détection des petits objets [9, 46, 122]. Une solution élégante consiste à utiliser la convolution dilatée (aussi appelée convolution à trous ou *sparse convolution*), car le champ réceptif peut être étendu sans perte de résolution et de ce fait capturer davantage de contexte sans perdre d'information spatiale [57, 196].

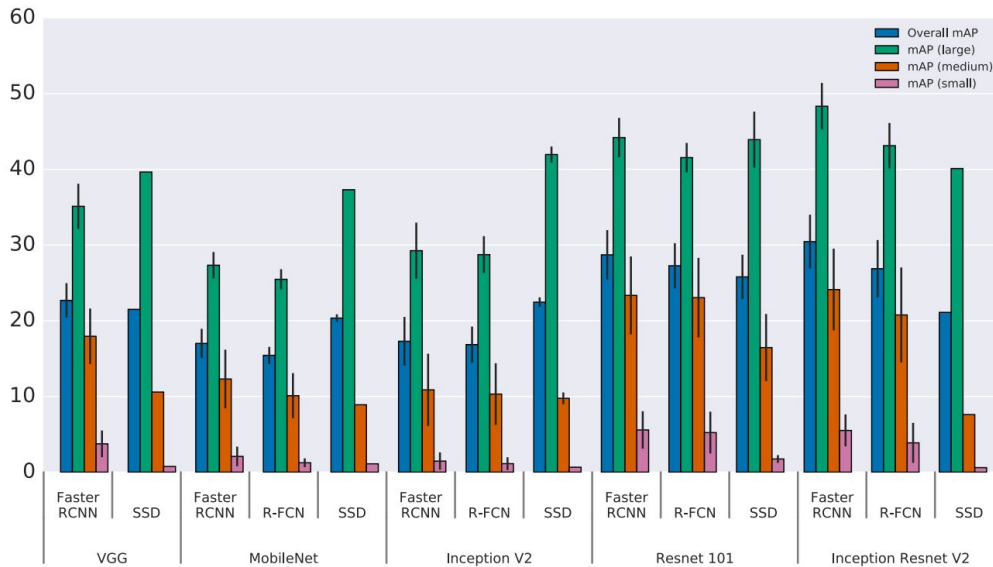


FIGURE 2.1 – Précision (*mean average precision (mAP)*) de la détection par réseaux de neurones d’objets en fonction de leur taille. En vert les grands objets, en orange les objets de taille moyenne et en rose les petits objets. L’analyse a été effectuée sur le jeu de données COCO pour *Common Objects in Context*, l’un des principal jeu de données de benchmark de réseau de neurones. Figure reprise de [68].

Architecture du réseau de neurones

Au lieu de considérer la détection elle-même, nous nous intéressons ici plus particulièrement au rehaussement du signal, une technique de traitement d’image visant à augmenter le signal uniquement dans les régions de l’image où se trouvent les objets d’intérêt. Dans notre cas, les objets d’intérêt sont les spots d’ARNm au sein des images smFISH. Pour cela, nous avons défini et implémenté un réseau de neurones permettant d’agrèger l’information contextuelle et spatiale.

Comme il a été démontré que l’utilisation d’éléments d’image au-delà de l’étendue de l’objet (contexte) améliore les résultats de la détection de petits objets [67], nous nous en sommes inspirés pour développer un module d’agrégation de contexte autour des spots d’ARNm pour rehausser leur signal. Ce module est un réseau à chemins multiples qui vient se placer en avant de la partie centrale du réseau (voir figure 2.2). Il prend en entrée une image et la traite le long de trois chemins différents.

1. Le premier chemin est composé de blocs de convolution traditionnels (convolution 2D, normalisation par lots, activation et pooling max). Ces blocs, souvent utilisés dans les CNN, sont particulièrement efficaces pour renforcer l’information

sémantique au détriment de l'information spatiale.

2. Le deuxième chemin n'utilise que des convolutions, sans opération de *max pooling* pour éviter que les spots faibles ne soient éliminés en ne gardant que l'information maximale.
3. Le troisième chemin utilise la convolution à trous [57] afin d'étendre le champ réceptif du réseau et intégrer davantage de contexte.

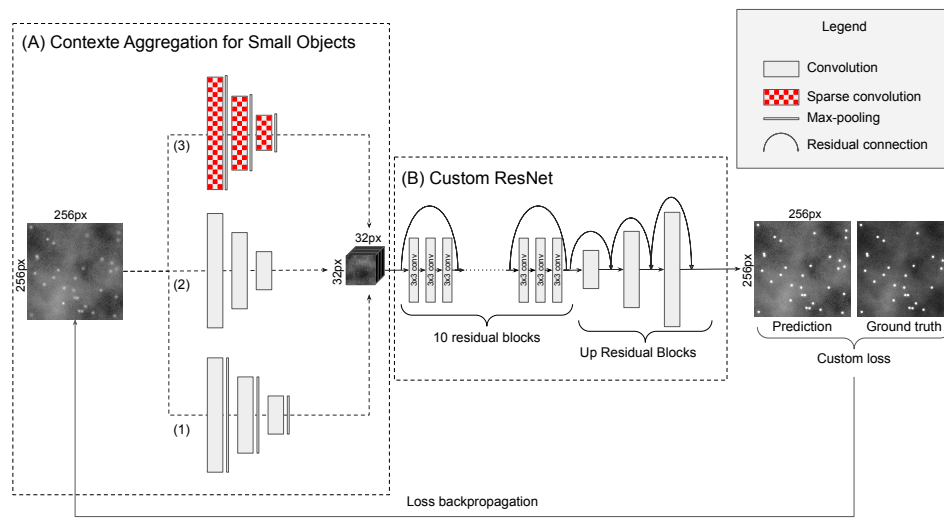


FIGURE 2.2 – L'architecture du réseau DeepSpot est composée du module *Context Aggregation for Small Objects*, constitué d'un réseau à chemins multiples (A) et d'un réseau central étant un ResNet sur mesure (B). Une fonction de coût sur mesure est utilisée pour l'entraînement du réseau.

En ce qui concerne la partie centrale (panneau B sur la figure 2.2) et terminale, nous avons modifié l'architecture du ResNet pour créer une solution sur mesure qui permet de propager les informations récoltées par le module d'agrégation de contexte tout au long du réseau, avant de reconstituer l'image avec un signal des spots rehaussé, via l'utilisation de blocs de convolution transposées résiduels.

Fonction de coût

Nous avons défini une fonction de coût sur mesure, définie comme une combinaison des fonctions d'entropie croisée binaire (BCE) et d'erreur quadratique moyenne (MSE) (voir le chapitre 1, section 1.5 pour plus de détails). Le terme principal de la fonction de coût est la BCE, définie par $\mathcal{L}_{\text{BCE}}(x, \hat{x}) = -(x \log(\hat{x}) + (1 - x) \log(1 - \hat{x}))$ qui mesure la différence entre les images prédites par le réseau \hat{x} et les images de référence (*ground truth*) x . Bien que cette fonction de coût soit principalement utilisée pour la classification,

elle peut également être utilisée pour la segmentation et pour le rehaussement de signal grâce à ses performances dans la comparaison pixel à pixel [75].

À ce terme principal BCE , nous avons ajouté un terme de régularisation défini par l'erreur quadratique moyenne (MSE), $\mathcal{L}_{MSE} = \frac{1}{n} \sum_{i=1}^n (\max(x_i) - \max(\hat{x}_i))^2$, calculée entre la valeur maximale de l'image prédite $\max(\hat{x}_i)$ et la valeur maximale de l'image de référence $\max(x_i)$. Cette régularisation pousse le réseau à produire des spots dont l'intensité est proche de 255 (valeur maximale des images sur 8 bits), et uniformise donc l'intensité du rehaussement du signal dans les images de sortie, ce qui facilite par la suite la détection automatique des spots. L'intuition derrière l'utilisation de la MSE est dirigée vers les images issues d'acquisition 2D ou d'images 3D aplaties, dans lesquelles des spots hors focus sont présents et dont le signal est plus faible que les spots dans le plan focal. La fonction de coût complète est alors $\mathcal{L}_{Totale} = \mathcal{L}_{BCE} + \mathcal{L}_{MSE}$.

2.2.2 Rehaussement du signal des ARNm dans les images smFISH

Jeux de données

Nous avons utilisé le jeu de données expérimentales smFISH de la publication [26], pour lequel nous disposons des coordonnées de spots, car les auteurs ont effectué leur détection de spots à l'aide de l'outil BIG-FISH [72]. Le jeu de données résultant est composé de 1553 images représentant 27 gènes différents.

Nous avons également utilisé le jeu de données de la publication deepBlink [39], composé de 129 images smFISH acquises à partir de 4 conditions de cultures cellulaires différentes. Les auteurs ont fourni leur annotation *ground truth* des localisations des spots d'ARNm réalisée à l'aide de TrackMate. Ce jeu de données n'est pas utilisé pour l'entraînement de DeepSpot, il est uniquement utilisé à des fins de test.

En plus des jeux de données expérimentaux, nous avons construit 20 jeux de données simulés avec des bruits de fond de motif et d'intensité aléatoire et des spots de taille, intensité et densité également aléatoires.

Méthode d'évaluation

Nous désignons par $D(X) = \{p_1, \dots, p_n\}$ les points détectés par le *spot detector* du logiciel d'analyse d'image ICY [124] à partir d'une image X ou le signal des spots a été rehaussé par DeepSpot et par $A(X) = \{q_1, \dots, q_m\}$, nous désignons l'annotation *ground truth* des coordonnées de spots. Pour prendre en compte le fait que m n'est pas nécessairement égal à n et que même pour un spot bien détecté, les coordonnées dans A et D peuvent légèrement différer, nous avons utilisé l'algorithme k -d tree [11]. Ce dernier

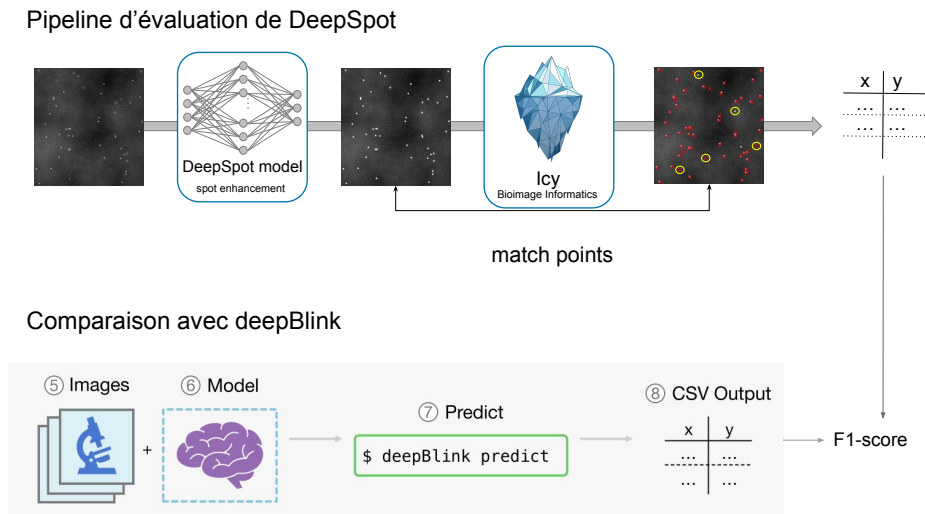


FIGURE 2.3 – Processus d'évaluation de DeepSpot. Les images originales sont données en entrée de DeepSpot, qui rehausse le signal des spots. Sur l'image où le signal a été rehaussé, les spots sont détectés par le *spot detector* de ICY [124]. Le résultat issu de ICY est ensuite comparé au *ground truth*, par un algorithme de *point matching*. ICY fournit également la liste des positions (x, y) des spots détectés. Les mêmes images originales sont fournies à deepBlink [39] qui renvoie également une liste des positions (x, y) des spots détectés. Afin de comparer les résultats des deux méthodes, le F1-Score par rapport au *ground truth* est calculé pour chaque liste de positions de spots.

permet d'obtenir une correspondance pour tous les points détectés, dans un rayon t (que nous avons fixé ici à $t = 3$) et de détecter également le nombre de faux négatifs et de faux positifs par rapport aux erreurs de correspondances entre $D(X)$ et $A(X)$. Les vrais négatifs sont définis comme étant tous les pixels p de la matrice de confusion tels que $p \in X \setminus \{A(X) \cup D(X)\}$, soit tout le fond de l'image, impliquant donc une très large sur-représentation des vrais négatifs dans la matrice de confusion. Pour cette raison, nous n'utiliserons pas les mesures de précision telles que l'AUC, mais plutôt des mesures comme le F1-score, qui ne prend pas en compte les vrais négatifs.

Intensité des spots

Pour éviter la sélection manuelle des paramètres de détection, il est impératif d'avoir une intensité homogène des spots sur l'ensemble du jeu de données, afin de n'utiliser qu'un seul jeu de paramètres pour toutes les images.

Dans notre évaluation, l'ensemble des jeux de données, les intensités des spots dont le signal a été rehaussé se situent entre 241 et 255, ce qui représente une variation de seulement 5,4 % par rapport à l'intensité maximale. Par ailleurs, au sein d'un même

jeu de données, les intervalles de confiance à 95% montrent une très faible dispersion, inférieure à 0.5% de l'intensité maximale. Ceci est suffisant pour séparer complètement les spots smFISH du fond de l'image dans les images où le signal a été rehaussé, et ce, pour la grande majorité des méthodes de détection.

2.2.3 Amélioration de la précision de détection des spots d'ARNm

À titre d'exemple, les résultats respectivement obtenus par DeepSpot et deepBlink sont visibles sur la figure 2.4.

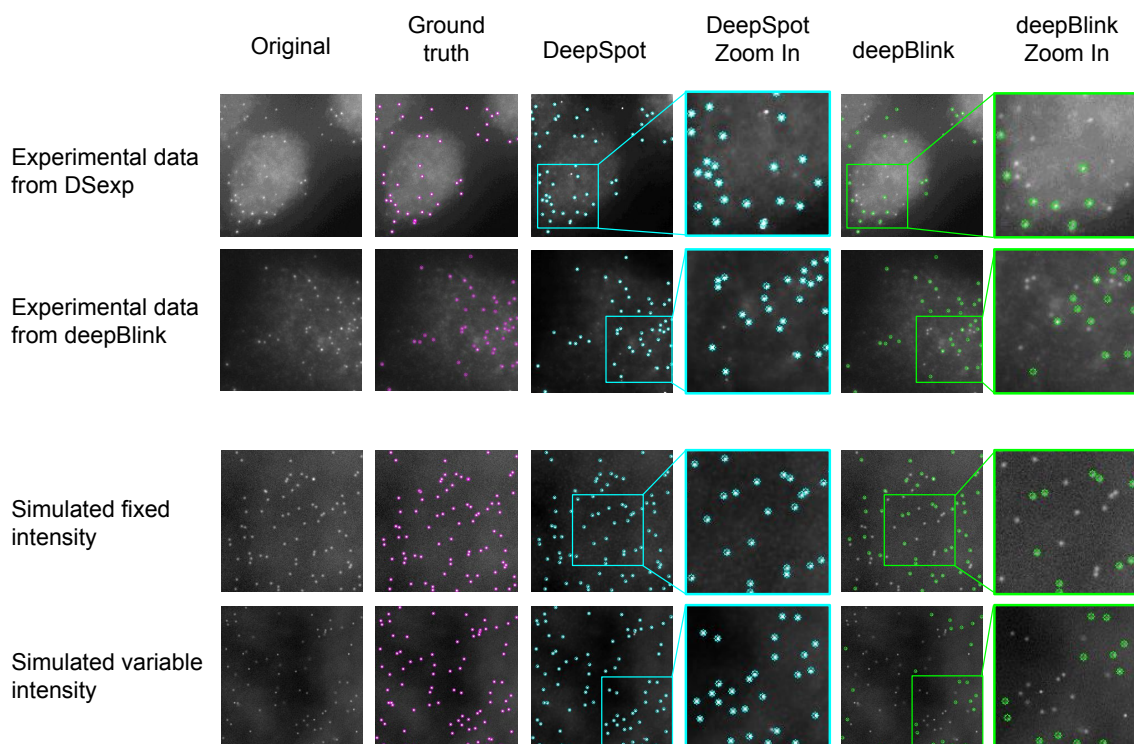


FIGURE 2.4 – Exemples de résultats obtenus avec DeepSpot et deepBlink sur les jeux de données de test expérimentaux (première ligne) et deepBlink (deuxième ligne), et les spots simulés avec des jeux de données à intensité fixe variable pour les troisième et quatrième lignes. Les cercles colorés indiquent les endroits où les spots ont été détectés par DeepSpot et deepBlink (en bleu et vert, respectivement). Dans la colonne *Ground truth*, les cercles roses indiquent les spots qui ont été précédemment annotés comme *Ground truth* par des méthodes indépendantes. Les deux dernières colonnes montrent un agrandissement aux positions indiquées par les rectangles de couleur pour DeepSpot et deepBlink.

Le F1 score global pour l'ensemble des jeux de données est légèrement supérieur à 97%, et aucune différence majeure dans la performance de DeepSpot ne peut être observée en fonction des conditions expérimentales ou des gènes étudiés. Cette constance dans la performance montre que DeepSpot a la capacité de rehausser le signal des spots de manière consistante dans des images aux caractéristiques variées.

Nous avons comparé la précision de la détection des spots par le modèle entraîné et proposé par les auteurs de deepBlink, avec celle de DeepSpot, à la fois sur nos jeux de données et sur le jeu de données fourni par les auteurs de deepBlink. Le tableau 2.1 montre les F1-score pour chaque jeu de données de test, expérimental et simulé.

TABLE 2.1 – Performance des modèles deepBlink et DeepSpot pour la détection des spots smFISH. Les F1-score sont calculés en faisant la moyenne des valeurs obtenues pour chaque image des jeux de données de test correspondant. Les valeurs entre parenthèses indiquent l'intervalle de confiance à 95%. Les valeurs les plus élevées sont indiquées en gras.

Model \ F1-score(%)	Données expérimentales	Données simulées	Données deepBlink
deepBlink	78.7 [75.26, 82.14]	57.68 [56.57, 58.79]	94.2 +- 0.141(std)
DeepSpot	94.43 [93.41, 95.45]	97.35 [97.27, 97.43]	87.92 [85.36, 90.49]

DeepSpot donne de manière générale un F1-Score plus élevé que celui de deepBlink sur nos jeux de données (Tableau 2.1). Cependant, nous avons constaté que deepBlink a de meilleures performances générales sur les données expérimentales que sur les données simulées, probablement parce que le modèle deepBlink n'a été entraîné que sur des données expérimentales. Les résultats de DeepSpot sont en revanche cohérents sur l'ensemble des jeux de données, y compris celui de deepBlink sur lequel DeepSpot n'a jamais été entraîné.

Comme attendu, les performances du modèle DeepBlink sont meilleures sur ses propres images smFISH, mais DeepSpot est parvenu à obtenir un score F1 de près de 88 %, un résultat globalement bon étant donné que DeepSpot n'a pas été entraîné sur les données de deepBlink. Ensemble, les résultats du tableau 2.1 indiquent que DeepSpot est une méthodologie robuste qui offre un modèle généraliste pour le rehaussement du signal des spots d'ARNm et garantit une détection de haute qualité des spots en aval sans avoir à effectuer de réglages de paramètres.

2.3 Discussion

Les méthodes récentes de microscopie smFISH sont capables de générer des milliers d’images, il est donc devenu impératif d’introduire des algorithmes capables de simplifier la détection des spots d’ARNm et en particulier d’éviter le réglage manuel des outils de détection.

Dans ce chapitre, nous avons présenté DeepSpot, une nouvelle architecture CNN spécialement conçue pour rehausser le signal des spots d’ARNm dans les images smFISH, permettant ainsi l’utilisation en aval d’algorithmes de détection de spots courants, tels que le détecteur de spots ICY, sans avoir à régler de paramètres. L’architecture de notre réseau introduit le module *Context Aggregation for Small Objects* qui s’appuie sur une convolution à trous pour fournir plus de contexte pour le rehaussement du signal des petits objets correspondant aux spots d’ARNm. Le réseau DeepSpot a été entraîné et testé sur 20 jeux de données simulés, ayant tous des caractéristiques de signal et de bruit différentes, ainsi que sur un jeu de données expérimental précédemment publié qui a été annoté pour les emplacements des spots. Nous avons montré (i) que notre approche obtient de meilleures performances lorsque l’entraînement est effectué sur des données dont l’intensité est très variable et (ii) que l’entraînement sur une combinaison de données expérimentales et simulées est une approche viable dans un contexte réel.

En outre, nous avons comparé les performances de la combinaison de DeepSpot et ICY à celles de la méthode deepBlink également basée sur de l’apprentissage profond et nous avons montré qu’en moyenne, DeepSpot permet une meilleure détection des spots d’ARNm que deepBlink. Nous avons constaté que la combinaison DeepSpot / ICY permet une bonne détection des spots sur les jeux de données de test correspondant aux jeux de données sur lesquels il a été entraîné, avec un score F1 moyen supérieur à 97%, mais a également obtenu des résultats de précision élevés sur des jeux de données totalement inconnus avec un score F1 de 88% pour les données fournies avec la publication deepBlink. Dans l’ensemble, les bons résultats obtenus sur des données connues et inconnues indiquent que DeepSpot est un modèle plus généraliste que deepBlink et qu’il atteint un bon équilibre entre le sur-apprentissage et le sous-apprentissage. Nous supposons que cette capacité de généralisation est probablement due à la fois à une forte régularisation au sein du réseau et à la diversité du signal fourni par la diversité des données d’entraînement.

Cependant, l’impact des fluorophores non spécifiquement liés n’a pas été évalué dans ce travail. Par conséquent, si DeepSpot est entraîné à distinguer les multiples variations d’intensité du signal, dans les cas où l’intensité du signal des sondes non spécifiquement liées est trop similaire en forme et en intensité au signal des sondes spécifiquement liées,

alors DeepSpot peut ne pas être en mesure de rehausser uniquement les spots correspondant aux sondes spécifiquement liées. Cette question, commune à toutes les méthodes de détection, mérite d'être approfondie.

En outre, il reste à déterminer si DeepSpot serait adapté à des systèmes plus complexes tels que les tissus et quelles adaptations en termes d'architecture et de données d'entraînement seraient nécessaires. De plus, même si l'architecture de DeepSpot a été conçue pour être facilement extensible à la 3D en changeant les couches de convolution 2D en couches 3D dans Tensorflow, les performances de DeepSpot en 3D n'ont pas été évaluées.

En conclusion, nous avons montré que le rehaussement du signal par DeepSpot permet une détection automatisée et une localisation précise des spots d'ARNm pour les méthodes d'analyse en aval et peut donc être un outil utile pour rationaliser non seulement la détection des spots, mais aussi les études d'enrichissement localisé en ARNm dans les cellules.

2.4 Article

RESEARCH ARTICLE

DeepSpot: A deep neural network for RNA spot enhancement in single-molecule fluorescence in-situ hybridization microscopy images

Emmanuel Bouilhol^{1,2*} , Anca F. Savulescu³, Edgar Lefevre², Benjamin Dartigues², Robyn Brackin⁴ and Macha Nikolski^{1,2*}

¹CNRS, IBGC, UMR 5095, Université de Bordeaux, Bordeaux, France

²Bordeaux Bioinformatics Center, Université de Bordeaux, Bordeaux, France

³IDM, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

⁴Advanced Medical Bioimaging CF, Charité—Universitätsmedizin, Berlin, Germany

*Corresponding authors. E-mails: emmanuel.bouilhol@gmail.com, macha.nikolski@u-bordeaux.fr

Received: 16 December 2021; **Revised:** 25 February 2022; **Accepted:** 31 March 2022

Keywords: Deep learning; mRNA spot detection; smFISH image analysis

Abstract

Detection of RNA spots in single-molecule fluorescence in-situ hybridization microscopy images remains a difficult task, especially when applied to large volumes of data. The variable intensity of RNA spots combined with the high noise level of the images often requires manual adjustment of the spot detection thresholds for each image. In this work, we introduce DeepSpot, a Deep Learning-based tool specifically designed for RNA spot enhancement that enables spot detection without the need to resort to image per image parameter tuning. We show how our method can enable downstream accurate spot detection. DeepSpot's architecture is inspired by small object detection approaches. It incorporates dilated convolutions into a module specifically designed for context aggregation for small object and uses Residual Convolutions to propagate this information along the network. This enables DeepSpot to enhance all RNA spots to the same intensity, and thus circumvents the need for parameter tuning. We evaluated how easily spots can be detected in images enhanced with our method by testing DeepSpot on 20 simulated and 3 experimental datasets, and showed that accuracy of more than 97% is achieved. Moreover, comparison with alternative deep learning approaches for mRNA spot detection (deepBlink) indicated that DeepSpot provides more precise mRNA detection. In addition, we generated single-molecule fluorescence in-situ hybridization images of mouse fibroblasts in a wound healing assay to evaluate whether DeepSpot enhancement can enable seamless mRNA spot detection and thus streamline studies of localized mRNA expression in cells.

Impact Statement

Our paper introduces DeepSpot, a Deep Learning-based tool specifically designed to enhance RNA spots which enables downstream spot detection without the need to resort to image per image parameter tuning. DeepSpot's architecture is inspired by small object detection approaches by integrating dilated convolutions into a module specifically designed for context aggregation for small object and using Residual Convolutions to propagate this information along the network.

1. Introduction

Single cell microscopy together with RNA single-molecule fluorescence in-situ hybridization (smFISH) technologies allows gene expression profiling at subcellular precision for determining molecular states of various cell types⁽¹⁾ and that in high-throughput fashion⁽²⁾. The repertoire of mRNA expression quantification methods is large and includes smFISH, clamp-FISH, amp-FISH, and multiplexed versions, such as MerFISH, all allowing the localization of RNA at subcellular level. There are technological differences between these methods in terms of number of detected RNAs and number of processed cells; however, all produce imaging data with mRNA spots that can be further matched to the spots' x and y coordinates. With such increased image acquisition automation and the consequent growing number of high-throughput projects focused on spatially resolved transcriptomics, the need for automated and highly accurate detection of mRNA spots in fluorescent microscopy images has become increasingly important.

smFISH is a method for visualizing individual RNA transcripts in fixed cells. smFISH is based on targeting RNA molecules with a set of 24–48 oligonucleotide probes, each individually labeled with one fluorophore. The combined fluorescence intensity level obtained from this high number of probes makes each RNA transcript visible as a spot that can be computationally identified and quantified⁽³⁾.

Despite the progress made in recent years, it is still difficult to detect the localization of spots corresponding to different mRNAs in a fully automated manner. First, background intensity is often irregular due to various factors, including autofluorescence that can be caused by intrinsic cell properties and fixative-induced fluorescence phenomenon⁽⁴⁾ or growth medium and buffers. An additional contributor to the background noise is off-target binding of probes, which depends on a various number of parameters, including the length of the transcript, its sequence, the cell type used, and others. Second, spot detection is affected by the nonhomogeneous intensity distribution and indistinct spot boundaries relative to the background. Moreover, fluorescence in-situ hybridization (FISH) images may have a low signal-to-noise ratio (SNR). Additionally, the boundary between background (noise) and signal (spots) is usually not evident⁽⁵⁾.

The main drawback of classical mRNA spot detection methods is the requirement of a strong human input to determine the best parameters to handle variable image to image properties such as SNR and presence of artifacts. Even small differences in these characteristics lead to the necessity for parameter fine-tuning⁽⁶⁾. Other than being time-consuming, the quality of detection largely depends on the capacity of the user to correctly choose the method's parameters according to each image properties (contrast, spots, artifacts, and noise). Some recent deep-learning-based approaches for mRNA spot detection try to circumvent this limitation, such as deepBlink⁽⁷⁾.

Here, we introduce DeepSpot, a convolutional neural network (CNN) method dedicated to the enhancement of fluorescent spots in microscopy images, thus enabling downstream mRNA spot detection by conventional widely used tools without need for parameter fine-tuning. With DeepSpot, we show that it is possible to avoid the manual parameter tuning steps by enhancing the signal of all spots so that they have the same intensity throughout all images regardless of the contrast, noise, or spots shape. All the code as well as the pretrained model is available on GitHub, and a plug-in for the image analysis tool Napari is also distributed (<https://github.com/cbib/DeepSpot>).

DeepSpot gives a new twist to the residual network (ResNet) network architecture and learns to automatically enhance the mRNA spots, bringing them all to the same intensity. In parallel, a multinet-work architecture is integrated, trained by minimizing the binary cross-entropy (BCE) while providing context for mRNA spots thanks to the atrous convolutions. We evaluated the impact of the spot enhancement on the downstream mRNA spot detection, by performing spot detection using Icy with fixed parameters on both simulated images and experimental images manually annotated. Moreover, we compared the quality of mRNA spot detection from images enhanced by DeepSpot with deepBlink, and have shown that our method achieves greater generalization to effectively handle full variability of smFISH data. Finally, to illustrate the end-to-end use of DeepSpot in projects where detecting subcellular localization with high precision is essential, we generated smFISH images from mouse fibroblasts in a

wound healing assay, where enrichment of expression of β -Actin toward the location of the wound is expected in the migrating 3T3 mouse fibroblasts.

2. Related Work

Methodologically mRNA spot detection can be related to the detection of small objects topic in image analysis. The goal of spot detection is to find small regions in an image that differ from the surroundings with respect to certain properties, such as brightness or shape, more precisely, regions with at least one local extremum⁽⁸⁾. Spots can be considered as a particular case of more or less circular objects of small extent. Object detection has been one of the key topics in computer vision which goal is to find the position of objects. However, small object detection, such as mRNA spots, remains difficult because of low-resolution and limited pixels⁽⁹⁾.

In this work, we propose a deep learning network inspired by small object detection approaches for mRNA spot enhancement and we show how it can enable the downstream accurate detection of spots.

2.1. mRNA spot detection

In the case of FISH images, mRNA spots are small, compact, and smaller than the resolution limit of the microscope⁽¹⁰⁾; therefore, images of mRNA spots correspond to the maximum intensity pixel surrounded by the diffraction of the fluorescence signal defined by the point spread function (PSF), which can be modeled by a Gaussian within small radius disk (see Figure 1b). This radius depends on several imaging parameters and optical properties of the microscope such as the diffraction limit or the excitation state of the fluorophore.

While there is no universal solution to the detection of small objects such as mRNA spots in fluorescent cellular images, a large plethora of work is available on the subject. A number of approaches have gained wide popularity thanks to the development of software tools embedding the algorithms and providing user with a graphical interface. In particular, ImageJ/Fiji^(11,12) is widely used, largely due to the plug-in-based architecture, recordable macro language, and programmable Java API. Another popular tool is CellProfiler⁽¹³⁾, based on similar paradigms and FISH-Quant⁽¹⁴⁾. A more recent platform, Icy⁽¹⁵⁾, also provides the possibility to develop new algorithms as well as a user interface for image analysis, including the Icy spot detector for the detection of mRNA spots, a method based on wavelet transform decomposition⁽¹⁰⁾.

Deep learning networks have been introduced for mRNA spot detection^(16,17), and more recently deepBlink⁽⁷⁾. The latter focuses on a fully convolutional neural network based on the U-Net architecture. deepBlink not only provides the code, but also annotated smFISH data and implements a threshold-independent localization of spots.

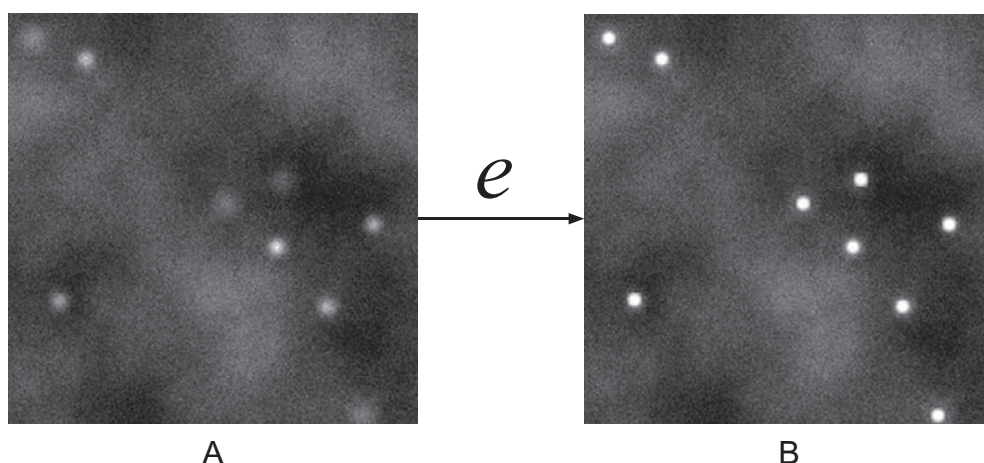


Figure 1. (a) RNA spots on a noisy background. (b) Spots' intensity is increased after the enhancement by e .

2.2. Detection and enhancement of small objects

Consistently with the mRNA spot detection difficulty, detection of small objects remains a challenging part of the general object detection problem due to the limited information contained in small regions of interest. For instance, it has been shown that the object size has a major impact on the accuracy of Deep Learning object detection networks such as VGG, ResNet, or Inception V2⁽⁹⁾. Indeed, small objects do not contain sufficient semantic information⁽¹⁸⁾, and thus the challenge is to capture semantic features while minimizing spatial information attenuation⁽¹⁹⁾.

Expectedly, adding more context improves the detection of small objects^(19–21). An elegant solution is to use the dilated convolution (a.k.a. atrous convolution), because the receptive field can be expanded without loss of resolution and thus capture additional context without loss of spatial information^(22,23).

Of particular interest to our work is the signal enhancement, an image processing technique aiming to reinforce the signal only in those regions of the image where the actual objects of interest are, and potentially to weaken the noise or the signal from other structures⁽²⁴⁾. In our case, the objects of interest are mRNA spots.

Image enhancement is the transformation of one image X into another image $e(X)$ (see Figure 1). Pixel values (intensities) of X at spot locations are modified according to the transformation function e , the resulting pixel values in image $e(X)$ measuring the certainty of mRNA presence at that position. Thus, $e(X)$ can be considered as a probability map that describes possible mRNA spot locations⁽²⁴⁾.

Small object enhancement has been developed in other fields than mRNA spots in fluorescent imaging, such as in astronomy to enhance stars or galaxies over the cosmic microwave background⁽²⁵⁾ or in the biomedical imaging to facilitate the human detection of larger spots such as nodules⁽²⁶⁾. In the microscopy field, Deep-STORM⁽²⁷⁾ enhances the resolution of single-molecule localization microscopy mRNA spots using deep learning. However, these images do not have the same characteristics as smFISH data in terms of noise or signal. For example, human nodules are much larger objects than typical mRNA spots. As for the star enhancement method proposed by Sadr *et al.*⁽²⁵⁾, it is not suited for low-intensity spots⁽²⁸⁾, which is a major concern in spot detection for smFISH images.

3. Materials and Methods

3.1. Materials

3.1.1. smFISH datasets with alternatively established spots' localization

We constructed a dataset DS_{exp} of 1,553 images from the experimental smFISH data acquired in Reference (29), by applying 256×256 pixels patches to better fit in the GPU memory (see Table 1). The authors have performed spot detection using image analysis techniques, such as local maximum detection, implemented in the BIG-FISH pipeline⁽³⁰⁾. Along with the scripts, the authors provided a list of 57 parameter combinations that were used to detect the spots in 57 different image acquisition series of 32 genes used in

Table 1. List of datasets used for training (22 training datasets) and evaluation (21 test datasets) of the DeepSpot network. Images in the DS_{var}^i datasets have spot intensities between 160 and 220 for each image. For the fixed intensity datasets DS_{fixed}^i , the spot intensity is set to one value within [160...220] for a given image, but varies from image to image. Ten variable intensity DS_{var}^i and 10 fixed intensity DS_{fixed}^i datasets are named according to the number of spots in the images, i . The dataset DS_{hybrid} combines DS_{exp} with 25% simulated images.

Datasets	#Spots per image	Intensity	#Images train	#Images test	#Datasets train/test
DS_{exp}	$10 \leq i \leq 150$	Variable	1,453	100	1/1
DS_{var}^i	$i \in \{[10 \dots 100] \bmod 10\}$	Variable	2,000	400	10/10
DS_{fixed}^i	$i \in \{[10 \dots 100] \bmod 10\}$	Fixed	2,000	400	10/10
DS_{hybrid}	$10 \leq i \leq 150$	Variable	2,000	NA	1/NA

their study. We ran the pipeline with these parameters and performed an additional manual curation to keep patches with number of spots between 10 and 150 and remove those with a visually obvious over- or under-identification of spots. The resulting DS_{exp} dataset of 1,553 images representing 27 different genes is thus experimentally generated and guarantees high confidence in the ground truth annotation of mRNA spot coordinates.

We have also downloaded the dataset DS_{dB} from the deepBlink publication⁽⁷⁾, composed of 129 smFISH images acquired from four different cell culture conditions of the HeLa-11ht cell line. The authors provided their annotation of mRNA spots' locations, which was performed using TrackMate and curated by experts.

To estimate the variability of spot intensity, we also calculated the coefficient of variation (CV) of spot intensities for each gene of DS_{exp} as well as for the entire DS_{exp} and DS_{dB} datasets (Supplementary Figure S1). While, for most images and genes, the CV ranges from 0.25 to 0.50, some have a very high CV. The global CV for all the spots of all the images in the DS_{exp} dataset is 0.72 and 0.22 for DS_{dB} .

3.1.2. Novel smFISH dataset from a wound healing assay

To evaluate whether DeepSpot enables precise mRNA spot detection in a biological context, we made use of the wound healing assay⁽³¹⁾ to generate a second experimental dataset DS_{wound} . In the wound healing assay, migrating cells, such as fibroblasts, are grown on a coverslip and serum starved for synchronization. A scratch in the middle of the coverslip is then generated, mimicking a wound, followed by induction of the cells to polarize and migrate toward the wound, to generate wound closure, done using replacement of serum starved medium with 10% FBS-containing medium. We used 3T3 mouse fibroblasts in a wound healing assay, followed by cell fixation. Fixed samples were taken for smFISH experiments to visualize and quantify β -Actin mRNA and imaged on a custom-built Nikon Ti Eclipse widefield TIRF microscope. β -Actin has been previously shown to be enriched in neuronal growth cones of extending axons, as well as the leading edges of migrating cells, and this enrichment has typically been associated with cell polarity and neuronal plasticity^(32–34). Based on this, we hypothesized that β -Actin would be enriched in the leading edge of migrating 3T3 fibroblasts. The dataset is composed of 96 images, 48 images of nonmigrating cells (control), and 48 images of migrating cells. Each image was divided into four patches, yielding a total of 384 patches of 256×256 pixels size.

3.1.3. Simulated and hybrid datasets

In addition to the experimental datasets, we have built 20 simulated datasets with images 256×256 pixels of width and height, the same as the patches of DS_{exp} . Briefly, the background was generated by a combination of Poisson noise and Perlin noise with a random intensity between [80, 150]. Elastic transformations were added to this noise to approximate the variety of textured background noise in the experimental images. Spots were generated as circles, randomly placed in the image, and then convolved with a Gaussian function that approximates PSF. Their size randomly ranges from 4 to 9 pixels in diameter, including Gaussian smoothing.

Two different types of simulated datasets were generated DS_{fixed} and DS_{var} , each containing 10 datasets defined according to the number of spots per image $i \in \{[10 \dots 100] \bmod 10\}$ (see Table 1). For example, in the DS_{fixed}^{20} dataset, each image contains 20 spots, and in the DS_{var}^{70} dataset, each image contains 70 spots. Each image in the DS_{fixed} dataset has the same fixed spot intensity for all spots randomly chosen in the interval [160, 220], whereas in the DS_{var} dataset, the spot intensity is randomly chosen from the same interval for each spot, resulting in images with variable spot intensity.

In addition to the experimental and simulated datasets, we have built a hybrid dataset DS_{hybrid} where the experimental data from DS_{exp} are augmented by appending an additional 25% of simulated images generated with both variable spots' intensity and variable spots' number per image, within the [160, 220] and [10, 100] intervals, respectively. The intensity of the spots is calculated with respect to the intensity of the background noise, so that the generated images have an SNR between 10 and 40. These values

correspond to the minimum and median SNR values in our experimental images, respectively (Supplementary Figure S2).

3.1.4. Ground truth

Since the goal of our network is to learn to transform an image X into $e(X)$, where intensity at spots' location is enhanced, the training step has to be provided with the enhanced counterpart of each image in the training set. That is, training sets include pairs of images $\langle X, e(X) \rangle$ where e is the procedure that is used to produce ground truth enhanced images: for each spot (Figure 1a), the ground truth enhancement procedure e is applied at the spots' locations $A(X)$, resulting in images where spots are enhanced as shown in Figure 1b.

In this work, we implement e as a kernel of 3×3 pixels at all locations where the spots were annotated in the experimental dataset DS_{exp} or generated for DS_{fixed} , DS_{var} , and DS_{hybrid} (see Table 1). The kernel has the same pixel values for all the spots, in order to drive the network to learn to enhance all spots up to the same level of intensity, regardless of the initial intensity in the acquired data. The enhancement kernel of DeepSpot is smaller than the smallest spot size in our datasets; therefore, it is not expected to augment the size spot. This is particularly important for spatially close spots. Moreover, background is kept the same between X and $e(X)$, so the transformation does not affect the background.

3.2. Method

In this section, we present the DeepSpot enhancement network in detail. We first overview the network architecture and then we discuss the custom loss function.

3.2.1. Network architecture

DeepSpot network is composed of two main components as shown in Figure 2. The first component is a multipath network shown in Panel A. The second component is an adapted residual network as shown in Panel B.

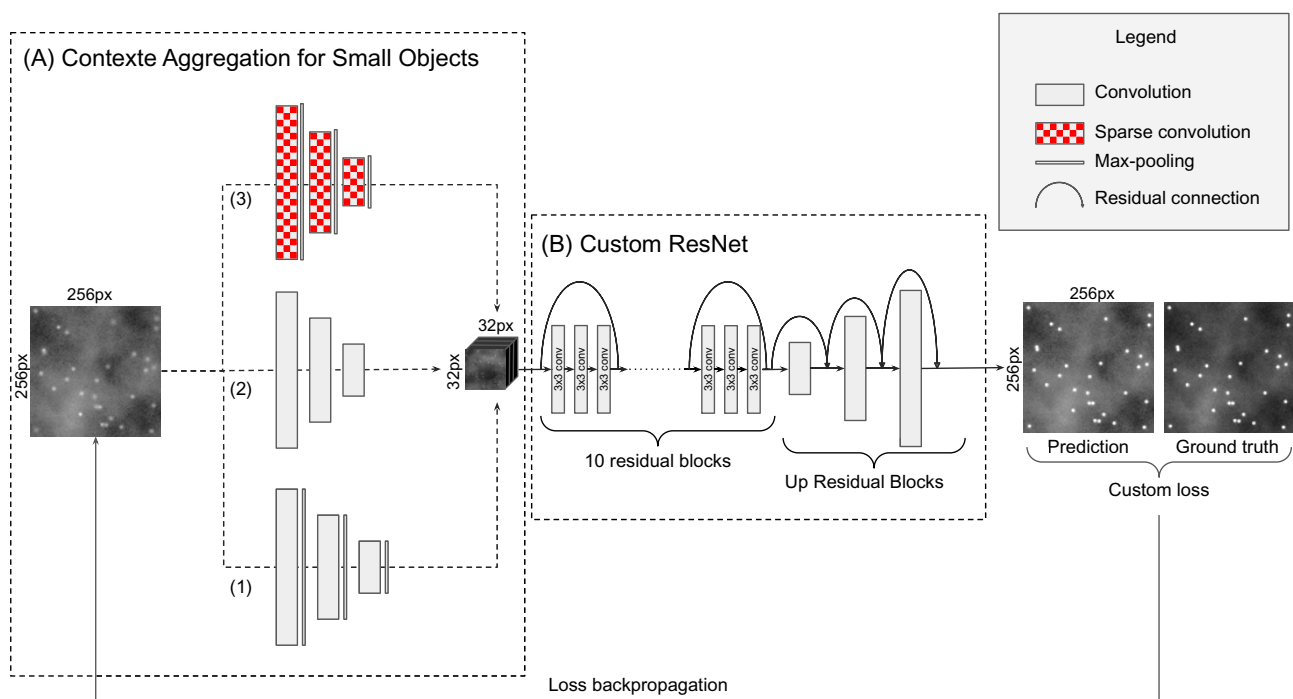


Figure 2. DeepSpot network architecture is composed of the context aggregation for small objects module constituted of a multipath network (Panel A) and a customized ResNet component (Panel B). A custom loss function is used for training the network.

Context aggregation module. As pointed out in Reference (22), finding small objects is fundamentally more challenging than large objects, because the signal is necessarily weaker. To solve this problem in the context of mRNA spot detection, we developed a new module, that we called the context aggregation for small objects (CASO) module⁽³⁵⁾, demonstrated that using image evidence beyond the object extent (context) always enhances small object detection results, and we therefore developed our CASO module to aggregate context around the mRNA spots. The CASO module is a multipath network as shown in Panel A of Figure 2. It takes the input image and processes it along three different paths each with different types of convolution blocks to collect specific information from the input image. Each path contains three convolution blocks.

1. The first path is composed of traditional convolution blocks (2D convolution, batch normalization, activation, and max pooling). These blocks, often used in CNNs, are particularly efficient to reinforce the semantic information at the expense of spatial information.
2. The second path uses only 2D convolutions, batch normalization, and activation. As Max Pooling is known to keep mostly the maximum intensities in images, some of the faint spots may be eliminated during the max pooling operation. In this path, we used strided 2D convolutions instead of a max pooling layer, to keep the information of low intensity of spots. To make sure to end up with the same receptive field as the first path, we set the stride to 2.
3. The third path makes use of the atrous convolution pooling⁽²²⁾, implemented as a 2D convolution with the dilatation rate of 2. The following layers are batch normalization, activation, and max pooling.

The CASO module is a multipath neural network and can learn more comprehensive and complementary features than a single path. In particular, the goal of the atrous convolution is to bring more context around the small spots (see Section 2.2), while the two other paths aggregate the semantic information of the bright spots and faint spots for the first and second paths, respectively. The results of the three encoding paths are then concatenated to construct a longer feature vector containing information extracted by each path. For all convolutional blocks, the activation function is the rectified linear unit (ReLU). The number of filters for the 2D convolutions in the CASO module are 32, 64, and 128 for the first, second, and third paths, respectively.

Custom ResNet. For the second component (Panel B of Figure 2), we customized the ResNet architecture to create a residual neural network composed of 10 consecutive convolutional residual blocks (ResBlock), using full pre-activation blocks described in Reference (36), where the authors suggested that the better results obtained by the full pre-activation blocks are due to the pre-activation by the batch normalization that improves regularization of the model due to the fact that the inputs to all weight layers have been normalized. Each ResBlock is composed of three subblocks as shown in Figure 3. A subblock is constituted of a batch normalization followed by an activation (ReLU) and a 2D convolution. After the three subblocks, a spatial dropout layer with the rate of 0.2 is applied. Each ResBlock ends by the residual connection.

To obtain an output image with the same size as the input image, we used a particular type of convolutional block. Recently, Reference (37) demonstrated that the use of up residual blocks (UpResBlocks) instead of classic up-convolution blocks improves the performance of generative networks by preserving the effective features from the low-dimensional feature space to the high-dimensional feature space. Our decoding path is composed of three UpResBlocks and reconstitutes an output with the same size as the input while propagating low-dimensional feature information of the enhanced spots from the custom ResNet to the last layer. Each UpResBlock is constituted of three subblocks containing a 2D transposed convolution, batch normalization, and activation (ReLU). The three subblocks are followed by a spatial dropout layer with the rate of 0.2. UpResBlocks then end with a residual connection. A sigmoid activation function is applied to the last convolution, so that all the pixels

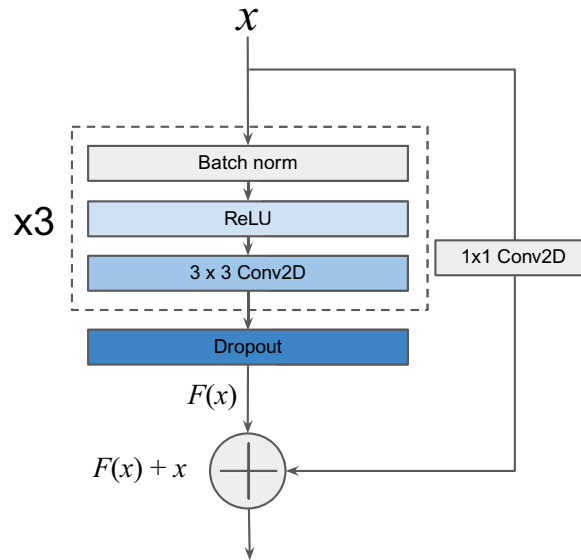


Figure 3. Full pre-activation residual block, composed of batch normalization, activation, and convolution, repeated three times before dropout and residual connection.

have values in the $[0, 1]$ interval. The final image is obtained by normalizing the pixel intensities between 0 and 255.

3.2.2. Loss

We defined our custom loss function as a combination of BCE and mean squared error (MSE) functions. The main term of the loss function is the BCE loss, defined by $\mathcal{L}_{\text{BCE}}(x, \hat{x}) = (x \log(\hat{x}) + (1 - x) \log(1 - \hat{x}))$ that measures the difference between the images predicted by the network \hat{x} and the ground truth images x . While mostly used for classification, it can also be used for segmentation and enhancement due to its performance for pixel-level classification⁽³⁸⁾.

To this main \mathcal{L}_{BCE} term, we added a regularization term defined by MSE, $\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\max(x_i) - \max(\hat{x}_i))^2$ that is computed between the maximum value of the predicted image $\max(\hat{x}_i)$ and the maximum value of the ground truth image $\max(x_i)$. This regularization drives the network to produce spots whose intensity is close to 255 (see Table 2), and therefore standardizes the signal enhancement intensity in the output images, which in its turn facilitates the downstream automatic detection of the spots. The total loss function is $\mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{MSE}}$.

4. Results

We trained the DeepSpot network on our 20 simulated training datasets DS_{fixed} and DS_{var} as well as on the experimental and hybrid datasets DS_{exp} and $\text{DS}_{\text{hybrid}}$, resulting in 22 models M_{fixed} , M_{var} , M_{exp} , and M_{hybrid} . Training parameters were optimized with HyperOpt algorithm⁽³⁹⁾ and the ASHA scheduler⁽⁴⁰⁾. The best configuration obtained and used for further trainings had the learning rate of 0.0001, the dropout rate of 0.2, and the batch size of 32 and 128 filters per convolution.

Each of the resulting 22 models was evaluated on the 21 test datasets from Table 1, yielding 462 enhanced test datasets (8,100 images in total). To assess whether DeepSpot enhancement enables easy spot detection, we applied the Icy spot detector⁽¹⁰⁾ to the images enhanced by different models. Moreover, we defined a unique set of Icy parameters that matches the shape and intensity of the enhancing kernel of the DeepSpot network: scale 3 and sensitivity 20, and scale 7 and sensitivity 100. We then evaluated whether the detected spots from the enhanced images matched well with the annotated ground truth of spots' coordinates.

Table 2. Spot enhancement performance in terms of resulting spot intensity. The measures displayed correspond to the spot intensity between $[0, 255]$ after enhancement by the neural network and averaged by category of models and datasets. Between brackets are shown the 95% confidence intervals. Model categories are listed in rows, whereas columns correspond to the dataset categories on which the different models were applied.

Train \ Test	DS _{exp}	DS _{fixed}	DS _{var}
M _{exp}	249.55 [249.29, 249.82]	250.81 [250.68, 250.93]	247.07 [246.68, 247.45]
M _{hybrid}	241.15 [240.79, 241.52]	248.71 [248.57, 248.85]	243.3 [243.14, 243.45]
M _{var}	248.92 [248.47, 249.37]	251.99 [251.93, 252.05]	254.57 [254.53, 254.62]
M _{fixed}	242.93 [242.16, 243.71]	254.73 [254.71, 254.75]	249.33 [249.09, 249.57]

4.1. Evaluation procedure

We denote by $D(X) = \{p_1, \dots, p_n\}$ the point pattern detected by Icy from the enhancement of an image X by DeepSpot and by $A(X) = \{q_1, \dots, q_m\}$ the ground truth annotation of spot coordinates. Notice that m is not necessarily equal to n , corresponding to under- or over-detection, and even for a well-detected spot, the coordinates in A and D may slightly differ. To account for these remarks, we used the k -d tree algorithm⁽⁴¹⁾ to query the detection D for nearest neighbors in A as proposed in References (42,43). The number of neighbors was set to 1 and the matching radius t to 3 (coherent with the enhancement kernel for ground truth images).

This allows to establish a matching for all annotated points under $t = 3$ and thus also defines the number of False Negatives or False Positives corresponding to the missing matches from $D(X)$ or $A(X)$, respectively. True Negatives are defined by all pixels p of the confusion matrix such that $p \in X \setminus \{A(X) \cup D(X)\}$. However, given that $|X| \gg |X \setminus \{A(X) \cup D(X)\}|$ implies inflated TN values, this makes measures such as accuracy, AUC, and ROC curve irrelevant.

The drawback of matching m versus n points is the possibility of ambiguous matching. With the k -d tree approach, it happens when two points $q_i, q_j \in A$ can match to one $p_k \in D$ (see Figure 4). This can happen if annotated spots q_i, q_j are close and the detected matching point for both of them, p_k , lies within the same distance t , which can correspond to an over-enhancement and thus blurring between the two spots in the enhanced image. While alternative solutions such as the Linear Assignment Problem can be used, they do not avoid the problem of matching two different numbers of points. The k -d tree approach has the advantage to keep the ambiguous matches (AMs) explicitly to measure this effect. We thus also report the number of AMs.

4.2. DeepSpot enhances the mRNA spot signal to the same intensity

To avoid the manual selection of the detection threshold, it is imperative to have a homogeneous spot intensity for the whole dataset in order to use a unique set of parameters for all images. Table 2 summarizes the intensities obtained after enhancement by DeepSpot for each category of datasets described in Table 1 (experimental, hybrid, simulated with variable, and fixed intensities).

As expected, intensities were closer to 255 when training and test datasets belong to the same category. For example, models trained on data with fixed intensities M_{fixed} and applied to data with fixed intensities DS_{fixed} produced enhanced spot intensities very close to 255. Similarly, enhancement close to 255 could be observed when models M_{var} were evaluated on DS_{var}. Of particular interest are the enhancement results

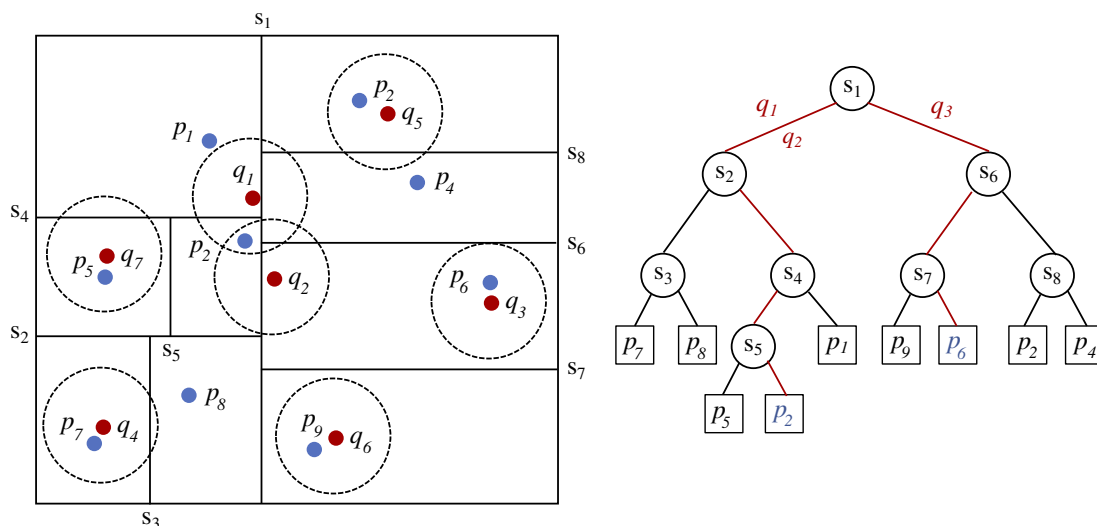


Figure 4. Spot matching by 1-neighbor k-d tree between the detected mRNA spots $\{p_1, \dots, p_9\}$ depicted in blue and annotated spots $\{q_1, \dots, q_7\}$ depicted in red. The k-d tree construction for $\{p_1, \dots, p_9\}$ is shown on the left. Using the matching radius depicted by circles, the k-d tree queries for q_1 and q_2 , shown in red, lead to the same leaf p_2 and correspond to an ambiguous match, while query for q_3 leads to a unique match. mRNA spots p_1, p_4 , and p_8 are the False Negatives.

from the M_{exp} training that are close to 250 for every dataset, experimental or simulated. Hybrid model enhances intensities to 241 on the experimental dataset. In general, the enhanced spot intensities were between 241 and 255, representing a variation of only 5.4% from the maximum intensity, which is sufficient to fully separate smFISH spots from the background in the enhanced images.

4.3. DeepSpot enables accurate mRNA spot detection

The summary statistics of performance of each model type are reported in Table 3 including the mean F1-score, precision, recall, and AMs, with 95% confidence interval. For a given model M , each metric was computed for all enhanced images (8,100 in total). The mean metric value \bar{x} and 95% confidence interval C were then calculated separately for each model type. Due to the high prevalence of True Negatives, instead of the Accuracy measure, we calculated the F1-score, which gives an indication of the model accuracy with a better balance between classes than the actual accuracy measure, by not

Table 3. Models' performance per model type. Metrics (F1-score, precision, recall, and ambiguous matches [AMs]) were calculated by averaging the values obtained for each image of the 21 test datasets. Top values in cells correspond to the mean value, whereas bottom values between brackets show the 95% confidence interval. Best values are highlighted in bold.

Model	F1-score (%)	Precision (%)	Recall (%)	AM (mean/img)
M_{exp}	87.8 [87.6, 88.1]	99.6 [99.5, 99.7]	80.1 [79.8, 80.5]	2.7 [1.4, 4.0]
M_{hybrid}	97.3 [97.2, 97.3]	99.3 [99.2, 99.4]	95.5 [95.4, 95.6]	0.94 [0.56, 1.32]
M_{var}	94.3 [94.3, 94.4]	99.5 [99.5, 99.5]	91.4 [91.3, 91.5]	1.19 [1.03, 1.35]
M_{fixed}	91.6 [91.5, 91.7]	99.9 [99.9, 99.9]	85.7 [85.6, 85.8]	2.05 [1.77, 2.32]

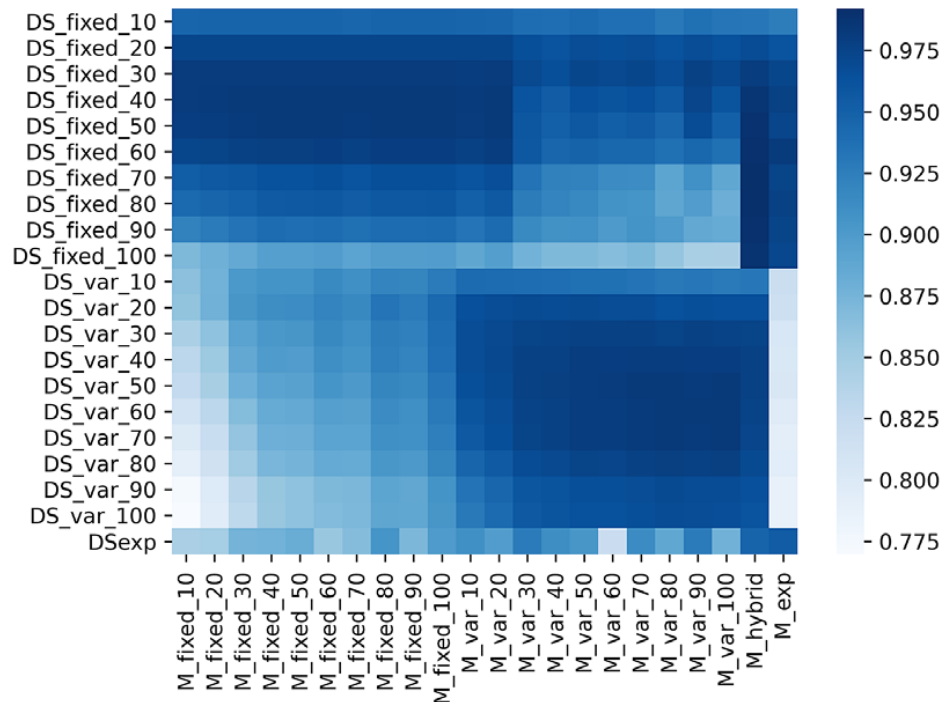


Figure 5. Heat map of the F1-scores obtained by each of the 22 models when evaluated on the 21 test datasets described in Table 1.

including True Negatives. We compared the F1-scores for each of the 14 genes present in the test dataset; no major difference in DeepSpot performance can be observed for these genes (Supplementary Figure S3). This consistency in performance across different SNR values (Supplementary Figure S2) shows that DeepSpot has the capacity to enhance spots in images with varying characteristics.

The results in Table 3 indicate that the number of FP is very low for all models, given that both precision and recall are high. Importantly, M_{hybrid} has shown best overall performance in terms of precision, recall, and F1-score, thus indicating that the mRNA spot enhancement by M_{hybrid} leads to the least FP and FN counts in the downstream mRNA spot detection.

Figure 5 shows the mean F1-scores for each of the 22 models evaluated on the 21 datasets. This heat map indicates that (a) models trained on the DS_{fixed}^i training sets perform better on the corresponding DS_{fixed}^i test sets rather than on variable intensity test sets, and (b) models trained on DS_{var}^i training sets show better performance on the corresponding test sets rather than on fixed intensity test sets. It also shows that M_{var} models are globally better than M_{fixed} models. A plausible hypothesis is that training on variable intensities makes models better at generalizing on other data. Finally, M_{hybrid} is the model that has the best overall performance, including the experimental dataset. Again, the diversity of training data drives this model to be more robust to newly encountered data.

More generally, given that the F1-score is above 90% for all the models, we can conclude that the architecture of the DeepSpot neural network is particularly suited for the task of mRNA spot enhancement.

4.4. DeepSpot enables more accurate spot detection compared with deepBlink

The main objective of our DeepSpot method is to circumvent the parameter fine-tuning and enable the downstream spot detection with a unique parameter set. As such, this objective fits well with the one that the authors of deepBlink⁽⁷⁾ have pursued, despite the fact that the latter proposes a new spot detection method, while our goal is to fit a spot enhancement step into commonly used workflows. Consequently, deepBlink constitutes a relevant comparison target.

Table 4. Models' performance for deepBlink and DeepSpot for smFISH spot detection. Overall F1-scores are calculated by averaging the values obtained for each image of the test datasets corresponding to each dataset category. Top values in each cell correspond to the mean value, whereas bottom values between brackets show the 95% confidence interval. Best values are highlighted in bold.

F1-score (%)				
Model	DS _{exp}	DS _{var}	DS _{fixed}	DS _{dB}
M _{dB}	78.7 [75.26, 82.14]	47.97 [46.86, 49.07]	67.4 [66.28, 68.51]	94.2 ± 0.141 (std)
M _{hybrid}	94.43 [93.41, 95.45]	96.67 [96.58, 96.75]	98.04 [97.96, 98.12]	87.92 [85.36, 90.49]

We compared the accuracy of spot detection by the model M_{dB} made available on deepBlink associated GitHub with that of DeepSpot when trained on hybrid data M_{hybrid}, both on our datasets as well as on the dataset provided by the authors of deepBlink, DS_{dB}. Table 4 shows the F1-scores for each dataset category. It should be noted that deepBlink's precision is close to 99%, while its recall is very low for certain datasets (Supplementary Figure S4), meaning that the main drawback of deepBlink is in terms of false negatives (missing true spots). We illustrate this in Figure 6, where we provide a comparison of deepBlink and DeepSpot on both experimental and simulated data examples.

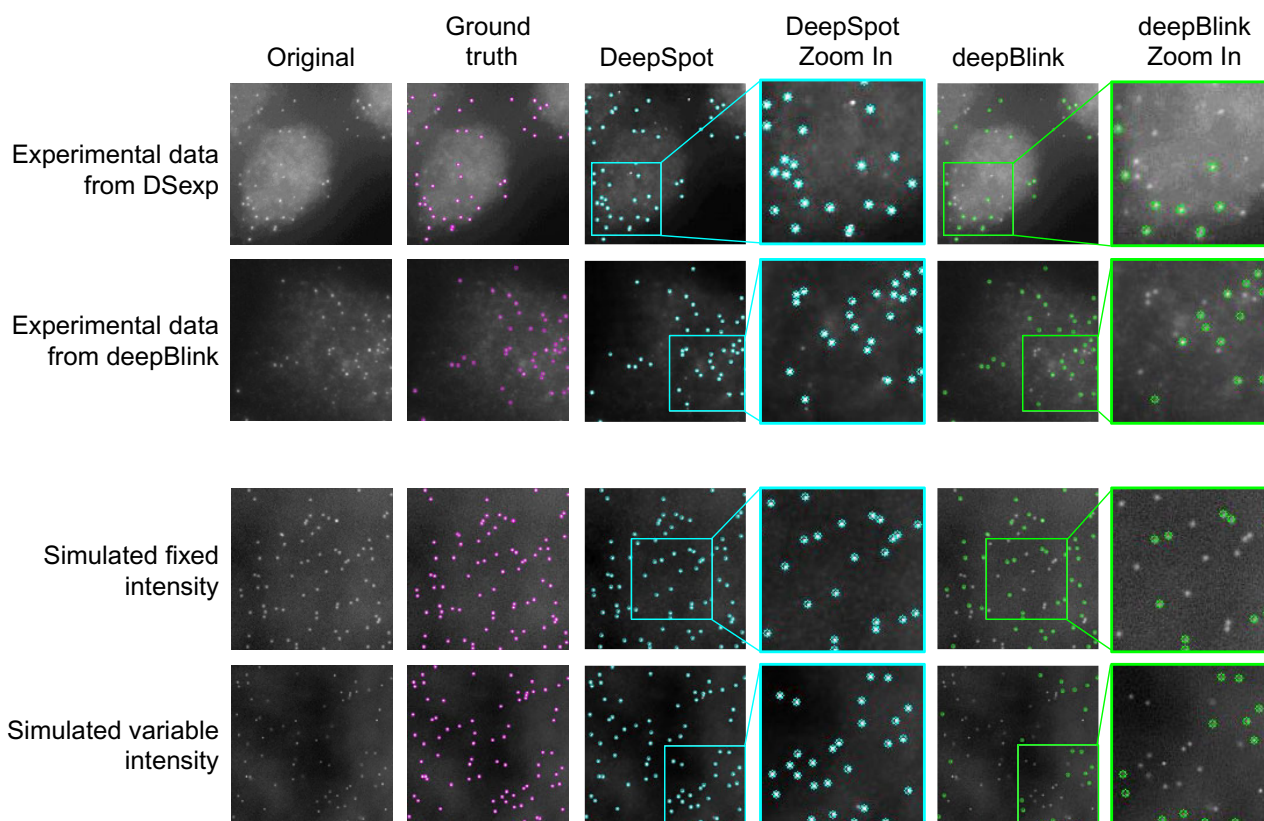


Figure 6. Examples of the results obtained with DeepSpot and deepBlink on the experimental test datasets DS_{exp} (first row) and deepBlink (second row), and the simulated spots with fixed (DS_{fixed}) and variable (DS_{var}) intensity datasets for the third and fourth rows, respectively. Colored circles indicate where the spots were detected by DeepSpot and deepBlink (blue and green, respectively). In the ground truth column, pink circles indicate the spots that were previously annotated as ground truth by alternative methods. The last two columns show the magnification at the positions indicated by the colored rectangles for DeepSpot and deepBlink, respectively.

DeepSpot clearly outperformed deepBlink on our datasets (Table 4). However, we found that M_{dB} performed better on experimental data than on simulated data, presumably because deepBlink model has only been trained on experimental data. M_{hybrid} results were consistent on all of our datasets. We have also applied our M_{hybrid} model to the smFISH test dataset DS_{dB} provided by the authors of deepBlink. Results reported in Table 4 for the detection of spots by M_{dB} from the images of the DS_{dB} dataset are those obtained by the authors and originally reported in Reference (7). Not surprisingly, deepBlink model performance is better on their own smFISH images; however, DeepSpot managed to have an F1-score of nearly 88%, a noticeable achievement since the model has not been trained on the DS_{dB} data. Together, the results of Table 4 indicated that DeepSpot is a robust methodology that offers a generalist model for mRNA spot enhancement and ensures high-quality downstream spot detection without parameter tuning.

4.5. DeepSpot's use in an end-to-end smFISH experiment

To evaluate whether DeepSpot can be effectively used in an end-to-end smFISH experiment, we have performed a wound healing assay in which cells migrate toward a wound to close it. To investigate whether β -Actin was enriched at the leading edges of 3T3 migrating fibroblast cells (see Section 3.1.2), the wound location was manually annotated as shown in a typical image example in Panel A of Figure 7.

All the cell images were segmented by manually, yielding cell and nucleus masks. Wound location defined the cell migration direction as schematically shown in Panel C of Figure 7. We partitioned the cell masks into three sections by computing 120° section centered at the nucleus centroid, and anchoring one of these sections as oriented toward the wound location S_{wound} at 60° angle to the left and to the right of the line between the nucleus centroid and the wound location. This cell segmentation allowed us to compute the normalized number of detected mRNA spots in the cytoplasmic part of each section S_1, S_2, S_3 .

Using the DypFISH framework⁽⁴⁴⁾, we further compared the cytoplasmic mRNA relative density in S_{wound} (light blue) and in sections that are not oriented toward the wound (dark blue), as shown in Panel B of Figure 6. We used the Polarity Index⁽⁴⁴⁾, that measures the enrichment of mRNA in different sections. Briefly, the Polarity Index measures how frequently the relative concentration within the wound section is higher than in the non-wound section. The Polarity Index lies between $[-1, 1]$: a positive value implies a wound-correlated enrichment of RNA transcripts, whereas a negative value implies enrichment away from the wound, and a value of zero implies no detectable enrichment.

β -Actin mRNA was highly enriched in the leading edge of migrating cells, whereas almost no detectable enrichment of β -Actin was found in the leading edge of control cells. This is in line with previously published data, showing enrichment of β -Actin in leading edges of migrating fibroblasts⁽⁴⁵⁾.

5. Discussion and Conclusion

Recent FISH microscopy methods are capable of generating thousands of images, and it has thus become imperative to introduce algorithms capable to streamline the detection of mRNA spots and in particular to avoid manual fine-tuning of numerous parameters.

5.1. Limitation

The impact of the nonspecifically bound probes was not evaluated in this work. Consequently, even if DeepSpot is trained to distinguish between multiple signal intensity variations, in cases where the signal intensity of the nonspecifically bound probes is too similar in shape and intensity to the signal of the specifically bound probes, then DeepSpot will not be able to enhance only spots corresponding to specifically bound probes. This issue, common to all detection methods, deserves further investigation.

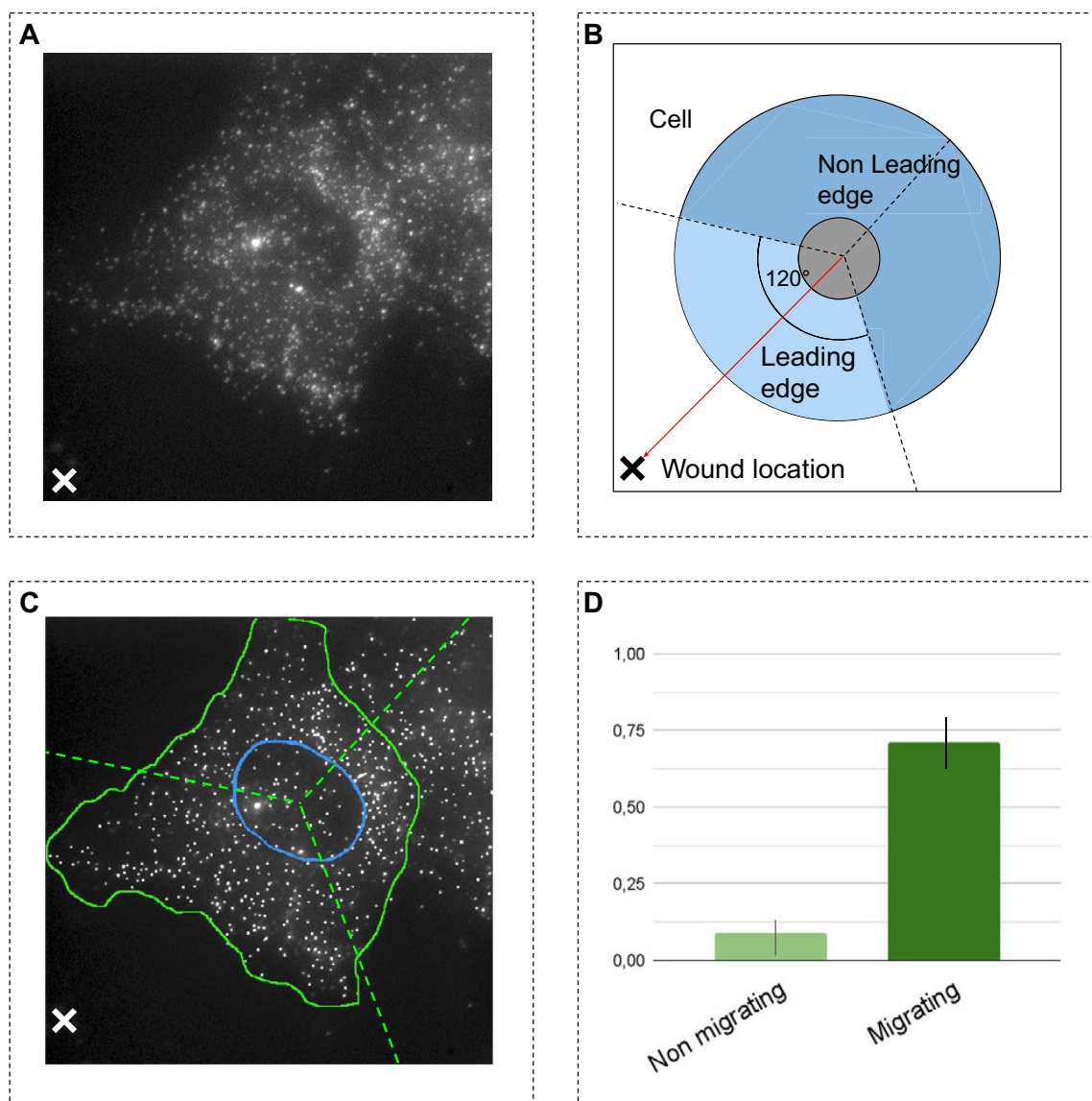


Figure 7. Processing steps for the end-to-end wound healing assay. Panel A shows a typical smFISH image of the wound healing assay. Panel B shows the cell quantization procedure in three sections and the direction of the wound (red arrow). S_{wound} section is oriented toward the wound and is shown in light blue, and other sections are in dark blue. Panel C represents how the cell and nucleus were manually segmented and mRNA spots were counted after enhancement within the cytoplasmic portion of each section. Panel D presents the Wound Polarity Index (WPI) of cytoplasmic mRNA transcripts in S_{wound} compared to other sections for the β -Actin RNA. WPI was calculated in migrating and nonmigrating cells. The bars correspond to the median and the error bars to the standard deviation from the median for 100 bootstrapped WPI estimates.

Moreover, it remains to be investigated whether DeepSpot would be suitable for more complex systems such as tissues and which adaptations in terms of its architecture and training data such improvement would require. Moreover, even if DeepSpot architecture was designed to be easily extensible to 3D by changing the 2D convolution layers to 3D layers in TensorFlow, the performance of DeepSpot in 3D has not been evaluated.

5.2. Discussion

In this work, we introduced DeepSpot, a novel CNN architecture specifically designed to enhance RNA spots in FISH images, thus enabling the downstream use of well-known spot detection algorithms, such as

the Icy spot detector, without parameter tuning. In particular, the architecture of our network introduces the CASO module that relies on sparse convolution to provide more context for enhancement of small objects corresponding to mRNA spots. DeepSpot network has been trained and tested on 21 simulated datasets, all with different signal and noise characteristics, as well as on a previously published experimental dataset that was annotated for spot locations. We have shown that (a) our approach achieves better performance when the training is performed on data with highly variable intensity and (b) performing training on a combination of experimental and simulated data is a viable approach in real-life setting.

Furthermore, we compared the performance of combining DeepSpot and Icy to that of the state-of-the-art deep-learning-based method deepBlink and have shown that, on average, DeepSpot enables a substantially better detection of mRNA spots than deepBlink. We found that DeepSpot/Icy workflow provided excellent quality spot detection on the test datasets corresponding to the datasets on which it has been trained, with the average F1-score of above 97%, but also achieved high-precision results on fully unknown datasets with the F1-score of 88% for the datasets provided with the deepBlink publication. Taken together, the good results on both known and unknown data indicate that DeepSpot is a more generalist model than deepBlink and that it achieves a good balance between overfitting and underfitting. We hypothesize that this generalization capacity is possibly due to both strong regularization within the network and the diversity of signal provided by the carefully constructed training data.

To evaluate how well our method is suited for end-to-end biological investigations, we have shown the efficiency of the DeepSpot model trained on the combination of experimental and simulated data in the context of an independent study of cell migration. We have performed smFISH to detect β -Actin in mouse fibroblasts in a wound healing assay and enhanced the resulting images using our combination model, which allowed us to detect that the β -Actin mRNA enrichment is specific to leading edge of migrating cells as contrasted by its expression in nonmigrating cells.

To conclude, we have shown that DeepSpot enhancement enables automated detection and accurate localization of mRNA spots for downstream analysis methods and can thus be a useful tool to streamline not only spot detection, but also studies of localized mRNA enrichment within cells.

Acknowledgment. We thank Dr. Arthur Imbert for sharing the experimental smFISH data and helping to run the BIG-FISH pipeline.

Competing Interests. The authors declare no competing interests exist.

Authorship Contributions. Conceptualization: E.B. and M.N.; Data curation: A.F.S.; Investigation: E.B. and M.N.; Methodology: E.B. and M.N.; Resources: A.F.S. and R.B.; Software: E.B., E.L., and B.D.; Supervision: M.N.; Validation: A.F.S., E.L., B.D., and M.N.; Visualization: E.B.; Writing: E.B., A.F.S., and M.N. All authors approved the final submitted draft.

Funding Statement. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Data Availability Statement. DeepSpot network along with the code for training and for mRNA spot enhancement is fully open source and available on GitHub at <https://github.com/cbib/DeepSpot>. Our pretrained models used in this study are also available on our GitHub page, as well as the Napari plug-in. Data for simulated images for DS_{var}, DS_{fixed}, and DS_{hybrid} are available on Xenodo at <https://doi.org/10.5281/zenodo.5724466>.

Hardware and Framework All the analysis presented in this paper are performed on 1 GPU Tesla T4 with 16 GB of memory, on a dedicated machine with two 2 CPU Intel Xeon Silver 4114 and 128 GB RAM. Training curves and comparison with a vanilla U-Net architecture are available in Supplementary Figure S5 and Supplementary Table S1. This work is implemented in Python 3.8, and we used TensorFlow 2.4 for creating and training the neural networks.

Supplementary Materials. To view supplementary material for this article, please visit <http://doi.org/10.1017/S2633903X22000034>.

References

1. Ke R, Mignardi M, Pacureanu A, *et al.* (2013) In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* **10**(9), 857–860.

2. Battich N, Stoeger T & Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat Methods* **10**(11), 1127–1133.
3. Raj A, Van Den Bogaard P, Rifkin SA, Van Oudenaarden A & Tyagi S (2008) Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**(10), 877–879.
4. Rich RM, Stankowska DL, Maliwal BP, et al. (2013) Elimination of autofluorescence background from fluorescence tissue images by use of time-gated detection and the azadioxatriangulenium (adota) fluorophore. *Anal Bioanal Chem* **405**(6), 2065–2075.
5. Zhang M (2015) *Small Blob Detection in Medical Images*. Tempe, AZ: Arizona State University.
6. Caicedo JC, Cooper S, Heigwer F, et al. (2017) Data-analysis strategies for image-based cell profiling. *Nat Methods* **14**(9), 849.
7. Eichenberger BT, Zhan Y, Rempfler M, Giorgetti L & Chao JA (2021) deepBlink: threshold-independent detection and localization of diffraction-limited spots. *Nucleic Acids Res* **49**(13), 7292–7297.
8. Lindeberg T (1993) Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention. *Int J Comput Vis* **11**(3), 283–318.
9. Huang J, Rathod V, Sun C, et al. (2017) Speed/accuracy trade-offs for modern convolutional object detectors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, Hawaii, pp. 7310–7311. IEEE.
10. Olivo-Marin J-C (2002) Extraction of spots in biological images using multiscale products. *Pattern Recognit* **35**(9), 1989–1996.
11. Abràmoff MD, Magalhães PJ & Ram SJ (2004) Image processing with imageJ. *Biophotonics Int* **11**(7), 36–42.
12. Schindelin J, Arganda-Carreras I, Frise E, et al. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**(7), 676.
13. Kametsky L, Jones TR, Fraser A, et al. (2011) Improved structure, function and compatibility for cellprofiler: modular high-throughput image analysis software. *Bioinformatics* **27**(8), 1179–1180.
14. Mueller F, Senecal A, Tantale K, et al. (2013) FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat Methods* **10**(4), 277–278.
15. De Chaumont F, Dallongeville S, Chenouard N, et al. (2012) Icy: an open bioimage informatics platform for extended reproducible research. *Nat Methods* **9**(7), 690.
16. Gudla PR, Nakayama K, Pegoraro G & Misteli T (2017) SpotLearn: convolutional neural network for detection of fluorescence in situ hybridization (FISH) signals in high-throughput imaging approaches. In *Cold Spring Harbor Symposia on Quantitative Biology*, vol. **82**, pp. 57–70. New York: Cold Spring Harbor Laboratory Press.
17. Mabaso MA, Withey DJ & Twala B (2018) Spot detection in microscopy images using convolutional neural network with sliding-window approach. In *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*, Funchal, Portugal. SciTePress.
18. Liu Z, Gao G, Sun L & Fang Z (2021) HRDNet: high-resolution detection network for small objects. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, Virtual, pp. 1–6. IEEE.
19. Fu K, Li J, Ma L, Mu K & Tian Y (2020) Intrinsic relationship reasoning for small object detection. Preprint, arXiv: 2009.00833.
20. Bell S, Zitnick CL, Bala K & Girshick R (2016) Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp. 2874–2883. IEEE.
21. Noh J, Bae W, Lee W, Seo J & Kim G (2019) Better to follow, follow to be better: towards precise supervision of feature super-resolution for small object detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, Korea (South), pp. 9725–9734. IEEE.
22. Hamaguchi R, Fujita K, Nemoto A, Imaizumi T & Hikosaka S (2017) Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. Preprint, arXiv:1709.00179.
23. Yu F & Koltun V (2016) Multi-scale context aggregation by dilated convolutions. In *International Conference on Learning Representations (ICLR) San Juan*, Puerto Rico. ICLR.
24. Smal I, Loog M, Niessen W & Meijering E (2009) Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Trans Med Imaging* **29**(2), 282–301.
25. Sadr A, Vos EE, Bassett BA, Hosenie Z, Oozeer N & Lochner M (2019) DeepSource: point source detection using deep learning. *Mon Not R Astron Soc* **484**(2), 2793–2806.
26. Liu J, White JM & Summers RM (2010) Automated detection of blob structures by Hessian analysis and object scale. In *2010 IEEE International Conference on Image Processing*, Hong Kong, China, pp. 841–844. IEEE.
27. Nehme E, Weiss LE, Michaeli T & Shechtman Y (2018) Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* **5**(4), 458–464.
28. Pino C, Sortino R, Sciacca E, Riggi S & Spampinato C (2021) Semantic segmentation of radio-astronomical images. In *Proceedings of the Progress in Artificial Intelligence and Pattern Recognition*, pp. 393–403. Cham: Springer International.
29. Chouaib R, Safieddine A, Pichon X, et al. (2020) A dual protein-mRNA localization screen reveals compartmentalized translation and widespread co-translational RNA targeting. *Dev Cell* **54**(6), 773–791.e5.
30. Imbert A, Ouyang W, Safieddine A, et al. (2022) FISH-quant v2: a scalable and modular analysis tool for smFISH image analysis. *RNA* **28**(6), 786–795.

31. Moutasim KA, Nystrom ML & Thomas GJ (2011) Cell migration and invasion assays. In *Cancer Cell Culture*, pp. 333–343. Springer.
32. Zhang H, Singer R & Bassell GJ (1999) Neurotrophin regulation of β -actin mRNA and protein localization within growth cones. *J Cell Biol* **147**(1), 59–70.
33. Lapidus K, Wyckoff J, Mouneimne G, *et al.* (2007) ZBP1 enhances cell polarity and reduces chemotaxis. *J Cell Sci* **120**(18), 3173–3178.
34. Condeelis J & Singer RH (2005) How and why does β -actin mRNA target?. *Biol Cell* **97**(1), 97–110.
35. Hu P & Ramanan D (2017) Finding tiny faces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, pp. 951–959. IEEE.
36. He K, Zhang X, Ren S & Sun J (2016) Identity mappings in deep residual networks. In *European Conference on Computer Vision*, Amsterdam, The Netherlands, pp. 630–645. Springer.
37. Wang Y, Guo X, Liu P & Wei B (2021) Up and down residual blocks for convolutional generative adversarial networks. *IEEE Access* **9**, 26051–26058.
38. Jadon S (2020) A survey of loss functions for semantic segmentation. In *IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Via del Mar, Chile, pp. 1–7. IEEE.
39. Bergstra J, Yamins D & Cox D (2013) Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In *International Conference on Machine Learning*, Atlanta, USA, pp. 115–123. PMLR.
40. Li L, Jamieson K, Rostamizadeh A, *et al.* (2020) A system for massively parallel hyperparameter tuning. Preprint, arXiv: 1810.05934.
41. Bentley J (1975) Multidimensional binary search trees used for associative searching. *Commun ACM* **18**(9), 509–517.
42. Samet H (1990) *The Design and Analysis of Spatial Data Structures*, vol. **85**. Reading, MA: Addison-Wesley.
43. De Berg M, Cheong O, Van Kreveld M & Overmars M (2008) *Computational Geometry: Algorithms and Applications*. Springer.
44. Savulescu AF, Brackin R, Bouilhol E, *et al.* (2021) Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH. *Cell Rep Methods* **1**(5), 100068.
45. Kislauskis EH, Li Z, Singer RH & Taneja KL (1993) Isoform-specific 3'-untranslated sequences sort alpha-cardiac and beta-cytoplasmic actin messenger RNAs to different cytoplasmic compartments. *J Cell Biol* **123**(1), 165–172.

Cite this article: Bouilhol E, Savulescu A. F, Lefevre E, Dartigues B, Brackin R and Nikolski M (2022). DeepSpot: A deep neural network for RNA spot enhancement in single-molecule fluorescence in-situ hybridization microscopy images. *Biological Imaging*, 2: e4. doi:<https://doi.org/10.1017/S2633903X22000034>

Supplementary Information (SI)

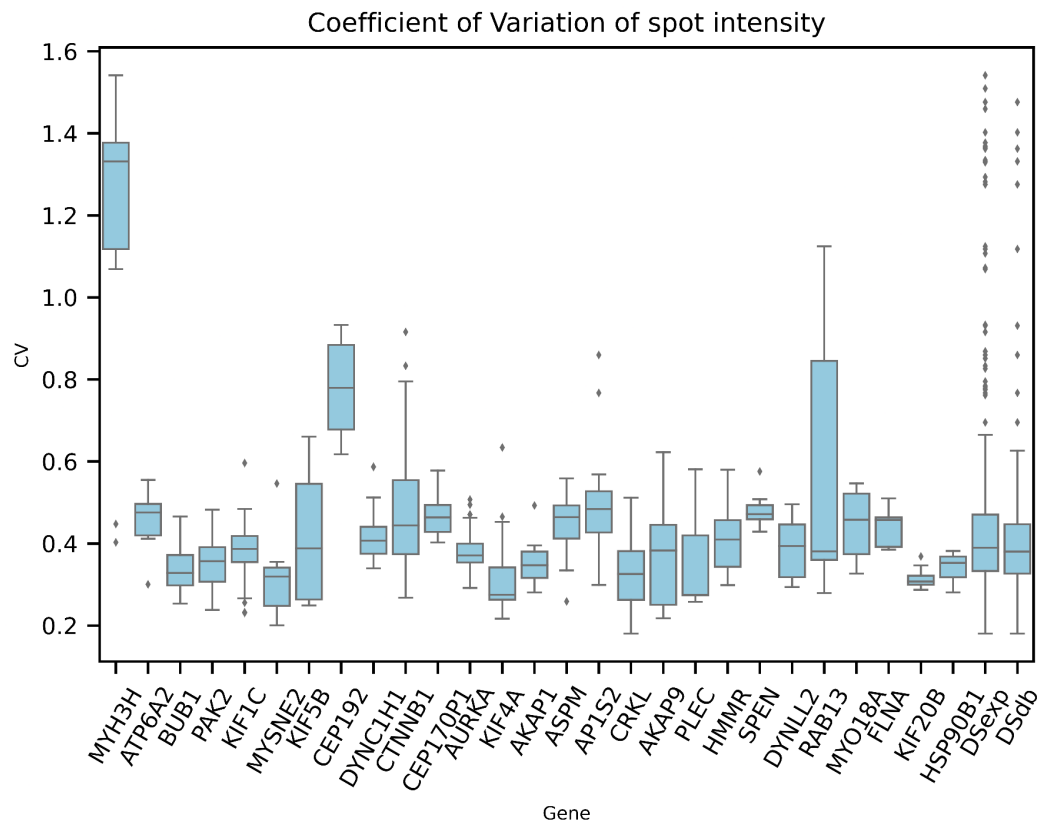
DeepSpot: a deep neural network for RNA spot enhancement in smFISH microscopy images

Authors : Emmanuel Bouilhol*, Anca Flavia Savulescu, Edgar Lefevre, Benjamin Dartigues, Robyn Brackin and Macha Nikolski*

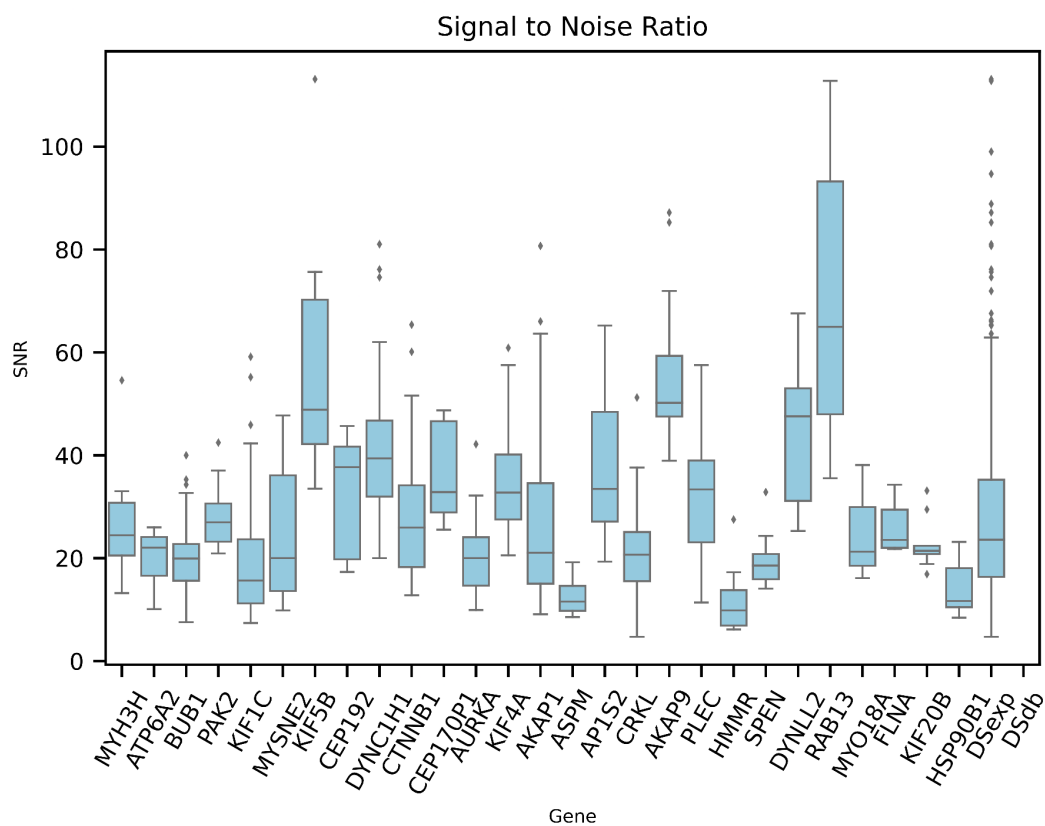
*To whom correspondence should be addressed: emmanuel.bouilhol@gmail.com;
macha.nikolski@u-bordeaux.fr

Supplementary Figures S1-5

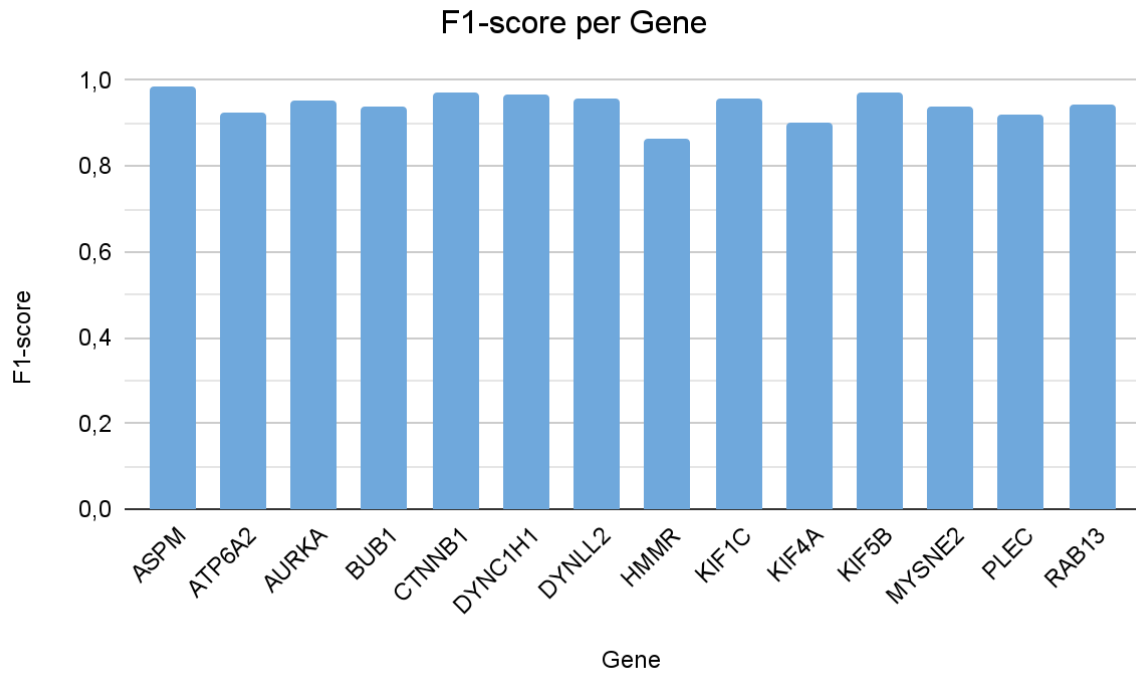
Supplementary Table S1



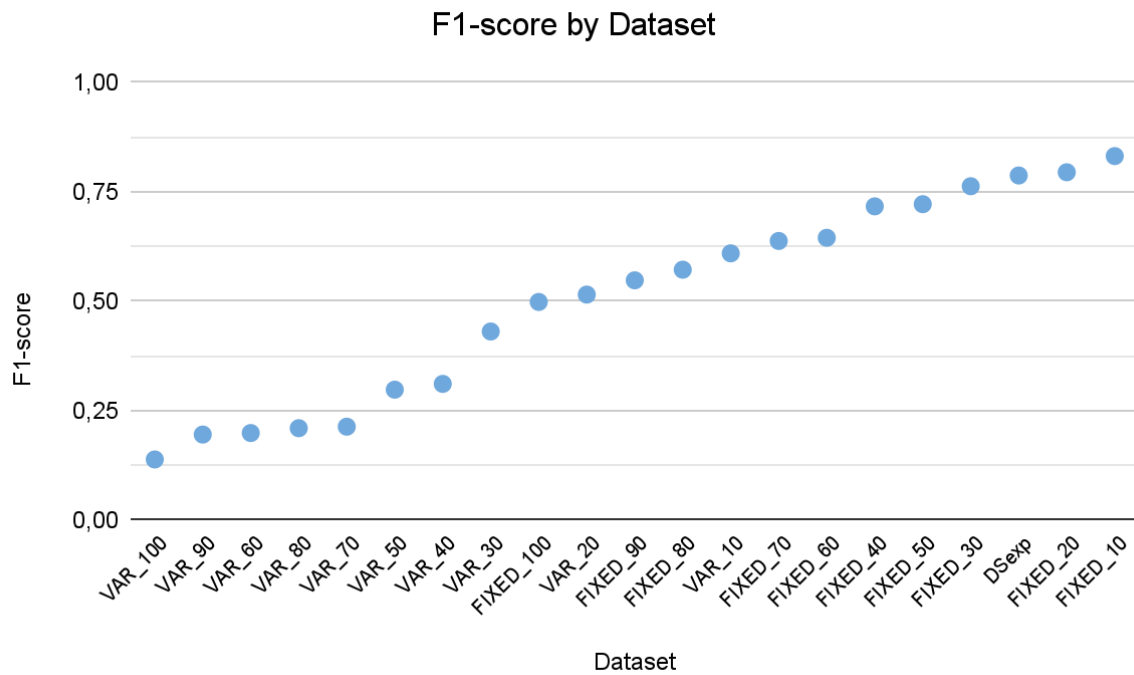
Supplementary Figure S1: Boxplots of the Coefficients of Variation of the spots intensities (defined as the ratio $\frac{\sigma}{\mu}$) for 27 genes and for the whole dataset DSexp (Chouaib et al., 2020). Maximum intensity of each spot from DSexp annotated by Chouaib et al. (2020) and DSdb annotated by Eichenberger et al., (2021) is used to compute the Coefficient of Variation (CV) of spot intensities.



Supplementary Figure S2: Signal to Noise Ratio (SNR) distributions. SNR is measured by $SNR = (S - B) / \sigma$, where S is the maximum pixel intensity of spots in an image, B is the mean value of the background and σ is the standard deviation of the background. SNR is computed for each image; boxplots represent the distribution of SNR for each of the 27 genes, for the entire curated dataset (DSexp) from the Chouaib et al., (2020) publication and the dataset (DSdb) from Eichenberger et al., (2021).

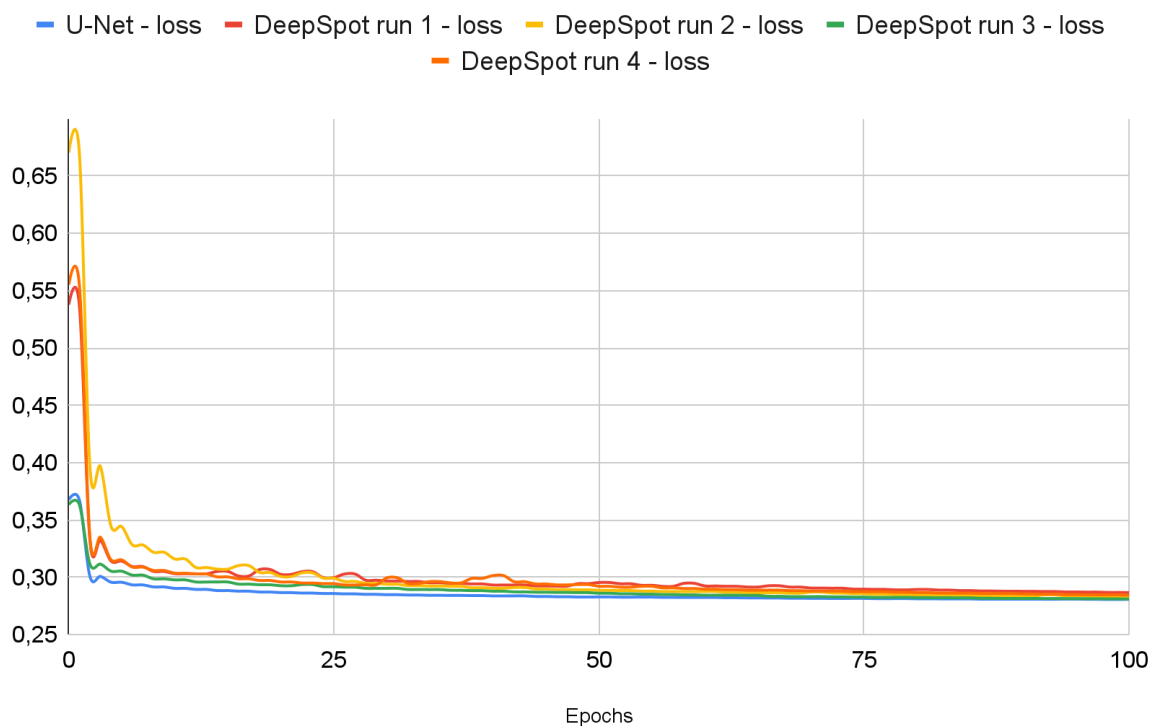


Supplemental Figure S3: F1-score of spot detection after DeepSpot enhancement for 14 genes belonging to the test dataset from Chouaib et al. (2020). For each gene, F1-scores were calculated by averaging the F1-score obtained for each image.



Supplemental Figure S4: F1-Scores of deepBlink spot detection on all simulated and experimental datasets. Scores were obtained by averaging the F1-score values of all images belonging to each dataset.

Training losses



Supplemental Figure S5: Loss values per epoch for 4 different trainings of DeepSpot and one training of a vanilla U-Net. The training curve of the U-Net is shown in blue. Differences in the number of epochs are due to the early-stopping. All trainings were performed with the same parameters, namely a patience of 50 epochs on the Nvidia T4 GPU cards.

Training	Epochs	Execution time (minutes)
U-Net	342	232
DeepSpot run 1	268	397
DeepSpot run 2	264	380
DeepSpot run 3	213	316
DeepSpot run 4	265	386

Supplemental Table S1: Benchmarking execution time for DeepSpot and a vanilla U-Net. All five trainings were performed with the same setup described in the manuscript in the "Hardware and Framework" section. Differences in the number of epochs are due to the use of early-stopping. All trainings were run with the same parameters, namely a patience of 50 epochs on the Nvidia T4 GPU cards.

Chapitre 3

Rehaussement du signal des noyaux cellulaires dans l'imagerie à faible intensité

3.1 Contexte scientifique

La segmentation des noyaux cellulaires présente un intérêt pour un certain nombre d'applications telles que la détection, le décompte, le suivi des cellules (*tracking*), leur analyse morphologique ainsi que pour pouvoir par la suite définir des régions de l'image dans lesquelles il sera possible de quantifier l'expression moléculaire, par exemple quantifier l'expression à la périphérie des noyaux. La segmentation automatique et précise des noyaux est particulièrement importante dans le cas de l'imagerie microscopique à haut débit, pour laquelle cela constitue souvent la première étape des pipelines d'analyses quantitatives.

Des outils récents basés sur l'apprentissage profond, tels que Cellpose [166] et Stardist [154], ont considérablement réduit la difficulté de segmentation des noyaux, notamment en supprimant l'étape consistant à adapter les paramètres de détection à chaque image ou à chaque série d'images. Cependant, malgré ces avancées méthodologiques importantes, aucune combinaison unique de méthodes ne peut être adoptée pour effectuer automatiquement la segmentation des noyaux dans toutes les images en raison de l'hétérogénéité des échantillons biologiques [64]. En particulier, l'imagerie *live-cell* représente un écueil pour ces techniques, car ces images sont souvent acquises avec de faibles niveaux de lumière et de ce fait, l'intensité du signal qui en résulte est assez faible.

De plus, ces modèles suivent le paradigme d'apprentissage supervisé et nécessitent donc des ensembles de données bien annotées. Cependant, de tels jeux de données sont difficiles à acquérir. Premièrement, obtenir des annotations expertes pour de grands volumes de données est coûteux en temps, ce qui implique qu'il n'en existe pas beaucoup.

Deuxièmement, l'annotation experte peut être elle-même source de biais dans les données, posant un challenge pour la qualité de l'annotation. De ce fait, il est compliqué d'avoir des jeux de données suffisamment grands, variés et de haute qualité, pour entraîner des modèles.

Dans ce chapitre, nous proposons de contourner cette difficulté intrinsèque à l'apprentissage supervisé en choisissant une approche non supervisée. De plus, comme dans le chapitre précédent, au lieu de s'attaquer directement au problème de segmentation, nous proposons une solution de rehaussement du signal, s'inscrivant aisément comme une étape dans des pipelines classiques. Plus précisément, nous tirons parti des récentes avancées dans le domaine des réseaux génératifs adverses non supervisés et plus particulièrement, du transfert de style. En ce qui concerne le problème de la segmentation des noyaux, il est possible de définir deux styles différents pour les images : (i) les images de noyaux acquises avec une faible intensité lumineuse, que nous utiliserons comme le style "source", (ii) les images acquises dans des conditions usuelles avec une intensité normale, que nous utiliserons comme le style "cible" (voir les exemples sur la figure 3.1. L'idée que nous développerons dans ce chapitre est alors de traduire les images du style source au style cible de façon à ce que les outils de segmentation entraînés sur des images ayant des caractéristiques du style cible puissent aisément segmenter les images après transfert de style.

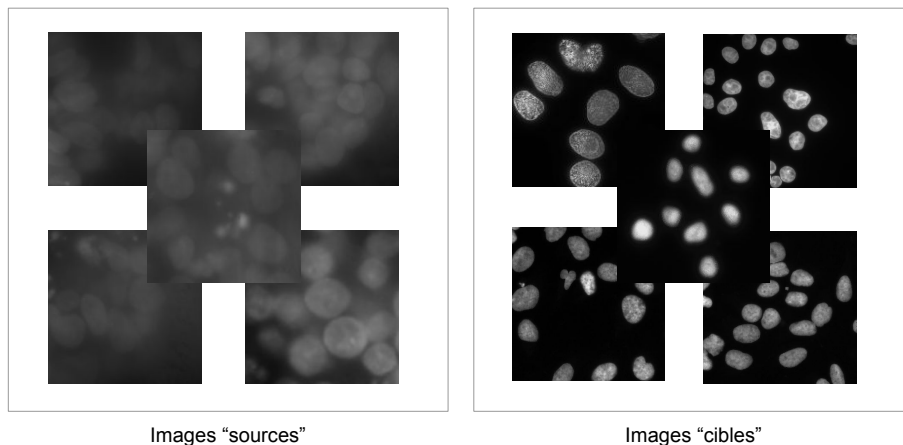


FIGURE 3.1 – Les deux styles d'images définis pour le CycleGAN. À Gauche le style "source", pour lequel les caractéristiques principales des images sont un SNR faible et une très légère saillance des noyaux. À droite le style "cible" contenant des images où le SNR est élevé et la saillance des noyaux très visible.

Nous présentons dans ce chapitre la méthode SaliencyNet, une approche non supervisée

basée sur l'apprentissage profond pour augmenter la saillance (*saliency*) des noyaux dans les images de microscopie. Nous appliquerons cette méthode au transfert de style depuis des images d'organoïdes acquises avec une faible intensité lumineuse en images où les noyaux sont visiblement saillants, sans qu'une annotation préalable des images ne soit requise.

Pour ce faire, nous avons implémenté un CycleGAN (voir le chapitre 1 section 1.5 pour plus de détails sur le CycleGAN) avec une fonction de coût sur mesure dédiée à l'augmentation du signal des noyaux, où l'intensité des noyaux, et en particulier leurs bords, sont rendus plus saillants indépendamment du contraste, de l'intensité, des textures ou des formes des noyaux dans les données d'entrée. Nous avons ensuite entraîné ce réseau de neurones, puis nous avons évalué l'impact du rehaussement du signal des noyaux sur leur segmentation en aval, à l'aide de méthodes conventionnelles ainsi qu'avec des méthodes basées sur l'apprentissage. Nous montrons par la suite que l'incorporation de SaliencyNet dans un pipeline de segmentation de noyaux, permet d'éviter les étapes manuelles de réglage fin des paramètres.

3.2 Contribution scientifique

3.2.1 Segmentation des noyaux

La segmentation des noyaux cellulaires ayant été précédemment traitées dans le chapitre 1 section 1.3, nous ne rappellerons ici que les principes nécessaires à ce chapitre.

Les méthodes de segmentation des noyaux peuvent être divisées en deux grands groupes : celles qui reposent sur des approches classiques de traitement d'images et celles qui utilisent des modèles d'apprentissage profond. Pour une revue approfondie des différentes méthodes existantes, nous nous référons à la très récente et très complète revue de Reka Hollandi [64].

Les pipelines de traitement d'images contiennent généralement un certain nombre d'étapes de filtrage et de seuillage combinées et, si nécessaire, à des opérateurs morphologiques pour séparer les noyaux [104, 91]. La plupart de ces méthodes sont disponibles sous forme de plugins dans les principaux outils d'analyses d'images biologiques, tels que Fiji [152], ICY [33], QuPath [4], CellProfiler [114] ou encore Napari [28].

Un des pipeline les plus utilisés dans la segmentation de noyaux [54], disponible comme plugin dans ICY, FiJi et Napari, consiste à débruiter les images avec un flou gaussien, avant de définir des régions en appliquant une tessellation de Voronoï, puis d'appliquer un seuillage d'Otsu pour obtenir le masque des noyaux et donc leur segmentation.

Cependant, ces étapes sont pour la plupart paramétriques, notamment l'utilisation des filtres gaussiens ou l'utilisateur doit définir la taille du noyau de convolution ainsi que la hauteur de la gaussienne, ce qui peut fortement impacter le résultat de la segmentation [64].

La nécessité d'une solution automatisée capable de segmenter les noyaux dans des images aux caractéristiques différentes a poussé à l'adoption de méthodes basées sur l'apprentissage profond. L'architecture U-Net [140] est utilisée dans le cadre de méthodes récentes de segmentation des noyaux/cellules par apprentissage profond, telles que Cellpose [166] et StarDist [154]. Une autre architecture performante est le Mask R-CNN, qui a été récemment adapté à la segmentation des noyaux par les auteurs de nucleAIzer [63]. ImageJ a récemment proposé des plugins de segmentation basés sur le Deep Learning, et des modèles pré-entraînés sont disponibles sur DeepImageJ [109].

3.2.2 Méthode d'évaluation

Pour évaluer la capacité de SaliencyNet à faciliter la segmentation en aval, nous avons choisi d'utiliser comme point de référence, deux méthodes de segmentation très utilisées, à savoir StarDist [154] et le seuillage d'Otsu, dans sa version non paramétrique. Nous avons comparé les résultats de la segmentation par ces deux méthodes sur nos images de test, avant et après rehaussement du signal des noyaux par SaliencyNet (voir figure 3.2).

Le modèle SaliencyNet et le modèle CycleGAN classique ont été appliqués aux données de test pour augmenter la saillance des noyaux. Les images originales et leurs équivalents ont ensuite été segmentés, et les masques obtenus ont ensuite été comparés au *ground truth*. À cette fin, l'annotation a été effectuée par des experts sur l'ensemble des données de test.

Pour mesurer la qualité de la segmentation obtenue, nous avons calculé l'intersection sur l'union (IoU) pour chaque image afin de quantifier le chevauchement (en nombre de pixels) entre la segmentation et la vérité terrain : $IOU = \frac{|S \cap G|}{|S \cup G|}$, où S est le masque résultant de la segmentation et G est le masque du *ground truth*.

3.2.3 Modèle d'apprentissage non supervisé pour le rehaussement de signal

Architecture du réseau de neurones

Dans cette section, nous présentons en détail le réseau de neurones permettant de rendre les noyaux plus saillants : SaliencyNet. Ce réseau de neurones met en œuvre le transfert de style pour les images de microscopie de noyaux avec l'architecture CycleGAN :

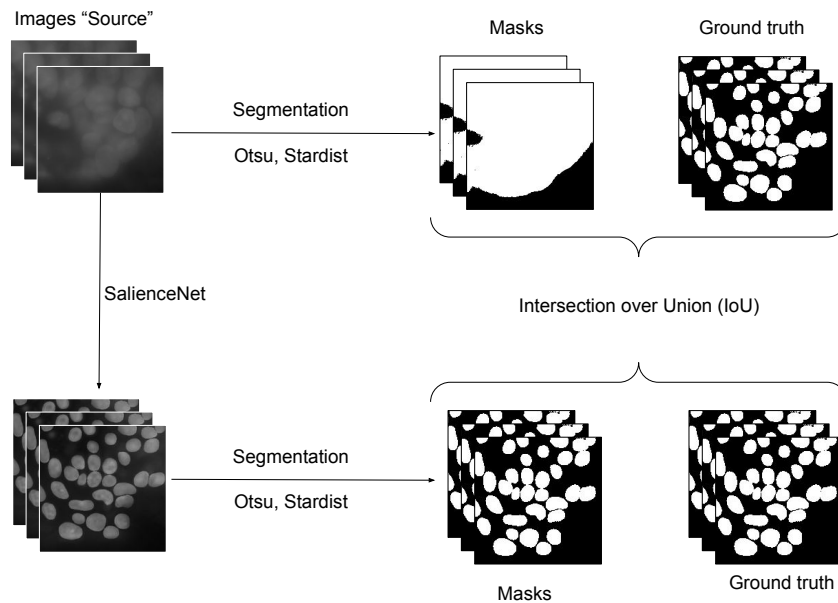


FIGURE 3.2 – Procédure d'évaluation de SaliencyNet. Les images "sources" sont dans un premier temps segmentées à l'aide de StarDist et du seuillage d'Otsu. Les images correspondantes, dont le signal a été rehaussé par SaliencyNet sont à leur tour segmentées avec les deux mêmes méthodes. Les masques obtenus sont comparés au *ground truth* et l'IoU est calculée afin de comparer la qualité des segmentations.

le réseau est composé de deux réseaux génératifs adverses (GAN), comme présenté sur la figure 3.3.

Soit X le domaine des images de noyaux de faible intensité et Y le domaine de style des images avec une forte saillance des noyaux. Le CycleGAN traduit une image du domaine source X au domaine cible Y en apprenant une correspondance $G : X \rightarrow Y$ telle que la distribution des images de $G(X)$ est indiscernable de Y par une fonction de coût adversariale.

L'architecture est basée sur l'apprentissage simultané de deux générateurs et de deux discriminants. Premièrement, le générateur G prend en entrée le domaine X et produit des images pour le domaine de style cible Y , et deuxièmement, le générateur F prend en entrée le domaine Y et génère des images pour le domaine X . Des modèles discriminants adversariaux sont utilisés pour piloter l'apprentissage en estimant dans quelle mesure les images générées correspondent au domaine : D_Y distingue les sorties de $G(X)$ du domaine Y ; de la même manière, D_X distingue les sorties de $F(Y)$ du domaine X .

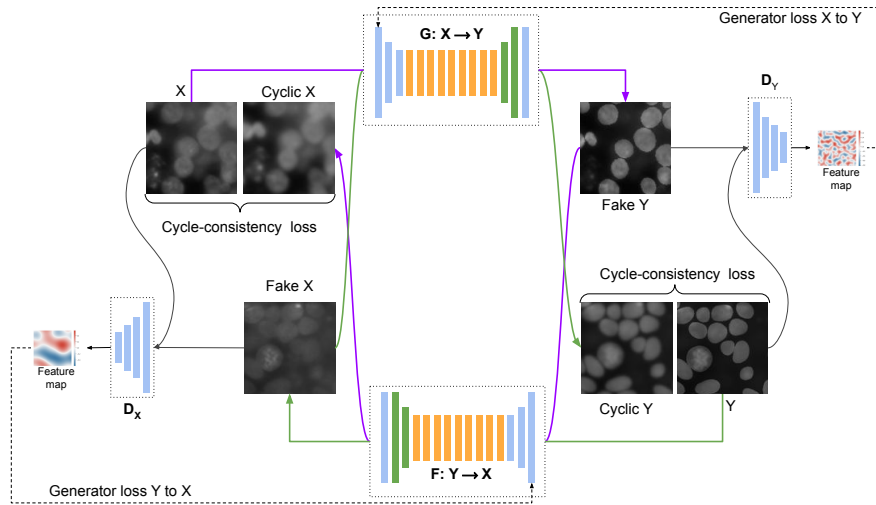


FIGURE 3.3 – Architecture de SaliencyNet. Le réseau est composé de 2 GANs, chaque GAN ayant un Générateur G et un Discriminateur D . L'élément principal du générateur est un réseau résiduel, tandis que le discriminateur est un PatchGAN dont la sortie est une *feature map* à partir de laquelle la fonction de coût du générateur est calculée. La fonction de coût cyclique est calculée entre l'image originale X et son image reconstruite $F(G(X))$ et entre Y et son image reconstruite $G(F(Y))$.

Dans notre modèle, les discriminateurs D_X et D_Y sont implémentés comme des classifieurs PatchGAN, composés de 4 blocs de convolution. Les générateurs G et F sont implémentés comme des ResNets, tous deux ayant la même structure, avec 3 convolutions descendantes, suivies de 9 blocs résiduels, avant d'appliquer 2 convolutions transposées.

Fonction de coût

Pour SaliencyNet, nous avons défini la fonction de coût du générateur comme une combinaison de trois termes : (i) l'erreur quadratique moyenne (MSE), (ii) l'erreur de gradient moyenne (MGE) et (iii) la Mean Structural SIMilarity index (MSSIM).

La MSE calcule la moyenne des différences quadratiques entre les valeurs réelles et les valeurs prédites $\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N ((x_i) - (\hat{x}_i))^2$. Ce terme garanti que le générateur ne produise pas de valeurs aberrantes trop éloignées du style cible.

Dans le cas de la segmentation des noyaux, la netteté des bords est un enjeu majeur dans la mesure où de nombreux algorithmes de segmentation utilisent des fonctions calculant le gradient pour définir les contours des noyaux, tel que l'opérateur LoG. Si le bord du noyau est flou, alors le gradient est faible et la segmentation peut être imprécise. Dans le but de renforcer les contours, nous avons ajouté le terme de MGE (Mean Gradient Error)

qui mesure les différences de bords des objets entre deux images, dans le but d'apprendre des bords nets. Il est basé sur les opérateurs de Sobel vertical et horizontal [76] G_v et G_h :

$$G_v = Y * \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3.1)$$

$$G_h = Y * \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}. \quad (3.2)$$

où $*$ est l'opérateur de convolution.

Ces gradients sont combinés pour définir une carte globale des gradient pour chaque pixel $G = \sqrt{G_v^2 + G_h^2}$. Cette carte de gradient est également calculée de la même manière pour les images prédites \hat{G} . La \mathcal{L}_{MGE} est alors définie comme suit :

$$\mathcal{L}_{\text{MGE}} = \frac{1}{n} \frac{1}{m} \sum_{i=1}^n \sum_{j=1}^m \left(G(i, j) - \hat{G}(i, j) \right)^2 \quad (3.3)$$

Pour amener le réseau à produire des images dont la structure est similaire à la structure d'entrée, nous avons ajouté la MSSIM [186]. Cette fonction de coût compare deux images en fonction de la luminance l , du contraste c et de l'information structurelle s .

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (3.4)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (3.5)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (3.6)$$

où μ_x et μ_y sont l'intensité moyenne pour l'image d'entrée et l'image prédite respectivement ; σ_x et σ_y sont les écart-type pour les images originales et prédites ; c_1 , c_2 et c_3 sont des constantes utilisées pour éviter l'instabilité lorsque les dénominateurs sont proches de 0.

Dans le cas de SaliencyNet, l'intuition derrière l'utilisation de la fonction de coût MSSIM est que dans le sens $X \rightarrow Y$ cela pousse le générateur à rehausser la luminance, tandis que dans le sens $Y \rightarrow X$ cela pousse le générateur à préserver la structure des images.

La fonction de coût totale du modèle SaliencyNet est définie comme la somme pondérée des trois termes présentés ci-dessus :

$$\mathcal{L} = \alpha MSE + \beta MGE + \gamma MSSIM$$

où α , β et γ sont les pondérations de chaque terme correspondant de sorte que la somme $\alpha + \beta + \gamma = 1$.

3.2.4 Jeux de données

Pour entraîner et évaluer SaliencyNet, nous avons collecté différents jeux de données afin d'avoir un jeu de données pour le style "source" et un jeu de données pour le style "cible". Les données ne nécessitent pas d'être appariées (paires source-cible), ce qui facilite la construction des jeux de données.

Pour entraîner et évaluer notre méthode de renforcement de SaliencyNet, nous avons rassemblé de multiples jeux de données. Tout d'abord, deux ensembles de données acquises expérimentalement et segmentées par des experts, qui ont été précédemment largement utilisées pour l'entraînement de méthodes de segmentation (voir ci-dessous). Deuxièmement, nous avons acquis un jeu de données d'images d'organoïdes dans des conditions de faible éclairage qui représente spécifiquement le défi de segmentation que nous adressons ici, et nous avons généré les jeux de données synthétiques correspondants. Ces images appartiennent à l'un des deux styles :

1. **Style source** avec une faible saillance des noyaux (ensembles de données organoïdes et synthétiques à faible saillance),
2. **Style de la cible** avec une forte saillance des noyaux (deux ensembles de données expérimentales et l'ensemble de données synthétiques à forte saillance).

Jeux de données "source"

Pour évaluer si SaliencyNet permet une segmentation précise des noyaux, nous avons acquis un ensemble de données de cultures cellulaires 3D avec la technique soSPIM. soSPIM est une approche de microscopie à feuille de lumière à objectif unique capable de rationaliser les cultures cellulaires 3D avec une imagerie 3D en direct rapide à des vitesses allant jusqu'à 300 cultures 3D par heure [47, 8]. Nous avons sélectionné 11 organoïdes de neuroectoderme présentant une grande variété de formes et de densités. Ces organoïdes ont été immunocolorés avec du DAPI et imagés à l'aide d'un microscope soSPIM, produisant 1056 coupes 2D qui composent le jeu de données DS_{org} .

Pour augmenter le jeu de données de style source, nous avons également généré des images synthétiques. Tout d'abord, nous avons effectué une segmentation experte des noyaux sur chaque tranche d'image 2D individuelle. Ensuite, les images couplées avec leurs masques annotés ont été utilisées pour entraîner un CycleGAN classique. Enfin, ce modèle CycleGAN a été appliqué pour transformer des formes elliptiques placées de façon aléatoire (approximant les formes des noyaux) en images ressemblant à des organoïdes. Nous avons généré 1 500 images synthétiques à faible contraste qui composent le jeu de données DS_{synth} .

Jeux de données "cible"

Afin de fournir le jeu de données de style cible pour l'entraînement du réseau Saliency-Net, nous avons collecté deux jeux de données expérimentaux où la saillance des noyaux était déjà suffisante pour être segmentés par des pipelines classiques et pour lesquels les masques de segmentation des noyaux sont disponibles. Nous les avons complétés par un jeu de données synthétique.

En 2018, une compétition Data Science Bowl organisée par Kaggle a publié un jeu de données pour la compétition " Identification and Segmentation of Nuclei in Cells" d'images acquises dans différentes conditions, à partir de différents types de cellules, variant en taille, agrandissement et méthode d'imagerie (fond clair et fluorescence). Les masques de noyaux ont été créés manuellement par des spécialistes et sont fournis avec le jeu de données. Dans ce chapitre, seules les images de cultures cellulaires de type DAPI ont été conservées, les colorations de type H&E ne faisant pas partie de notre problématique. Ces images composent le jeu de données TS_{DB} , contenant 551 images.

Nous avons également collecté des images de noyaux acquises expérimentalement à partir de lignées cellulaires humaines [26] et qui ont été utilisées pour définir le jeu de données TS_{LS} . Il est composé de 568 images provenant de 57 conditions d'acquisition différentes. Les masques de noyaux ont été acquis avec NucleAIzer par les auteurs.

Par ailleurs, nous avons généré un jeu de données synthétique TS_{synth} , de la même manière que le jeu de données synthétique des images source, mais avec une forte saillance des noyaux et contenant 2000 images.

La composition totale des divers jeux de données utilisés dans ce chapitre est présentée dans la table 3.1.

Pour entraîner SaliencyNet, nous avons divisé chacun des jeux de données sources et cibles en sous-ensembles d'entraînement et de test, dans des proportions de 90% et 10%. Les jeux de données de test sont désignés par la suite par les abréviations DS_{org}^t , DS_{synth}^t , TS_{DB}^t et TS_{LS}^t .

TABLE 3.1 – Nombre d'images et de noyaux dans chaque jeu de données (colonne DS) utilisé pour l'entraînement et le test de SaliencyNet.

DS	#Images	#Nuclei	Style
DS _{org}	1056	43633	Source
DS _{synth}	1500	128962	Source
TS _{LS}	568	20754	Cible
TS _{DB}	551	23121	Cible
TS _{synth}	2000	171915	Cible

3.2.5 Rehaussement de la saillance des noyaux

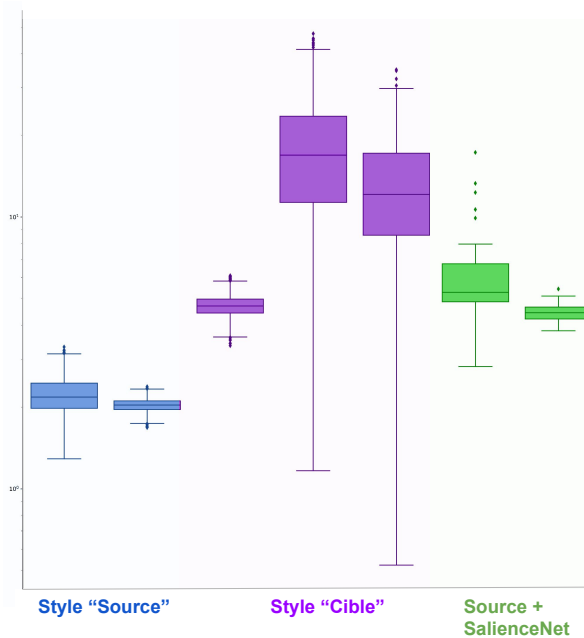


FIGURE 3.4 – Distributions du rapport signal/bruit (SNR). Les boxplots représentent la distribution du SNR pour les images de style source en bleu, les images de style cible en violet. Les distributions SNR des images de test après rehaussement par SaliencyNet sont présentées pour les ensembles de données à faible luminosité, en vert.

Pour évaluer si SaliencyNet a rendu les noyaux plus saillants, nous avons calculé sa mesure indirecte - le rapport signal/bruit (SNR) comme $SNR = (m - \mu_B) / \sigma$, où m est l'intensité maximale des pixels dans les masques des noyaux dans une image, μ_B est la valeur moyenne du fond et σ est l'écart type du fond. Le rapport signal/bruit a été calculé pour les ensembles de données de style source et les ensembles de données de style cible. Nous avons également mesuré le rapport signal/bruit des jeux de données de test après rehaussement du signal par SaliencyNet.

La figure 3.4 permet d’observer que les images ayant été acquises en faible intensité, avec un SNR faible, sont transformées en images ayant un SNR équivalent à celui de l’ensemble des données du style cible.

3.2.6 Amélioration de la segmentation des noyaux

Pour mesurer les performances du réseau, nous avons dans un premier segmenté les images rehaussées à l’aide de deux méthodes : StarDist et le seuillage adaptatif d’Otsu, pour obtenir des masques de segmentation. Nous avons, dans un second temps, segmenté les mêmes images, non rehaussées, à l’aide des mêmes méthodes et finalement, nous avons comparé les masques ainsi obtenus avec les masques issus des annotations expertes disponibles pour les différents jeux de données.

TABLE 3.2 – Valeurs des IoU obtenues pour la segmentation avec Otsu ou StarDist pour les images originales et les images dont le signal à été rehaussé avec SaliencyNet, pour les 4 jeux de données de test. Les valeurs indiquées correspondent à la moyenne géométrique de l’IoU de chaque image, par jeu de données.

IOU DS	Otsu	SN + Otsu	StarDist	SN + Star- Dist
DS_{org}^t	0.45	0.75	0.49	0.75
DS_{synth}^t	0.62	0.90	0.62	0.86
TS_{LS}^t	0.82	0.86	0.90	0.90
TS_{DB}^t	0.69	0.78	0.83	0.83

Pour comparer les masques, nous avons mesuré l’IoU (*Intersection Over Union*), il en résulte que la moyenne géométrique de l’IoU de toutes les images non rehaussées était de 0,49 pour la segmentation StarDist et de 0,45 pour la segmentation Otsu, tandis que la moyenne géométrique de l’IoU après rehaussement par SaliencyNet allait de 0,54 à 0,75 pour StarDist et de 0,48 à 0,75 pour la segmentation Otsu. Les meilleurs résultats ont été obtenus par le modèle SaliencyNet avec $\alpha = 0.2$, $\beta = 0.2$, et $\gamma = 0.6$, avec une moyenne géométrique de l’IoU de 0.75 pour StarDist et pour Otsu. Les résultats complets sont présentés dans la table 3.2.

La figure 3.5 illustre les résultats de la segmentation par StarDist pour les images originales et pour celles dont le signal des noyaux a été rehaussé par SaliencyNet pour les images des jeux de données de test DS_{org}^t et DS_{synth}^t .

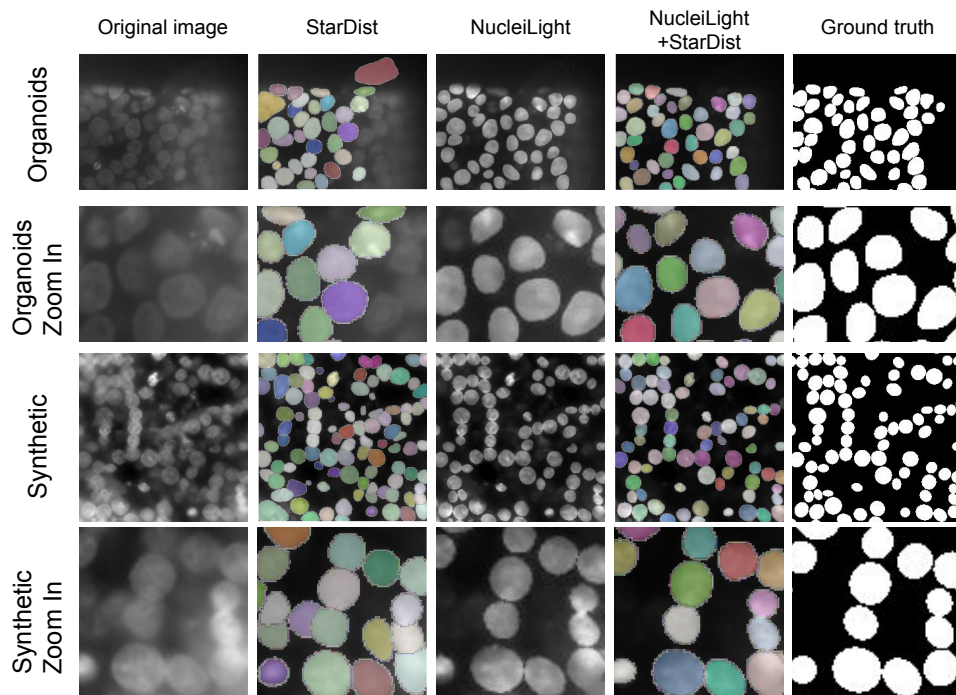


FIGURE 3.5 – Exemples de résultats de segmentation par StarDist obtenus sans rehaussement et après rehaussement du signal par le modèle SaliencyNet. Les exemples d’images proviennent de deux jeux de données de test à faible saillance : DS_{org}^t pour les deux rangées supérieures, et DS_{synth}^t pour les deux rangées inférieures. Les rangées 2 et 4 présentent un zoom avant sur les rangées respectives situées juste au-dessus.

3.3 Discussion

Dans ce chapitre, nous avons montré qu’il est possible d’employer des réseaux de neurones, de manière non supervisée, afin de rehausser le signal des noyaux dans des images de microscopie où le signal d’intérêt est faible. En implémentant une version sur mesure du CycleGAN et en l’appliquant à des données de microscopie, nous avons pu modifier le style des images pour lesquelles la segmentation était difficile, en images où la segmentation est plus aisée.

Pour implémenter SaliencyNet, nous avons combiné trois fonctions de coût avec des propriétés différentes et nous avons montré que notre adaptation de CycleGan améliorerait la qualité de segmentation des algorithmes de segmentation de noyaux classiques et basés sur l’apprentissage profond pour les images de noyaux à faible rapport signal/bruit.

Comme toute méthode basée sur l’apprentissage, le modèle SaliencyNet pourrait bénéficier d’être entraîné sur de plus grands volumes d’images et surtout sur des images plus diversifiées ayant des caractéristiques différentes. Néanmoins, dans l’ensemble, nos résultats indiquent que SaliencyNet pourrait être intégré comme une nouvelle étape dans

les pipelines de segmentation de noyaux, avant l'étape de segmentation, dans l'optique de mettre en place des pipelines qui ne nécessitent pas d'intervention humaine et ainsi faciliter l'analyse de grands jeux de données.

3.4 Article

SaliencyNet: An Unsupervised Image-to-Image Translation Method for Nuclei Saliency Enhancement in Microscopy Images

Bouilhol Emmanuel^{1,2}, Edgar Lefevre², Thierno Barry³, Florian Levet^{3,4}, Anne Beghin^{5,8,9}, Virgile Viasnoff^{5,6,7}, Xareni Galindo³, Rémi Galland³, Jean-Baptiste Sibarita³ and Macha Nikolski^{1,2}

¹*Université de Bordeaux, CNRS, IBGC, UMR 5095, 33000, Bordeaux, France*

²*Université de Bordeaux, Bordeaux Bioinformatics Center, 33000, Bordeaux, France*

³*University of Bordeaux, CNRS, IINS, UMR 5297, Bordeaux, France*

⁴*University Bordeaux, CNRS, INSERM, Bordeaux Imaging Center, BIC, UAR 3420, US 4, Bordeaux, France*

⁵*Mechanobiology Institute, National University of Singapore, Singapore, Singapore*

⁶*IRL 3639 CNRS, Singapore, Singapore*

⁷*Department of Biological Sciences, National University of Singapore, Singapore, Singapore*

⁸*Immunology Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore*

⁹*Department of Microbiology and Immunology, National University of Singapore, Singapore, Singapore
emmanuel.bouilhol@u-bordeaux.fr; macha.nikolski@u-bordeaux.fr*

Keywords: Bioimaging, Deep Learning, Microscopy, Image Processing, Nuclei Segmentation.

Abstract: Automatic segmentation of nuclei in low-light microscopy images remains a difficult task, especially for high-throughput experiments where the need for automation is strong. Low saliency of nuclei with respect to the background, variability of their intensity together with low signal-to-noise ratio in these images constitute a major challenge for mainstream algorithms of nuclei segmentation. In this work we introduce SaliencyNet, an unsupervised deep learning-based method that uses the style transfer properties of cycleGAN to transform low saliency images into high saliency images, thus enabling accurate segmentation by downstream analysis methods, and that without need for any parameter tuning. We have acquired a novel dataset of organoid images with soSPIM, a microscopy technique that enables the acquisition of images in low-light conditions. Our experiments show that SaliencyNet increased the saliency of these images up to the desired level. Moreover, we evaluated the impact of SaliencyNet on segmentation for both Otsu thresholding and StarDist and have shown that enhancing nuclei with SaliencyNet improved segmentation results using Otsu thresholding by 30% and using StarDist by 26% in terms of IOU when compared to segmentation of non-enhanced images. Together these results show that SaliencyNet can be used as a common preprocessing step to automate nuclei segmentation pipelines for low-light microscopy images.

1 INTRODUCTION

Segmentation of cell nuclei is of particular interest for a number of applications such as cell detection, counting or tracking, morphology analysis and quantification of molecular expression. Being able to automatically segment cell nuclei with high precision is particularly important in the case of high-throughput microscopy imaging, where it is often the first step for downstream quantitative data analysis workflows. Indeed, the quality of downstream quantitative analyses is heavily dependent on the accuracy of segmentation, making precise nuclei segmentation essential for

drawing meaningful biological conclusions.

Many solutions have been developed, as exemplified by computational competitions such as reported in (Caicedo et al., 2019). Among popular classical image analysis methods used for nuclei segmentation are thresholding and watershed algorithm (Malpica et al., 1997) as well as active contour (Li et al., 2007).

Challenges for automatizing this process are due to a number of image characteristics that can strongly vary between biological and image acquisition conditions. Among them, we can cite such aspects as morphological differences between nuclei from different tissues, heterogeneity of intensity and texture,

variation in spatial organization such as the presence of both sparse or dense images with touching nuclei, as well as imaging artifacts (e.g., low signal-to-noise ratio or out-of-focus signal) (Zhou et al., 2019). This results in the necessity to fine-tune numerous parameters between different image acquisitions, or even between individual images.

Recent deep-learning based tools such as Cellpose (Stringer et al., 2021) and Stardist (Schmidt et al., 2018) have greatly reduced the necessity to choose specific parameters. However, despite these important methodological advances, no single combination of methods and parameters can be adopted to automatically perform nuclei segmentation in all images, due to the aforementioned heterogeneity of biological samples and technical artifacts (Hollandi et al., 2022). In particular, live-cell imaging represents a stumbling block for these techniques, since these images are often acquired with low light levels and thus have very low SNR and artifacts. Moreover, current supervised Deep Learning models for nuclei segmentation follow the supervised paradigm and thus require well annotated datasets, which (i) engenders bias due to inaccuracy and incompleteness of available segmentation, where nuclei are improperly annotated and unevenly distributed across images (He et al., 2020) and (ii) limits application to datasets with different image characteristics.

In this paper, instead of focusing on the segmentation itself, we propose to tackle this problem by enhancing the nuclei prior to segmentation step, making the task easy for classical nuclei segmentation tools. Specifically, we take advantage of recent advances in the field of unsupervised generative adversarial networks, aiming to translate images from the source domain to the target domain and alleviating the image annotation requirement. For nuclei enhancement task, the target domain corresponds to images with highly salient nuclei, where strong signal difference between the nuclei and background make segmentation straightforward.

In this work, we introduce SaliencyNet, a novel unsupervised Deep Learning-based approach for nuclei saliency enhancement in microscopy images that does not require image annotation when there is need to train the network on new data with different characteristics. We showcase how this can be achieved for translating organoid images acquired with low light into contrasted output images, by training the SaliencyNet without providing the network with prior annotation of newly acquired low contrast images. SaliencyNet gives a new twist to automatic nuclei segmentation by adapting the domain style transfer framework to this specific task, and thus does not

require extensive annotation. We trained a ResNet-based CycleGAN with a custom loss function dedicated to the task of nuclei enhancement, where the intensity of the nuclei and in particular their borders, are made more salient regardless of the contrast, intensity, textures, or shapes of the nuclei in the input data. Furthermore, we evaluated the impact of the obtained nuclei enhancement on the downstream nuclei segmentation by performing segmentation using conventional methods on the enhanced images and have shown that such a pipeline achieves better performance than segmenting the nuclei directly on the original images. We demonstrate here that incorporating SaliencyNet in a standard segmentation pipeline, makes it possible to avoid the manual parameter fine-tuning steps.

2 RELATED WORK

2.1 Nuclei Segmentation

Nucleus segmentation methods can be partitioned in two major groups: those that rely on classical image processing approaches and those that propose Deep Learning models. For a thorough review, we refer the reader to (Hollandi et al., 2022).

Image processing pipelines usually contain a number of filtering and thresholding steps combined, if needed, with basic morphological operators to differentiate nuclei (Malpica et al., 1997; Li et al., 2007). A number of such methods are available as plugins of the main biological analyses open-source software tools such as Fiji (Schindelin et al., 2012), ICY (De Chaumont et al., 2012), QuPath (Bankhead et al., 2017) or CellProfiler (McQuin et al., 2018). The development of classical image processing methods for nuclei segmentation is still an active field. For example, in a recently published image processing library, CLIJ2 (Haase et al., 2020), the authors proposed a nuclei segmentation pipeline "Voronoi Otsu Labeling" in which they first denoise the images with Gaussian blur, second to separate regions using Voronoi tessellation, and to finally obtain a binary mask by applying an Otsu thresholding to obtain the segmentation. However, time-consuming parameter fine-tuning is required from the user at different steps of such classical image processing pipelines, making processing large amount of data impractical (Hollandi et al., 2022).

The need for an automatized solution capable to segment the nuclei in images with different characteristics, pushed for the adoption of methods based on Deep Learning. The U-Net architecture (Ronneberger et al., 2015) is used as part of recent Deep

Learning nucleus/cell segmentation methods, such as Cellpose (Stringer et al., 2021) and StarDist (Schmidt et al., 2018). Another successful architecture is Mask R-CNN, that has been recently adapted for nuclei segmentation by the authors of nucleAIzer (Hollandi et al., 2020). ImageJ has recently proposed Deep Learning-based segmentation plugins, and pre-trained models are available through DeepImageJ (Gómez-de Mariscal et al., 2021).

The success of the aforementioned Deep Learning methods for nuclei segmentation is in particular, due to the use of large and relatively varied training datasets, with images acquired using different microscopy modalities. Nevertheless, establishing a general solution is still an unmet need, especially for images acquired with novel microscopy techniques, such as for example the live-cell imaging (Ettinger and Wittmann, 2014) that reduces the intensity of the light sources of the microscope to a minimum in order to limit the photo-damage to the cells thus be able to observe them over long periods of time. Resulting images have a reduced signal intensity and low SNR. Importantly, having both (i) not been part of the training and evaluation datasets of the aforementioned methods and (ii) having different characteristics, such images represent a yet unsolved challenge for nuclei segmentation.

Moreover, the performance of the existing supervised deep-learning methods depends on the amount of high-quality annotated data available for training. Despite the large effort that was put to produce publicly available labels for nuclei segmentation, such as the 2018 Data Science Bowl competition, such data is often partially or even incorrectly labeled (He et al., 2020).

2.2 Image Preprocessing

A frequently used approach to overcome the difficulty of segmentation is to preprocess the images to improve their quality. In the case of nuclei, such enhancement mainly concerns the contrast between the nuclei and the background. Most traditional image enhancement techniques rely on filtering (low pass, high pass) or on naive noise removal such as Gaussian blur. Other methods are based on normalization of image intensity, such as histogram equalization or contrast stretching (see for review (Qi et al., 2021)). However, in the same way as the segmentation methods themselves, these image preprocessing techniques lack the generalization ability. For example, filtering or signal normalization is not applicable to images with low SNR, as it cannot distinguish well enough the signal from the background.

Deep Learning has been also applied at the pre-processing step, in particular to estimate the transformation function between sets of acquired images and their enhanced counterparts through supervised learning. One of the first and most successful methods was introduced with the CARE network (Weigert et al., 2018), designed to restore fluorescence microscopy data without the need to generate manual training data. The authors showed that it is possible to learn the mapping between low-intensity and high-intensity image pairs using a U-Net based neural network. In the case of live-cell imaging, this makes possible to restore the image quality. However, two characteristics of this network limit the generalization capacity of CARE to new types of images. First, CARE network follows the supervised training paradigm and thus requires matching pairs of the same image and the corresponding nuclei masks, which is time-consuming. Second, CARE comports 5 separately trained networks and uses a disagreement score between the individual network predictions to eliminate unreliable results, which implies that images with characteristics that strongly differ from those in the training set will not be well restored (Weigert et al., 2018).

2.3 Image to Image Translation

Image quality enhancement has also been approached through image to image translation deep learning methods. The goal is to transform an image having a particular style (source style) into a desired target style. The most efficient models are based on GANs (Pang et al., 2021; Wang et al., 2020). Authors of pix2pix (Isola et al., 2017) were the first to apply a GAN-based architecture to perform the image to image translation. It is a fully supervised method that requires large paired image datasets to train the translation model that transforms the source images to the desired target images. In the context of nuclei segmentation, pix2pix has been used by the authors of nucleAIzer for data augmentation of their training nuclei datasets. Specialized image enhancement models have been since proposed, such as Cycle-CBAM (You et al., 2019) for retinal image enhancement and UW-CycleGAN (Du et al., 2021) for underwater image enhancement, both based on the CycleGAN architecture. Moreover, enhancement of objects of interest has been proposed by the authors of DE-CycleGAN (Gao et al., 2021) to enhance the weak targets for the purpose of accurate vehicle detection.

3 PROPOSED METHOD

In this section, we present the SaliencyNet nuclei saliency enhancement network in detail. We first present the network’s architecture, and then we discuss the custom generator loss function that drives the saliency enhancement.

3.1 Network Architecture

SaliencyNet implements the image style transfer for nuclei microscopy images with CycleGAN architecture (Zhu et al., 2017) where the network is composed of two Generative Adversarial Network (GAN) blocks that exchange information during training as shown in figure 1.

Let X be the domain of acquired nuclei images and Y the style domain of images with enhanced nuclei saliency. Images do not have to be paired. SaliencyNet translates an image from domain X to the target domain Y by learning a mapping $G : X \rightarrow Y$ such that the distribution of images from $G(X)$ is indistinguishable from Y by an adversarial loss.

The architecture is based on the simultaneous training of two generator models and two discriminator models (see figure 1). First, generator G takes input from the domain X and outputs images for the target style domain Y , and second, generator F takes input from the domain Y and generates images for the domain X . Adversarial discriminator models are used to drive the training by estimating how well the generated images fit the domain: D_Y distinguishes the outputs of $G(X)$ from domain Y ; in the same manner, D_X distinguishes the outputs of $F(Y)$ from domain X .

In our model the discriminators D_X and D_Y are implemented as PatchGAN classifiers, composed of 4 convolution blocks (see figure 2), each containing a convolution layer, an instance normalization layer and an activation layer (LeakyReLU). The generators G and F are implemented as ResNets having the same structure with 3 down convolutions, followed by 9 residual blocks, before applying 2 transpose convolutions and one last convolution layer with a Tanh activation (see figure 2).

The discriminator is implemented as a PatchGAN model that outputs a square feature map of values, each value encoding the probability that the corresponding patch in the input image is real. These values are further averaged to generate the global likelihood.

Since the mapping $G : X \rightarrow Y$ is highly under-constrained, CycleGAN couples it with an inverse mapping $F : Y \rightarrow X$: the output "fake Y " from the $X \rightarrow Y$ generator is used as input to the $Y \rightarrow X$ gener-

ator, whose output "cycle X " should match the original input image X (and vice versa). This is enforced through the cycle consistency loss to obtain $F(G(X)) \approx X$ and $G(F(Y)) \approx Y$.

3.2 Generator Loss Function

An additional generator loss is used to enforce the cycle consistency and to measure the difference between the generated output "cycle X " and X as well as between the "cycle Y " and Y . This regularization makes possible to constrain the generation process to image translation.

For SaliencyNet we defined the generator loss function as a combination of three terms: (i) the Mean Squared Error (MSE), (ii) the Mean Gradient Error (MGE) and (iii) the Mean Structural SIMilarity index (MSSIM).

The Mean Squared Error (MSE) computes the mean of the squared differences between true and predicted values $\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^N ((x_i) - (\hat{x}_i))^2$. This term ensures that the generator does not produce outliers too far from the target domain. However, MSE used alone is known to lead to blurring due to the averaging between possible outputs, which in image-to-image translation can lead to low-quality blurred results.

In the case of nuclei segmentation, blurring can yield images where nuclei boundaries are particularly difficult to accurately segment. To solve this gradient problem, we added the Mean Gradient Error (Lu and Chen, 2022) term \mathcal{L}_{MGE} that measures the differences in edges of objects between two images, with the aim to learn sharp edges. It is based on vertical and horizontal Sobel operators (Kanopoulos et al., 1988), G_v and G_h :

$$G_v = Y * \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$G_h = Y * \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}.$$

where $*$ is the convolution operator.

These gradients are combined to define a global pixel-wise gradient map $G = \sqrt{G_v^2 + G_h^2}$. The gradient map for predicted images \hat{G} is computed in the same way. The \mathcal{L}_{MGE} is defined as:

$$\mathcal{L}_{MGE} = \frac{1}{n} \frac{1}{m} \sum_{i=1}^n \sum_{j=1}^m (G(i, j) - \hat{G}(i, j))^2$$

Finally, to drive the network to produce images with a structure similar to the input structure, we added \mathcal{L}_{MSSIM} the Mean Structural SIMilarity index

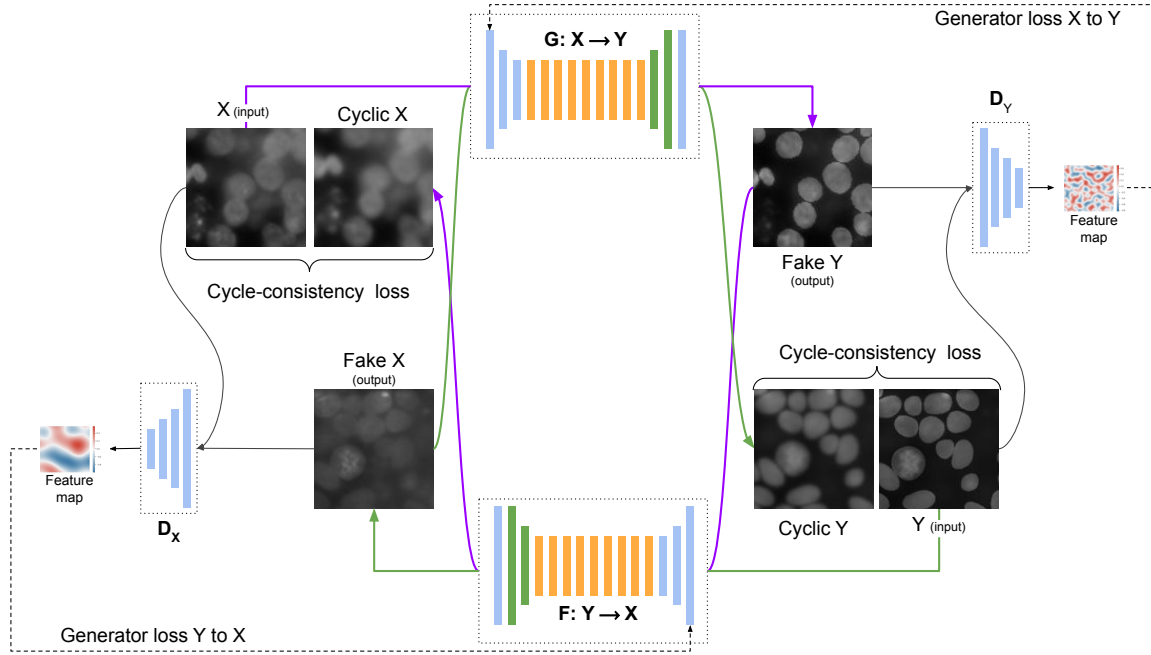


Figure 1: Architecture of SaliencyNet. The network is composed by 2 GANs, each GAN having a Generator G and a Discriminator D . The main element of the Generator is a residual network, while the discriminator is a PatchGAN whose output is a feature map. The generator loss is computed based in this feature map. The inputs to the network are X (input) and Y (input), and the outputs are Fake X and Fake Y , corresponding to $G(F(Y))$ and $F(G(X))$, respectively. Cycle consistency loss is computed between the original image X and it's reconstructed image $F(G(X))$ and between Y and it's reconstructed image $G(F(Y))$.

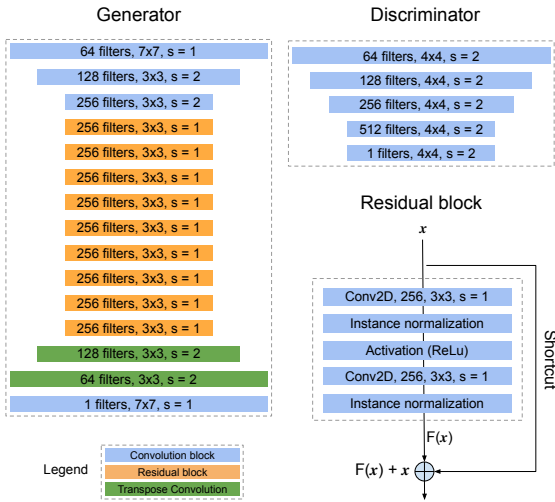


Figure 2: Composition of the networks constituting SaliencyNet. Generators embed a residual network composed of 9 residual blocks, each block being itself composed of 2 convolution layers. Convolution blocks are composed of a convolution layer, an instance normalization layer and an activation layer which is ReLU for the generator and LeakyReLU for the discriminator.

(MSSIM) (Wang et al., 2004) as the last term. This loss function compares two images based on lumi-

nance l , contrast c and structural information s :

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

where μ_x and μ_y denote the mean intensity for the input and generated image respectively; σ_x and σ_y are standard deviation for the original and generated images; c_1, c_2 and c_3 are constants used to avoid instability when the denominators are close to 0.

The mean SSIM can be obtained over the entire image using a local window as follows:

$$\mathcal{L}_{\text{MSSIM}}(x, y) = \frac{1}{M} \sum_{i=1}^M l(x_i, y_i) \cdot c(x_i, y_i) \cdot s(x_i, y_i)$$

where x and y denote the input and the generated image, respectively, while x_i and y_i are the images at the i -th window when the local window slides over the original and generated images, and M is the number of local windows in the image.

In the case of SaliencyNet the intuition for the MSSIM loss for the $X \rightarrow Y$ generator is to enforce the

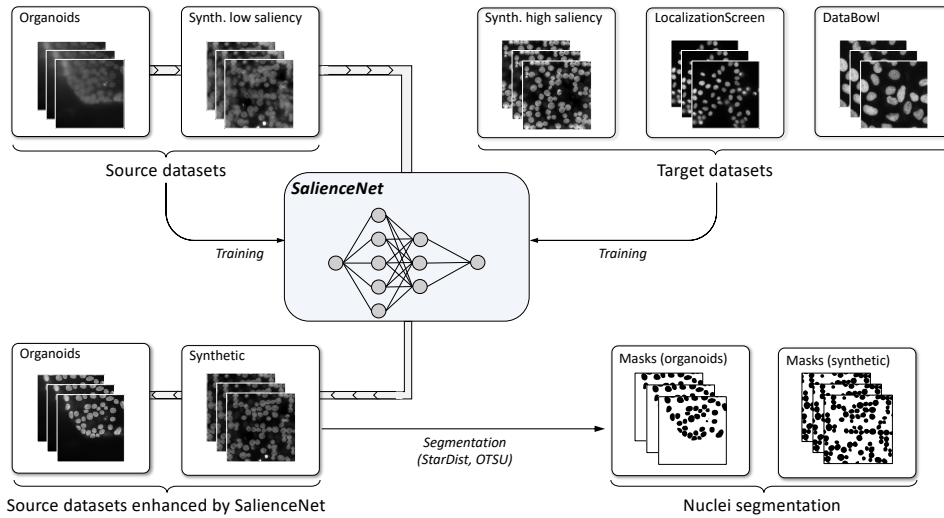


Figure 3: Training and evaluation of SaliencyNet. 90% of source style datasets and of the experimental target style datasets were used to train the network. It was then applied to the remaining 10% of the source style datasets to obtain enhanced images. Nuclei in the images enhanced by SaliencyNet were segmented using classical methods with fixed parameters.

luminance enhancement, while for the $Y \rightarrow X$ generator to preserve the structure.

The total loss function is defined as the weighted sum of the three terms:

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{MSE}} + \beta \mathcal{L}_{\text{MGE}} + \gamma \mathcal{L}_{\text{MSSIM}}$$

where α , β and γ are the weights of each corresponding term so that the sum $\alpha + \beta + \gamma = 1$.

4 DATASETS

To train and evaluate our SaliencyNet enhancement method, we have collected different datasets (see figure 3). First, two experimentally acquired and expertly segmented datasets (see section 4.2), which have been previously extensively used for training segmentation models. Second, we have acquired a dataset of organoid images with low-light conditions that specifically represents the segmentation challenge that we want to address, as well as generated the corresponding synthetic datasets (see sections 4.1 and 4.2). These images belong to one of the two styles:

1. **Source style** with low saliency of nuclei (organoid and synthetic low saliency datasets),
2. **Target style** with high saliency of nuclei (two experimental datasets and the synthetic high saliency dataset).

4.1 Source Style Datasets

To evaluate whether SaliencyNet enables precise nuclei segmentation, we acquired a 3D cell culture

dataset with the soSPIM technique. soSPIM is a single objective light-sheet microscopy approach capable of streamlining 3D cell cultures with fast 3D live-imaging at speeds up to 300 3D cultures per hour (Galland et al., 2015; Beghin et al., 2022). We selected 11 neuroectoderm organoids exhibiting a wide variety of shapes and densities. These organoids have been differentiated from hESCs, fixed at day 8, immunostained with DAPI and imaged using soSPIM, yielding 1056 2D slices. These 2D image slices are composing the DS_{org} dataset.

To augment the source style dataset, in addition to the experimental 1056 2D image slices, we generated synthetic images. First, we performed an expert segmentation of nuclei on each individual 2D slice. Second, images paired with their annotated masks were used to train a simple CycleGAN. Finally, this CycleGAN model was applied to transform randomly placed elliptical shapes (roughly approximating nuclei shapes) into organoid look-alike images. The elliptical shapes provide “nuclei” masks in a trivial way. We generated 1500 synthetic low-saliency images, denoted by DS_{synth} .

4.2 Target Style Datasets

The goal of our network is to learn to transform an image i into $e(i)$, where saliency (Kim and Varshney, 2006) at nuclei location is enhanced. To provide the target style dataset for training the SaliencyNet network, we have collected two experimental datasets where the nuclei saliency was already satisfactory for segmentation by classical pipelines and for which the nuclei segmentation masks are available. We comple-

Table 1: Number of images and nuclei in each dataset (DS column) used for the training and testing of SaliencyNet.

DS	#Images	#Nuclei	Style
DS _{org}	1056	43633	Source
DS _{synth}	1500	128962	Source
TS _{LS}	568	20754	Target
TS _{DB}	551	23121	Target
TS _{synth}	2000	171915	Target

mented them by a synthetic high-saliency dataset.

In 2018, a Data Science Bowl competition organized by Kaggle released a dataset for a challenge of "Identification and Segmentation of Nuclei in Cells" of images acquired under different conditions and of different cell types and that vary in size, magnification, and imaging method (brightfield and fluorescence). Nuclei masks have been manually created by specialists and are provided with the dataset. For the purpose of this paper, only grayscale cell culture images were kept, yielding the TS_{DB} with 551 images.

Experimentally acquired nuclei images from human cell lines (Chouaib et al., 2020) were used to define the TS_{LS} dataset. It is composed of 568 images from 57 different acquisition conditions of 32 gene expression measured in the study for the purpose of performing a localization screen. Nuclei masks have been acquired with NucleAIzer.

The synthetic high-saliency dataset, TS_{synth}, was generated following the same procedure as DS_{synth} (see section 4.1) with 2000 images, but with enhanced saliency. Nuclei masks are provided by the input generation procedure (elliptical shapes).

Taking these 3 datasets together (summarized in table 1), the target style dataset contains 3119 images.

5 RESULTS

To train the SaliencyNet models, we split each of the two source datasets as well as the two experimental target datasets (see figure 3) into train and test subsets, in 90% and 10% proportions. The test datasets are denoted DS_{org}^t, DS_{synth}^t, TS_{DB}^t and TS_{LS}^t, respectfully.

We then performed the hyperparameter search by exploring all possible combinations of α , β , and γ (weights of the loss components, see section 3.2) with step of 0.1 in order to estimate which combination of parameters yielded the best model. This resulted in 42 models, denoted by (α, β, γ) combinations in figure 5. Moreover, for comparison purposes we have trained a vanilla CycleGAN model, without any modification with respect to the original CycleGAN network.

All the 42 SaliencyNet models and the vanilla Cy-

cleGAN were applied to the 4 test datasets DS_{org}^t, DS_{synth}^t, TS_{DB}^t and TS_{LS}^t to perform saliency enhancement. The original images and their enhanced counterparts were then segmented, without any parameter tuning, using two widely used segmentation methods:

1. the non-parametric version of the classical segmentation Otsu thresholding method with an adaptive threshold,
2. StarDist, a deep-learning based segmentation, with the 2D fluo versatile model as provided by (Schmidt et al., 2018), without re-training on our data or supplementary fine-tuning.

The resulting masks were then compared with ground truth. For this purpose, expert ground truth annotation was performed on the DS_{org}^t test dataset; nuclei masks (ground truth) were already available for the 3 other test datasets (see sections 4.1 and 4.2).

To measure the quality of the resulting segmentation, we computed the intersection over union (IOU) for each image to quantify the overlap (in pixel count) between the segmentation and the ground truth: $IOU = \frac{|S \cap G|}{|S \cup G|}$, where S is the mask resulting from segmentation and G is the ground truth mask.

SaliencyNet Enables Accurate Nuclei Segmentation. First, we determined the most performant model with respect to our goal of segmenting low SNR images from live-cell imaging by looking at the enhancement performance on organoid images. Figure 5 shows the IOU scores for segmentation by StarDist of both enhanced (by 42 models) and non-enhanced images, and that for each image of the experimental organoid test dataset DS_{org}^t. Individual model results for Otsu segmentation being very similar (albeit slightly worse), are not shown in this figure.

The IOU is shown in figure 5 for each image with respect to the ground truth. The geometric mean of all non-enhanced image's IOU was 0.49 for StarDist segmentation and 0.45 for Otsu segmentation, while the geometric mean for IOU after saliency enhancement ranged from 0.54 to 0.75 for StarDist and from 0.48 to 0.75 for Otsu segmentation. Best results were obtained for segmentation after enhancement by the SaliencyNet model with $\alpha = 0.2$, $\beta = 0.2$, and $\gamma = 0.6$, with a geometric mean of IOU of 0.75 for both StarDist and Otsu. We denote this model by M .

The impact of this model M on the quality of the downstream segmentation was computed as the ratio of IOU for images enhanced with M over the IOU of non-enhanced images. Impact values range between 1.08 and 2.57 for Otsu segmentation and between 1.09 and 10.73 for StarDist segmentation; notice that the lower bound is > 1 in both cases.

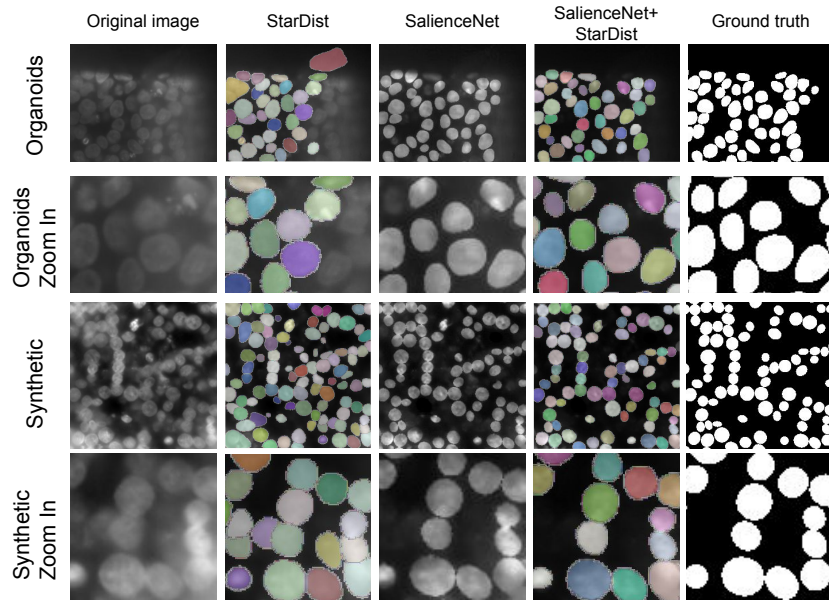


Figure 4: Examples of segmentation results by StarDist obtained without enhancement and after enhancement by SaliencyNet M model. Sample images come from two low saliency test datasets: DS_{org}^t for the two upper rows, and DS_{synth}^t for the two lower rows. Rows 2 and 4 show a zoom-in of the respective rows right above.

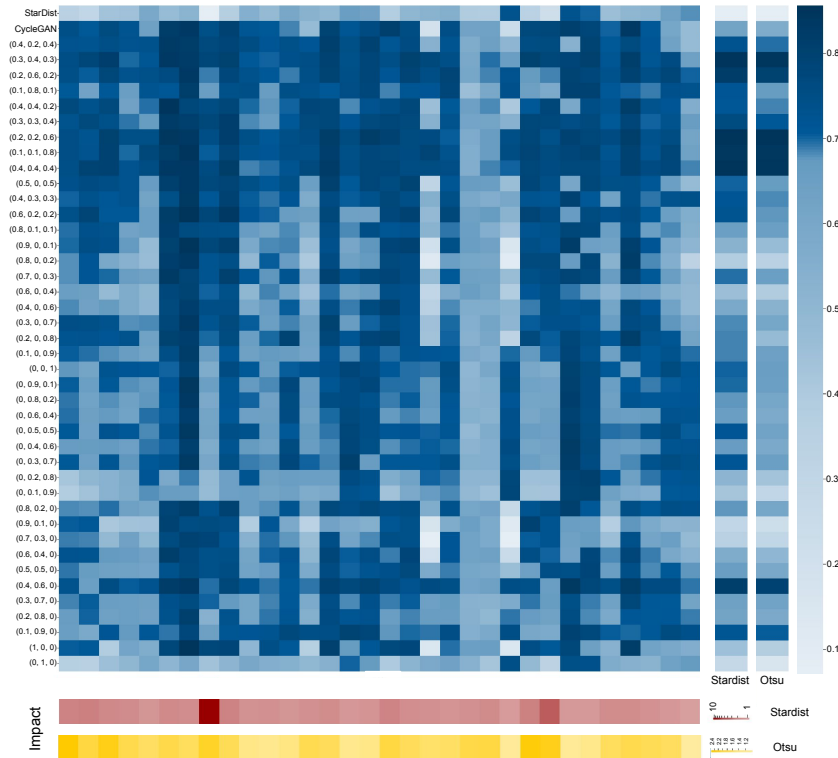


Figure 5: Heatmap of the IOU values for each image segmentation of DS_{org}^t , after enhancement and without enhancement. Average IOU values are shown in the two right columns (StarDist and Otsu). First line: IOUs obtained for non-enhanced images; second line: IOUs after enhancement by the vanilla CycleGAN and segmentation by StarDist, all the other lines: IOUs for segmentation by StarDist after enhancement by SaliencyNet for different α , β and γ values. Bottom-most rows (red and yellow color-scale) show the impact of M enhancement on segmentation quality as IOU ratio (log scale).

Table 2: IOU values for segmentation with Otsu or StarDist of non-enhanced and enhanced with M SaliencyNet (SN) model images in 4 test datasets. All reported values are the geometric mean, per dataset, of the IOU of individual images. Top values in each cell correspond to the geometric mean, bottom values between brackets show the values for the 0.25 and 0.75 percentiles.

IOU DS	Otsu	SN + Otsu	StarDist	SN + StarDist
DS _{org} ^t	0.45 [0.38, 0.51]	0.75 [0.74, 0.79]	0.49 [0.38, 0.59]	0.75 [0.74, 0.79]
DS _{synth} ^t	0.62 [0.61, 0.67]	0.90 [0.89, 0.90]	0.62 [0.59, 0.65]	0.86 [0.84, 0.87]
TS _{LS} ^t	0.82 [0.80, 0.90]	0.86 [0.83, 0.91]	0.90 [0.88, 0.91]	0.90 [0.88, 0.91]
TS _{DB} ^t	0.69 [0.63, 0.86]	0.78 [0.74, 0.87]	0.83 [0.79, 0.88]	0.83 [0.79, 0.88]

An illustration of StarDist segmentation results for non-enhanced images and for those enhanced by M for low saliency organoid images is provided in figure 4.

Table 2 shows the IOU values for the best model M of SaliencyNet applied to the 4 test datasets before and after enhancement by SaliencyNet. On one hand, we observed that SaliencyNet indeed improved the accuracy of nuclei segmentation in low-light source datasets. On the other hand, this table shows that for the already salient images, SaliencyNet did not degrade the quality of segmentation.

Together, these results show that saliency enhancement by SaliencyNet enables accurate downstream nuclei segmentation by widely used methods without need for parameter tuning.

SaliencyNet Enhances Nuclei Saliency. To evaluate whether SaliencyNet improved saliency, we computed its indirect measure - the Signal to Noise Ratio (SNR) as $SNR = (m - \mu_B) / \sigma$, where m is the maximum pixel intensity within the nuclei masks in an image, μ_B is the mean value of the background and σ is the standard deviation of the background. The SNR was computed for the source style datasets DS_{org} and DS_{synth}, and the target style datasets TS_{synth}, TS_{LS} and TS_{DB}. We also measured the SNR of the test datasets DS_{org}^t and DS_{synth}^t after enhancement by SaliencyNet.

We observed (see figure 6) that SaliencyNet enhanced SNR in low-light images of DS_{org}^t and DS_{synth}^t close to the level of SNR of the already salient experimental target style images TS_{LS} and TS_{DB}, and up to the SNR level of the synthetic salient dataset TS_{synth}.

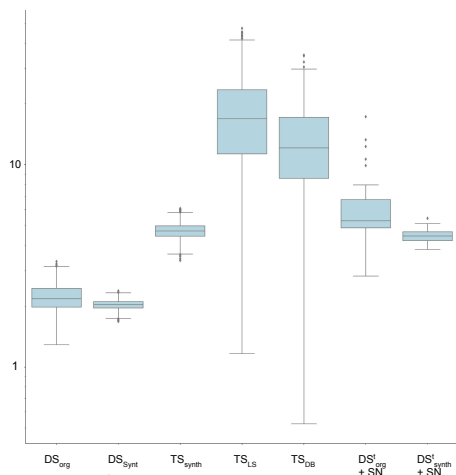


Figure 6: Signal to Noise Ratio (SNR) distributions. Box-plots represent the distribution of SNR for the source style images DS_{org}^t and DS_{synth}^t and the target style images TS_{LS}, TS_{DB}, TS_{synth}. SNR distributions of the test images after enhancement with SaliencyNet M model are shown for the low-light DS_{org}^t and DS_{synth}^t datasets.

6 SUMMARY

In this work, we introduced SaliencyNet, a CycleGAN-based network specifically designed to enhance nuclei’s saliency in low SNR images that does not require annotation for training on new data.

We used the soSPIM light-sheet microscopy, a technique that allows to illuminate the biological sample with little light compared to other methods, to acquire organoid images. The result is that the illumination and the SNR are lower in these images and the nuclei are less salient. We used these organoid images as source style for training our network and further for testing. To implement SaliencyNet we combined three loss functions with different properties and have shown that our adaptation of CycleGan improved segmentation results after enhancement relative to both segmentation of non-enhanced images and of those enhanced with the vanilla CycleGan.

We compared the segmentation results of widely used non-parametric Otsu thresholding and StarDist on both raw images and images enhanced with SaliencyNet of our novel organoid live-cell imaging dataset. We have shown that using SaliencyNet improved the segmentation quality of both classical and deep learning based nuclei segmentation algorithms in low SNR nuclei images. It should be noted that adding the SaliencyNet enhancement step prior to nuclei segmentation did not degrade the quality of results for the already salient datasets.

Taken together, these results show that

SaliencyNet is a useful new step for nuclei segmentation workflows.

7 CODE AVAILABILITY

SaliencyNet network's code for training and testing nuclei enhancement is fully open source and available on GitHub at <https://github.com/cbib/SaliencyNet>. Our best pre-trained model M used in this study is also available from the same GitHub page.

REFERENCES

- Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., et al. (2017). Qupath: Open source software for digital pathology image analysis. *Scientific reports*, 7(1):1–7.
- Beghin, A., Greci, G., Sahni, G., Guo, S., Rajendiran, H., Delaire, T., Mohamad Raffi, S. B., Blanc, D., de Mets, R., Ong, H. T., et al. (2022). Automated high-speed 3d imaging of organoid cultures with multi-scale phenotypic quantification. *Nature Methods*, 19(7):881–892.
- Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghghi, M., Heng, C., Becker, T., Doan, M., McQuin, C., et al. (2019). Nucleus segmentation across imaging experiments: the 2018 data science bowl. *Nature methods*, 16(12):1247–1253.
- Chouaib, R., Safieddine, A., Pichon, X., Imbert, A., Kwon, O. S., Samacoits, A., Traboulsi, A.-M., Robert, M.-C., Tsanov, N., Coleno, E., Poser, I., Zimmer, C., Hyman, A., Le Hir, H., Zibara, K., Peter, M., Mueller, F., Walter, T., and Bertrand, E. (2020). A dual protein-mrna localization screen reveals compartmentalized translation and widespread co-translational rna targeting. *Developmental Cell*, 54(6):773–791.e5.
- De Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Meas-Yedid, V., Pankajakshan, P., Lecomte, T., Le Montagner, Y., et al. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nature methods*, 9(7):690.
- Du, R., Li, W., Chen, S., Li, C., and Zhang, Y. (2021). Unpaired underwater image enhancement based on cyclegan. *Information*, 13(1):1.
- Ettinger, A. and Wittmann, T. (2014). Fluorescence live cell imaging. *Methods in cell biology*, 123:77–94.
- Galland, R., Greci, G., Aravind, A., Viasnoff, V., Studer, V., and Sibarita, J.-B. (2015). 3d high-and super-resolution imaging using single-objective spim. *Nature methods*, 12(7):641–644.
- Gao, P., Tian, T., Li, L., Ma, J., and Tian, J. (2021). Decyclegan: An object enhancement network for weak vehicle detection in satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:3403–3414.
- Gómez-de Mariscal, E., García-López-de Haro, C., Ouyang, W., Donati, L., Lundberg, E., Unser, M., Muñoz-Barrutia, A., and Sage, D. (2021). Deepimagej: A user-friendly environment to run deep learning models in imagej. *Nature Methods*, 18(10):1192–1195.
- Haase, R., Royer, L. A., Steinbach, P., Schmidt, D., Dibrov, A., Schmidt, U., Weigert, M., Maghelli, N., Tomancak, P., Jug, F., et al. (2020). Clij: Gpu-accelerated image processing for everyone. *Nature methods*, 17(1):5–6.
- He, J., Wang, C., Jiang, D., Li, Z., Liu, Y., and Zhang, T. (2020). Cyclegan with an improved loss function for cell detection using partly labeled images. *IEEE Journal of Biomedical and Health Informatics*, 24(9):2473–2480.
- Hollandi, R., Moshkov, N., Paavolainen, L., Tasnadi, E., Piccinini, F., and Horvath, P. (2022). Nucleus segmentation: towards automated solutions. *Trends in Cell Biology*.
- Hollandi, R., Szkalitsy, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., Grexa, I., Molnar, J., Balind, A., Gorbe, M., et al. (2020). nucleaizer: a parameter-free deep learning framework for nucleus segmentation using image style transfer. *Cell Systems*, 10(5):453–458.
- Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134.
- Kanopoulos, N., Vasanthavada, N., and Baker, R. L. (1988). Design of an image edge detection filter using the sobel operator. *IEEE Journal of solid-state circuits*, 23(2):358–367.
- Kim, Y. and Varshney, A. (2006). Saliency-guided enhancement for volume visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):925–932.
- Li, G., Liu, T., Tarokh, A., Nie, J., Guo, L., Mara, A., Holley, S., and Wong, S. T. (2007). 3d cell nuclei segmentation based on gradient flow tracking. *BMC cell biology*, 8(1):1–10.
- Lu, Z. and Chen, Y. (2022). Single image super-resolution based on a modified u-net with mixed gradient loss. *signal, image and video processing*, 16(5):1143–1151.
- Malpica, N., De Solórzano, C. O., Vaquero, J. J., Santos, A., Vallcorba, I., García-Sagredo, J. M., and Del Pozo, F. (1997). Applying watershed algorithms to the segmentation of clustered nuclei. *Cytometry: The Journal of the International Society for Analytical Cytology*, 28(4):289–297.
- McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., Doan, M., Ding, L., Rafelski, S. M., Thirstrup, D., et al. (2018). Cellprofiler 3.0: Next-generation image processing for biology. *PLoS biology*, 16(7):e2005970.
- Pang, Y., Lin, J., Qin, T., and Chen, Z. (2021). Image-to-image translation: Methods and applications. *IEEE Transactions on Multimedia*.

- Qi, Y., Yang, Z., Sun, W., Lou, M., Lian, J., Zhao, W., Deng, X., and Ma, Y. (2021). A comprehensive overview of image enhancement techniques. *Archives of Computational Methods in Engineering*, pages 1–25.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–682.
- Schmidt, U., Weigert, M., Broaddus, C., and Myers, G. (2018). Cell detection with star-convex polygons. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 265–273. Springer.
- Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature methods*, 18(1):100–106.
- Wang, L., Chen, W., Yang, W., Bi, F., and Yu, F. R. (2020). A state-of-the-art review on image synthesis with generative adversarial networks. *IEEE Access*, 8:63514–63537.
- Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612.
- Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., et al. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature methods*, 15(12):1090–1097.
- You, Q., Wan, C., Sun, J., Shen, J., Ye, H., and Yu, Q. (2019). Fundus image enhancement method based on cyclegan. In *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pages 4500–4503. IEEE.
- Zhou, Y., Onder, O. F., Dou, Q., Tsougenis, E., Chen, H., and Heng, P.-A. (2019). Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In *International conference on information processing in medical imaging*, pages 682–693. Springer.
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232.

Chapitre 4

Analyses de la distribution spatiale subcellulaire de l'ARNm et des protéines dans les données smFISH

4.1 Contexte scientifique

Comme nous l'avons introduit dans le chapitre 1, section 1.4, la localisation subcellulaire des molécules est un intérêt majeur pour décrire la fonction des cellules [108, 119, 198, 13, 26] et bien que des études aient tenté de quantifier la distribution spatiale des ARNm [19, 129, 167, 192], la majorité des études portant sur la localisation subcellulaire des ARNm et protéines sont plutôt qualitatives, manquant de quantifications détaillées pour décrire systématiquement les positions des ARNm et des protéines. Elles se sont généralement limitées à des systèmes dans lesquels l'hétérogénéité spatiale est maîtrisée et où les régions subcellulaires sont facilement définies, par exemple, les modèles de développement [170, 102], les systèmes neuronaux [22, 6], ainsi que les cellules polarisées [111, 116, 193]. Ainsi, des approches capables de quantifier la localisation subcellulaire de l'ARN, permettraient d'identifier et de caractériser davantage les différents états et sous-populations cellulaires. De ce fait, pour élucider les mécanismes de distribution spatiale des molécules, il faut développer des outils d'analyse quantitative qui permettent d'analyser les localisations subcellulaires de manière statistiquement robuste.

Dans ce chapitre, nous présenterons DypFISH, une librairie d'analyse de données spatiales, qui permet de surmonter certaines des limites mentionnées précédemment en analysant quantitativement les distributions spatiales d'ARNm et de protéines. Entre autres, DypFISH introduit des techniques unifiées d'analyse qui permettent l'évaluation conjointe des données ponctuelles d'ARNm dans les images smFISH et des données continues de protéines par immunofluorescence (IF).

Cette librairie se veut généraliste et extensible et se base sur deux éléments essentiels :

le signal à analyser (ARNm ou protéines) et la grille de lecture spatiale associée. À partir de ces deux éléments, DypFISH permet d'analyser (i) des patterns de distribution de molécules, (ii) la corrélation de la localisation des molécules avec des repères cellulaires choisis, par exemple des organites et (iii), les corrélations des distributions des ARNm et des protéines.

Ainsi, en aval des analyses de DypFISH, les images doivent être traitées pour extraire des descripteurs dits "primaires" (figure 4.1, panneau A) des images, notamment la localisation des ARNm (voir chapitre 2) et la mesure de l'intensité du signal protéique (voir chapitre 1, section 1.3) et pour définir la grille de lecture construite sur la base de la segmentation de cellules et de noyaux (voir chapitre 3) ainsi que de la localisation d'organites, par exemple du centrosome (MTOC). Ensuite, les analyses spatiales (voir chapitre 1, section 1.4) peuvent être appliquées en utilisant les différentes briques proposées dans la librairie DypFISH, tel que présenté dans la figure 4.1.

La première étape dans l'utilisation de la librairie DypFISH est donc la définition de la grille de lecture spatiale individuellement pour chaque image de cellule. Ensemble avec le signal à analyser (localisation des ARNm ou intensité protéique), cela constitue ce que nous avons appelé des "descripteurs secondaires". Plusieurs types de grilles de lecture peuvent être définis, premièrement un découpage en isolignes, allant de la périphérie de la cellule jusqu'à la bordure du noyau dénommé *peripheral distance map*, deuxièmement un découpage en quadrants qui correspondent à des rayons partant du noyau et allant jusqu'à la membrane cellulaire, chaque quadrant étant numéroté dans le sens des aiguilles d'une montre avec le quadrant contenant le centrosome pour point de départ.

Ces grilles de lecture, assemblées au signal, permettent de quantifier ce dernier dans chacun des segments définis par les isolignes et les quadrants de chaque cellule. Dans notre étude, le repère choisi pour aligner ces grilles de lecture de différentes cellules et rendre les quantifications dans différents segments des cellules comparables est le centrosome. Ces quantifications par régions peuvent être calculées en 2D et en 3D.

La seconde partie de la librairie implémente les fonctions de statistiques spatiales sur la distribution des ARNm (figure 4.1, panneau A (ii)), sur la base des descripteurs secondaires précédemment définis. À la différence de descripteurs cellulaires définis pour une cellule, ces statistiques s'appliquent à un ensemble de cellules - par exemple, correspondant à l'acquisition d'images dans une condition d'intérêt (ex. drogue / sans drogue). Parmi les analyses possibles, on trouve par exemple l'enrichissement de la densité des molécules à la périphérie de la cellule ou bien, à l'inverse, à proximité du noyau. Par ailleurs, les mesures de dispersion du signal sont complétées par la fonction *Ripley K* (voir chapitre 1, section 1.4) que nous avons généralisée pour estimer le degré de clustering des signaux ARNm et protéine, malgré les différences de la nature du signal (ponctuel pour les ARNm

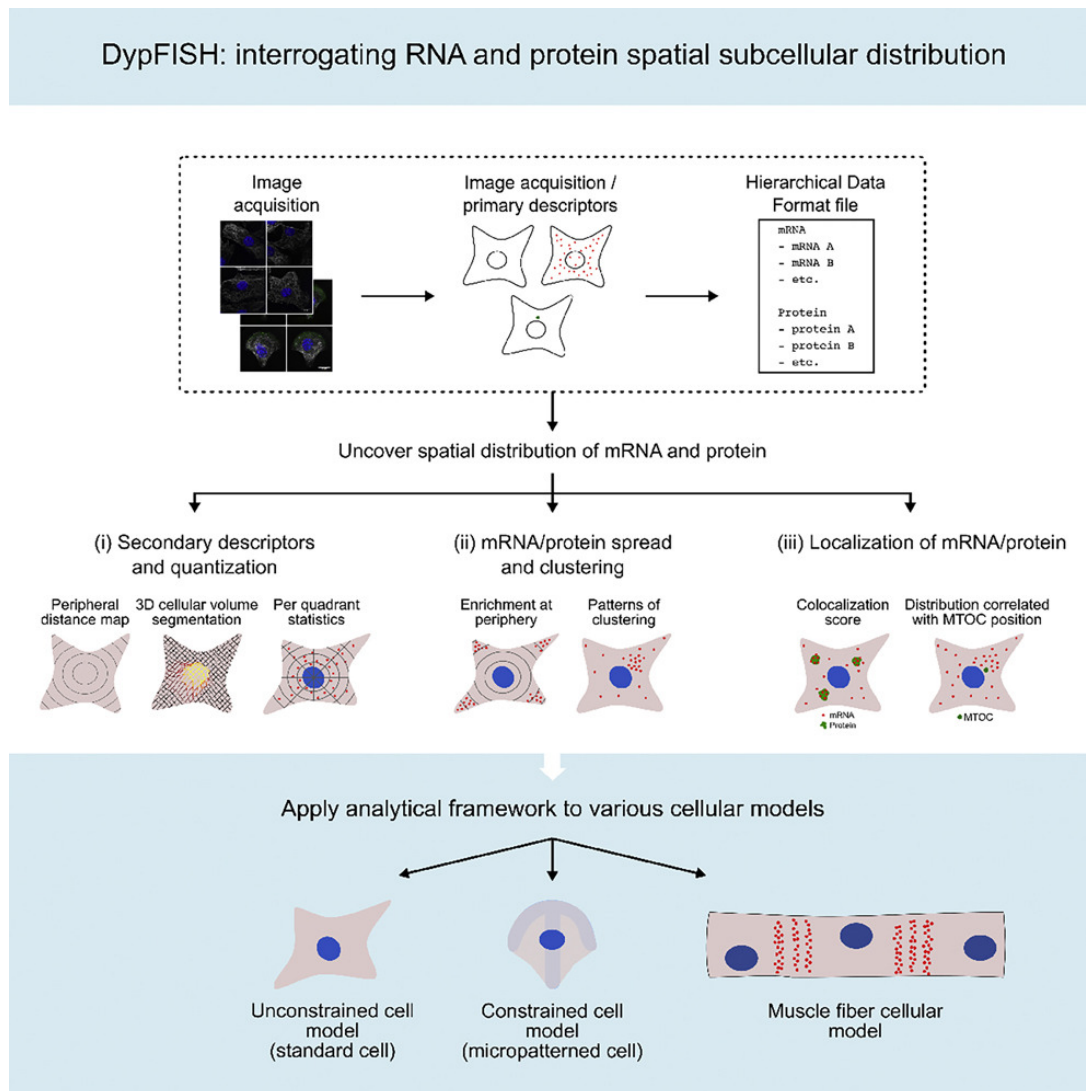


FIGURE 4.1 – Boîte à outils DypFISH. (A) Traitement des images de sm-FISH et IF, segmentation des cellules et des noyaux, localisation des ARNm et des protéines. Cette étape est suivie du calcul d'un ensemble de statistiques pour (i) des images individuelles, (ii) et (iii) pour des ensembles d'images. Le panneau (B) présente les différents modèles cellulaires sur lesquels DypFISH a été appliqué.

et continu pour les protéines).

Finalement, les fonctions permettant d'étudier la façon dont le signal des ARNm/protéines est organisé dans les cellules (figure 4.1 panneau A (iii)) sont aussi proposées dans la librairie et s'appliquent également à des ensembles d'images. Pour ce faire, grâce à la localisation d'un repère d'intérêt, combiné à la quantification précédemment mentionnée des ARNm et des protéines, il est possible de mesurer leur enrichissement dans les quadrants contenant ce repère par rapport au reste de la cellule. Par ailleurs, il est possible d'établir un score de colocalisation qui mesure si deux ensembles différents d'images, par

exemple un ARNm et une protéine, présentant des profils similaires dans la distribution de leur densité au sein des deux groupes d'images correspondants.

4.2 Contribution scientifique

La librairie DypFISH se voulant généraliste, nous avons appliqué les différentes fonctionnalités précédemment décrites à plusieurs types de cellules :

1. des fibroblastes de souris dont la morphologie a été contrainte par un "micropattern", un plastique adhésif en forme d'ancre, permettant de réduire la variabilité entre les cellules [172, 151],
2. des fibroblastes de souris standard, non contraints par un micropattern, utilisées comme cellules témoins,
3. des cellules de muscles squelettiques de souris.

Pour les deux types de fibroblastes, contraints et non contraints, la polarisation a été ensuite induite pour être étudiée. Une cellule est dite polarisée lorsqu'elle a développé un axe d'organisation principal suite à la réorganisation de son cytosquelette et de ses organites, aboutissant à une migration ou une croissance directionnelle.

Dans la suite de ce chapitre, nous présentons des exemples d'analyses spatiales qu'il est possible d'effectuer avec la librairie DypFISH.

4.2.1 Enrichissement localisé

Ce premier exemple illustre le type d'analyse représenté par la figure 4.1 panneau A (ii), *mRNA/protein spread and clustering*.

Enrichissement périphérique

Nous avons calculé la fraction périphérique des ARNm et des protéines à un pourcentage donné p de la distance radiale entre le noyau et la membrane plasmique. La densité périphérique est définie comme le ratio du nombre de molécules (le décompte des ARNm ou la somme des intensités pour les protéines) présents à la périphérie, sur le nombre de molécules présents dans tout le cytoplasme. Pour cela, nous définissons un vecteur $C = (c_1, \dots, c_{100})$ contenant les décomptes des ARNm à une distance inférieure ou égale à $i \in [0, \dots, 100]$ de la paroi cellulaire. Ce vecteur C est divisé par les valeurs de volume périphérique $V^{periphery}$ correspondantes pour obtenir la densité dans des régions concentriques depuis la périphérie de la cellule et pour chaque isoligne (c_i/V^i). Ces densités périphériques

locales sont ensuite normalisées par la densité totale du cytoplasme $d_{\text{cytoplasme}}$, donnant un vecteur F de densités relatives pour chaque région tel que $F = \left(\frac{c_i/V^i}{d_{\text{cytoplasme}}}\right)$.

En utilisant cette approche, nous avons étudié les profils d'enrichissement périphérique des ARNm et des protéines correspondantes de plusieurs ARNm, *Arhgdia*, *Pard3*, β -*Actin*, *Gapdh*, *Rab13* et *Pkp4*. Pour ce faire, nous avons calculé un vecteur F pour les ARNm et pour les protéines à partir desquels il est alors possible de mesurer s'il y a un enrichissement dans la périphérie des cellules polarisées en calculant la part des ARNm situés à la périphérie de la cellule. Nous avons constaté que certains transcrits, dont *Rab13*, *Pkp4* et *Pard3*, sont enrichis en périphérie et jusqu'à 20 % de la distance radiale (voir figure 4.2). Les distributions de l'ARNm et de la protéine de *Gapdh* ont été utilisées comme témoins, car il a été montré précédemment que ni son ARNm ni sa protéine ne présentent de profils d'enrichissement périphériques [116].

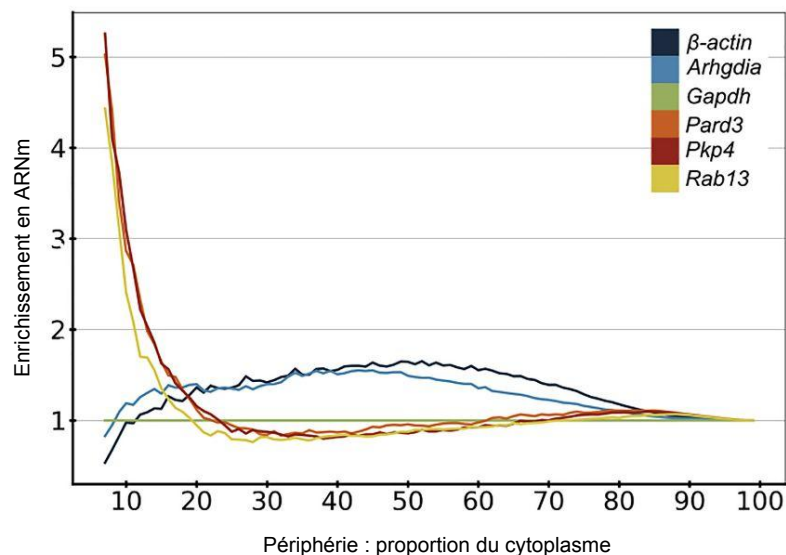


FIGURE 4.2 – Comparaison de l'enrichissement de cinq ARNm par rapport à l'ARNm *Gapdh* dans la région périphérique de la cellule, dont la largeur varie de 0% à 100% de la distance radiale depuis la membrane plasmique jusqu'au noyau.

Enrichissement localisé autour des stries Z dans les cellules musculaires

Nous avons également estimé la densité localisée d'ARNm dans des cellules de muscles squelettiques de souris, dans des cellules matures et immatures. Pour cela, nous avons analysé la distribution de l'ARNm *actn2* qui code pour une protéine présente au niveau des stries Z lors de la différenciation musculaire. Nous avons défini une grille de lecture adaptée à ce type de cellules longitudinales (voir figure 4.3), en découpant la cellule en

quadrants verticaux dans lesquels nous avons mesuré la densité locale d'ARNm. Pour une cellule M , nous la découpons en n segments verticaux q_i, \dots, q_n . Soit t_i le nombre de spots appartenant à un segment q_i , la concentration relative c_i en ARNm dans q_i est définie de la façon suivante :

$$c_i = \frac{t_i/a_i}{T_{ARNm}/A} \quad (4.1)$$

où A correspond à l'aire du masque cellulaire et a_i à l'aire du masque correspondant au segment q_i . Nous avons produit des graphiques de distribution et des heatmaps représentant la densité locale d'ARNm entre deux noyaux, pour $n = 20$ et $n = 80$.

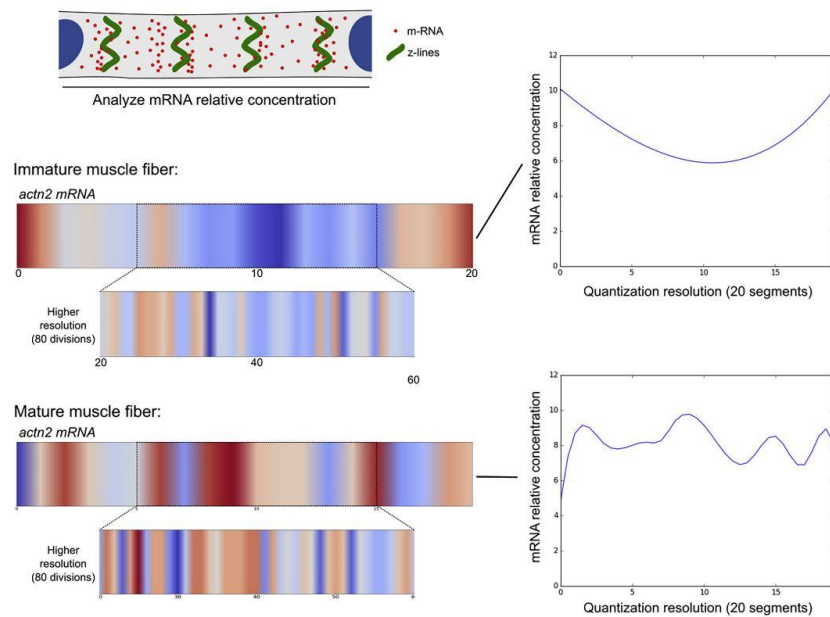


FIGURE 4.3 – Chaque cellule a été quantifiée en quadrants verticaux, et la densité relative d'ARNm dans chaque quadrant a été calculée. Une densité locale en forme de vague est observé pour *actn2* dans les fibres matures par rapport aux fibres immatures.

Nous avons observé un haut degré de clustering en forme de vague pour l'ARNm *actn2* dans les myofibres matures par rapport au myofibres immature, suggérant que la distribution de l'ARNm *actn2* suit spécifiquement l'organisation de sa protéine au lieu de la précéder (voir figure 4.3) et donc que la localisation de l'ARNm *actn2* précède l'organisation de la protéine formant les stries Z.

4.2.2 Degré de clustering des ARNm et des protéines

Ce second exemple illustre également le type d'analyse représenté par la figure 4.1 panneau A (ii), *mRNA/protein spread and clustering*.

Nous avons analysé le clustering des ARNm et des protéines en utilisant une généralisation de la fonction Ripley K (voir chapitre 1, section 1.4). La fonction Ripley K est couramment utilisée pour caractériser l'étendue du clustering de points tels que les ARNm. Nous avons utilisé une généralisation de l'algorithme Ripley-K [187] pour évaluer la présence ou l'absence de clustering dans les distributions spatiales des ARNm et des protéines. Pour ce faire, nous avons défini le "degré de clustering", une mesure sans unité qui peut être utilisée pour comparer le clustering entre différentes molécules, dont la mesure peut être discrète ou continue.

Nous avons pour cela normalisé la fonction \hat{K} sous l'hypothèse du processus homogène de Poisson, pour obtenir ce qui est communément appelée la fonction H de Ripley. Ceci permet de définir l'indice de clustering H^* dans un cercle (ou une sphère dans le cas d'images 3D) de radius r , comme un estimateur de $\hat{H}(r)$ en comparant la fonction H de Ripley calculée empiriquement à une distribution spatiale aléatoire complète (CSR) tel que :

$$\hat{H}^*(r) = \hat{H}(r) \quad (4.2)$$

$$\hat{H}^*(r) = \begin{cases} \hat{H}(r)/\hat{H}_{95}(r) & \text{si } \hat{H}_r \geq 0 \\ -\hat{H}(r)/\hat{H}_5(r) & \text{sinon} \end{cases} \quad (4.3)$$

Où $\hat{H}_{95}(r)$ et $\hat{H}_5(r)$ sont les 95^e et 5^e percentiles de $\hat{H}(r)$.

Le degré de clustering est alors défini comme l'aire de \hat{H}^* au-dessus de 1 (voir figure 4.4). Avec cette mesure, il est possible d'évaluer le degré de clustering de tous les transcrits et protéines.

Nous avons calculé le degré de clustering à chaque heure d'acquisition, après l'induction de la polarisation, pour les ARNm et les protéines, en normalisant ensuite les valeurs obtenues par la moyenne pour pouvoir rendre les échelles de valeur compatibles entre l'ARNm et la protéine. Les résultats sont présentés sur la figure 4.5.

Un pic dans le profil de l'ARNm suivi d'un pic dans le profil de la protéine correspondante peut être observé pour *Pard3*, *β -Actin*, et *Gapdh*, montrant une présence de clustering décalée dans le temps (figure 4.5). Une interprétation plausible de cette relation temporelle pourrait être un clustering local de la traduction et de la transcription,

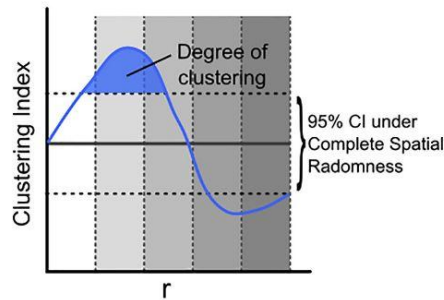


FIGURE 4.4 – Pour calculer le degré de clustering, nous avons défini un estimateur de la fonction Ripley K, en normalisant la fonction aux 95e et 5e percentiles sous le processus homogène de Poisson. Un clustering significatif d'ARNm ou de protéines a lieu lorsque la fonction de l'estimateur est au-dessus de l'intervalle de confiance de 95 %. Le degré de clustering correspond à l'aire sous la courbe de cet estimateur, supérieure au 95e percentile de la distribution spatiale aléatoire.

produisant une concentration locale élevée de protéine. En revanche, une légère diminution de l'enrichissement périphérique et du degré de clustering a été observée pour les protéines des ARNm *Arhgdia* et *Pard3*. Dans l'ensemble, ces données indiquent que dans les fibroblastes où la traduction des protéines est inhibée, les ARNm *Arhgdia* et *Pard3* sont localisés de la même façon que dans les cellules témoins. Cependant, le fait que l'enrichissement périphérique et que le degré de clustering des protéines correspondantes soient moindres, pourrait indiquer une traduction locale de ces gènes à la périphérie de la cellule.

4.2.3 Corrélacion de localisation des ARNm et des protéines avec la position du centrosome

Ce troisième exemple illustre le type d'analyses représentées sur la figure 4.1, panneau A (iii), *Localisation of mRNA/protein*.

Dans la majorité des cellules polarisées, le centrosome est positionné entre le noyau et le bord antérieur de la cellule, avant la migration [56, 50]. Ayant démontré que plusieurs transcriptions d'ARNm et protéines présentent un enrichissement périphérique et un clustering spatial dans les cellules polarisées, nous avons cherché à déterminer si ces observations étaient liées au positionnement du centrosome. Pour ce faire, nous avons divisé les cellules en quadrants, puis calculé la densité relative des ARNm et des protéines pour chaque quadrant. Nous avons ensuite localisé le centrosome dans un de ces quadrants. Nous avons observé un enrichissement plus élevé de tous les transcrits d'ARNm cytoplasmiques dans le quadrant contenant le centrosome par rapport au quadrant ne contenant pas le centrosome. Toutes les protéines observées ont également été trouvées plus enrichies

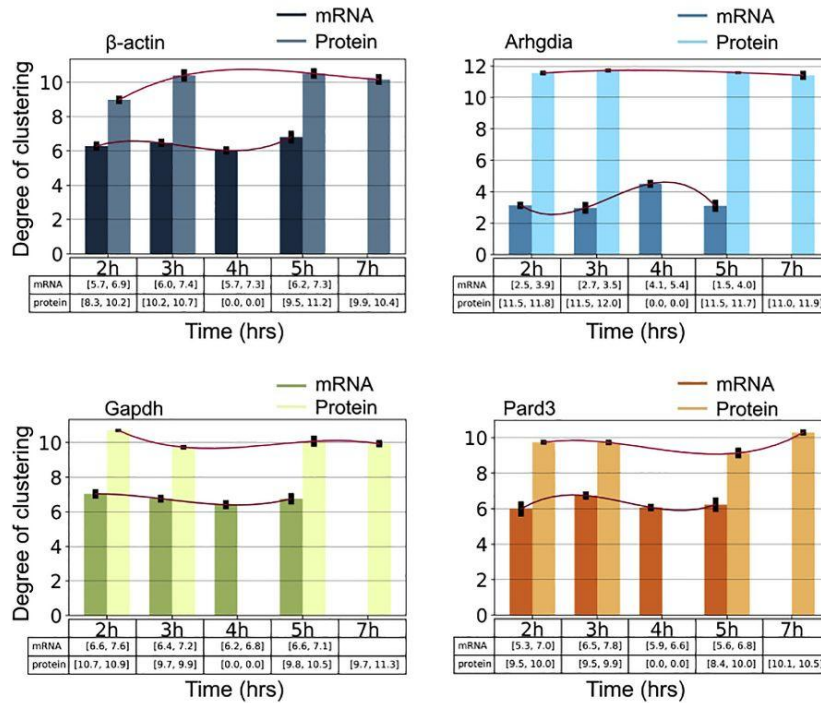


FIGURE 4.5 – Dynamique de clustering pour quatre paires ARNm-protéine en utilisant le degré de clustering normalisé. Les valeurs nulles indiquent l'absence de données sur les protéines ou les ARNm pour l'instant concerné. Les barres d'erreur indiquent l'erreur standard de la moyenne.

dans le quadrant contenant le centrosome par rapport au quadrant ne contenant pas le centrosome, à l'exception de Pard3 (voir figure 4.6).

Sur la base des observations ci-dessus, nous avons introduit un indice de polarité, le *MTOC polarity index (MPI)*, $PI_M \in [-1, 1]$ qui mesure l'enrichissement du signal des ARNm ou des protéines à proximité du centrosome, pour une série d'acquisition d'images donnée.

Pour une série S d'images étudiées, nous notons $S_M = S_i$ l'ensemble de tous les quadrants contenant le centrosome et $S_{-M} = S_j$ l'ensemble des quadrants ne contenant pas le centrosome. Le MPI mesure alors la fréquence à laquelle la concentration relative dans les quadrants MTOC est plus élevée que dans les quadrants non MTOC. Plus formellement, il est défini tel que :

$$PI_M = \frac{2|\{S_i | c_i > m\}|}{|S|} - 1 \quad (4.4)$$

ou m est la médiane des concentrations relatives du signal, c_i est la concentration relative des ARNm dans les quadrants S_M et c_j la concentration relative des ARNm pour

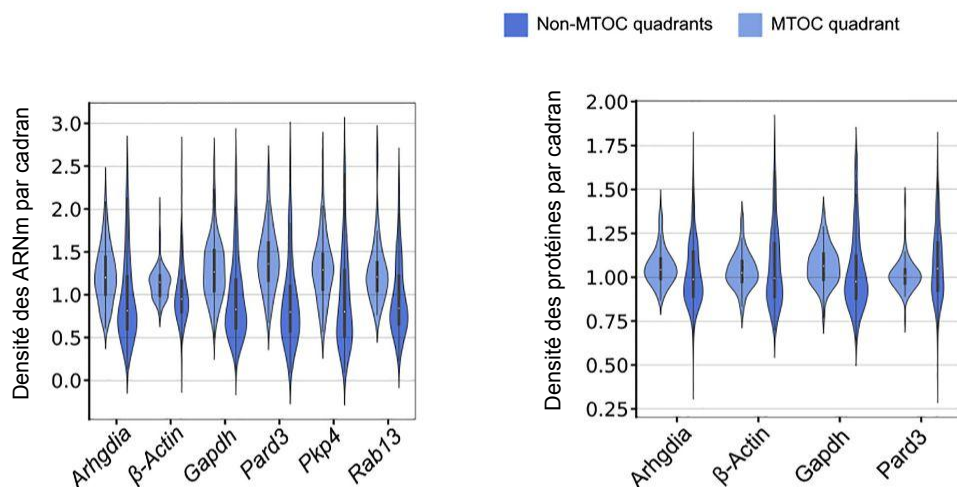


FIGURE 4.6 – Densité relative des ARNm et des protéines cytoplasmiques dans les quadrants ne contenant pas de centrosome (MTOC) (bleu foncé) et dans les quadrants contenant les centrosomes (bleu clair). L'enrichissement dans le quadrant contenant le centrosome est défini par les différences entre les moyennes de ces distributions.

les quadrants S_{-M} .

Des valeurs positives de MPI impliquent un enrichissement des transcrits d'ARNm dépendant du centrosome, des valeurs négatives impliquent un enrichissement éloigné du centrosome, et une valeur de zéro n'implique aucune relation détectable avec le centrosome.

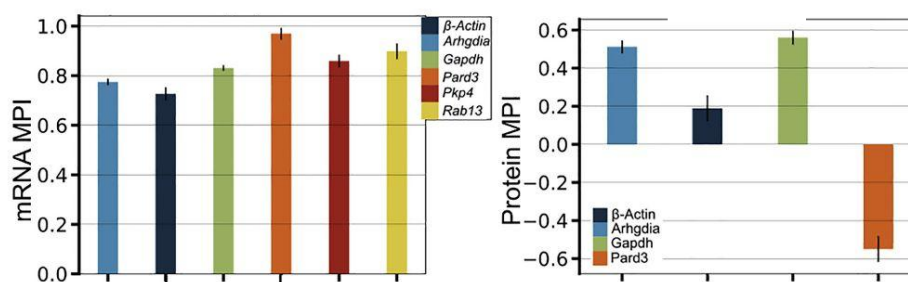


FIGURE 4.7 – Valeurs des MPI pour les ARNm et les protéines présents dans le cytoplasme des cellules.

Nous avons calculé le MPI pour tous les transcrits et toutes les protéines, en calculant les valeurs pour l'ensemble de la population cytoplasmique. Les ARNm ainsi que les protéines pour *Gapdh*, *Arhgdia* et *beta-Actin* présentent des scores de MPI élevés (voir figure 4.7). En revanche, les score de MPI sont faibles pour la protéine *Pard3*, ce qui est cohérent avec le faible enrichissement de *Pard3* au niveau du centrosome (figure 4.6).

4.2.4 Colocalisation des distributions des ARNm et des protéines

Ce quatrième et dernier exemple illustre également le type d'analyses représentées sur la figure 4.1, panneau A (iii), *Localisation of mRNA/protein*.

Pour étudier les corrélations dans les distributions spatiales des ARNm et des protéines, nous avons calculé des scores de colocalisation (CS) pour quatre paires ARNm-protéine *Arhgdia*, *Gapdh*, β -*Actin* et *Pard3*.

Analyse des positions ARNm-protéines à différents intervalles de temps

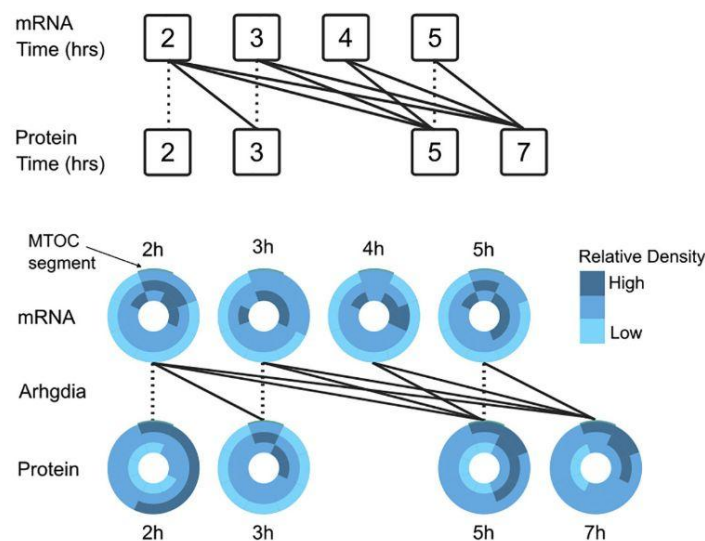


FIGURE 4.8 – Différentes paires d'intervalles de temps sont définies de manière à ce que $t_{ARNm} < t_{protéines}$, afin de calculer le score de colocalisation. Les grilles de densité, représentant les vecteurs de densité relative basés sur une quantification par quadrant pour l'ARNm et la protéine *Arhgdia* sont représentés ici, et ce, pour chaque intervalle de temps. Les quadrants, représentant les régions cellulaires, sont en bleu foncé si la densité relative locale est supérieure à la densité moyenne de tous les quadrants additionnée de l'écart-type, et en bleu clair si la densité relative locale est inférieure à la moyenne de tous les quadrants moins l'écart-type ; les valeurs intermédiaires correspondent à la densité relative dans cet intervalle. Le quadrant contenant le centrosome est mis en évidence par une bordure turquoise.

Le score de colocalisation pour une paire ARNm-protéine donnée est calculé sur la base des vecteurs \bar{C}_{ARNm} et $\bar{C}_{protéin}$, descripteurs de distribution des ARNm et des protéines et prend des valeurs entre 0 et 1. Pour l'analyse des colocalisation ARNm/protéines, nous calculons le Colocalisation Score pour une double mesure Φ où ϕ est un processus discret représentant les emplacement des ARNm et ψ une mesure aléatoire générale représentant les concentrations de protéines. Le CS est défini pour des paires de données à un instant

t tel que $S = (\phi_t) \times (\psi_t)$ en utilisant une mesure de similarité γ . En pratique γ est implémenté par la corrélation de Pearson entre les vecteurs de densité.

Une colocalisation est alors "élevée" ou "faible" pour une paire ARNm/protéine donnée, en fonction de la valeur de la corrélation de Pearson obtenue. Les valeurs élevées (distributions spatiales corrélées) suggèrent une traduction locale dans des régions subcellulaires spécifiques, bien qu'elles n'excluent pas des mécanismes alternatifs tels que des voies de localisation distinctes de l'ARNm et de la protéine avec un transport plus tardif de la protéine. Pour explorer la cohérence des corrélations dans les distributions spatiales entre un ARNm et une protéine à différents intervalles de temps, comme présenté sur la figure 4.8.

Nous avons mesuré le score de colocalisation pour tout le cytoplasme, ainsi que pour la périphérie de la cellule (figure 4.9).

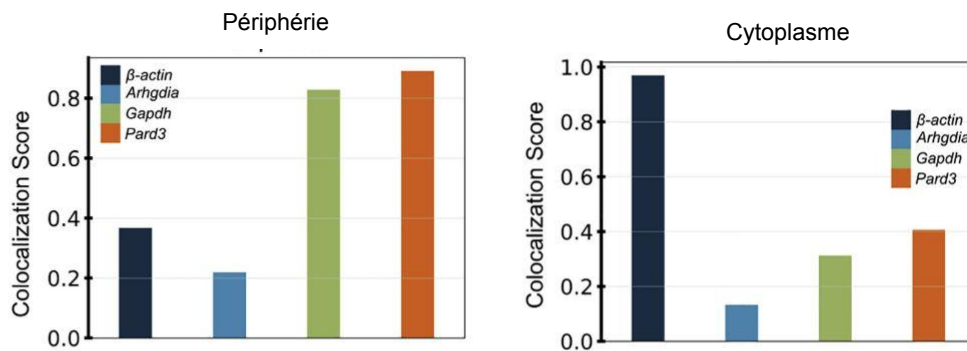


FIGURE 4.9 – Les valeurs de CS ont été calculées en utilisant les corrélations globales de tous les quadrants de la zone cytoplasmique ainsi que les corrélations locales de sous-ensembles de quadrants dans les régions périphériques.

Nous avons également calculé le score de colocalisation pour les paires ARNm-protéine cytoplasmiques, à différents intervalles de temps (figure 4.10). En général, des valeurs CS modérées à élevées ont été observées pour tous les gènes aux différents intervalles de temps.

Nous pouvons observer une variabilité entre les différents gènes dans les valeurs de CS à divers intervalles de temps. Par exemple, l'ARNm β -*Actin* montre des valeurs de CS élevées avec sa protéine correspondante pour la plupart des intervalles de temps, tandis que les valeurs de CS pour l'ARNm *Arhgdia* dans les derniers intervalles sont généralement inférieures aux valeurs de CS pour les premiers intervalles (figure 4.10).

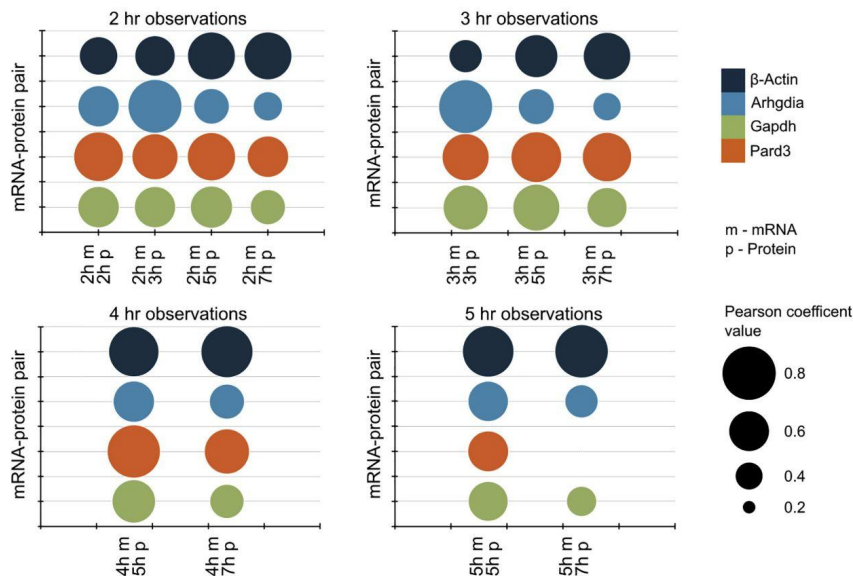


FIGURE 4.10 – Les corrélations des distributions de densité relative pour toutes les paires ARNm-protéine cytoplasmiques ont été observées pour certains intervalles de temps.

4.3 Discussion

Dans ce chapitre, nous avons proposé diverses méthodes pour étudier la localisation de l'ARNm et des protéines avec une précision subcellulaire, sur la base des données de microscopie smFISH et IF. Nous avons proposé une méthode computationnelle quantitative, reproductible et réutilisable pour étudier la distribution spatiale des ARNm et des protéines. Le développement de la librairie d'analyses spatiales DypFISH, a permis de révéler des relations dans la localisation spatiale et temporelle des paires ARNm-protéines entre elles et par rapport à la polarisation des cellules. DypFISH permet de détecter quantitativement les changements dans les profils de localisation et donc de mettre en avant l'existence de corrélations de distributions spatiales pour les paires ARNm-protéines et de les associer à la localisation du centrosome.

DypFISH fournit une boîte à outils pour caractériser divers aspects de la localisation subcellulaire des molécules d'ARNm et de protéines. Un nombre important d'analyses ont été développées et sont disponibles dans cette boîte à outils. La table 4.1 récapitule les différentes analyses spatiales implémentées dans cette librairie.

On peut relever quelques limites à l'utilisation des méthodes implémentées dans DypFISH. Premièrement, la localisation subcellulaire d'une molécule est définie par rapport à un marqueur cellulaire donné et il ne s'agit pas d'un paramètre absolu. Deuxièmement, les cellules micropatternées ont une topologie spécifique et bien définie, ce qui n'est pas le cas des cellules standard en culture, et les définitions de "périphérie", de "centre", etc.

sont donc moins évidentes. Troisièmement, l'annotation du centrosome a été faite manuellement ; l'automatisation est difficile, car le signal utilisant l'anticorps anti-tubuline peut être flou, il y a un certain pourcentage de cellules avec plus d'un centrosome et l'annotation nécessite souvent de regarder l'image en trois dimensions. Finalement, l'effet potentiellement induit par la modification des paramètres de quantification des cellules (tels que le nombre de quadrants et de cercles concentriques) sur les résultats des analyses statistiques n'a pas été évalué.

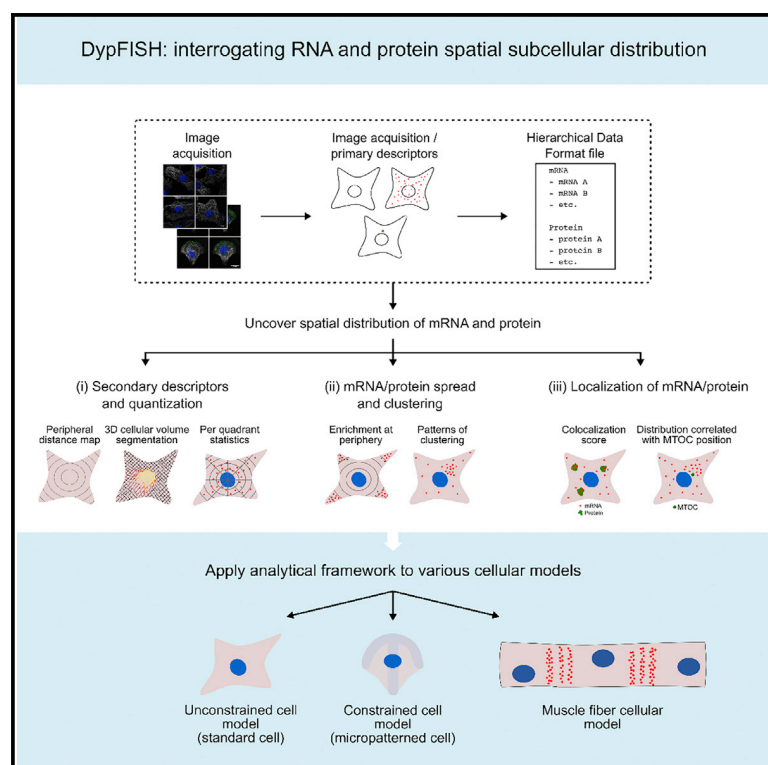
TABLE 4.1 – Liste des méthodes d’analyses spatiales présentes dans la boîte à outils DypFISH.

Analyse	Description
Zero level	Dans une image 3D, détermine quelle slice contient la base de la cellule
Height-map	Segmente la cellule en 3D à partir du zero level
cell volume	Calcul le volume de la cellule sur la base de la height-map
Protein intensities	Mesure l’intensité des protéines, par slice, dans une image 3D
Cytoplasmic total counts	Calcul le nombre total de transcrits présents dans le cytoplasme
Peripheral distance map	Découpe une cellule en cercles concentriques dont le rayon varie entre le noyaux et la membrane cellulaire
2D cell quantization	Génère une tessellation de la cellule à partir de quadrants centrés sur le noyaux
Fine-grained quantization	Génère une grille de lecture permettant une haute précision d’analyse
3D cell quantization	Génère une tessellation en trois dimensions de la cellule dans
Peripheral fraction and enrichment	Calcul la fraction périphérique des ARNm et des protéines à un pourcentage donné de la distance radiale entre la périphérie de la cellule et le noyau
Volume corrected noise measure	Correction des niveaux d’expression d’ARNm en fonction du volume des cellules
Cytoplasmic spread	Mesure la régularité de la répartition d’une molécule dans la cellule
MTOC polarity index	Mesure l’enrichissement en ARNm ou en protéine à proximité du MTOC.
mRNA / protein distribution profile	Crée un vecteur de densités de molécules contenant les densités mesurées dans chaque quadrant
Colocalization score	Mesure l’interdépendance entre les distributions spatiales des ARNm et des protéines
Degree of clustering (Ripley-K)	Compare le clustering entre différentes molécules et différentes conditions cellulaires. Généralisé au signaux continus

4.4 Article

Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH

Graphical abstract



Authors

Anca F. Savulescu, Robyn Brackin, Emmanuel Bouilhol, ..., Edgar R. Gomes, Macha Nikolski, Musa M. Mhlanga

Correspondence

ankasi100@gmail.com (A.F.S.), macha.nikolski@u-bordeaux.fr (M.N.), belenus@mhlantalab.org (M.M.M.)

In brief

Savulescu et al. present DypFISH, a method to quantitatively investigate the subcellular spatial localization of RNA and proteins. DypFISH introduces a range of analytical techniques and statistics to interrogate single-molecule RNA FISH data in combination with protein immunolabeling.

Highlights

- Micropatterning of cells enhances reproducibility of mRNA subcellular distributions
- The DypFISH method is suited to study clustering patterns of mRNA and protein
- By using DypFISH, mRNA peripheral enrichment and in relation to the MTOC was found
- DypFISH revealed colocalization patterns of biomolecules in relation to each other



Article

Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH

Anca F. Savulescu,^{1,*} Robyn Brackin,⁵ Emmanuel Bouilhol,^{2,7} Benjamin Dartigues,² Jonathan H. Warrell,⁶ Mafalda R. Pimentel,⁴ Nicolas Beaume,¹ Isabela C. Fortunato,⁴ Stephane Dallongeville,³ Mikaël Boule,^{8,9} Hayssam Soueidan,² Fabrice Agou,^{8,10} Jan Schmoranzler,⁵ Jean-Christophe Olivo-Marin,³ Claudio A. Franco,⁴ Edgar R. Gomes,⁴ Macha Nikolski,^{2,7,*} and Musa M. Mhlanga^{11,12,13,14,*}

¹Division of Chemical, Systems & Synthetic Biology, Institute for Infectious Disease & Molecular Medicine, Faculty of Health Sciences, University of Cape Town, 7295 Cape Town, South Africa

²Université de Bordeaux, Bordeaux Bioinformatics Center, 33000 Bordeaux, France

³Unité D'Analyse D'Images Biologiques, Institut Pasteur, 75015 Paris, France

⁴Instituto de Medicina Molecular, Faculdade de Medicina Universidade de Lisboa, 1649-028 Lisbon, Portugal

⁵Advanced Medical Bioimaging, Charité – Universitätsmedizin, 10-117 Berlin, Germany

⁶Molecular Biophysics & Biochemistry, Yale University, New Haven, CT 06520, USA

⁷Université de Bordeaux, CNRS, IBGC, UMR 5095, 33077 Bordeaux, France

⁸Chemogenomic and Biological Screening Core Facility, C2RT, Department of Structural Biology and Chemistry, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France

⁹Université de Paris, Sorbonne Paris Cité, Paris, France

¹⁰Department of Structural Biology and Chemistry, URA 2185, Pasteur Institute, Paris, France

¹¹Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University Medical Center, 6525 GA Nijmegen, the Netherlands

¹²Epigenomics & Single Cell Biophysics Group, Department of Cell Biology, FNWI, Radboud University, 6525 GA Nijmegen, the Netherlands

¹³Department of Human Genetics, Radboud University Medical Center, 6525 GA Nijmegen, the Netherlands

¹⁴Lead contact

*Correspondence: ankasi100@gmail.com (A.F.S.), macha.nikolski@u-bordeaux.fr (M.N.), belenus@mhlangalab.org (M.M.M.)
<https://doi.org/10.1016/j.crmeth.2021.100068>

MOTIVATION Subcellular spatial localization of RNAs and proteins, in addition to their expression levels, is widely acknowledged to be important for characterization of cell types and states. Alongside developments in single-cell RNA sequencing (scRNA-seq) techniques, there is a growing need to develop technologies that are able to provide quantitative information on the spatial location of molecules on a single-cell level. Thus, we developed DypFISH, an experimental and computational approach providing quantitative spatial measurements of molecules on a subcellular level, including clustering statistics, their location in relation to specific organelles, and colocalization between transcripts and their corresponding proteins. DypFISH provides a versatile toolbox and aims to enable the investigation of subcellular localization of molecules at fine-grained spatial resolution in single cells for a variety of cell types and topologies.

SUMMARY

Advances in single-cell RNA sequencing have allowed for the identification of cellular subtypes on the basis of quantification of the number of transcripts in each cell. However, cells might also differ in the spatial distribution of molecules, including RNAs. Here, we present DypFISH, an approach to quantitatively investigate the subcellular localization of RNA and protein. We introduce a range of analytical techniques to interrogate single-molecule RNA fluorescence *in situ* hybridization (smFISH) data in combination with protein immunolabeling. DypFISH is suited to study patterns of clustering of molecules, the association of mRNA-protein subcellular localization with microtubule organizing center orientation, and interdependence of mRNA-protein spatial distributions. We showcase how our analytical tools can achieve biological insights by utilizing cell micropatterning to constrain cellular architecture, which leads to reduction in subcellular mRNA distribution variation, allowing for the characterization of their localization patterns. Furthermore, we show that our method can be applied to physiological systems such as skeletal muscle fibers.



INTRODUCTION

The need to incorporate subcellular spatial information of central dogma molecules into traditional omics approaches has led to the call for spatially resolved omics of various kinds (Crosetto et al., 2015). This has become more urgent as projects such as the Human Cell Atlas use technologies such as single-cell RNA sequencing (scRNA-seq) to characterize subtypes of cells on the basis of their molecular signatures by counting the number of RNA transcripts. Although much progress has been made in spatially resolved transcriptomics (reviewed in Medioni and Besse, 2018; Strell et al., 2019), incorporating spatial information into omics approaches carries with it several difficulties, such as coping with biological heterogeneity and noise. Standard approaches such as RNA-seq to measure gene expression typically neglect spatial information. By revealing three-dimensional (3D) subcellular positions of RNA transcripts, cell states as defined by scRNA-seq might be altered or redefined, thus gaining higher resolution and information content.

The importance of subcellular localization of mRNA transcripts as a means to spatially and temporally restrict translation has been demonstrated in a wide variety of cell types (Bashirullah et al., 1998; Besse and Ephrussi, 2008; Jansen, 2001; Kloc et al., 2002; Martin and Ephrussi, 2009; Zappulo et al., 2017; Chouaib et al., 2020). Localizing specific mRNA transcripts to distinct subcellular localizations and subsequent local translation serves as an important determinant of protein localization and is often influential to cell function (Mardakheh et al., 2015; Moor et al., 2017; Zappulo et al., 2017). Although the number of fully characterized localized mRNAs is currently small, emerging studies have demonstrated that the RNA localization phenomenon is more widespread than previously assumed and might in fact be relevant for the majority of mRNA transcripts, including long non-coding RNAs (Bouvette et al., 2018; Cabili et al., 2015; La Manno et al., 2018; Lécuycer et al., 2007; Moor et al., 2017; Sharp et al., 2011; Weis et al., 2013; Zappulo et al., 2017). Although a few studies have attempted to quantify spatial distribution of RNA (Briley et al., 2015; Park et al., 2012; Samacoits et al., 2018; Stueland et al., 2019; Yamagishi et al., 2009), the majority of studies investigating subcellular localization of numerous RNAs and proteins have been generally qualitative, lacking detailed quantitative approaches to systematically describe the positions of RNAs and proteins. They have typically been limited to systems in which spatial heterogeneity is controlled and subcellular partitions are easily defined, such as developmental models (Macdonald and Struhl, 1988; Tautz and Pfeifle, 1989), neuronal systems (Batish et al., 2012; Buxbaum et al., 2014), and polarized cells (Martin and Ephrussi 2009; Mili et al., 2008; Clatterbuck-Soper et al., 2017). Thus, a method that is able to capture and quantify dynamic RNA subcellular positioning, complementary to scRNA-seq, would help in the further identification and characterization of different cell states and subpopulations. In sum, to unravel the mechanisms of RNA spatial and temporal distribution, quantitative analytical tools that probe these relationships systematically need to be developed.

Here, we describe DypFISH, a spatially resolved high-throughput computational approach overcoming the aforemen-

tioned limitations by quantitatively measuring and analyzing the spatial colocalization of mRNA and protein distributions at fine-grained spatial resolution in single cells. On the experimental side, we make use of micropatterning of cells, which have been shown to lead to reproducible cell size and shape (reviewed in Théry, 2010), as well as spatial organization of organelles (Schauer et al., 2010), allowing the averaging of high numbers of cells. By using micropatterns, we are able to showcase the analysis of distinct patterns of subcellular localization of molecules of interest.

On the computational side, DypFISH introduces analytical techniques that allow joint analysis of discrete point-based small-molecule fluorescence *in situ* hybridization (smFISH) mRNA data and continuous intensity immunofluorescence (IF) protein data. The computational techniques include a generalized approach to identifying clustering patterns across different time points, an approach to identifying dependencies of mRNA and protein spatial distributions on organelle positioning, and an approach to identifying correlated spatial distributions between mRNA transcripts and their corresponding protein products globally and at specific subcellular locations. The DypFISH computational framework allows the uncovering of fine-grained aspects of localization patterns for different time points of RNA and proteins. DypFISH probes the dependencies uncovered through perturbation studies, thus allowing one to test for possible mechanisms underlying subcellular localization dynamics.

RESULTS

Analytical method

DypFISH is designed as a library for writing analysis scripts to study subcellular spatial distributions of molecules. The preliminary step for using the DypFISH framework is processing cellular images and storing the extracted information in an hdf5 file format, structured in a hierarchical way (Figure 1A, top box). In our case the relevant categories were the molecule name and type, acquisition time points, and the images themselves. Each image in the hdf file is required to have primary image descriptors (Figure 1A). These include cell and nucleus masks as well as position of a landmark of interest, in our case the microtubule organizing center (MTOC), which is indicative of the cell polarity; descriptors encoding the signal (spots' positions for smFISH images and signal intensities for IF images); and for 3D images, the 3D volume segmentation, as well as the zero level that indicates the last Z slice within the focus of the confocal microscope.

The DypFISH analytical framework consists of three major components and is presented schematically in Figure 1A.

- (1) The first functionality of the DypFISH pipeline is computation of secondary descriptors (Figure 1Ai), stored in a separate hdf5 file. First, we compute the peripheral distance map composed of equally spaced isolines between the cell periphery and the nucleus. Second, each cell is subdivided into quadrants that are given a standard order, starting from the MTOC-containing quadrant. Finally, isolines and quadrants taken together allow the quantization of the cell in two and three dimensions and computation

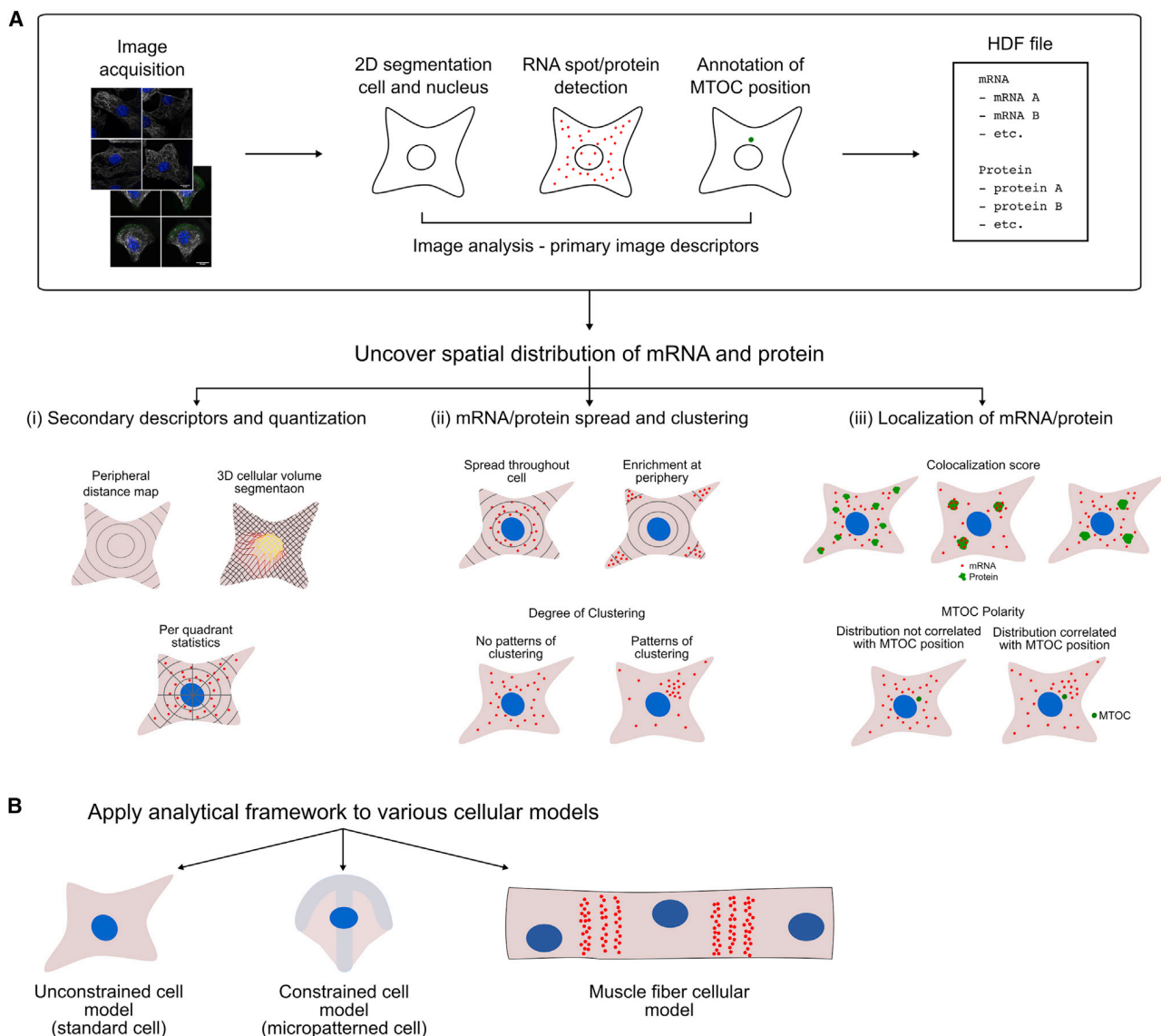


Figure 1. Analytical steps of the DypFISH method

(A) Steps in processing of smFISH and IF microscopy images, followed by various statistics computed by the DypFISH analytical framework either for (i) individual processed images or (ii) and (iii) sets of processed images.

(B) Cellular models in which the DypFISH analytical framework has been applied.

of relative signal densities for each of the segments (Figure 1Ai). Labeling of the MTOC enables cell-to-cell alignment, allowing for comparison of the resulting relative density vectors and segment-per-segment combination in the downstream statistical analyses.

- (2) mRNA spread and clustering are statistics computed for sets of images based on secondary descriptors (Figure 1Aii). Among possible implemented analyses is the enrichment of molecule density at the cell periphery and, conversely, the measure of molecule spread toward the nucleus envelope (centrality). Basic entropy Koza-chenko-Leonenko statistics, which measure how evenly

the mRNA/protein signal is spread, are complemented by our generalized Ripley K function, which allows us to estimate the degree of clustering for mRNA and protein signal in a uniform way, despite differences in the nature of the signal (point-based and continuous).

- (3) Specific analyses are implemented to study how mRNA/protein clusters are organized in the cells (Figure 1Aiii). First, labeling of the MTOC combined with cell quantization allows us to measure the enrichment of mRNAs and proteins in the MTOC-containing quadrants in relation to the rest of the cell. Second, colocalization score measures whether two different sets of images, for example

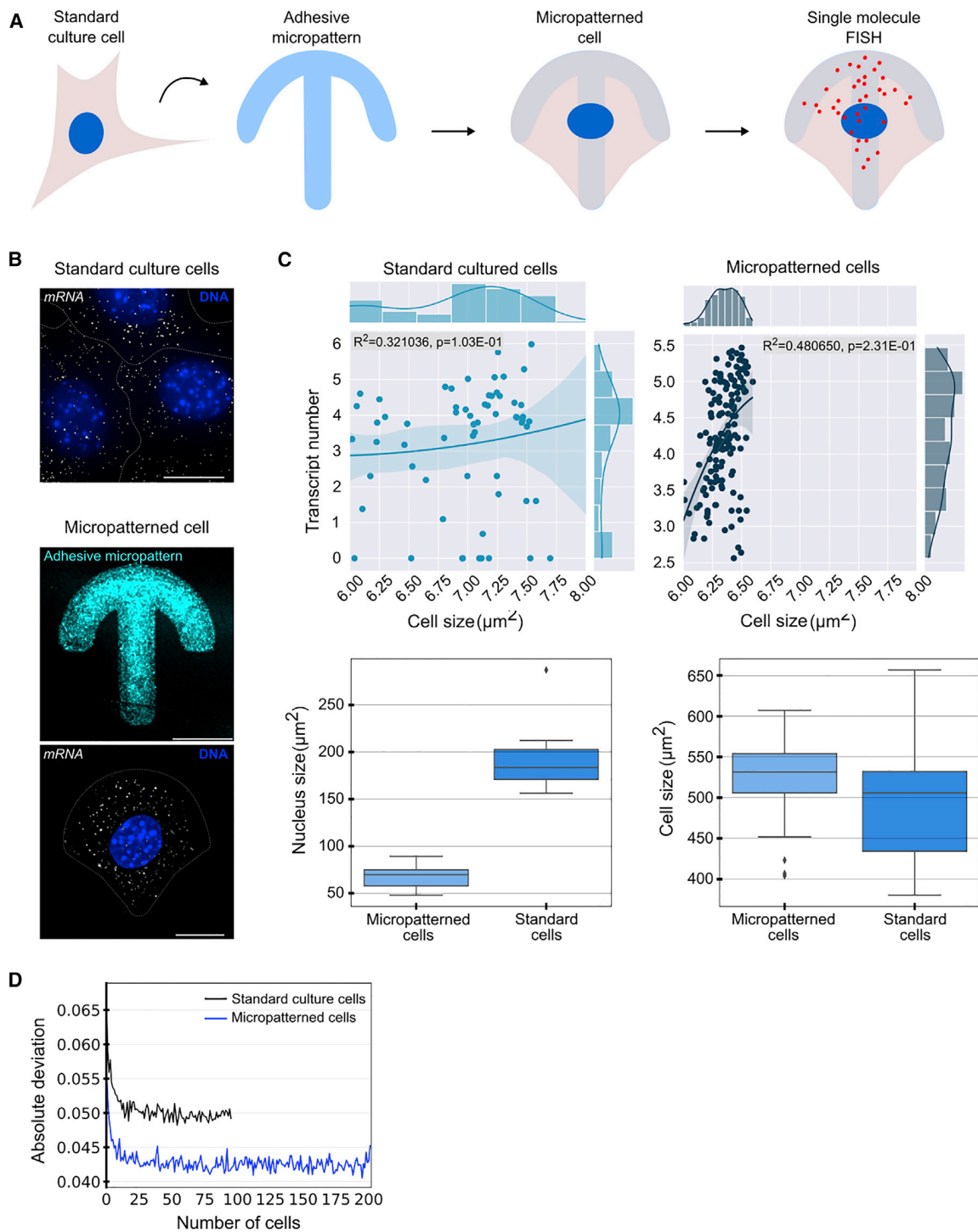


Figure 2. Reproducibility of mRNA and protein distributions in micropatterned cells

(A) Mouse fibroblasts were plated on fibronectin-coated micropatterns and induced to polarize by addition of serum. Cells were fixed and smFISH was performed to target mRNAs of interest.

(B) *Arhgdia* mRNA was visualized by using smFISH (gray) in standard and crossbow-shaped micropatterned mouse fibroblasts, micropatterns were visualized by coating with labeled fibrinogen (cyan), and DNA was stained with Hoechst (blue). Scale bar, 10 μm.

(legend continued on next page)

mRNA at a given time point and its encoded protein at a later time point, exhibit similar patterns of their respective cellular density distributions.

The DypFISH analytical framework was applied to several cellular models, including those with constrained geometry—micropatterned mouse fibroblasts and skeletal muscle fibers—as well as standard unconstrained mouse fibroblasts for several analyses (Figure 1B).

Micropatterning of cells enhances reproducibility of mRNA subcellular distributions

We were interested in the ability to characterize subcellular positioning of RNA and protein in various cellular contexts and states. We selected a fibroblast system that allows the investigation of RNA positioning in relation to its polarity state (Mili et al., 2008) and combined it with micropatterning of cells to reduce cell-to-cell variability. We chose mRNA transcripts that had previously been identified as enriched in lamellipodia of fibroblasts upon polarization and cell migration (Hengst et al., 2009; Mili et al., 2008; Schmoranzler et al., 2009).

Micropatterning has been shown to lead to stereotypical localization of organelles, such as the centrosome, early endosomes, lysosomes, and the Golgi apparatus (Théry et al., 2005; Théry et al., 2006; Schauer et al., 2010). We were interested in establishing whether micropatterning can similarly reduce variation in mRNA spatial distributions and enable us to construct a quantitative framework for measuring reproducible subcellular spatial localization of RNA and protein. To this end, mouse fibroblasts were induced to polarize on crossbow-shaped micropatterns shown to be suitable for the study of polarizing cells (Théry et al., 2006; Schauer et al., 2010) and fixed at different time points after induction (Figure 2A). Each slide contained multiple 12 × 12 grids of crossbow-shaped micropatterns to which the cells adhered (Figure S1A). We developed an autonomous image acquisition and semi-automated image analysis pipeline (Figure 1A) that was able to scan each microfabricated slide and autonomously acquire images of individual cells. We used standard wide-field fluorescence microscopy, spinning-disk confocal microscopy, or a StellarVision microscope using synthetic aperture optics technology to acquire a 3D stack of images for each cell. SmFISH and IF were performed to label mRNAs and corresponding proteins of interest, respectively. We also labeled the microtubule (MT) cytoskeleton, the nucleus, and the micropattern. Representative images of the micropattern and smFISH are shown in Figure 2B. A complete list of the acquired images and descriptions of image acquisition characteristics is presented in Table S1.

To investigate the subcellular localization and colocalization patterns over time, we followed mRNAs and corresponding proteins in polarized fibroblasts at different time points (2, 3, 4, and 5 h for mRNA and 2, 3, 5, and 7 h for protein) after induction of polarization by serum. These time points were chosen as

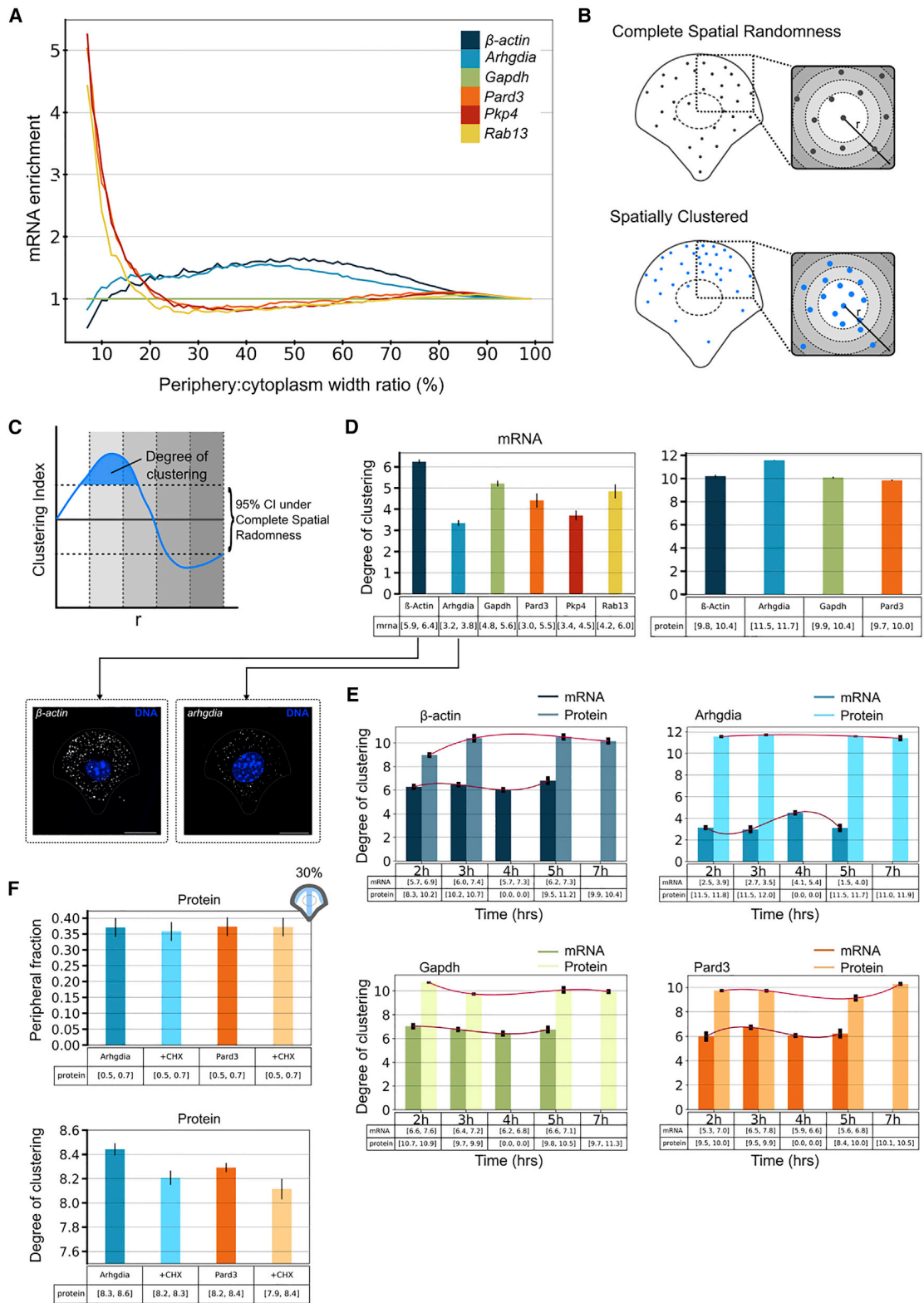
fibroblasts polarize over this timescale on crossbow-shaped micropatterns. To determine whether micropatterning leads to reduced heterogeneity in mRNA distributions, we compared fibroblasts grown in standard culture with those grown on micropatterns. Representative images of standard cultured and micropatterned cells are shown for the *Arhgdia* mRNA in Figure 2B.

First, we compared the reproducibility of *Arhgdia* mRNA distributions in standard cultured and micropatterned cells by using circular quantization in isolines from the periphery to the nucleus envelope. The absolute deviation of the quantized distribution of a randomly selected cell from a pooled average is reduced in micropatterned cells for all pool sizes up to ~40 cells (Figure 2D). The error profiles are concordant with a previous study, which estimated that ~20 micropatterned cells were necessary to establish reproducible organelle positions by using the AMISE metric (Schauer et al., 2010). We further investigated the effect of micropatterning on the volume-corrected noise measure *Nm* introduced in (Padovan-Merhar et al., 2015), which corresponds to the coefficient of variation of mRNA density corrected for the cell's volume. Following this study and others (Fomina-Yadlin et al., 2014; Gupta et al., 2012; Jain et al., 2013; Kempe et al., 2015), we modeled the relationship between transcript number and cell size by using polynomial regression, as demonstrated for *Arhgdia* mRNA (Figure 2C). Our data show reduction in transcript number variation between micropatterned cells in comparison with standard cultured cells, in line with Battich et al. (2015), in which constraining the cell phenotypic state resulted in reduction in cell-to-cell variability in cytoplasmic RNA concentration. The dispersion around the linear model fit line in the micropatterned cells as attested by the Pearson square coefficient r^2 (Figure 2C) is lower, leading to a lower *Nm* value (Figure S1C). This comparison revealed a tighter distribution of cell and nuclear sizes in the micropatterned cells, consistent with a mechanism to compensate for cell size fluctuations and maintain the concentration of mRNA in the cell, as proposed previously (Figure 2C; Padovan-Merhar et al., 2015).

Next, we investigated the profiles over time of the *Nm* for a series of mRNA transcripts including *Pkp4* and *Rab13*, which are enriched at the leading edge in polarized fibroblasts (Mili et al., 2008; Moissoglu et al., 2019), *Pard3*, which translates into the Par3 protein that regulates polarity in various cell types and is enriched in developing axons (Hengst et al., 2009; Schmoranzler et al., 2009), β -*Actin*, a well-studied localized mRNA in various cell types (Bassell et al., 1998; Buxbaum et al., 2014; Katz et al., 2012; Kislauskis et al., 1997; Park et al., 2012; Park et al., 2014; Yoon et al., 2016; and others), and *Gapdh*, which to the best of our knowledge is not known to localize to specific subcellular domains. We found for a number of transcripts a reduction in *Nm* over time, up to the 4 h time point (Figure S1B). These data indicated that micropatterns conveyed an advantage for quantitative analysis of RNA subcellular position over standard cell culture.

(C) Relationship between cell size and *Arhgdia* transcript copy number in standard cultured and micropatterned cells. Solid lines in upper graphs show the least-squares fit. Pearson square coefficient values R^2 are 0.48 and 0.32 for micropatterned and cultured cells, respectively. Lower graphs compare cell and nucleus size in the two conditions.

(D) Absolute deviation of *Arhgdia* mRNA distribution of a randomly selected cell from a pooled average of up to ~40 cells for cultured and micropatterned cells.



(legend on next page)

A subset of mRNAs and corresponding proteins shows peripheral enrichment and correlated clustering between time points

We investigated the joint localization patterns for mRNA and corresponding protein of four mRNA transcripts (*Arhgdia*, *Pard3*, β -*Actin*, and *Gapdh*) and the mRNA localization dynamics for the remaining two transcripts (*Pkp4* and *Rab13*). We developed a cell quantization method for computing local density statistics and used the MTOC cellular landmark to obtain consistent and alignable cellular segments. We initially characterized whether the mRNA and corresponding proteins, respectively, are enriched in the periphery of polarized cells. To do so, we calculated the fraction of cytoplasmic transcripts found at the boundary of the cell, within an area whose width is a fixed proportion of the radial distance to the nucleus edge (STAR Methods, “peripheral distance map”). Certain transcripts, including *Rab13* and *Pkp4*, as well as *Pard3*, are peripherally enriched for up to 20% of the radial distance (Figure 3A) in polarized cells on microfabricated patterns. *Gapdh* mRNA and protein distributions were used as controls, as neither the mRNA nor protein were expected to show patterns of enrichment (Mili et al., 2008; Figure S2A). Note that values in the extreme periphery are not stable, due to imprecision in segmentation of the cell mask. Representative smFISH images of β -*Actin* and *Gapdh* are shown in Figure S2B.

We next analyzed clustering of mRNAs and proteins by using a generalization of Ripley’s K analysis (Lee et al., 2013; Ripley, 1977) that we first theoretically investigated in Warrell et al., 2016). Ripley’s K is a commonly used algorithm to describe the extent of clustering of points, such as mRNAs (Figure 3B). For each mRNA spot and each distance d the number of transcripts lying within a sphere of radius d is counted. Spatial clustering can then be calculated by estimating the probability distribution of this function under a null hypothesis of complete spatial randomness (CSR) and comparing it with the function calculated from observed transcripts. We adjusted the algorithm (STAR Methods, “degree of clustering (Ripley-K)”) on the basis of the generalized Ripley K function for evaluating the extent of clustering of both mRNA and protein spatial distributions. This was done by computing the degree of clustering, a unitless measure that can be used to compare clustering between different molecules and conditions. We summed the area where the normal-

ized Ripley K function deviates from the 95% confidence interval of the random distribution (Figure 3C).

We evaluated the degree of clustering of all transcripts and proteins across all time points, revealing high overall values for all proteins and various values for the different mRNA transcripts (Figure 3D). Representative smFISH images for β -*Actin* and *Arhgdia* are shown in Figure 3D. We further calculated the degree of clustering at each individual time point post induction of polarization for *Arhgdia*, *Gapdh*, β -*Actin*, and *Pard3* mRNAs and proteins, and normalized it by mean for value’s scale compatibility between mRNA and protein (Figure 3E). A clear peak in the mRNA profile followed by a peak in the corresponding proteins can be seen for *Pard3*, β -*Actin*, and *Gapdh*, suggesting temporal correlation (Figure 3E). A plausible interpretation of this temporal correlation might be local translation—clustering and translation of the transcript, yielding a high local concentration of protein that is observed as a cluster. To test this hypothesis, we used the translation inhibitor cycloheximide (CHX) and compared the peripheral enrichment and the degree of clustering of *Arhgdia* and *Pard3* in inhibited and control cells (see description of image acquisition characteristics for CHX [Table S1] and STAR Methods). As expected, no significant reduction in peripheral enrichment for both *Arhgdia* and *Pard3* mRNAs was observed (Figure S2C), and there was a slightly increased degree of clustering for both mRNAs (Figure S2E). Taken together, these data indicate that *Arhgdia* and *Pard3* are localized to the periphery similarly in inhibited and control cells and cluster slightly more throughout the cytoplasm in inhibited cells. In contrast, a slight decrease was observed in peripheral enrichment and degree of clustering for both RhoGDIa (the protein translated from *Arhgdia*, to which we will refer to as *Arhgdia* for simplicity) and *Pard3* proteins (Figures 3F and S2D). Altogether, these data indicate that in fibroblasts in which protein translation is inhibited, *Arhgdia* and *Pard3* mRNAs localize similarly to control cells; however, the reduced peripheral fraction and degree of clustering of the *Arhgdia* and *Pard3* proteins hint toward potential local translation of these genes at the periphery. As transport of molecules is not expected to be affected in these cells, we do not anticipate the reduction in enrichment and clustering of *Arhgdia* and *Pard3* to be attributable to impaired transport of the proteins.

Figure 3. Peripheral enrichment and clustering dynamics of mRNA-protein pairs

- (A) Comparison of the enrichment of five mRNAs with respect to *Gapdh* mRNA in a peripheral cellular region whose width varies from 0% to 100% of the radial distance from the plasma membrane to the nucleus.
- (B) Clustering is characterized by comparing observed transcript and protein distributions with complete spatial randomness. For mRNAs, the Ripley K function is estimated for an observed distribution and samples from a homogeneous Poisson process by counting the number of pairs of points lying within a radius r of each event.
- (C) To compute the degree of clustering, we defined an estimator of the Ripley K function, called clustering index, by normalizing the observed Ripley K function to the 95th and 5th percentiles under the Poisson process. Significant clustering of mRNAs or proteins is found at radius r where the estimator function is over the 95% confidence interval under the CSR assumption. Values below the 95% confidence interval indicate the dispersion of the molecules. The degree of clustering is the area under the estimator’s curve that is above the 95th percentile of the random distribution.
- (D) Comparison of log degree of clustering for mRNAs and proteins (all time points, log values shown after scaling by natural log). Error bars show the standard error of the median. Representative single-molecule images for β -*Actin* and *Arhgdia* are shown. Scale bar, 10 μ m.
- (E) Clustering dynamics for four mRNA-protein pairs by using normalized degree of clustering. Zero values indicate the absence of protein or mRNA data for the given time point. Error bars show the standard error of the mean.
- (F) Log degree of clustering (bottom graph), as well as enrichment of *Arhgdia* and *Pard3* proteins in cells treated with CHX compared with control cells at 30% peripheral regions (top graph). Differences in degree of clustering were assessed by using a Mann-Whitney test. p values: *Arhgdia* versus *Arhgdia* + CHX = 0.2; *Pard3* versus *Pard3* + CHX = 0.04.

Localization of a subset of mRNAs and corresponding proteins shows correlation with the MTOC position

In the majority of polarized cell systems, the MTOC is positioned between the nucleus and the leading edge of the cell prior to migration or local cell growth (Gomes et al., 2005; Hale et al., 2011). Whether mRNAs and their corresponding proteins are subject to reorientation in relation to the MTOC is unknown. Having demonstrated that several mRNA transcripts and proteins exhibit peripheral enrichment and spatial clustering in polarized cells, we sought to determine whether these observations were related to MTOC positioning. We divided the cell into quadrants (Figure 4A), computed per-quadrant mRNA and protein relative density, and determined the MTOC localization within these quadrants (Figure 4A). We observed higher enrichment of all cytoplasmic mRNA transcripts in the MTOC-containing quadrant compared with the non-MTOC-containing quadrant, considering differences in their respective density (Figure 4B). Similarly, all peripheral transcripts at 30% of the radial distance, other than *Pkp4*, showed higher enrichment in the MTOC-containing quadrant compared with the non-MTOC-containing quadrant (Figure S3A). All cytoplasmic proteins were also found to be more enriched in the MTOC-containing quadrant compared with the non-MTOC-containing quadrant, with the exception of *Pard3* (Figure 4B). All proteins showed enrichment in the MTOC-containing quadrant at 30% of the radial distance (Figure S3A).

On the basis of the above observations, we introduced an MTOC polarity index (MPI) to analyze MTOC-dependent enrichment in mRNA and protein distributions. This indicator lying between -1 and $+1$ is derived by normalizing the differences of signal concentration between the MTOC-associated quadrant and the other quadrants. Positive MPI values imply MTOC-dependent enrichment of RNA transcripts, negative values imply enrichment away from the MTOC, and a value of zero implies no detectable relationship with the MTOC. Representative smFISH images for β -Actin and *Pard3* are shown in Figure 4C.

We calculated the MPI for all transcripts and proteins, computing values for both the whole cytoplasmic population (Figure 4C), and the peripheral population at 10% radial distance (Figure S3B), by using all time points. All mRNAs as well as *Gapdh*, *Arhgdia*, and β -Actin proteins show high MPI scores in the cytoplasmic population. The low MPI score for *Pard3* protein is in line with low MTOC enrichment for cytoplasmic *Pard3* (Figure 4B). Similarly to the cytoplasmic transcripts, all mRNAs show high MPI scores in the peripheral population (Figure S3B). Additionally, we calculated MPI scores across time (Figures 4D, S3C, and S3D) and observed similar wave-like profiles for mRNA and corresponding protein (with a lag in time for the protein) for cytoplasmic *Arhgdia*, cytoplasmic and peripheral *Pard3*, and peripheral β -Actin, *Arhgdia*, and *Gapdh*, and cytoplasmic β -Actin and *Gapdh* to lesser extent (Figures S3C and S3D). This suggested a temporal association between mRNA, protein, and MTOC orientation.

We next applied the method presented above to examine the biological relevance of the spatial distribution of a given mRNA. We chose to focus on *Arhgdia* mRNA and characterized changes in its subcellular location, followed by measuring the effects of these changes on the localization and function of *Arhgdia*'s en-

coded protein. The *Arhgdia* transcript is translated into RhoGDI α , which regulates the Rho family of guanosine triphosphatases (GTPases) (reviewed in Garcia-Mata et al., 2011; Xie et al., 2017). We hypothesized that changes in the subcellular localization of *Arhgdia* would have an effect on the location and/or translation of the transcript to RhoGDI α and subsequent effects on Rho GTPases and related cellular processes, including cellular polarization and migration.

We first used a biochemical pull-down assay to pull down *Arhgdia* and identify its protein binding partners by mass spectrometry (Savulescu et al., 2020). By this process we identified a number of candidate protein-interacting partners. Of these we chose to focus on the PRRC2C protein, given that the PRRC2C subcellular localization at the ER (data not shown) is linked to cell polarity and in addition contains a polyQ motif (Figure S5A), suggesting interactions with RNA (Kunde et al., 2011; Lee et al., 2013; Langdon et al., 2018). We depleted PRRC2C (Figure S5B) to see whether it had convergent phenotypes with *Arhgdia*. We hypothesized that if PRRC2C was important to *Arhgdia* mRNA biology, this would result in quantitative changes detectable by DypFISH. PRRC2C-depleted cells displayed a lower *Arhgdia* mRNA count (Figure S4A), a slightly reduced peripheral fraction (Figure S5C), similar centrality and cytoplasmic spread (Figure S5D), and a slightly lower degree of *Arhgdia* mRNA clustering (Figure S4B). Depletion of PRRC2C also led to reduced peripheral MTOC enrichment, peripheral MPI score (Figure S5E), and a substantially reduced cytoplasmic MPI score (Figures S4C and S5E) for *Arhgdia*. A higher concentration of RhoGDI α was observed in PRRC2C-depleted cells (Figure S4A), suggesting a higher rate of translation in these cells as well as a decreased peripheral fraction and slightly increased centrality and cytoplasmic spread (Figures S5C and S5D). RhoGDI α retains Rho GTPases in an inactive, stable form in the cytosol, preventing their degradation (reviewed in Garcia-Mata et al., 2011). We hypothesized that an elevated concentration of RhoGDI α would induce an increased pool of inactive cytosolic Rho GTPases, leading to inhibition of cell migration. Consistent with our hypothesis, depletion of PRRC2C led to inhibition of migration of 3T3 mouse fibroblasts and human umbilical vein endothelial cells (HUVECs) in a wound assay (Figure S4D; Videos S1 and S2). We then tested whether the inhibition in migration was due to impaired velocity in PRRC2C-depleted cells. Velocity measurements of individual cells in the monolayer did not show differences between control and PRRC2C-depleted cells (Figure S5F); however, we observed a slight decrease in the ability of cells to maintain their directionality, termed directness, in PRRC2C-depleted cells (Figure S5F). Direction of migration is defined by the subcellular positions of the Golgi apparatus and MTOC in the cell, known as axial polarity (Ridley et al., 2003). To test whether it was abolished in the absence of PRRC2C, we performed an axial polarity assay (Carvalho et al., 2019). Loss of Golgi orientation in PRRC2C-depleted cells was observed (Figure S4E), suggesting the loss of axial polarity and, thus, migration defects. In conclusion, our data suggest that PRRC2C-depleted cells are unable to migrate properly.

Altogether, our data suggest that the absence of PRRC2C results in loss of *Arhgdia*'s enrichment in the MTOC vicinity and

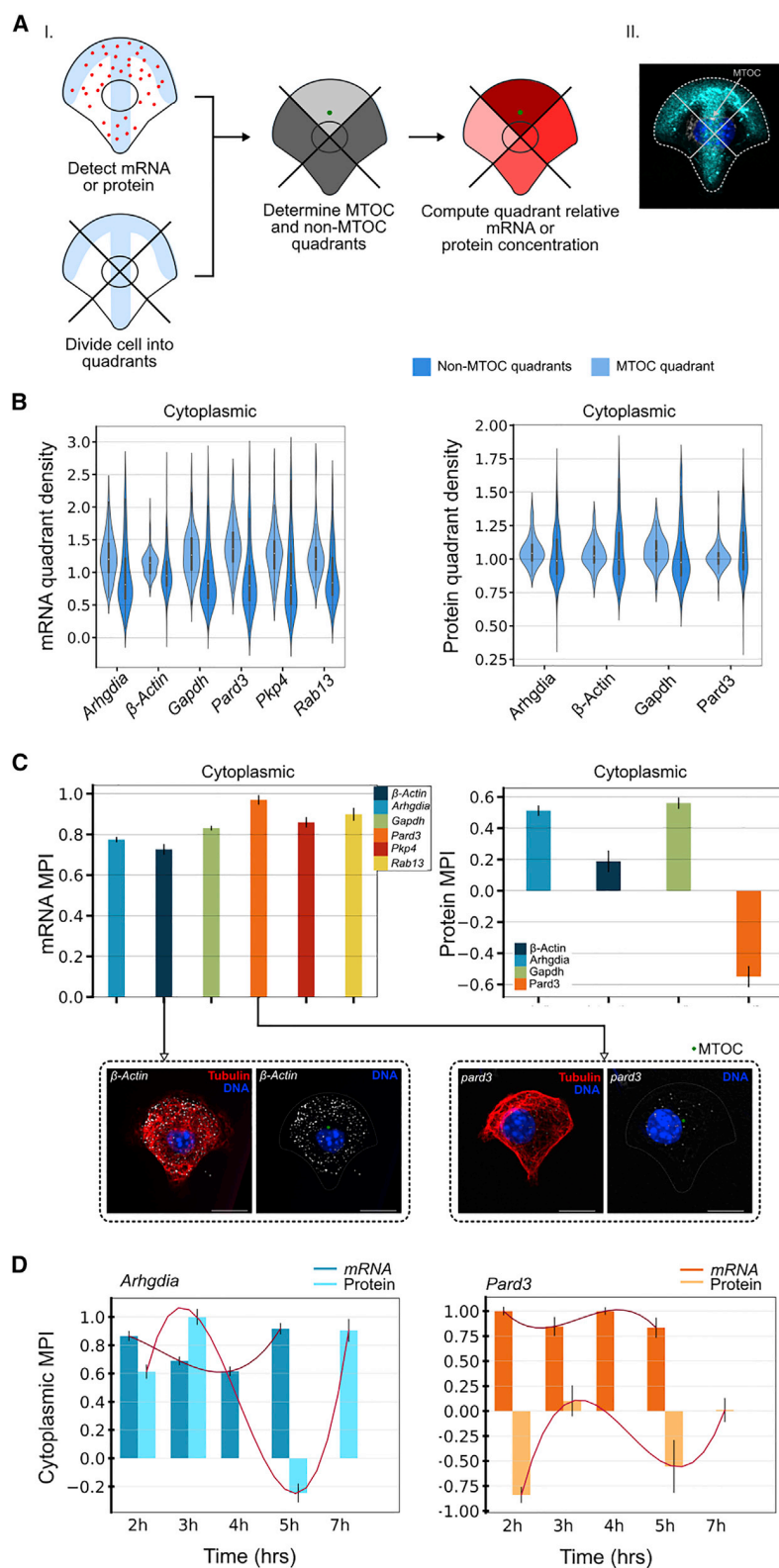


Figure 4. Relationship between cytoplasmic mRNA and protein distributions and MTOC position

(A) (i) Schematic of the MTOC correlative influence analysis and of the analysis to determine the MPI value. (ii) The MTOC position is annotated in projected 3D tubulin IF images. The MTOC polarity index (MPI) measures how frequently the relative concentration within the MTOC quadrants is higher than in the non-MTOC quadrants. MPI takes values between -1 and $+1$, with positive values indicating enriched mRNA or protein concentration in the MTOC quadrant and negative values indicating enrichment away from the MTOC quadrant.

(B) Cytoplasmic mRNA and protein relative density in non-MTOC-containing quadrants (dark blue) and MTOC-containing quadrants (light blue). Enrichment in the MTOC-containing quadrant is defined by differences in means of these distributions.

(C) MPI values for mRNAs and proteins in cytoplasmic populations (all time points). Representative smFISH images are shown for β -Actin and *Pard3* (right image), as well as the corresponding tubulin stain (left image). Scale bar, $10 \mu\text{m}$.

(D) MPI dynamics for *Arhgdia* and *Pard3* mRNA-protein pairs.

Bar graphs in (C) show median and standard deviation from the median error bars for 100 bootstrapped MPI estimates. Graphs in (D) show median surrounded by envelope indicating standard deviation from the median error bars of 100 bootstrapped estimates fitted to cubic splines.

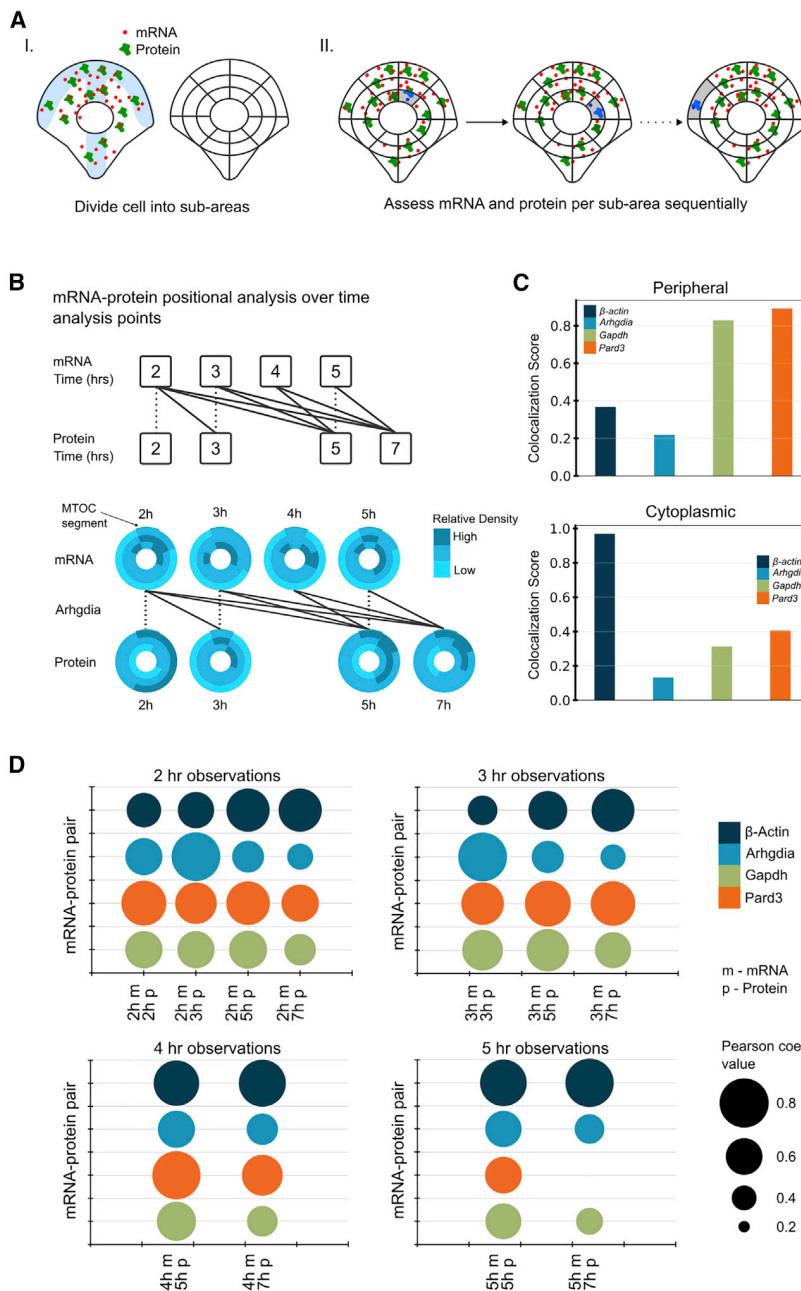


Figure 5. Interdependency of localization dynamics for corresponding mRNAs and proteins

(A) Colocalization score (CS) is a correlation between mRNA and protein spatial distributions for image acquisitions at several time points.

(B) Forward-leading time-point pairs, defined as $t_1 < t_2$ (connected by solid lines), were used for calculating CS values. Cytoplasmic density maps representing relative density vectors based on fine-grained quantization for Arhgdia mRNA and protein for each time point are shown. Cellular regions are dark blue if the local relative density is greater than the mean density of all segments in the cytoplasm plus standard deviation, and light blue if the local relative density is smaller than the mean of all segments in the cytoplasm minus standard deviation; intermediate values correspond to the relative density within the $[-\sigma, \sigma]$ interval. The MTOC-containing quadrant is highlighted by a thin turquoise stripe.

(C) CS values were computed by using global correlations of all voxels/segments across the cytoplasmic area and local correlations across subsets of voxels/segments within peripheral regions.

(D) Correlations of relative density distributions for all cytoplasmic mRNA-protein pairs were observed for certain time points by using the fine-grained quantization scheme.

Colocalization of mRNA-protein distributions is consistent with MTOC-related patterns of localized translation

As previously described, several mRNA-protein pairs showed similarities in distributions of basic descriptors, clustering indices, and MTOC polarity indices. Correlations in spatial distributions at different time points could reflect spatially and temporally restricted translation (local translation) (Besse and Ephrussi, 2008) and/or separate localization pathways for mRNAs and proteins to common subcellular locations.

To investigate such scenarios, we performed fine-grained quantization of cells and were able to compute subcellular spatial distribution profiles of mRNAs and proteins (see STAR Methods, “colocalization score” and Figure 5Ai), visualized using density maps for each time point for

changes in its subcellular localization and degree of clustering pattern, as well as leading to seemingly elevated levels of translation, resulting in a higher concentration of RhoGDIa protein. We suggest that an elevated RhoGDIa concentration, as well as higher MPI score of the RhoGDIa protein in PRRC2C-depleted cells, leads to an imbalanced Rho GTPase cycle and subsequent inhibition in cell migration (Figure S4F). In summary, and most significantly, by using DypFISH, we unearthed variations in polarity and cell migration due to quantifiable changes in the MTOC-related subcellular localization of Arhgdia mRNA.

four corresponding mRNA-protein pairs (Arhgdia, Gapdh, β -Actin, and Pard3) (Figures 5B and S6B).

For a pair of mRNA and protein spatial distributions, their colocalization is measured by the Pearson correlation of subcellular density vectors. We consider “high” or “low” colocalization for an mRNA/protein pair depending on the value of the Pearson correlation. High values (interdependent spatial distributions) might suggest local translation within specific subcellular localizations, although they do not rule out alternative mechanisms such as separate mRNA

and protein localization pathways with delayed protein transport.

To explore the consistency of potential correlations in spatial distributions between mRNA and the corresponding protein over different time points, we derived a colocalization score (CS), which is a value between 0 and 1, computed as the stochastic effect size based on the correlations between the subcellular distributions of mRNA and corresponding protein in a ranking across “forward-leading” time points (for details see STAR Methods, “colocalization score” and Figure 5B). These time points were chosen because we considered the additional time for translation to occur once the mRNA is localized.

The CS provides a global measure of whether the two molecules are distributed in a similar way in cells across all “forward-leading” time points as compared with the other time points. We measured the similarities of relative density vectors (1) globally across the cytoplasmic area by computing global CS values (Figure 5C) and (2) at the periphery using relative density vectors restricted to peripheral regions (Figure 5C). We observed correlated spatial distributions for cytoplasmic mRNA-protein pairs, with various temporal patterns among the different genes (Figures 5D and S6A). Generally, medium to high CS values were observed for all genes at the different time points (Figure 5D); however, we observed variability between the genes in CS values at various time points. For example, *β-Actin* mRNA at all time points showed high CS values with its corresponding protein at most time points, whereas the CS values for *Arhgdia* mRNA at late time points were generally lower than the values at early time points (Figure 5D).

Perturbation of various cytoskeletal components disrupts characteristic mRNA-protein localization and interdependency patterns, and hints at local translation

Our analysis revealed mRNA-protein colocalized distributions for certain time points, which suggested local translation. As we could not rule out independent localization of mRNA and corresponding proteins, we introduced two perturbations to inhibit potential transport pathways of the different molecules. We disrupted microtubule polymerization by using nocodazole (Figure 6A), which we reasoned would lead to disruption of the correlation with the MTOC of spatial distributions of selected mRNA and proteins, as well as potential loss of local translation. We selected *Arhgdia* and *Pard3* mRNA-protein pairs to test this hypothesis, collected data at 3 h and 5 h after exposure to nocodazole, and compared it with untreated data from the equivalent time points.

We calculated the effects of nocodazole treatment on the total mRNA count (Figure S7A), the centrality of the transcript and protein (Figures 6B and S7B), and the 10% peripheral fraction (Figure 6C). *Arhgdia* centrality was increased upon treatment with nocodazole and its peripheral fraction was not significantly modified (Figures 6B and 6C), suggesting that *Arhgdia* mRNA and protein do not make exclusive use of the microtubule network for intracellular transport. *Arhgdia* protein centrality was slightly decreased and the peripheral fraction of the protein was reduced (Figures 6B and 6C), indicating a partial dependency on the microtubule network for subcellular transport. The centrality of *Pard3* mRNA was increased, whereas its pe-

ripheral fraction was slightly decreased upon treatment with nocodazole (Figures 6B and 6C), suggesting that *Pard3* mRNA does not make exclusive use of the microtubule network for intracellular transport. A reduction was detected in the centrality of *Pard3* protein in the presence of nocodazole (Figure 6B), and there was a slight increase in the peripheral fraction of the protein (Figure 6C). Interestingly, an increase in the *Pard3* mRNA and corresponding protein concentration was observed in the presence of nocodazole (Figure S7A).

Next, we probed the effects of nocodazole treatment on the relationship between the polarization of the cell and the subcellular localization of the mRNA transcripts and their corresponding proteins (Figures 6D and 6E). The MTOC enrichment of *Arhgdia* in nocodazole-treated cells compared with control cells was not substantially modified (Figure 6D) and the MPI of *Arhgdia* in these cells was slightly reduced (Figure 6E), indicating that the orientation of the mRNA subcellular localization was partially lost. An interesting effect was observed for *Pard3*, whereby no significant change was detected in the MTOC enrichment, with a slight increase in the MPI value for the mRNA in the nocodazole-treated cells (Figures 6D and 6E), indicating that the perturbation to the MTOC had the effect of tightening the relationship between the MTOC position and the RNA distribution. Representative smFISH images for *Arhgdia* in control and nocodazole-depleted cells are shown in Figure 6E.

We then probed the effects of nocodazole on mRNA-protein pairs by calculating local CS and maps for control and nocodazole-treated cells, restricted to the 3–5 h time points (Figure 6F). The interdependency of the *Arhgdia* mRNA-protein pair was disturbed, as observed in altered CS values for all time points in nocodazole-treated compared with control cells (Figure 6F). Nocodazole also affected the interdependency for the *Pard3* mRNA-protein pair, mainly observed in the 3 h mRNA to 3 h protein and 5 h protein time points (Figure 6F).

To probe whether other cytoskeletal networks could influence RNA subcellular positioning, we treated the cells with the actin polymerization inhibitor cytochalasin D (cytoD) for 1 h before fixation of cells and compared FISH and IF data for the *Arhgdia* mRNA-protein pair. *Arhgdia* mRNA's centrality and cytoplasmic spread were slightly increased (Figure S7B), whereas the concentration of the mRNA was slightly lower on the periphery of cytoD-treated cells compared with control cells (Figure S7C). The centrality and cytoplasmic spread of *Arhgdia* protein were not significantly affected in the presence of cytoD (Figure S7B); however, the peripheral concentration of the protein was slightly decreased in the cytoD-treated cells (Figure S7C). Taken together, these data suggest that a lower *Arhgdia* mRNA peripheral concentration leads to reduced peripheral *Arhgdia* protein concentration or that mechanisms anchoring *Arhgdia* mRNA and protein to the periphery are lost in the presence of the drug.

Sarcomeric mRNAs cluster in a striated pattern in differentiated myofibers

To further validate Dyp-FISH, we sought a cellular model in which RNA subcellular localization and potential local protein translation linked to changes in cells' state could be interrogated. We focused on the unusually large multinucleated muscle cells

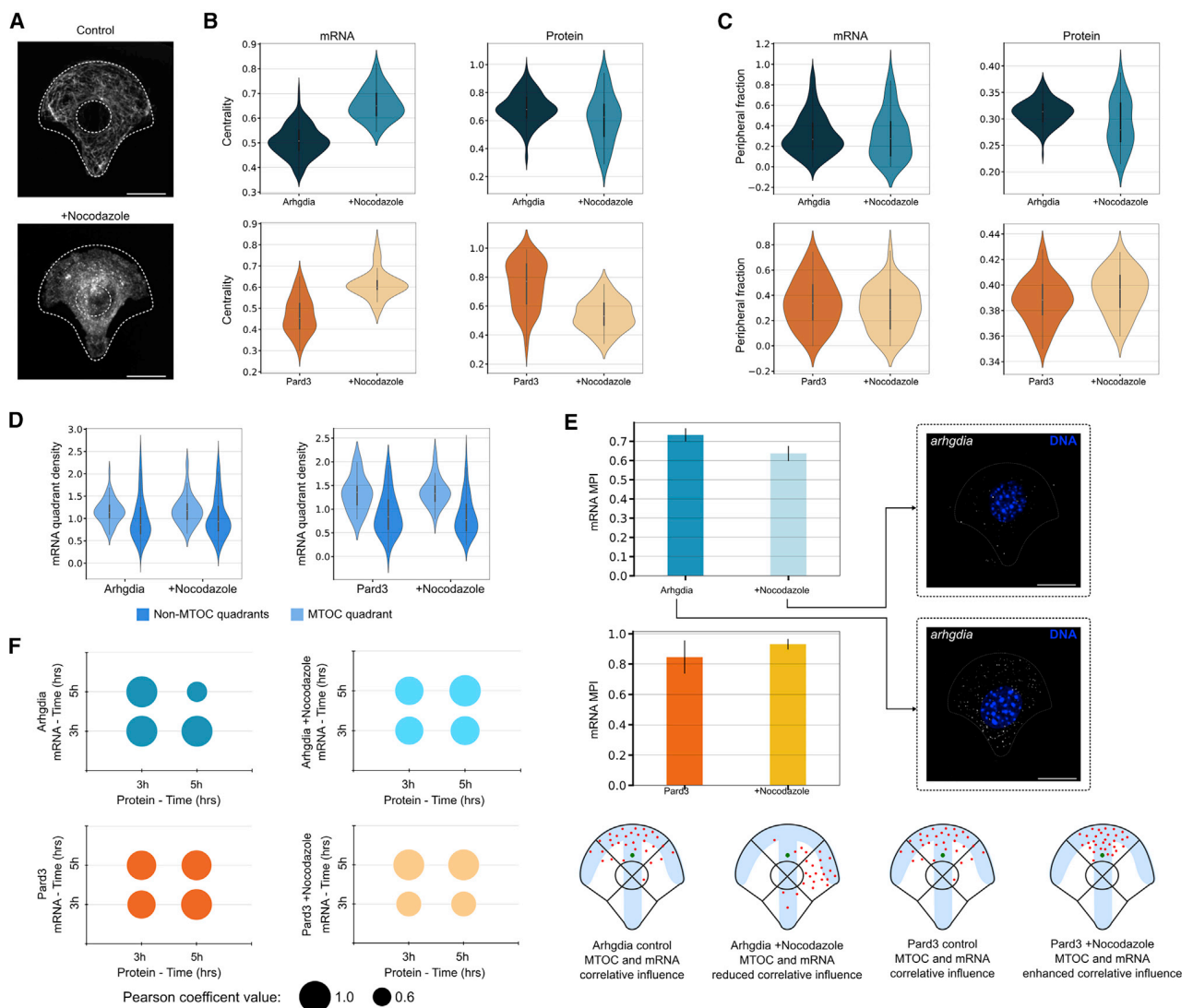


Figure 6. Effects of cytoskeleton disturbance on mRNA-protein localization and interdependent dynamics

(A) Nocodazole was added to cells seeded on micropatterns, inducing inhibition of microtubule polymerization (tubulin IF images). Scale bars, 10 μ m. (B) Centrality of *Arhgdia* and *Pard3* transcripts and proteins (at 3 h and 5 h combined) is defined as a statistic measuring the evenness of a molecule spread across the cell, with the value 1 for even distribution. Error bars show standard error of the mean. (C) The peripheral fraction of *Arhgdia* and *Pard3* transcripts and proteins (at 3 h and 5 h combined) were calculated similarly to Figure 3B. (D) mRNA quadrant density (MTOC enrichment profiles) for *Arhgdia* and *Pard3* shown for nocodazole-treated and control cells at 3 h and 5 h combined. (E) MPI scores shown for control and nocodazole-treated cells for *Arhgdia* and *Pard3* transcripts at 3 h and 5 h combined. Representative smFISH images of *Arhgdia* in control and nocodazole-treated cells are shown. Scale bars, 10 μ m. A cartoon representing the effect of nocodazole treatment on the MPI scores is shown in the bottom panel. (F) Interdependent dynamics for *Arhgdia* and *Par3* in the presence of nocodazole and in control cells.

termed myofibers that are formed by fusion of mononucleated cells, with the myonuclei being spaced at regular intervals, their main function being to generate mechanical force via contraction (Bruusgaard et al., 2003, 2006). Muscle contraction is achieved by the shortening of sarcomeres that are organized along the length of the myofiber and each sarcomere is flanked by a Z line, the site of anchoring of the actin filaments, resulting in the striation of the myofiber (Franzini-Armstrong and Peachey, 1981). We used skeletal muscle, given that the myofiber has an

invariable tubular shape and a highly predictable cytoplasmic organization. *In vitro* differentiation of myofibers allows for high-resolution imaging throughout developmental stages, including the formation of patterned sarcomeres with well-defined Z-line striations, similar to what is observed in neonatal myofibers (Falcone et al., 2014; Pimentel et al., 2017; Vilmont et al., 2016). Thus, the size and regularity of myofiber architecture make them an excellent candidate to compute spatial distribution profiles of mRNAs and proteins.

We analyzed the distribution of an mRNA that encodes a protein found at the Z lines during muscle differentiation. We chose the *actn2* mRNA, which encodes for α -actinin, the main component of Z lines in relation to the sarcomeric Z lines (Figure 7A). The majority of the *actn2* mRNA was found in the vicinity of the Z line in mature myofibers (Figure 7B). To understand whether this depends on the developmental stage, we imaged immature myofibers in which the Z lines are less organized. We also probed the *Gapdh* mRNA distribution as a non-sarcomeric control in both immature and mature myofibers. The degree of mRNA proximity was lower in both cases, suggesting that possibly *actn2* mRNA localization precedes protein organization (Figure 7B) of the Z line. To address this question, we quantized each cell perpendicularly to the cell axis in vertical quadrants by using a number of various divisions (Figures 7C and S8A), similarly to sarcomeres' organization. The mRNA local density was computed in each quadrant by normalizing the counts by the relevant area (Figures 7C, S8A, and S8B). We observed a wave-like clustering for *actn2* in mature compared with immature fibers by using various divisions (Figures 7C, S8A, and S8B). No clustering was observed for *Gapdh* (Figures 7C, S8A, and S8B). Curiously, in adult mouse myofibers, it was recently described that *Gapdh* mRNA also clusters at the Z line, suggesting that myofiber hypertrophy alters mRNA clustering properties (Denes et al., 2021).

The highest degree of clustering was observed for *actn2* mRNA in mature myofibers when compared with the immature counterpart or with *Gapdh* mRNA. These data suggest that the *actn2* mRNA distribution specifically follows the respective protein organization instead of preceding it (Figure 7D). These data shed light on a long-standing question in the field and produce a basis of testable hypotheses for how *actn2* mRNA is directed to the Z line.

DISCUSSION

A wide variety of methods have been proposed for studying RNA localization with subcellular accuracy, including microscopy-based methods, such as those based on FISH (Battich et al., 2013; Chen et al., 2015; Lécuyer et al., 2007). We set out to develop a quantitative, reproducible, and reusable computational method for investigating the spatial distribution of RNA and protein. To achieve this, we hypothesized that the best showcase of our approach would be to use micropatterning of cells in order to reduce cellular heterogeneity and enhance the reproducibility of spatial distributions. We combined this with smFISH, IF labeling, and automated high-content imaging, as well as the development of our DypFISH computational framework. We revealed dependence of mRNA-protein localization and dynamics on MTOC orientation and cell polarization in polarized fibroblasts. Through perturbation studies, we have demonstrated DypFISH's ability to quantitatively detect changes in localization behavior, confirming both the robustness of the approach and its ability to test mechanistic hypotheses. This approach can be applied to other systems in which (1) there is a clear constrained architecture, such as the skeletal muscle shown in Figure 7, or/and (2) cellular landmarks can be stained, similarly to the MTOC landmark in our study.

RNA subcellular localization patterns and the role of the MTOC in subcellular spatial distribution

The association of different mRNA species with MTOC orientation indicates specific subcellular distribution patterns in accordance with broader processes, which are also responsible for controlling the nucleus and MTOC relative orientation during polarization (Kim et al., 2014; Razafsky et al., 2014). We posed the question as to whether proteins translated from mRNAs that consistently localize to similar subcellular locations are translated in these locations. By revealing the existence of correlations of spatial distributions for mRNA-protein pairs at different time points, which are associated with the MTOC orientation and can be impaired by applying perturbations including inhibition of protein translation, our results suggest local translation. Additionally, our methodology can be applied to study variability of localized gene expression in morphologically constrained biological contexts.

We revealed various aspects of the spatial distribution of specific mRNA transcripts and proteins of interest in time, such as *Arhgdia* and its protein product RhoGDIa. We have shown that the majority of the *Arhgdia* mRNA population relates to the MTOC position, with a small fraction being localized to the periphery. The reduced cytoplasmic MPI of *Arhgdia* mRNA, altered CS plots for *Arhgdia* mRNA-protein in the presence of a cytoskeletal polymerization inhibitor, and the reduction in peripheral enrichment and clustering of *Arhgdia* protein when translation is inhibited suggest MTOC-dependent peripheral local translation. Furthermore, depletion of PRRC2C, an *Arhgdia* mRNA binding protein, led to changes in the subcellular localization of *Arhgdia* in relation to the MTOC, changes in translation of the corresponding protein, and further downstream effects related to RhoGDIa function as a regulator of Rho GTPases, including inhibition of cell migration. Taken together, these data exemplify the importance of characterizing the spatial distribution of mRNA quantitatively and how measurable changes in the subcellular distribution of the transcript are linked to RNA and protein function.

Our analyses also revealed important aspects of *Pard3* spatial and temporal distribution. The transcript and protein show a correlated temporal degree of clustering and a reduction in peripheral enrichment and clustering of the *Pard3* protein in the presence of a translational inhibitor. In contrast to *Arhgdia*, there is an increase in the MPI of *Pard3* mRNA upon disruption with nocodazole, indicating MTOC-MT influence on *Pard3* transport. Localization of the *Pard3* protein in peripheral clusters at cell-cell adhesions was previously shown in fibroblasts when grown in culture (Schmoranzner et al., 2009), which is concordant with the peripheral enrichment of *Pard3* mRNA and protein in our system.

Resolving spatial distribution of RNA and protein in a quantitative manner across time points

We have shown our approach to be suitable for the identification of patterns of RNA spatial and temporal distribution and RNA-protein interdependent localization, both globally and locally. We demonstrated how our analytical tools can be applied to micropatterned cells for studying RNA subcellular distribution and how it can be used for other constrained cellular geometries.

Several aspects of our approach are suitable as a basis for diverse kinds of spatially resolved omics (Crosetto et al., 2015).

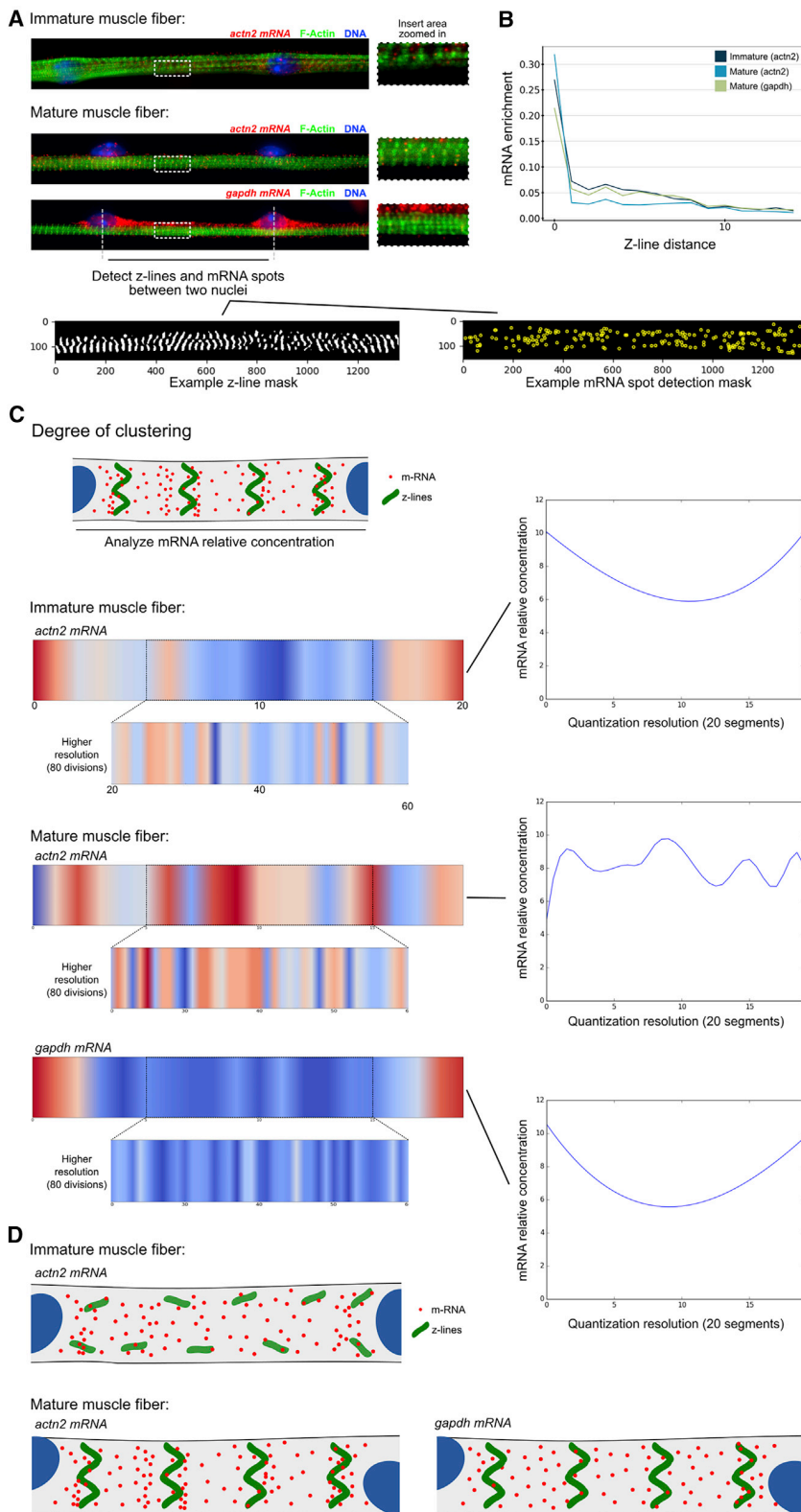


Figure 7. Sarcomeric mRNAs cluster in a striated pattern in differentiated myofibers

(A) Typical epifluorescent images of immature and mature muscle fibers. The DNA was stained with DAPI (blue), F-actin was visualized by using IF (green), and *Actn2/Gapdh* were visualized by using smFISH (red). Z-line and RNA spot detection masks were extracted by using IF and single-molecule data, respectively.

(B) mRNA distance profiles. For each mRNA we computed its distance to the closest Z lines, which allowed us to count the number of mRNAs having a certain distance to Z lines. Normalized median counts are represented on the y axis. A higher number of *actn2* immature mRNA falls inside or close to Z lines in comparison with mature fibers, suggesting greater clustering of mRNA between Z lines for mature *actn2*. The Z-line distance is a Euclidean distance.

(C) The mRNA local density was computed between two nuclei. Each cell was quantized in vertical quadrants, and relative concentration of mRNA in each quadrant was computed by normalizing the counts by the relevant surface. A wave-like clustering is observed for *actn2* in mature compared with immature fibers. No clustering is observed for *Gapdh*.

(D) Model describing *actn2* mRNA distribution in immature and mature fibers.

The quantitative nature of the computational analyses, autonomous image acquisition, and automated features of our data processing make DypFISH highly scalable to high-throughput studies and in mammalian tissue cells. Particularly, the capacity of the approach to investigate changes in localization under perturbation make DypFISH suitable for inferring spatially organized regulatory networks (Crosetto et al., 2015) and capable of being combined with multiplexing techniques (Chen et al., 2015) to reveal dynamic changes in cell state. As well as correlated mRNA-protein localization, our generalized clustering approach can be used to detect interdependent clustering patterns between RBPs and different kinds of RNAs (long non-coding RNAs, mRNAs, microRNAs) or interdependent protein-protein clustering patterns. Finally, we believe that techniques such as those presented here will help to make possible further development of integrated approaches and thus contribute to large-scale multi-omics studies at subcellular resolution, such as the Human Cell Atlas.

Limitations of study

This study provides a toolbox to characterize various aspects of subcellular localization of RNA and protein molecules. These include subcellular location, relation to a cellular marker, clustering patterns, and colocalization between molecules, among others. The limitations of this study include the following: (1) The subcellular localization of a biomolecule is defined in relation to a chosen cellular marker and it is not an absolute parameter. (2) Micropatterned cells and muscles have a specific, defined topology but standard cells in culture do not, and thus the definitions of “peripheral,” “central,” and so forth are less clear-cut. (3) The annotation of the MTOC has been done manually; automation is difficult, as the signal using the anti-tubulin antibody might be blurry, there is a certain percentage of cells with more than one MTOC, and the annotation often requires looking through a few Z slices. However, we do expect the annotation to be automated in the future. (4) Although many cell lines and some primary cells spread well on micropatterns, not all primary cell types and cell lines spread and adopt the shape of the micropattern, limiting the number of cell types that can be used. (5) The effect of modification of cell quantization parameters (such as number of quadrants and isolines) on the resulting statistical results has not been evaluated.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Cell culture and medium
- **METHOD DETAILS**
 - Cell culture and treatments
 - RNA isolation and analysis
 - Cell micropatterning
 - RNA probes and reagents

- Immuno-RNA FISH staining
- Immunofluorescence staining
- Polarity and migration assays
- Image acquisition
- Materials
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Primary image descriptors
 - Secondary image descriptors
 - Statistical analysis
 - Additional methods for muscle data analysis
 - Additional methods for CHX inhibition analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crmeth.2021.100068>.

ACKNOWLEDGMENTS

We thank all members of the Gene Expression and Biophysics Laboratory (Mhlanga Lab). We thank N. Crosetto, and members of the L. Pelkmans lab, members of the J.-C. Olivo-Marin lab, K. Schauer, and members of the B. Goud lab for comments on the manuscript. A.F.S. is a recipient of a postdoctoral fellowship from the Claude Leon Foundation, South Africa. This research has been supported by the following grants, all to M.M.M. V2YGE81 (to A.F.S.) and PG-V2KYPO7, TA 2011 011 from the Council for Industrial and Scientific Research (CSIR) (South Africa) and by a grant from the Emerging Research Area Program of The Department of Science and Technology (South Africa) Department of Science & Technology Center of Competence Grant, SA Medical Research Council SHIP grant, and CSIR Parliamentary Grant. M.M.M. and M.N. are Chan Zuckerberg Investigators of the Chan Zuckerberg Initiative – Human Cell Atlas program.

AUTHOR CONTRIBUTIONS

Conceptualization, M.M.M., M.N., and A.F.S.; methodology, A.F.S., R.B., J.H.W., M.R.P., N.B., I.C.F., S.D., M.B., M.N., E.B., B.D., F.A., C.A.F., E.R.G., and J.-C.O.-M.; software, M.N., E.B., B.D., H.S., J.-C.O.-M., J.H.W., M.B., and S.D.; formal analysis, M.N., E.B., and B.D.; investigation, A.F.S., R.B., M.R.P., J.S., and I.C.F.; data curation, A.F.S., M.N., E.B., and B.D.; writing – original draft, A.F.S., R.B., J.H.W., M.N., M.R.P., and M.M.M.; writing – review and editing, M.N., M.M.M., and A.F.S.; supervision, M.M.M., C.A.F., E.R.G., and M.N.; project administration, M.M.M., M.N., and A.F.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. The author list of this paper includes contributors from the location where the research was conducted who participated in the data collection, design, analysis, and/or interpretation of the work.

Received: February 8, 2021

Revised: June 15, 2021

Accepted: August 3, 2021

Published: September 13, 2021

REFERENCES

Abràmoff, M.D., Magalhães, P.J., and Ram, S.J. (2004). Image processing with ImageJ. *Biophoton. Int.* 11, 36–42.

- Azioune, A., Storch, M., Bornens, M., Thery, M., and Piel, M. (2009). Simple and rapid process for single cell micro-patterning. *Lab. Chip* 9, 1640–1642.
- Bashirullah, A., Cooperstock, R.L., and Lipshitz, H.D. (1998). RNA localization in development. *Annu. Rev. Biochem.* 67, 335–394.
- Bassell, G.J., Zhang, H., Byrd, A.L., Femino, A.M., Singer, R.H., Taneja, K.L., Lifshitz, L.M., Herman, I.M., and Kosik, K.S. (1998). Sorting of beta-actin mRNA and protein to neurites and growth cones in culture. *J. Neurosci.* 18, 251–265.
- Batish, M., van den Bogaard, P., Kramer, F.R., and Tyagi, S. (2012). Neuronal mRNAs travel singly into dendrites. *Proc. Natl. Acad. Sci. U S A* 109, 4645–4650.
- Battich, N., Stoeger, T., and Pelkmans, L. (2013). Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* 10, 1127–1133.
- Battich, N., Stoeger, T., and Pelkmans, L. (2015). Control of transcript variability in single mammalian cells. *Cell* 163, 1596–1610.
- Besse, F., and Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. *Nat. Rev. Mol. Cell Biol.* 9, 971–980.
- Bouvrette, L.P., Cody, N.A.L., Bergalet, J., Lefebvre, F.A., Diot, C., Wang, X., Blanchette, M., and Lécuyer, E. (2018). CeFra-seq reveals broad asymmetric mRNA and noncoding RNA distribution profiles in *Drosophila* and human cells. *RNA* 24, 98–113.
- Briley, W.E., Bondy, M.H., Randeria, P.S., Dupper, T.J., and Mirkin, C.A. (2015). Quantification and real-time tracking of RNA in live cells using Sticky-flares. *Proc. Natl. Acad. Sci. U S A* 112, 9591–9595.
- Bruusgaard, J.C., Liestøl, K., Ekmark, M., Kollstad, K., and Gundersen, K. (2003). Number and spatial distribution of nuclei in the muscle fibres of normal mice studied in vivo. *J. Physiol.* 551, 467–478.
- Bruusgaard, J.C., Liestøl, K., and Gundersen, K. (2006). Distribution of myonuclei and microtubules in live muscle fibers of young, middle-aged, and old mice. *J. Appl. Physiol.* 100, 2024–2030.
- Buxbaum, A.R., Wu, B., and Singer, R.H. (2014). Single β -actin mRNA detection in neurons reveals a mechanism for regulating its translatability. *Science* 343, 419–422.
- Cabili, M.N., Dunagin, M.C., McClanahan, P.D., Biasesch, A., Padovan-Merhar, O., Regev, A., and Raj, A. (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* 16, 20.
- Carvalho, J.R., Fortunato, I.C., Franco, C.A., Fonseca, C.G., Pezzarossa, A., Barbacena, P., Dominguez-Cejudo, M.A., Vasconcelos, F.F., Santos, N.C., Carvalho, F.A., and Franco, C.A. (2019). Non-canonical Wnt signaling regulates junctional mechanocoupling during angiogenic collective cell migration. *eLife* 8. <https://doi.org/10.7554/eLife.45853>.
- Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348, aaa6090.
- S.N. Chiu, D. Stoyan, W. Kendall, and J. Mecke, eds. (2013). *Stochastic Geometry and its Applications*, 3rd Edition (Wiley-Blackwell).
- Chouaib, R., Safieddine, A., Pichon, X., Imbert, A., Kwon, O.S., Samacoits, A., Traboulsi, A.-M., Robert, M.-C., Tsanov, N., Coleno, A., Poser, I., et al. (2020). A dual protein-mRNA localization screen reveals compartmentalized translation and widespread co-translational RNA targeting. *Dev. Cell* 54, 773–791.e5.
- Clatterbuck-Soper, S.F., Jackrel, M.E., Shorter, J., and Mili, S. (2017). FUS inclusions disrupt RNA localization by sequestering kinesin-1 and inhibiting microtubule deetyrosination. *J. Cell Biol.* 216, 1015–1034.
- Crosetto, N., Bienko, M., and van Oudenaarden, A. (2015). Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* 16, 57–66.
- de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Vannary, M.-Y., Pankajakshan, P., Lecomte, T., Le Montagner, Y., Lagache, T., Dufour, A., and Olivo-Marin, J.-C. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9, 690–696.
- Denes, L.T., Kelley, C.P., and Wang, E.T. (2021). Microtubule-based transport is essential to distribute RNA and nascent protein in skeletal muscle. *bioRxiv*. <https://doi.org/10.1101/2021.02.26.433059>.
- Edelstein, A.D., Tsuchida, M.A., Amodaj, N., Pinkard, H., Vale, R.D., and Stuurman, N. (2014). Advanced methods of microscope control using μ Manager software. *J. Biol. Methods* 1, e10.
- Falcone, S., Roman, W., Hnia, K., Gache, V., Didier, N., Lainé, J., Auradé, F., Marty, I., Nishino, I., Charlet-Berguerand, N., Romero, N.B., et al. (2014). N-WASP is required for Amphiphysin-2/BIN1-dependent nuclear positioning and triad organization in skeletal muscle and is involved in the pathophysiology of centronuclear myopathy. *EMBO Mol. Med.* 6, 1455–1475.
- Fomina-Yadlin, D., Du, Z., and McGrew, J.T. (2014). Gene expression measurements normalized to cell number reveal large scale differences due to cell size changes, transcriptional amplification and transcriptional repression in CHO cells. *J. Biotechnol.* 189, 58–69.
- Franzini-Armstrong, C., and Peachey, L.D. (1981). Striated muscle-contraction and control mechanisms. *J. Cell Biol.* 91, 166s–186s.
- Garcia-Mata, R., Boulter, E., and Burridge, K. (2011). The ‘invisible hand’: regulation of RHO GTPases by RHOGDIs. *Nat. Rev. Mol. Cell Biol.* 12, 493–504.
- Gomes, E.R., Jani, S., and Gundersen, G.G. (2005). Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells. *Cell* 121, 451–463.
- Gupta, S., Marcel, N., Sarin, A., and Shivashankar, G.V. (2012). Role of actin dependent nuclear deformation in regulating early gene expression. *PLoS One* 7, e53031.
- Hale, C.M., Chen, W.C., Khatau, S.B., Daniels, B.R., Lee, J.S., and Wirtz, D. (2011). SMRT analysis of MTOC and nuclear positioning reveals the role of EB1 and LIC1 in single-cell polarization. *J. Cell Sci.* 124, 4267–4285.
- Hengst, U., Deglincerti, A., Kim, J.K., Jeon, N.L., and Jaffrey, S.R. (2009). Axonal elongation triggered by stimulus-induced local translation of a polarity complex protein. *Nat. Cell Biol.* 11, 1024–1030.
- Jain, N., Iyer, K.V., Kumar, A., and Shivashankar, G.V. (2013). Cell geometric constraints induce modular gene-expression patterns via redistribution of HDAC3 regulated by actomyosin contractility. *Proc. Natl. Acad. Sci. U S A* 110, 11349–11354.
- Jansen, R.P. (2001). mRNA localization: message on the move. *Nat. Rev. Mol. Cell Biol.* 2, 247–256.
- Katz, Z.B., Wells, A.L., Park, H.Y., Wu, B., Shenoy, S.M., and Singer, R.H. (2012). β -Actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes Dev.* 26, 1885–1890.
- Kempe, H., Schwabe, A., Crémazy, F., Verschure, P.J., and Bruggeman, F.J. (2015). The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. *Mol. Biol. Cell* 26, 797–804.
- Kim, D.-H., Cho, S., and Wirtz, D. (2014). Tight coupling between nucleus and cell migration through the perinuclear actin cap. *J. Cell Sci.* 127, 2528–2541.
- Kislauskis, E.H., Zhu, X., and Singer, R.H. (1997). beta-Actin messenger RNA localization and protein synthesis augment cell motility. *J. Cell Biol.* 136, 1263–1270.
- Kloc, M., Zearfoss, N.R., and Etkin, L.D. (2002). Mechanisms of subcellular mRNA localization. *Cell* 108, 533–544.
- Kozachenko, L.F., and Leonenko, N.N. (1987). Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii. Problems Inform. Transmission* 23, 9–16.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastri, M.E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L.E., et al. (2018). RNA velocity of single cells. *Nature* 560, 494–498.
- Langdon, E.M., Qiu, Y., Ghanbari Niaki, A., McLaughlin, G.A., Weidmann, C.A., Gerbich, T.M., Smith, J.A., Crutchley, J.M., Termini, C.M., Weeks, K.M., Myong, S., and Gladfelter, A.S. (2018). mRNA structure determines specificity of a polyQ-driven phase separation. *Science* 360, 922–927.
- Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174–187.

- Lee, C., Zhang, H., Baker, A.E., Occhipinti, P., Borsuk, M.E., and Gladfelter, A.S. (2013). Protein aggregation behavior regulates cyclin transcript localization and cell-cycle control. *Dev. Cell* 25, 572–584.
- Macdonald, P.M., and Struhl, G. (1988). cis-acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos. *Nature* 336, 595–598.
- Kunde, S.A., Musante, L., Grimme, A., Fischer, U., Müller, E., Wanker, E.E., and Kalscheuer, V.M. (2011). The X-chromosome-linked intellectual disability protein PQBP1 is a component of neuronal RNA granules and regulates the appearance of stress granules. *Hum. Mol. Genet.* 20, 4916–4931.
- Mardakheh, F.K., Paul, A., Kümper, S., Sadok, A., Paterson, H., McCarthy, A., Yuan, Y., and Marshall, C.J. (2015). Global analysis of mRNA, translation, and protein localization: local translation is a key regulator of cell protrusions. *Dev. Cell* 35, 344–357.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719–730.
- Medioni, C., and Besse, F. (2018). The secret life of RNA: lessons from emerging methodologies. *Methods Mol. Biol.* 1649, 1–28.
- Mili, S., Moissoglu, K., and Macara, I. (2008). Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* 453, 115–119.
- Moissoglu, K., Yasuda, K., Wang, T., Chrisafis, G., Mili, S., et al. (2019). Translational regulation of protrusion-localized RNAs involves silencing and clustering after transport. *eLife* 8. <https://doi.org/10.7554/eLife.44752>.
- Moor, A.E., Golan, M., Massasa, E.E., Lemze, D., Weizman, T., Shenhav, R., Baydatch, S., Mizrahi, O., Winkler, R., Golani, O., Stern-Ginossar, N., and Itzkovitz, S. (2017). Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science* 357, 1299–1303.
- Oliivo-Marín, J.-C. (2002). Extraction of spots in biological images using multi-scale products. *Pattern Recogn.* 35, 1989–1996.
- Padovan-Merhar, O., Gautham, P.N., Biaisch, A.G., Mayer, A., Scarfone, S., Foley, S.W., Wu, A.R.L., Churchman, S., Singh, A., and Raj, A. (2015). Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol. Cell* 58, 339–352.
- Park, H.Y., Lim, H., Yoon, Y.J., Follenzi, A., Nwokafor, C., Lopez-Jones, M., Meng, X., and Singer, R.H. (2014). Visualization of dynamics of single endogenous mRNA labeled in live mouse. *Science* 343, 422–424.
- Park, H.Y., Trcek, T., Wells, A.L., Chao, J.A., and Singer, R.H. (2012). An unbiased analysis method to quantify mRNA localization reveals its correlation with cell motility. *Cell Rep.* 1, 179–184.
- Pimentel, M.R., Falcone, S., Cadot, B., and Gomes, E.R. (2017). In vitro differentiation of mature myofibers for live imaging. *J. Vis. Exp.* 119, 55141. <https://doi.org/10.3791/55141>.
- Raj, A., van den Boogard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* 5, 877–879.
- Razafsky, D., Wirtz, D., and Hodzic, D. (2014). Nuclear envelope in nuclear positioning and cell migration. *Adv. Exp. Med. Biol.* 773, 471–490.
- Ridley, A.J., Schwartz, M.A., Burridge, K., Firtel, R.A., Ginsberg, M.H., Borisy, G., Parsons, J.T., Horwitz, A.R., et al. (2003). Cell migration: integrating signals from front to back. *Science* 302, 1704–1709.
- Ripley, B.D. (1977). Modelling spatial patterns. *J. R. Stat. Soc.* 39, 172–192.
- Roman, W., Martins, J.P., Carvalho, F.A., Voituriez, R., Abella, J.V., Santos, N.C., Cadot, B., Way, M., and Gomes, E.R. (2017). Myofibril contraction and crosslinking drive nuclear movement to the periphery of skeletal muscle. *Nat. Cell Biol.* 19, 1189–1201.
- Samacoits, A., Chouaib, R., Safieddine, A., Traboulsi, A.-M., Ouyang, W., Zimmer, C., Peter, M., Bertrand, E., Walter, T., and Mueller, F. (2018). A computational framework to study sub-cellular RNA localization. *Nat. Commun.* 9, 4584.
- Savulescu, A.F., Stoychev, S., Mamputha, S., and Mhlanga, M.M. (2020). Biochemical pulldown of mRNAs and long noncoding RNAs from cellular lysates coupled with mass spectrometry to identify protein binding partners. *Bio Protoc.* 10, e3639.
- Schauer, K., Duong, T., Bleakley, K., Bardin, S., Bornens, M., and Goud, B. (2010). Probabilistic density maps to study global endomembrane organization. *Nat. Methods* 7, 560–566.
- Schmoranzner, J., Fawcett, J.P., Segura, M., Tan, S., Vallee, R.B., Pawson, T., and Gundersen, G.G. (2009). Par3 and dynein associate to regulate local microtubule dynamics and centrosome orientation during migration. *Curr. Biol.* 19, 1065–1074.
- Sharp, J.A., Plant, J.J., Ohsumi, T.k., Borowsky, M., and Blower, M.D. (2011). Functional analysis of the microtubule-interacting transcriptome. *Mol. Biol. Cell* 22, 4312–4323.
- Strell, C., Hilsche, M.M., Laxman, N., Svedlund, J., Wu, C., Yokota, C., and Nilsson, M. (2019). Placing RNA in context—methods for spatially resolved transcriptomics. *FEBS J.* 286, 1468–1481.
- Stueland, M., Wang, T., Park, H.Y., and Mili, S. (2019). RDI calculator: an analysis tool to assess RNA distributions in cells. *Sci. Rep.* 9, 8267.
- Tautz, D., and Pfeifle, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* 98, 81–85.
- Théry, M. (2010). Micropatterning as a tool to decipher cell morphogenesis and functions. *J. Cell Sci.* 123, 4201–4213.
- Théry, M., Racine, V., Pepin, A., Piel, M., Chen, Y., Jean-Baptiste, S., and Bornens, M. (2005). The extracellular matrix guides the orientation of the cell division axis. *Nat. Cell Biol.* 7, 947–953.
- Théry, M., Racine, V., Piel, M., Pepin, A., Dimitrov, A., Chen, Y., Jean-Baptiste, S., and Bornens, M. (2006). Anisotropy of cell adhesive microenvironment governs cell internal organization and orientation of polarity. *Proc. Natl. Acad. Sci. U S A* 103, 19771–19776.
- Vargha, A., and Delaney, H.D. (2000). A critique and improvement of the CL common language effect size statistics of McGraw and Wong. *J. Educ. Behav. Stat.* 25, 101–132.
- Vilmont, V., Cadot, B., Ouanounou, G., and Gomes, E.R. (2016). A system for studying mechanisms of neuromuscular junction development and maintenance. *Development* 143, 2464–2477.
- Warrell, J.H., Savulescu, A.F., Brackin, R., and Mhlanga, M.M. (2016). Generalized statistical tests for mRNA and protein subcellular spatial patterning against complete spatial randomness. *arXiv*. <http://arxiv.org/abs/1602.06429>.
- Weis, B.L., Schleiff, E., and Zerges, W. (2013). Protein targeting to subcellular organelles via mRNA localization. *Biochim. Biophys. Acta Mol. Cell Res.* 1833, 260–273.
- Xie, F., Shao, S., Aziz, A.U.R., Zhang, B., Wang, H., and Liu, B. (2017). Role of Rho-specific guanine nucleotide dissociation inhibitor α regulation in cell migration. *Acta Histochem* 119 (3), 183–189. <https://doi.org/10.1016/j.acthis.2017.01.008>.
- Yamagishi, M., Shirasaki, Y., and Funatsu, T. (2009). Size-dependent accumulation of mRNA at the leading edge of chicken embryo fibroblasts. *Biochem. Biophys. Res. Commun.* 390, 750–754.
- Yoon, Y.J., Wu, B., Buxbaum, A.R., Das, S., Tsai, A., English, B.P., Grimm, J.B., Lavis, L.D., and Singer, R.H. (2016). Glutamate-induced RNA localization and translation in neurons. *Proc. Natl. Acad. Sci. U S A* 113, E6877–E6886.
- Zappulo, A., van den Bruck, D., Ciolli Mattioli, C., Franke, V., Imami, K., McShane, E., Moreno-Estelles, M., Calviello, L., Filipchik, A., Peguero-Sanchez, E., Müller, T., Woehler, A., et al. (2017). RNA localization is a key determinant of neurite-enriched proteome. *Nat. Commun.* 8, 583. <https://doi.org/10.1038/s41467-017-00690-6>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Tubulin	Abcam	# ab6160; RRID:AB_305328
anti RhoGDI antibody	Santa Cruz	# sc-360; RRID:AB_2227516
Anti-PARD3	Abcam	# ab64646; RRID:AB_1603911
Anti β -Actin	Santa Cruz	# sc-81178; RRID:AB_2223230
Anti Gapdh	Santa Cruz	# sc-25778; RRID:AB_10167668
anti-GM130	BD Biosciences	# 610823; RRID:AB_398142
Goat Anti-RABBIT IgG ATTO 550 Conjugated	Rockland	# 611-154-122S; RRID:AB_10894121
Donkey anti Rat AlexaFluor647	Abcam	# ab150151
Donkey anti Mouse Alexa488	Thermo Fisher Scientific	# A21202; RRID:AB_141607
Chemicals, peptides, and recombinant proteins		
Collagenase type I	Sigma-Aldrich	Cat #C0130
Dispase	Roche	Cat #04942078001
Iscove's Modified Dulbecco's Medium (IMDM) with GlutaMAX	Invitrogen	Cat #31980022
Fetal bovine serum (FBS)	Eurobio	Cat #CVFVSF00-01
Penicillin/streptomycin (Penstrep)	Thermo Fisher Scientific	Cat #15140122
Matrigel	Corning	Cat #354230
Recombinant Rat Agrin	R&D systems	Cat #550-AG-100
Horse serum	Thermo Fisher Scientific	Cat #26050088
DMEM/F12 medium	Thermo Fisher scientific (Gibco)	# 11054001
EGM-2 Bulletkit	Lonza	# CC-3162
FBS	Biochrom	# S0613
Cytochalasin D	Sigma	# 22144-77-0
Cycloheximide	Sigma	# 66-81-9
Nocodazole	Sigma	# 31430-18-9
penicillin/streptomycin	Gibco	# 15140122
Hepes	Sigma	# H0887-100 ML
Human umbilical vein endothelial cells (HUVECs)	Lonza	# C2519A
PLL(20)-g[3.5]- PEG(2)	Surface Solutions (SuSoS)	N/A
PBS	Thermo Fisher Scientific (Gibco)	# 10010023
Fibronectin	Thermo Fisher Scientific (Gibco)	# 3010018
Fibrinogen-Alexa Fluor 488	Thermo Fischer Scientific (Invitrogen)	# F13191
NaHCO ₃	Sigma	# 144-55-8
Sodium tetraborate	Sigma	# 1330-43-4
ATTO-565 NHS-ester dye	ATTO-TEC	# AD 565
Formaldehyde	Sigma	# 50-00-0
SSC-Buffer	Sigma	# S6639-1L
Ethanol 99.8 %	Sigma	# 64-17-5
Dextran Sulfate	VWR	# 0198-50G
Ribonucleoside Vanadyl Complex	NEB (New England Labs)	# S1402S

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
tRNA from E. coli MRE 600	Sigma	# TRNAMRE-RO
Formamide	Sigma	# F9037-100ML
BSA	Thermo Fisher Scientific	# AM2616
D-(+)-Glucose solution	Sigma	# 50-99-7
Glucose oxidase	Sigma	# 9001-37-0
Catalase	Sigma	# 9001-05-2
DAPI	Life Technologies	# D1306
Critical commercial assays		
DharmaFECT one reagent	Dharmacon, GE Healthcare	# T-2005-01
RNeasy Mini Kit (Qiagen)	Qiagen	# 74104
Superscript IV First-Strand Synthesis System	Invitrogen	# 8091050
Experimental models: Cell lines		
NIH/3T3	ATCC, Cellonex	# CCL-92
Human umbilical vein endothelial cells (HUVECs)	Lonza	# C2519A
Experimental models: Organisms/strains		
Mouse: C57BL/6 Strain	Charles River	code: 027
Oligonucleotides		
ON-TARGETplus Human PRRC2C (23215) siRNA - SMARTpool, 10 nmol	Dharmacon	# L-014078-00-0010
On-Targetplus Non-targeting siRNA #1, 20 nM	Dharmacon	# D-001810-01-05
human GAPDH forward sequence GTCAAGGCTGAGAACGGGAA	This manuscript	N/A
human GAPDH reverse sequence TGGACTCCACGACGACTACTCA	This manuscript	N/A
human PRRC2C forward sequence GAAGCAGTCCAGTCAGCC	This manuscript	N/A
human PRRC2C reverse sequence GTTGCGTGACTGAAGAACC	This manuscript	N/A
<i>Gapdh</i> probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A
<i>Arhgdia</i> probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A
β -Actin probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A
<i>Rab13</i> probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Pkp4</i> probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A
<i>Pard3</i> probes - 24-48 20-mers, each 20-mer containing a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe	Biosearch Technologies	N/A
Deposited data		
HDF5 data files of primary and secondary image descriptors for FISH Micropatterned images, CHX, CytoD, Nocodazole, PRRC2C and Muscle data	http://dypfish.org/	https://doi.org/10.5281/zenodo.5155127
Software and algorithms		
Chemotaxis and Migration Tool	Chemotaxis and Migration Tool (https://ibidi.com/chemotaxis-analysis/171-chemotaxis-and-migration-tool.html)	N/A
Fiji/ImageJ (2.0.0)	Schindelin et al., 2012	N/A
GraphPad Prism	GraphPad Prism (https://graphpad.com)	N/A
MATLAB - polarity analysis script to calculate the Polarity Index (PI)	Carvalho et al., 2019	N/A
ICY and μ Manager	de Chaumont et al., 2012 https://icy.bioimageanalysis.org Edelstein et al., 2014 https://doi.org/10.14440/jbm.2014.36	N/A
SV Control software	Optical Biosystems	N/A
HDF5	The HDF Group, Hierarchical data format version 5, 2000-2010, http://www.hdfgroup.org/HDF5	N/A
DypFISH code	https://github.com/cbib/dypfish	https://doi.org/10.5281/zenodo.5153514

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Musa M. Mhlanga (belenus@mhlangalab.org).

Materials availability

This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture and medium

NIH/3T3 cells (mouse fibroblast sarcoma cell line) were purchased from ATCC, Cellonex and were cultured in DMEM-F12 (Thermo Fisher Scientific - Gibco, # 11054001) supplemented with 10 % fetal bovine serum (Biochrom, # S0613). Human umbilical vein endothelial cells (HUVECs) were purchased from Lonza, and cultured in EGM-2 Bulletkit (Lonza, # CC-316) supplemented with 1% penicillin/streptomycin (Gibco, # 15140122). Primary muscle cells were isolated from newborn mice and differentiated as described in [Pimentel et al., 2017](#).

METHOD DETAILS

Cell culture and treatments

NIH/3T3 cells were grown at 37°C to 100 % confluence in DMEM/F12 medium supplemented with 10 % FBS in a humidified atmosphere containing 5 % CO₂. Prior to micropatterning cells were serum-starved for 16 hr in DMEM/F12. For disruption of microtubule

polymerization, nocodazole (Sigma) was added to a final concentration of 50 ng/ml to the medium, post removal of unattached cells and incubated for 3/5 hours before fixation of cells. Cytochalasin D (Sigma) was added to a final concentration of 1 mg/ml to the medium post removal of unattached cells and incubated for 1 h before fixation of cells. Cycloheximide (Sigma) was added to a final concentration of 10 μ g/ml 24 h prior to seeding on micropatterns, as well as post washes of unattached cells, following standard practice. HUVECs were cultured following the manufacturer's guidelines with complete medium EGM-2 Bulletkit (CC-3162, Lonza) supplemented with 1% penicillin/streptomycin (#15140122, Gibco). HUVECs were transfected with 25 nM of siRNA using the DharmaFECT one reagent (Dharmacon, GE Healthcare) and following the Dharmacon siRNA Transfection Protocol. L-014078-00-0010 ON-TARGETplus Human PRRC2C (23215) siRNA - SMARTpool, 10 nmol and D-001810-01-05 On-Target^{plus} Non-targeting siRNA #1 (Dharmacon), 20 nM.

RNA isolation and analysis

RNA extraction from HUVECs was performed using RNeasy Mini Kit (Qiagen) as described by the manufacturer's protocol, followed by cDNA synthesis (Superscript IV First-Strand Synthesis System, Invitrogen) and subsequent quantitative real-time PCR (RT-qPCR). The following primers were used: human GAPDH forward sequence GTCAAGGCTGAGAACGGGAA and reverse sequence TGGACTCCACGACGTACTCA, human PRRC2C forward sequence GAAGCAGTTCAGTCAGCC and reverse sequence GTTGC GTGACTGAAGAACC, mouse PRRC2C forward sequence GAAGCAGTTCAGTCAGCC and reverse sequence GTTGC GTGACTGAAGAACC.

Cell micropatterning

Micropattern production was performed as previously described (Azioune et al., 2009). Briefly, glass coverslips were exposed to deep UV light using a UVO Cleaner (Jelight Company) for 5 mins. Cleaned coverslips were incubated with 0.1 mg/ml PLL-g-PEG (Surface Solutions) in 10 mM HEPES, pH 7.4 at RT for 1 hr. They were then rinsed once in PBS followed by one rinse in MilliQ water. The pegylated glass coverslips were then placed on a custom designed chromium photomask (Delta mask) (containing the desired micropatterns) and exposed to deep UV light for 5 mins. The patterned glass coverslips were then incubated with a fibronectin/fibrinogen-Alexa Fluor488 mixture (Life Technologies) in 100 mM NaHCO₃, pH 8.5, at RT for 1 hr. The coverslips were then rinsed in PBS and used immediately for cell seeding. Serum-starved NIH/3T3 cells were seeded on the micropatterned surfaces at a density of 10,000 cells/cm². After 30 mins, unattached cells were removed by gentle aspiration and replacement of the medium to DMEM/F12 medium supplemented with 10% FBS, which induced polarization of cells. Attached micropatterned cells were incubated at 37°C for 2 to 7 hours.

RNA probes and reagents

Design and manufacture of RNA FISH probes for use in the single molecule FISH method were performed according to the protocol by (Raj et al., 2008). Multiple 20-mer oligonucleotide probes targeting the following mRNAs: *Gapdh*, *Arhgdia*, β -*Actin*, *Rab13*, *Pkp4* and *Pard3* were purchased (Biosearch Technologies). Each 20-mer contains a mdC(TEG-Amino) 3' modification used to conjugate an NHS-ester ATTO-565 fluorescent dye (ATTO-TEC) to the probe. In brief, concentrated oligonucleotide probes were resuspended in 0.1 M Sodium tetraborate (Sigma) and mixed with resuspended 0.25 mg of the NHS-ester dye and incubated overnight at 37°C. This was followed by ethanol precipitation of the probes and purification by reverse phase HPLC on a XBRIDGETM OST C18 column to enrich for dye conjugated probes.

Immuno-RNA FISH staining

For experiments utilizing the *Gapdh*, *Arhgdia*, β -*Actin* RNA probes: micropatterned NIH/3T3 cells were fixed in 3.7 % formaldehyde for 10 min at 37°C followed by washes in PBS and overnight permeabilization in 70% ethanol at 4°C. For experiments utilizing the *Rab13*, *Pkp4*, *Pard3* RNA probes: micropatterned NIH/3T3 cells were fixed in pre-chilled methanol for 10 min, followed immediately by RNA FISH. The single molecule FISH method was modified from (Raj et al., 2008) to include immunofluorescence staining to detect the microtubule cytoskeleton. Cells were rehydrated in wash buffer (10 % formaldehyde, 2X SSC) for 5 min. Hybridization was conducted overnight in a humidified chamber at 37°C in Hyb buffer (10 % dextran sulfate, 1 μ g/ μ l E.coli tRNA, 2mM Vanadyl ribonucleoside complex, 0.02 % RNase-free BSA, 10 % formamide, 2X SSC) combined with 50 ng of the desired RNA probe along with primary antibody - rat monoclonal anti-tubulin antibody (Abcam). Cells were then washed 2X (30 min at room temperature) with antibody wash buffer (10 % formaldehyde, 2X SSC, anti-rat secondary antibody conjugated to Alexa Fluor 647 (Abcam)) followed by 1X wash with wash buffer. Cells were then incubated in equilibration buffer (0.4 % glucose, 2X SSC) for 5 mins and counter stained with 1 μ g/ml DAPI (4',6-diamidino-2-phenylindole; Life Technologies). Coverslips were mounted in imaging buffer (3.7 μ g/ μ l glucose oxidase and 1U catalase in equilibration buffer) and imaged.

Immunofluorescence staining

Micropatterned cells were fixed in 3.7 % formaldehyde for 10 min at 37°C, then washed with PBS followed by overnight incubation in 70 % ethanol at 4°C. The cells were then washed with FBS followed by permeabilization for 10 min in 0.25 % Triton-X at room temperature. Following this, the cells were washed thrice with PBS for 5 min each and incubated in blocking buffer (0.2 % BSA/PBS) for 30 min at room temperature. The cells were then incubated in the desired primary antibody solution (diluted in PBS) along with rat

monoclonal anti-tubulin antibody (Abcam) to detect the microtubule cytoskeleton for 1 hr at room temperature. RhoGDI α , Par3, β -Actin and Gapdh proteins were detected using rabbit polyclonal anti-Arhgdia (Santa Cruz), rabbit polyclonal anti-Pard3 (Abcam), mouse monoclonal β -Actin (Santa Cruz) and mouse monoclonal anti-Gapdh (Santa Cruz) respectively. Cells were then washed 3X with PBS following incubation with corresponding anti-rabbit secondary antibody conjugated to ATTO 550 (Rockland) together with anti-rat secondary antibody conjugated to Alexa Fluor 647 (Abcam) for 1 hr at room temperature. A further 3X wash with PBS was conducted followed by incubation in equilibration buffer (0.4 % glucose, 2X SSC) for 5 mins and counter stained with 1 μ g/ml DAPI (4',6-diamidino-2-phenylindole; Life Technologies). Coverslips were mounted in imaging buffer (3.7 μ g/ μ l glucose oxidase and 1U catalase in equilibration buffer) and imaged. For immunofluorescence in HUVECs, cells were fixed using 1 % Paraformaldehyde (PFA) supplemented with 1 M MgCl $_2$ and 1 M CaCl $_2$ (1 μ l/2 ml). Blocking and permeabilization was made with 3 % BSA in PBS-T (PBS with 0.1 % Triton X-100). Primary antibody - mouse anti-GM130 (BD Biosciences) was diluted in blocking solution and incubated for 2 h. Secondary antibody donkey anti-mouse Alexa488 (Thermo Fisher Scientific) was diluted in the blocking solution [3% BSA in PBS-T (PBS with 0.1 % Triton X-100)] and incubated for 1 h. For nucleus staining, HUVECs were incubated with 1x DAPI and then coverslips were mounted on microscopy glass slides using Mowiol DABCO (Sigma-Aldrich).

Polarity and migration assays

To calculate axial cell polarity and analyze cell migration, scratch-wound assays were performed. The wound was created on the surface of a 12-well-plate for cell migration assessment and on a microscopy glass-slide for axial polarity measurements with a 200 μ L pipette tip. HUVECs migration was followed for 16 hours, individual cells were tracked and migration behavior, directness and cell velocity, were analyzed using the FIJI TrackMate plug in and the Chemotaxis and Migration Tool (free software from Ibbidi). Tile-scan images of cells stained for Golgi (GM130) and nucleus (DAPI) were imported and analyzed in MATLAB using a polarity analysis script in order to calculate the Polarity Index (PI). PI shows the orientation strength of the cell monolayer and it varies from 0 to +1, where 0 corresponds to random localization of the Golgi and +1 corresponds to same directionality of the Golgi in all.

Image acquisition

Most samples were imaged on a custom built spinning disk confocal Revolution XD system (Andor) comprising of a Zeiss Axio Observer.Z1 microscope with a 63X Plan-Apochromat objective (numerical aperture 1.4) and a cooled EMCCD camera (Andor iXon 897). Z-dimension positioning and control was accomplished by a piezoelectric motor (NanoScanZ, Prior Scientific). Images were captured using a custom developed algorithm based on ICY and μ Manager that allowed autonomous image acquisition (Figure S1A). In brief, the position of the micropatterns on the micropatterned surface were determined autonomously using the grid detection, alignment and calibration algorithm. This was then followed by sequential autonomous stepping through the micropatterned grid to determine the presence of a cell on the micropattern. If a single cell was detected on the micropattern surface by the algorithm then a z-dimension series of images was captured every 0.3 μ m in four different fluorescence channels using emission filters for DAPI (DNA), Alexa Fluor 488 (micropatterns), ATTO 565 (mRNA/protein) and Alexa Fluor 647 (tubulin) and exposure times of 10 ms, 350 ms, 1 s (mRNA) or 500 ms (protein) and 350 ms respectively. A few samples were imaged on a custom built Nikon Ti Eclipse widefield TIRF microscope using a 100X N.A. 1.49 Nikon Apochromat TIRF oil immersion objective and equivalent fluorescent channels as above, followed by processing using an automated background noise subtraction algorithm using ImageJ (Abràmoff et al., 2004). Samples were also imaged on a StellarVision microscope using Synthetic Aperture Optics technology (Optical Bio-systems) and processed to obtain stacked single cell images for analysis. These are listed in the relevant supplementary tables. All images acquired on the Zeiss Axio Observer.Z1 microscope were 512x512 pixels in size, whereas the ones acquired on the StellarVision were 251x251 pixels. See Tables of image acquisition characteristics for image counts.

Materials

Images were acquired for *Arhgdia*, *Gapdh*, *β -Actin*, *Pard3*, *Pkp4*, and *Rab13* genes. Tables of image acquisition characteristics below recapitulate acquisition conditions, techniques and number of acquired images in each series of FISH and IF data.

The following table recapitulates all the image acquisition series characteristics and numbers for mouse fibroblast cells grown in micropatterned and standard cultures (indicated by *).

Gene	Molecular species	Technique	Time point (number of images)
Arhgdia*	mRNA	FISH	1 h (76); 3 h (24)
Arhgdia (control)	mRNA	FISH	2 h (59); 3 h (45); 4 h (52); 5 h (48)
Arhgdia (control)	protein	IF	2 h (31); 3 h (14); 5 h (61); 7 h (21)
β -Actin	mRNA	FISH	2 h (41); 3 h (42); 4 h (28); 5 h (29)
β -Actin	protein	IF	2 h (9); 3 h (19); 5 h (17); 7 h (24)
Gapdh	mRNA	FISH	2 h (58); 3 h (62); 4 h (50); 5 h (43)
Gapdh	protein	IF	2 h (27); 3 h (31); 5 h (21); 7 h (21)

(Continued on next page)

Continued

Gene	Molecular species	Technique	Time point (number of images)
Pard3 (control)	mRNA	FISH	2 h (25); 3 h (13); 4 h (15); 5 h (12)
Pard3 (control)	protein	IF	2 h (25); 3 h (20); 5 h (9); 7 h (26)
Pkp4	mRNA	FISH	2 h (17); 3 h (32); 4 h (11); 5 h (26)
Rab13	mRNA	FISH	2 h (16); 3 h (13); 4 h (22); 5 h (8)

Table describing Image acquisition characteristics for micropatterned and cultured cells.

Image acquisition series characteristics and numbers for mouse fibroblast cells grown in micropatterned cultures in control and drug-disrupted conditions for the CytoD experiment are detailed in the following table.

Gene (condition)	Molecular species	Technique	Time point (number of images)
Arhgdia (Control)	mRNA	FISH	1 h 15 (19)
Arhgdia (Control)	Protein	IF	1 h 15 (19)
Arhgdia (CytoD)	mRNA	FISH	1 h 15 (22)
Arhgdia (CytoD)	protein	IF	1 h 15 (25)
Pard3 (Control)	protein	IF	1 h 15 (24)
Pard3 (CytoD)	protein	IF	1 h 15 (20)

Table describing the image acquisition characteristics for the CytoD experiments.

Below are listed the image acquisition series characteristics and numbers for mouse fibroblast cells grown in micropatterned cultures in control and drug-disrupted conditions for the nocodazole experiment.

Gene (condition)	Molecular species	Technique	Time point (number of images)
Arhgdia (control)	mRNA	FISH	2 h (59); 3 h (45); 4 h (52); 5 h (50)
Arhgdia (control)	protein	IF	2 h (31); 3 h (14); 5 h (61); 7 h (21)
Arhgdia (nocodazole)	mRNA	FISH	3 h (41); 5 h (32)
Arhgdia (nocodazole)	protein	IF	3 h (20); 5 h (20)
Pard3 (control)	mRNA	FISH	2 h (29); 3 h (13); 4 h (15); 5 h (12)
Pard3 (control)	protein	IF	2 h (25); 3 h (20); 5 h (9); 7 h (26)
Pard3 (nocodazole)	mRNA	FISH	3 h (25); 5 h (21)
Pard3 (nocodazole)	protein	IF	3 h (14); 5 h (22)

Table describing the image acquisition characteristics for the Nocodazole experiments.

All image acquisition series characteristics and numbers for mouse fibroblast cells grown in micropatterned cultures in control and drug-disrupted conditions for the CHX experiment are presented in the following table.

Gene (condition)	Molecular species	Technique	Time point (number of images)
Arhgdia (Supp. control)	mRNA	FISH	2 h (19); 3 h (11); 5 h (11)
Arhgdia (Supp. control)	protein	IF	2 h (34); 3 h (25); 5 h (26)
Arhgdia (CHX)	mRNA	FISH	2h (5); 3h (2); 5h (4)
Arhgdia (CHX)	protein	IF	2h (16); 3h (16); 5h (16)
Pard3 (Supp. control)	mRNA	FISH	2h (15); 3h (12)
Pard3 (Supp. control)	protein	IF	2h (34); 3h (24); 5h (18)
Pard3 (CHX)	mRNA	FISH	2h (5); 3h (4); 5h (3)
Pard3 (CHX)	protein	IF	2h (10); 3h (12); 5h (3)

Table describing the image acquisition characteristics for the CHX experiments.

The following table recapitulates the image acquisition series characteristics and numbers for mouse fibroblast cells grown in micropatterned cultures in control and siRNA conditions for the PRRC2C experiment.

Gene	Molecular species	Technique	Time point (number of images)
Arhgdia (PRRC2C)	mRNA	FISH	PRRC2C (20); Control (38)
Arhgdia (PRRC2C)	protein	IF	PRRC2C (36); Control (46)

Table describing the image acquisition characteristics for the PRRC2C experiments.

Myofibers were differentiated and fixed as previously described (Pimentel et al., 2017; Roman et al., 2017).

Images were acquired for Actn2 and Gapdh genes. Acquisition conditions, techniques and number of acquired images in each series are recapitulated in the following table.

Gene	Element	Technique	Number of images
Actn2	Phalloidin mature fibers	FISH / IF	12
Actn2	Phalloidin immature fibers	FISH / IF	8
Gapdh	Phalloidin mature fibers	FISH / IF	14

Table describing the image acquisition characteristics for muscle cells.

In all of our experiments, we stained different cellular elements acquired at the same time as the FISH and IF signals, with different staining as detailed in the Feature staining table.

Feature	Staining
DNA	DAPI - 358/461
Micropatterns (only for micropatterned cells)	Fibrinogen - Alexa Fluor 488
Tubulin	Alexa Fluor 647

Table showing the summary of labeled cellular markers in the same images as the FISH and IF signals.

Images were acquired in the TIFF format. Our image processing pipeline transformed images into an HDF5 file (downloadable from the website www.dypfish.org).

QUANTIFICATION AND STATISTICAL ANALYSIS

All computational analysis performed in the DypFISH project and described below were implemented in Python.

Primary image descriptors

Given the TIFF files, we first computed primary image descriptors for each image and stored them in an HDF5 file for each acquisition series.

MTOC and nucleus centroid

FISH and IF images were manually annotated using the γ -Tubulin signal to obtain the coordinates (x, y) of the MTOC. The nucleus centroid was computed as the geometric center of the nucleus mask (see below).

Cell, nucleus and cytoplasm masks

Cell and nucleus masks were computed for all images (FISH and IF) using γ -Tubulin and DAPI signals, respectively.

For each image we obtained the maximum projection of the γ -Tubulin stained z-stack. A vignetting correction is further applied to each resulting image individually by performing a pixel wise multiplication between each pixel value and the vignetting function. The detected cells being in the microscope's focus, we assumed the optical center to be the center of the image and the intensity fall-off to be radially symmetric and the vignetting function is defined for each pixel x, y as $e^{-d / [(w/2)^2 + (h/2)^2]}$, where $d = (x - w/3)^2 + (y - h/2)^2$ and w and h are the image's width and height, respectively. For contrast enhancement we performed histogram stretching by applying a linear normalization in order to stretch the interval of the intensities to the $[0, 255]$ interval.

Cell Mask was detected from the γ -Tubulin channel. We applied a local entropy filter to each pixel i as follows: $e(i) = -\sum_{j \in N_i} p_j \log_2 p_j$, where p_j is the proportion of pixels in the neighborhood N_i having the same intensity as pixel j . The neighborhood

size was set to 30×30 . For certain noisy image series we further applied a percentile thresholding to the resulting entropy histogram. As a last step, we performed Canny edge detection, which detected edges by applying Sobel operators to the smoothed image, followed by hysteresis. However, the resulting edges were usually non-contiguous due to a weak γ -Tubulin signal or a high rate of noise. Thus, we successively applied mathematical morphological operators, such as dilation and closing, followed by erosion conventionally used to fill small gaps.

As the result we obtained a contiguous contour to which we applied the marching squares algorithm in order to obtain a 2D cellular segmentation mask, $M^{cell}(x,y)$, which is 1 for the cellular region and 0 otherwise.

Nucleus mask procedure was very similar using the DAPI signal and yielding $M^{nucleus}(x,y)$, except that the local entropy filter was in most instances replaced by an Otsu filter, depending on the quality of the DAPI signal. Mathematical morphology algorithms were applied to neighborhoods ranging from 16×16 to 20×20 depending on the image acquisition characteristics.

Binary cellular and nucleus masks above were used to define a binary cytoplasm mask of the cell, $M^{cytoplasm}(x,y) = M^{cell}(x,y) \wedge \neg M^{nucleus}(x,y)$

Zero level

An acquired image stack might contain irrelevant slices because the focal field of the microscope is outside the cell (above or below). To determine which slice contained the bottom of the cell and had to be considered as the first relevant bottom slice of the stack, we defined the *zero level* descriptor corresponding to the index of the slice having the maximum summed γ -Tubulin intensity. This zero level reference z-slice was used in further analysis such as e.g. the height-map computation or the degree of clustering.

Height-map and cell volume

The height-map was built by segmenting each z-slice of a stacked image, which generated the 3D segmentation of the cell. It was performed for all FISH and IF images using the γ -Tubulin signal. Given a z-slice above the zero level we applied the cell mask detection procedure previously described, which defined a z-slice mask $M_z(x,y)$ with values corresponding to the height of the slice (z) within the mask and 0 outside. This set of slice masks defined the 3D representation, called height-map and denoted $h(x,y)$ where the value at each coordinate (x,y) is the maximum over all slice masks, $\max_z(M_z(x,y))$.

Based on the height-map we defined the cell volume $V = \sum_{z=1}^n M_z(x,y)$ as the sum of volumes of all pixels within the height-map, where for each pixel $p \in M_z(x,y)$, its volume is $sv(p) = (1 \div 9.75 \mu m)^2 \times 0.3 \mu m$, where $9.75 \mu m$ is a size coefficient between pixel in μm , and $0.3 \mu m$ is the height of the slice (Specific constants are dependent on the microscope and camera settings).

Protein intensities

Protein signal was computed for each immunofluorescence (IF) image as the sum of intensities across all z-slices and denoted as $I(x,y)$.

mRNA spot detection

To detect transcript positions from FISH data we used the ICY spot detector (Olivo-Marin, 2002). For images having $\max(z) \leq 12$ we used the following parameters: 2D wavelets and sensitivity 70 at pixel-scale 2; otherwise the parameters were set to: 1 pixel and 2 pixel length-scales with sensitivity 80. These parameters have been adapted for the detection of transcripts in the PRRC2C analysis, the sensitivity has been set to 50 at pixel-scale 2 when the images have $\max(z) \leq 12$; otherwise we used the 2D wavelets and a sensitivity of 70 at the pixel-scale 2.

For cultured cells, as well as for CytoD micropatterned cell series, we have applied a custom-developed spot detection script (these images present a very high noise content preventing efficient use of ICY). First, we applied a background noise subtraction by using Sobel and Gaussian filters, successively. Second, we applied the white top-hat filter in order to enhance bright objects of interest (potential mRNA spots) on a dark background. Finally, we used the Laplacian of Gaussians filter for mRNA spot detection. Furthermore, for the 3D analyses spots detected below the zero level were eliminated. Cells with less than 10 spots and 100 for the PRRC2C data (due to acquisition noise in cytoplasm) were eliminated.

Secondary image descriptors

Based on the primary image descriptors we computed secondary descriptors that corresponded to per image statistics.

Cytoplasmic total counts

Let us denote M the set of all mRNA spots for a given FISH image, $|M| = N$. The cytoplasmic total mRNA descriptor was calculated as the number of transcripts within $M^{cytoplasm}$, that is $T_{mRNA} = |\{m \in M \mid M^{cytoplasm}(x,y) = 1\}|$. The cytoplasmic total IF intensity is the summed IF intensity across the $M^{cytoplasm}$ region for protein images: $T_{IF} = \sum \{I(x,y) \mid M^{cytoplasm}(x,y) = 1\}$.

Peripheral distance map

For a given image, the peripheral distance map corresponds to a collection of peripheral masks based on $M^{cytoplasm}(x,y)$, where the width of the periphery varies as a proportion of the cytoplasmic radial distance. We segmented $M^{cytoplasm}(x,y)$ into 100 isolines from the nucleus contour to the periphery by projecting a ray from the nucleus centroid to the cell border, which was then segmented in 100 equidistant points. The 100 isolines were then built by constructing polygons that connect 360 points (one ray per degree). These isolines define a symbolic distance map D , where $D(x,y)$ is the isobar value for (x,y) corresponding to the "distance" from the nucleus envelope, with 100 at the nucleus and 0 at the cell edge. Given a fixed percent p between 0 and 100, the mask $M^{periphery}(x,y,p)$ is 1 for $D(x,y) < p$ and 0 otherwise. Hence, the periphery mask for a given p contains a strip at the cell edge whose width is a fixed-proportion of the radial distance. Thus defined peripheral distance map allows to compute in 3D peripheral volumes $V^{periphery} = (V^p)$ at each isoline and consequently distances from any voxel to either cell periphery or nucleus wall in 3D.

Cell quantization

In order to compute localisation statistics over multiple micropatterned images, compatibility of these images is required. We have chosen the MTOC position to be the reference point for the 2D cell geometry.

2D quantization: quadrants and per quadrant statistics. As shown on the schematic in Figures 4A and 4B, given a cell mask $M^{cell}(x,y)$, we generate the tessellation of the image by centering two orthogonal axes at the nucleus centroid and rotating them over 360 degrees, each position of these axes defining a partition of the cell mask into four quadrants, one of them containing the MTOC, Q_M . We retained the orientation that maximizes the mRNA count within a quadrant containing the MTOC, that is $max_d(T_{mRNA} = |\{m \in M \mid M^{cytoplasm}(x,y) = 1 \wedge Q_M(x,y) = 1\}|)$, $d \in [0, 359]$. The resulting four quadrants Q_1, Q_2, Q_3, Q_4 are numbered so that Q_1 always corresponds to Q_M and the remaining three quadrants are numbered in a clockwise fashion.

For protein intensities, quadrants are defined in a similar fashion using T_{IF} . Definition of cell mask partitioning in quadrants q_1, q_2, q_3, q_4 enables cell's quantization in 2D in terms of per quadrant statistics of mRNA and protein signal. Quadrants' respective areas are denoted by a_1, a_2, a_3, a_4 . We denoted by t_i the total number of mRNA spots falling in q_i in the case of FISH data, or the summed intensity across q_i in the case of IF data.

Then the local mRNA density was computed as the relative concentration c_i of mRNA in quadrant i and is defined to be $c_i = \frac{t_i/a_i}{T_{mRNA}/A}$, where A is the cell mask area. In the case of protein signal we replaced T_{mRNA} by T_{IF} .

Fine-grained quantization. In the same fashion as for the peripheral distance map we defined an additional subdivision of the cellular mask in isolines, their number being defined by the percent p . Given the previously defined quadrants, we further subdivided each of them in 2, yielding the tessellation in 8 parts that divide the circle in 45 degree sectors. Using the isolines and the 8 sectors we quantized the cell masks into $8 \times p$ segments organized in a concentric fashion starting from the nucleus towards the cell periphery (see the schematic in Figure 5A). Quantization for thus obtained segments was computed in the same fashion as for quadrants, resulting in a $8 \times p$ vectors of per segment signal concentration statistics for each cell, that we denoted $C = (c_i)$.

3D quantization: 3D quadrants and per quadrant statistics. Cell mask's tessellation into quadrants as defined in II.3.A (axes position) is projected onto each z-slice, thus yielding the cell's partition into four 3D quadrants Q_1, Q_2, Q_3, Q_4 , their respective volumes being denoted by v_1, v_2, v_3, v_4 . The volume of each quadrant is calculated as the sum of volumes of pixels within it using the same coefficients as for the cell volume.

We denoted by t_i the total number of mRNA spots falling in Q_i in the case of FISH data, or the summed intensity across Q_i in the case of IF data.

Then the relative concentration c_i of mRNA in quadrant Q_i is defined to be $c_i = \frac{t_i/v_i}{T_{mRNA}/V}$. In the case of protein signal T_{mRNA} is replaced by T_{IF} .

In the same fashion as for 2D, we computed fine-grained quantization statistics in 3D. Specifically, signal densities (mRNA or protein) were computed for 8 quadrants and each $1/3$ of isolines, that is for 8×3 volumic segments, yielding signal concentration per volumic region vectors for each cell, that we denoted $C = (c_i)$. Notice that these vectors can be aligned between cells since all the quadrants are anchored by the MTOC position.

Statistical analysis

Primary and secondary image descriptors were used to compute statistics for image acquisition series and to compare them.

Peripheral fraction and enrichment

Based on the cytoplasm masks, we calculated the peripheral fraction of mRNA and proteins at a given percent p of the radial distance between the cell periphery and the cytoplasm. First, the peripheral density is defined as the ratio of the transcript counts (respectively, summed IF intensities) across the $M^{periphery}(x,y,p)$ and volumes of $M^{cytoplasm}$ regions. First thanks to the distance map D we define a vector $C = (c_1, \dots, c_{100})$ containing the counts of mRNAs (resp., summed intensities) at distances smaller or equal to $i \in [0, \dots, 100]$ from the cell edge. This count vector C is divided by values of corresponding peripheral volumes $V^{periphery}$ to obtain spot (or protein) densities in concentric regions from the periphery to each isoline (c_i/V^i). These local peripheral densities are further normalized by the total density of the cytoplasm $d_{cytoplasm} = N/V_{cytoplasm}$, resulting in a vector of relative densities for all regions from the periphery

$F = \left(\frac{c_i/V^i}{d_{cytoplasm}} \right)$. The mRNA (resp., protein) fractions for each gene and for each isobar were defined as vector of median relative densities over all FISH images for this gene over all time points.

Volume corrected noise measure

We have measured the cell-to-cell expression variability in mRNA levels that cannot be accounted for by cell-to-cell differences in volume by computing the volume corrected noise measure Nm for micropatterned and standardly cultured cells. It was calculated following the approach of a previous study (Padovan-Merhar et al., 2015):

$$Nm = \left(\frac{\sigma_N}{E(N)} \right)^2 - \left(\frac{b \times E(V)}{a + b \times E(V)} \right) \left(\frac{Cov(m, V)}{E(m) E(V)} \right)$$

where N is the total mRNA count, V is the cell volume, a, b are the offset and slope of the least-squares best-fit linear regression of $E(N)$ on V , and σ, E and Cov are the notations for standard deviation, expectation and covariance, respectively.

Cytoplasmic spread

Cytoplasmic spread is measured by two statistical parameters that estimate how evenly a molecule is spread across the cell: cytoplasmic centrality and cytoplasmic entropy. The *centrality* statistics for mRNAs corresponds to the average distance from the nucleus envelope of cytoplasmic mRNAs, normalised by the maximal distance from the nucleus (corresponding to the number of isolines used to compute the distance map D). For the protein intensities, computation is the same but for peaks of intensity that are above the average. The higher the value, the closer to the periphery of the cell is the signal (mRNA or protein) distribution, value of 1 corresponding to 100% of the signal being at the periphery. The uniformity of the spread of molecule distributions in the cytoplasm is measured by computing Kozachenko-Leonenko entropy estimates H of a spatial random variable in 2D or in 3D based on the k th-nearest neighbour distances between mRNA (intensity peaks, respectively) coordinates following [Kozachenko and Leonenko, 1987]. Entropy is a unitless measure, to make the results comparable all the values are normalised by the $\max(H)$ across all cells.

MTOC polarity index

We defined a polarity index $PI_M \in [-1, 1]$, termed the *MTOC polarity index*, that measures the enrichment of mRNA or protein signal for a given image acquisition series in the vicinity of the MTOC location.

For the set S of images from an acquisition series under study, we denoted by $S_M = \{S_i\}$ and $S_{-M} = \{S_j\}$ the sets of all MTOC containing quadrants and quadrants that do not contain the MTOC, respectively. Intuitively, the MTOC polarity index measures how frequently the relative concentration within the MTOC quadrants is higher than in the non-MTOC quadrants. Formally it is defined as follows:

$$PI_M = \frac{2 |\{S_i \mid c_i > m\}|}{|S|} - 1,$$

where m is the median of relative signal concentrations c_j for quadrants in S_{-M} .

Positive values of PI_M imply MTOC correlated enrichment of RNA transcripts or proteins, negative values imply enrichment away from the MTOC and a value of zero implies no detectable enrichment.

Statistical relevance of PI_M is measured using the null hypothesis that $\forall i, c_i = m$, which corresponds to the complete spatial randomness. Under this hypothesis the population value of PI_M is 0. However, we have shown in (Warrell et al., 2016) that the empirical distribution of PI_M follows the binomial distribution asymptotically. Thus, the binomial test was used to evaluate the statistical relevance of PI_M for a given set of images.

mRNA / protein distribution profile

In order to define a spatial distribution profile of mRNAs and proteins for images acquired at a given time point, we used the fine-grained quantization of the cells (see paragraphs IV.4.B and IV.4.C). A single vector was computed at each time point by averaging across the pool of acquired images, hence estimating its expected value at that time point. Recall, that for each cell we computed a vector $C = (c_i)$ of per segment signal (mRNA or protein) concentration statistics and that these vectors are comparable. Then for a given time point we computed a median spatial profile $\bar{C} = (\bar{c}_i)$ representative of this time point by averaging per segment all C_i for this acquisition series.

Colocalization score

The goal of this analysis is to measure the interdependence between the mRNA and protein spatial distributions. To do this, we defined the *Colocalization Score* (CS) as an effect size computed based on correlations between mRNA and protein spatial distributions for image acquisitions that we applied for several sets of time points: those where protein acquisition time is greater than mRNA acquisition time and the others. Colocalization Score for a given mRNA-protein pair is calculated based on mRNA and protein distribution descriptor vectors \bar{C}_{mRNA} and \bar{C}_P and takes values between 0 and 1.

Notice that CS can be calculated for any measure of correlation between mRNA and protein distributions, which allowed us to examine the interdependence of a molecule's dynamics within specifically defined subcellular regions and at different time points.

More formally, we supposed a 2-measure discrete time process ϕ , containing observations (ϕ_{t_1}) and (ψ_{t_2}) at time points $T_1 = \{t_1\}$ and $T_2 = \{t_2\}$, respectively. Then, the (empirical) Colocalization Score is defined for pairs of data $S = \{(\phi_{t_1}) \times (\psi_{t_2})\}$ at time points T_1 and T_2 using a similarity function γ and the Vargha and Delaney's A_{12} statistics [Vargha and Delaney, 2000], a measure of stochastic superiority when comparing two distributions (generalized effect size):

$$A_{12} = P(\alpha > \beta) + 0.5P(\alpha = \beta)$$

where α and β are the rank-sums $\alpha = \sum_{t_1, t_2} r_\gamma((\phi_{t_1}, \psi_{t_2}))$, $t_1 < t_2$ and $\beta = \sum_{t_1, t_2} r_\gamma((\phi_{t_1}, \psi_{t_2}))$, $t_1 > t_2$, respectively and where r_γ is the rank of the tuple where the order is given by the γ function. This γ function is the similarity between pairs of observations is computed as $\gamma'((t_1, t_2)) \rightarrow \gamma(\hat{E}[\sigma(\phi_{t_1})], \hat{E}[\sigma(\psi_{t_2})])$. The 'forward-leading' set S_1 is defined as $S_1 = \{(t_1, t_2) \in T_1 \times T_2 \mid t_1 < t_2\}$, and its complement S_2 contains all pairs of time-points such that $t_1 \geq t_2$. Thus, A_{12} can be understood as the effect size between the rank-sum of the similarities $\gamma'(t_1, t_2)$ across all 'forward-leading' time point pairs versus the rank-sum of the similarities $\gamma'(t_1, t_2)$ across all the other time-points. Notice that A_{12} statistics lies in the $[0, 1]$ interval.

In practice, for the analysis of mRNA / protein colocalizations, we considered computing the CS for a 2-measure discrete time processes ϕ in which ϕ is a point process and ψ a general random measure (representing mRNA locations and protein concentrations

respectively). In the current study we used relative density distributions of mRNA at $T_1 = \{2, 3, 4, 5\}$ and proteins at $T_2 = \{2, 3, 5, 7\}$ (the discrete time points representing time in hours), and γ is the Pearson Correlation Coefficient between two relative density distribution vectors (see “mRNA/protein distribution profile”). We note that these distributions can be computed based on a particular quantization of cells and can cover the whole cell (forming a global CS) or a some section of the cell (e.g. the periphery).

Notice that measuring a correlation between a local density of some mRNA in a given voxel and its protein counterpart at a later time point is dependent on the choice of timepoints and how dynamic the molecules are. To alleviate this issue, instead of considering isolated voxels, we consider their direct neighborhoods.

Moreover, from the rank-sum of the similarities $\gamma'(t_1, t_2)$ across S_1 defined previously, an exact permutation test can be derived to calculate significance levels for a given value of A_{12} and a steady-state null hypothesis.

Degree of clustering (Ripley-K)

The degree of clustering statistic has been previously introduced based on the framework of point processes by (Lee et al., 2013). It is a unitless measure that can be used to compare clustering between different molecules and conditions. In (Warrell et al., 2016) we generalized this definition to the framework of continuous random measures, which allows us to calculate the degree of clustering for both FISH and IF data, the former being modelled as point processes, and the latter modelled as a continuous-valued random measure. Our generalized algorithm for calculating the degree of clustering is summarized below. For theoretical considerations please see (Warrell et al., 2016).

A classical tool for the point process analysis is the Ripley’s K function defined as the mean number of events that occurred inside a ball of radius r around a randomly selected event normalized by λ , the number of events per unit area (Ripley, B. D. 1977). A classical estimator of the Ripley’s K function can be defined as in (Chiu et al., 2013; Ripley 1977):

$$\hat{K}(r) = \frac{1}{\lambda^2 \nu(w)} \sum_{i=1}^n N_i(r)$$

where N_i is the number of event points (mRNA transcripts) in a ball of radius r centered on i , λ is the density, ν is the volume or area (in 3D and 2D, respectively) of the observed region w , and n is the number of points.

We normalized \hat{K} under a homogeneous Poisson process, which is commonly known as Ripley’s H function $\hat{H}(r) = \sqrt[3]{(3\hat{K}(r))/(4\pi) - r}$, $d \in \{2, 3\}$ where d is equal to 3 in the case of volume-based computation and 2 in the case of 2D.

This in turn makes it possible to define the clustering index H^* as an estimator of $\hat{H}(r)$ by comparing the Ripley’s H function calculated empirically to its distribution under complete spatial randomness (CSR):

$$\hat{H}^*(r) = \begin{cases} \hat{H}(r)/\hat{H}_{95}(r) & \text{if } \hat{H}(r) \geq 0 \\ -\hat{H}(r)/\hat{H}_5(r) & \text{otherwise} \end{cases}$$

where $\hat{H}_{95}(r)$ and $\hat{H}_5(r)$ are the 95th and 5th percentiles respectively of $\hat{H}(r)$.

CSR is modeled using random permutations of actual data points (100 times in our study), which enabled us to compute the 95% and 5% confidence bounds of CSR. Spatial clustering is considered to be significant at radius r if the computed $\hat{K}(r)$ is over the upper (95%) or lower (5%) bounds of the random distribution.

In (Warrell et al., 2016) we have introduced a convolution-based \hat{H}^* estimator based on the random permutation-test. This estimator normalizes $\hat{H}(r)$ so that $|\hat{H}^*(r)| > 1$ only when $\hat{H}(r)$ falls outside the 95% confidence interval for a homogeneous Poisson Process. Moreover, we have shown that our permutation test using the convolution-based estimator reduced to the clustering index estimator used by (Lee et al., 2013) for the point process case. This enabled the implementation of a common consistent computational framework for both point and continuous processes. The degree of clustering $\hat{\delta}(r)$ is then defined as the area of \hat{H}^* above 1, that is $\hat{\delta}(r) = \int_{x \in (0,r)} \max(\hat{H}(x) - 1, 0) dx$.

Within this unified framework we can use the same computational approach for mRNA data as in (Lee et al., 2013) to compute the degree of clustering. Below we define the specific procedure for its computation in the case of protein data.

Cells are quantized into voxels V_1, \dots, V_n where each voxel t the value of the observed quantity $\varphi(V_1), \dots, \varphi(V_n)$ in the case of 3D analysis (or into pixels in the case of 2D). We denoted by I an array in which each element corresponded to the intensity value for each voxel.

The convolution-based estimator \hat{K}_c can be computed using the following formula: $\hat{K}_c(r) = \frac{1}{\lambda^2 V} \int [|\cdot| \leq r] (I * I')(x) dx - \frac{1}{\lambda}$ where λ is an estimate of average intensity per unit volume, $[\cdot]$ is the indicator function that is 1 for a true statement and 0 otherwise, $I'(x) = I(-x)$, $*$ is the convolution operator, and V is the volume of the window over which the cell is observed. In practice, given the fact that the cell thickness is quite low, we can approximate the 3D convolution by a 2D convolution.

Thus we have a common computational framework to evaluate the presence or absence of clustering for both mRNA and protein data.

Additional methods for muscle data analysis

In this section we report adaptations of the methods presented in “primary image descriptors,” “secondary image descriptors,” and “additional methods for muscle data analysis” to the case of muscle cells (see the table describing the Image acquisition characteristics for muscle cells).

Cell and nucleus masks, nucleus centroid

Cell and nucleus masks were computed for all muscle images, using γ -Tubulin and DAPI signals, respectively, using the same general principles as in “peripheral distance map.” As these acquisition series benefit from a better segmentation, only Otsu threshold method was necessary to obtain the binarized images.

After these steps, we obtained a contiguous contour with small white spots due to noise in the images. We applied mathematical morphological operators such as dilatation and closing to get a full mask of the muscle cell $M^{cell}(x,y)$. The nucleus mask $M^{nucleus}(x,y)$ detection followed the exact same steps and parameters as for the micropatterned images - the nucleus centroid was computed as the geometric center of the nucleus mask.

mRNA signal detection

mRNA spots detection was done using ICY spot detector (Olivo-Marín 2002) to find transcript positions, the parameters were set to: 1 pixel and 2 pixel length-scales with a fixed sensitivity of 80.

Z-lines' masks

The main component of Z-lines is the Alpha actinin protein. To facilitate the analysis, we have defined an additional secondary descriptor computed from the Phalloidin signal, called Z-lines mask $M^{Z-lines}(x,y)$.

For each z-slice of each image we performed the contour detection for the Z-lines. First, we applied a vertical Sobel operator, which detected the vertical edges of an image, followed by a Gaussian kernel to smooth artifacts of the Sobel filtering and reinforce the Z-line signal. An Otsu binarization was then processed. As a result we obtained a set of Z-lines masks $M^{Z-lines} = \{M_z^{Z-lines}(x,y)\}$ (Figure 7).

For further analysis we restricted the cell to z-slices containing more than 25 mRNA spots (to avoid false positives due to high noise). Notice that the spots falling in the eliminated slices were also excluded from the analysis.

We defined an additional descriptor called Z-line spacing Z that represented the median spacing between two lines. For each z-slice at each y coordinate we computed all the distances $d((x_i,y), (x_j,y))$ where x_i and x_j were 2 consecutive Z-lines contours. Z was defined as the median of all d for all acquired cells. For our data $Z = 15$ pixels.

Z-line mRNA distance profile

In order to evaluate the mRNA clustering in the vicinity of Z-lines, we computed the average 2D Euclidean distance of mRNA spots to their nearest Z-line for mature and immature cells.

Using the $M_z^{Z-lines}(x,y)$ masks, we computed for each mRNA spot $m \in M$ positioned at (x,y,z_m) the minimal 2D Euclidean distance to a Z-line falling within a disk of radius Z . This computation was performed within the $M_z^{Z-lines}(x,y)$ z-mask such that $z = z_m$. If m fall within $M_z^{Z-lines}(x,y)$, then the minimal distance was set to 0. Thus for each image we obtained a $D = \{d\}$, $|D| = N$ the set of all minimal distances between mRNA and Z-lines. In turn this allowed us to define for each image a count vector $\delta = (\delta_i)$ of size Z where δ_i is the number of mRNA spots at each distance $d = i$, $d \leq Z$ normalized by N .

For a given image acquisition we defined its Z-line mRNA distance profile to be $\bar{\delta} = (\bar{\delta}_i)$, where $\bar{\delta}_i$ is the median of all δ_i (Figure 7B).

Cell quantization

Given that muscle cells contained more than one nucleus, each cell mask was restricted to be between two consecutive nuclei centroids as shown on the schematic in Figure 7C). Given a cell mask $M^{cell}(x,y)$, definition of cell mask tessellation in n vertical segments $q_1 \dots q_n$ enables cells quantization in 2D in terms of per segment statistics of mRNA concentrations. Quantization was performed with $n = 20$ and $n = 80$, see results in Figure 8C.

mRNA spatial distribution

To estimate mRNA clustering along muscle cells, we computed local mRNA density for each cell using the cell quantization introduced in “cell quantization.” We denote by t_i the total number of mRNA spots falling in a given q_i . Then the local mRNA density is computed in the same way as for fibroblast cells (see “peripheral fraction and enrichment”) as the relative concentration c_i of mRNA in q_i and is defined to be $c_i = \frac{t_i/a_i}{T_{mRNA}/A}$, where A is the cell mask area. We produced distribution plots and heatmaps representing mRNA local density between two nuclei (Figure 7C).

Additional methods for CHX inhibition analysis

Image processing has been performed for Arhgdia (Supp. control), Arhgdia (CHX), Pard3 (Supp. Control), Pard3 (CHX) (see table describing image acquisition characteristics for CHX) as presented above for primary and secondary image descriptors except for the FISH signal being used to detect cell masks for mRNA images. The number of acquired mRNA images was quite low for these experiments, consequently we have only performed global analysis (compared to the datasets where we could analyse time point by time point). We performed peripheral fraction enrichment and degree of clustering analyses in the same way as described in “peripheral fraction and enrichment” and “degree of clustering (Ripley-K),” respectively, except for 2 points: (i) all the analysis was performed in 2D, (ii) for mRNA the computation was done using the continuous signal in the same way as for the IF, similarly to Stueland et al., 2019.

DypFISH software code (in Python) as well as data (in HDF5 format) used in this study are freely available from <http://www.dypfish.org> and <https://github.com/cbib/Dypfish>.

Chapitre 5

Discussion et perspectives

5.1 Synthèse

Le développement de méthodes de microscopie à haut-débit joue un rôle important pour la recherche en biologie, dont peut dépendre notre capacité à explorer en détail les processus cellulaires. Dans cette thèse, nous nous sommes intéressés au développement de nouvelles méthodes de traitement d'image utilisant le Deep Learning permettant d'améliorer le signal des images de microscopie et de leur quantification.

Nous avons présenté deux approches basées sur l'apprentissage profond, supervisé pour le chapitre 2, non supervisé pour le chapitre 3, ainsi qu'un cadre pour étudier la distribution spatiale des molécules dans les cellules au chapitre 4.

Rehaussement du signal pour les spots d'ARNm

Plus spécifiquement, dans le chapitre 2 nous avons présenté DeepSpot, une méthode computationnelle utilisant le Deep Learning pour rehausser le signal des spots correspondant aux ARNm dans les images de type smFISH. Avec DeepSpot, nous avons montré que rajouter une étape de rehaussement du signal dans les pipelines d'analyse d'images, préalablement à la détection des spots, présente deux avantages majeurs. Le premier est que cela permet d'améliorer la précision des algorithmes de détection et de ce fait, d'avoir une quantification plus précise. Le second avantage est que cette étape permet d'automatiser la détection de spots, même lorsqu'elle est effectuée par des algorithmes qui ne sont pas nativement non paramétriques, tels que le très largement utilisé Laplacien de Gaussien. Nous avons introduit une architecture utilisant des convolutions à trous, des réseaux à chemins multiples ainsi que les réseaux de neurones résiduels pour le traitement des images smFISH et montré que cela permet un traitement efficace des petits objets que sont les spots d'ARNm, par rapport à d'autres méthodes d'augmentation du signal qui ne sont pas adaptées pour ce type d'objets [188]. Nous avons également montré l'importance qui doit être accordée à la construction du jeu de données d'entraînement. L'annotation

manuelle des images smFISH est une tâche chronophage et comportant un certain degré d’incertitude sur la précision d’annotation, dépendant à la fois de la qualité des images et de l’interprétation du signal par l’annotateur. Pour cette raison, plusieurs méthodes récentes portant sur la détection de spots dans les images smFISH ont intégré des images issues de simulation smFISH [146, 72] dans leurs analyses. Dans cette étude, nous avons également ajouté de la variabilité dans notre jeu de données en simulant des données smFISH. Nous avons montré que la combinaison de données smFISH expérimentales et simulées a permis d’obtenir un modèle plus précis et plus robuste que les modèles entraînés uniquement sur des données expérimentales ou uniquement simulées. Et nous avons montré que le rehaussement du signal des spots suivi de sa détection par une méthode classique permettait d’obtenir globalement de meilleurs résultats qu’avec une méthode de détection entièrement basée sur le Deep Learning, telle que deepBlink.

DeepSpot étant une méthode basée sur l’apprentissage, nous supposons que la méthode pourrait être appliquée à d’autres types d’images de spots que celles issues de microscopie smFISH moyennant un entraînement sur ce type de données.

Rehaussement du signal pour les noyaux cellulaires

Dans le chapitre 3 nous avons proposé une méthode d’apprentissage non supervisée, SaliencyNet, utilisant le transfert de style pour rehausser le signal des noyaux présentant une faible intensité dans des images d’organoïdes. Dans ce chapitre également, nous avons pu montrer que les étapes de rehaussement du signal, en amont de la segmentation, permettent, soit d’automatiser cette dernière, même avec des méthodes non automatiques (Otsu), soit d’améliorer les résultats des méthodes automatiques (StarDist).

Une autre approche, CARE [188], a été amplement utilisée pour améliorer le rapport signal sur bruit dans de nombreux types d’images, de manière supervisée, sur des données appariées ou chaque image acquise en faible intensité doit avoir sa contrepartie acquise en haute intensité dans les données d’apprentissage, nécessitant l’acquisition de paires d’images. Avec SaliencyNet, nous montrons que les méthodes non supervisées peuvent également remplir ce rôle. L’emploi du CycleGAN permet, par rapport à CARE, de ne pas avoir à imager deux fois le même échantillon (une fois en basse intensité et une fois en haute intensité), nécessitant donc moins de temps d’acquisition, ce qui est particulièrement important dans le cas de l’imagerie *live-cell*.

Nous avons montré que la disponibilité de données d’apprentissage publiques avec une saillance suffisante permettait de rendre possible l’emploi d’une solution d’apprentissage non supervisée. Nous avons également montré qu’apporter des modifications dans la fonction de coût du CycleGAN classique permettait de gagner en précision.

Enfin, nous pensons que SaliencyNet peut s'appliquer à d'autres types d'images que celles des noyaux, par exemple à des images de cellules afin de faciliter la segmentation de celles-ci. La méthode étant générique, il est également envisageable de l'appliquer aux images de noyaux acquises avec d'autres techniques que l'imagerie soSPIM. Bien entendu, une évaluation rigoureuse de la performance restera à faire.

Application d'analyses spatiales

Dans le chapitre 4, nous avons proposé un cadre unifié pour étudier la distribution spatiale de molécules dans les cellules. Nous avons utilisé les résultats des chapitres précédents pour calculer des statistiques spatiales, dans un premier temps en définissant un cadre de lecture pour chaque cellule analysée, puis en définissant des repères cellulaires d'intérêt permettant d'aligner les cadres de lecture et donc d'appliquer des statistiques spatiales sur des ensembles de cellules.

Nous avons développé une librairie généraliste d'analyses quantitatives de distributions spatiales des molécules, DypFISH, permettant de quantifier la localisation subcellulaire des ARNm et des protéines et ainsi permettre de caractériser les différents états et sous-populations cellulaires. Nous avons introduit des techniques d'analyses qui permettent l'évaluation conjointe des données ponctuelles d'ARNm dans les images smFISH et des données continues de protéines par immunofluorescence (IF) de manière statistiquement robuste. À partir de ces deux types de données, DypFISH permet d'analyser des motifs de distribution de molécules et d'étudier la corrélation de la localisation des molécules entre elles, ou avec des repères cellulaires choisis.

5.2 Limites

Bien que le succès des méthodes utilisant du Deep Learning soit indéniable sur le plan scientifique et pratique de l'analyse d'images de microscopie de fluorescence, il est important de garder à l'esprit un certain nombre de limites inhérentes à la technique.

Premièrement, ces méthodes ne sont actuellement pas capables de remplacer tout le panel de méthodes classiques existantes. Le principal atout du Deep Learning est avant tout sa capacité à automatiser une tâche bien spécifique et à la généraliser à des nouvelles données. Si des méthodes telles que StarDist, CARE, DECODE ou CellPose, ont réussi à obtenir une très bonne capacité de généralisation pour une tâche spécifique, toutes ces méthodes ont cependant un ou plusieurs angles morts pour lesquels elles ne sont pas adaptées ou simplement pas entraînées. Il en va de même pour toutes les méthodes basées sur l'apprentissage. Nous avons montré par exemple dans le chapitre 2 que DeepSpot généralise bien sûr des images smFISH dont le signal provient de 27 gènes dont les acquisitions

donnent des rapports signal sur bruit très différents, ce qui n'est pas le cas pour deep-Blink. Nous avons également montré que DeepSpot généralise raisonnablement sur des images ayant des caractéristiques différentes que celles du jeu d'apprentissage. De ce fait, un bon modèle de DeepSpot devrait être capable de généraliser sur d'autres images de type smFISH. Cependant, il serait intéressant de pouvoir définir les limites exactes d'applications d'un tel modèle. Le même raisonnement s'applique bien entendu à SaliencyNet, dont les limites exactes des capacités de rehaussement du signal devraient être explorées davantage.

Deuxièmement, et de façon liée à la précédente, intervient la limite posée par la disponibilité de données d'entraînement. À ce sujet, plusieurs points cruciaux peuvent être abordés.

1. **La quantité des données** disponibles pour un apprentissage est absolument décisive. Derrières les algorithmes de classification d'objets qui arrivent aujourd'hui à généraliser sur des milliers de classes d'objets, on retrouve des jeux de données comportant des millions d'images, annotées par des milliers d'utilisateurs. Bien que des initiatives telles que Omero, la Broad Bioimage Benchmark Collection, les jeux de données du Allen Institute, The BioImage Archive (BIA), existent, il reste difficile de construire des jeux de données dépassant les quelques milliers d'images, ce qui limite énormément les possibilités en terme d'apprentissage.
2. **La qualité des données** disponibles peut également fortement impacter l'apprentissage et la prédiction d'un réseau de neurones. La variabilité expérimentale intrinsèque à la recherche sur des échantillons vivants implique parfois l'acquisition de jeux de données dont la qualité peut questionner sur la possibilité de les utiliser pour des analyses statistiques. C'est ici un sujet sensible, néanmoins la question d'utiliser des données sur lesquelles même un œil humain, expert, ne peut pas toujours discerner avec certitude le signal du bruit où des artefacts, devrait être posée systématiquement.
3. **La qualité des annotations** soulève deux questions principales en fonction de si l'on souhaite entraîner un modèle ou bien le tester. Si le contenu des images de microscopie de fluorescence est relativement simple (ex. spots, noyaux, cellules) par rapport à des images de la vie courante pouvant contenir plusieurs dizaines d'objets différents, la difficulté de l'annotation réside cependant dans la nécessité d'un œil "expert" et dans la qualité des images et des objets. De forts taux de bruits peuvent rendre des objets peu visibles et leurs contours flous. Partant de ce constat, pour un seul et même objet (ex. un noyau), annoté par deux experts, les labels résultants comporteront nécessairement des différences, dans le choix du tracé du

contour, en fonction des préférences personnelles de l'annotateur. Par ailleurs, un seul objet annoté deux fois de suite par la même personne présentera également des différences. Bien qu'il existe des méthodes de consensus ou des heuristiques pour réduire les erreurs d'annotations [178], en microscopie de fluorescence il est de toute façon rare de disposer de plusieurs labels à comparer. D'autres erreurs d'annotation peuvent apparaître comme des annotations incomplètes, par exemple lorsque seulement une partie des objets est annotée. La première question que l'on peut se poser est alors d'estimer ce que les modèles de Deep Learning vont pouvoir corriger par eux-mêmes et ce qu'ils vont intégrer comme erreur dans leur apprentissage [168] et donc quels biais vont être appris. La seconde question concerne la prédiction et plus spécifiquement la précision de celle-ci. Pour évaluer la performance d'un réseau de neurones, par exemple pour une tâche de segmentation, il est possible d'évaluer un certain nombre de métriques (précision, rappel, sensibilité, spécificité, IoU, F1-score, indice de Dice, etc.). La valeur d'une telle métrique dépend donc de la performance du réseau, mais également de l'exactitude de l'annotation. Si, par exemple, l'annotation d'un jeu de données contient 10% de pixels qui se sont vu attribuer le mauvais label au moment de l'annotation, les métriques de performance en seront impactées d'autant. Bien que cela n'empêche pas la comparaison relative entre deux modèles, il n'est alors pas possible de se faire une idée exacte de la performance d'un modèle et de savoir à qui attribuer l'erreur, au modèle ou à l'annotation. Certains réseaux de neurones, comme le U-Net, sont connus pour avoir une certaine propension à combler de petites lacunes dans l'annotation, là encore, il est difficile d'estimer si cette correction correspond bien à la réalité.

Une limite technique s'applique également à l'utilisation du Deep Learning, à savoir l'accès à des ressources de calcul suffisantes. Cependant, cette limite tend à se résoudre car beaucoup d'efforts ont été faits pour rendre plus accessible l'utilisation du Deep Learning, d'une part par rapport aux capacités de calcul, d'autre part par rapport aux compétences nécessaires à son emploi. Les initiatives telles que DeepCell, StarDist, CellPose, qui proposent toutes une interface graphique en ligne pour la segmentation de cellules et de noyaux, de même que ZeroCostDL4Mic qui permet d'utiliser dans le cloud un grand nombre de modèles de Deep Learning dédiés à la microscopie, DeepImageJ permettant d'utiliser du Deep Learning dans ImageJ/Fiji et plus récemment l'arrivée des plugins Napari qui peuvent embarquer nativement les réseaux de neurones développés en Python, toutes ces initiatives tendent à démocratiser l'utilisation du Deep Learning par un public de biologistes ou de bioimage analysts qui ne sont pas nécessairement spécialistes en programmation.

Enfin, concernant la librairie d'analyses spatiales DypFISH, une limite notable est que les grilles de lecture spatiales actuellement prédéfinies sont limitées et que l'analyse de nouvelles typologies cellulaires demanderait potentiellement la définition de nouvelles grilles de lectures. Ceci implique qu'un tel outil, pour durer dans le temps et pour être utilisé par la communauté scientifique, demande une forte implication du laboratoire pour le maintenir à jour et pour répondre aux attentes des potentiels utilisateurs, ce qui n'est pas toujours réaliste. Cette limite n'est pas spécifique à DypFISH et est le lot commun de tous les outils ayant été développé pour une recherche spécifique, la même interrogation se présente pour le maintien à disposition de la communauté, de versions de DeepSpot et SaliencyNet mises à jour.

5.3 Perspectives

Le Deep Learning est appliqué à l'automatisation de plus en plus de tâches pour les images de microscopie, telles que la reconstruction d'images, le débruitage, la déconvolution, la super-résolution ou encore la reconstruction de structures. Le nombre de méthodes disponibles augmente, signe du réel besoin d'automatisation de ces tâches de traitement et d'analyse d'images.

À l'heure actuelle et malgré les multiples initiatives de collection et de partage des données, il n'existe pas de jeux de données qui puissent servir de *Gold Standard*, à la manière d'un "microscopy ImageNet". Plusieurs choses peuvent expliquer ceci, d'une part la spécificité des modèles impliquerait d'avoir autant de jeux de données qu'il y a de tâches concernées, d'autre part cela impliquerait l'investissement de plusieurs dizaines de personnes pour annoter et établir ensuite des consensus. Néanmoins, une initiative telle que la compétition Kaggle Data Science Bowl 2018 a bien mis en évidence qu'il était possible de parvenir à établir des jeux de données s'approchant du *Gold Standard* grâce à l'ampleur de son utilisation sur une tâche bien précise, la segmentation de noyaux. Ce type de compétition, sous cette forme ou une autre, pourrait profiter à d'autres tâches, comme le débruitage, la reconstruction d'images, la super-résolution ou encore pour la détection de spots. Dans ce dernier cas, afin de mesurer la précision d'un modèle, il est courant d'utiliser des données simulées pour lesquelles l'on connaît avec exactitude la localisation des spots. Cependant, là encore, la diversité des simulateurs et des modalités de microscopie rend l'établissement d'un jeu de données de référence assez complexe et nécessiterait la collaboration à grande échelle de microscopistes, de bioanalystes et d'informaticiens. Nous pensons cependant que cela serait d'un grand bénéfice pour le développement de méthodes plus robustes.

Par ailleurs, définir des jeux de données standardisés permettraient également d'introduire un autre concept dans les algorithmes de détection ou de segmentation, celui de l'incertitude. En prenant l'exemple de la détection de spots dans les images smFISH, où certains spots peuvent être très brillants et d'autres très diffus et très flous (ex. focus/hors focus), il pourrait être intéressant de disposer d'un score de qualité à la manière de ceux utilisés en bioinformatique dans les fichiers fastq issus de séquençages, ou chaque base d'ADN séquencée se voit associer un score de qualité basé sur les probabilités d'erreurs de séquençage. Cela permet, par la suite, de supprimer une séquence ADN d'une analyse, si sa qualité est jugée trop faible. Si l'on disposait d'un score de qualité pour chaque localisation de spots à la sortie des algorithmes de détection, cela rendrait possible la décision d'inclure une image dans une analyse ou non en fonction de la qualité de la détection et, par exemple, de définir qu'une analyse ne porte que sur des images où la détection dépasse un indice de confiance donné. Ce genre de score pourrait s'obtenir par les méthodes de Deep Learning, il est en effet largement utilisé dans les tâches de classification (généralement avec une couche *softmax* en fin de réseau) et pourrait certainement être implémenté pour les tâches de détection et de segmentation en prenant en compte les caractéristiques des spots (brillance, contraste) et en utilisant la probabilité renvoyée par la couche *softmax* ou en utilisant des cartes de confiance dans la prédiction comme le fait DivNoising [132]. Ceci est une perspective qui reste bien entendu totalement à explorer.

Concernant le rehaussement du signal avec DeepSpot, un certain nombre de pistes peuvent être explorées pour rendre la méthode plus robuste. Actuellement, l'agrégation de contexte autour des spots ne se fait qu'à l'aide d'une seule taille de convolution à trous (3×3). Pour s'approcher d'une méthode de Deep Learning ressemblant au *Multi-scale wavelet detector* [124] présent dans le logiciel d'analyse d'image ICY, il serait intéressant d'explorer la possibilité d'utiliser plusieurs tailles de convolution à trous dans le réseau de convolution et potentiellement d'établir un consensus à partir de ces multiples tailles de convolution à trous. Cela permettrait peut-être une meilleure généralisation du modèle DeepSpot sur des tailles de spots non apprises avec le jeu de données, et le consensus permettrait potentiellement d'améliorer la capacité de DeepSpot à rehausser le signal des spots légèrement hors focus. Actuellement, la tâche apprise par DeepSpot consiste à rehausser le signal des spots par une gaussienne sur une surface de 3×3 pixels. Ceci est généralement suffisant pour des images smFISH mais si l'on souhaitait adapter la méthode à des images de SMLM, il faudrait étudier la possibilité de ne rehausser le signal que sur un seul pixel afin d'obtenir des images avec une super-résolution, la tâche se focalisant sur un objet plus petit pourrait potentiellement être plus dure à apprendre. La question se pose alors de savoir si l'architecture DeepSpot serait adaptée à une telle tâche.

Pour SaliencyNet et le rehaussement du signal des noyaux, nous nous sommes basés

sur l'architecture CycleGAN, qui à déjà fait l'objet de beaucoup d'études et d'approfondissements. Nous avons apporté des modifications sur la fonction de coût, mais il serait intéressant de tester le rehaussement du signal en utilisant des méthodes plus récentes de transfert de style telles que la Contrastive Unpaired Translation (CUT) qui permet des entraînements plus rapides, moins énergivores et qui nécessite moins d'images pour chaque style [130].

De manière plus générale, la quantité et la diversité des données et des outils disponibles laisse envisager l'emploi de méthodes transversales, intégrant des données hétérogènes, issues de l'imagerie et, par exemple, du séquençage, comme nous le présentons dans notre article présent dans l'Annexe C (Prédiction de la localisation subcellulaire des ARNm : Apprentissage à partir de sources de données hétérogènes). Une possibilité serait d'utiliser les réseaux de neurones *multi-input* pour construire des modèles intégrant des données hétérogènes dans un même espace latent à partir duquel pourront être extraites des prédictions [149].

Enfin, concernant le maintien à jour et à disposition de la communauté scientifique des outils développés durant cette thèse (DeepSpot, SaliencyNet, DypFISH) une solution envisageable serait le regroupement de plusieurs outils au sein d'un même *repository* global, ayant un support à plus long terme. C'est ce qui commence à être mis en place par exemple dans le HenriquesLab avec ZeroCostDL4Mic qui regroupe des modèles de réseaux de neurones issus de différents laboratoires, ainsi que du côté de CSBDeep, qui regroupe plusieurs outils tels que CARE, Noise2Noise et StarDist. Il serait utile de déposer DeepSpot et SaliencyNet auprès de ces ressources reconnues par la communauté. En revanche, concernant les outils d'analyses spatiales, plusieurs boîtes à outils et bibliothèques existent, mais il ne semble pas y avoir de recoupement entre elles. Une solution envisageable serait la convergence de la boîte à outils BIG-FISH avec la bibliothèque DypFISH. Étant toutes deux développées en Python, il serait possible d'intégrer les fonctions d'analyses spatiales de DypFISH, dans BIG-FISH dans le cas où elle n'y serait pas déjà, comme c'est le cas par exemple de la généralisation de Ripley K au signaux continus issus de l'immunofluorescence.

5.4 Conclusion générale

Au cours de cette thèse, nous nous sommes principalement attachés au rehaussement du signal dans les images de microscopie de fluorescence. Sur l'ensemble des méthodes de Deep Learning existantes, assez peu se concentrent sur ce sujet, la plupart étant orientées sur le débruitage ou sur les tâches de segmentation et de détection du signal. Nous pensons avoir montré que le rehaussement du signal a toute sa place dans le traitement d'images de microscopie et que des outils comme DeepSpot ou SaliencyNet pourraient être intégrés dans des pipelines d'analyse d'images, au même titre que les algorithmes de débruitage, pour obtenir des images dont le signal est facile à détecter ou segmenter. Au vu des différentes implémentations de réseaux de neurones disponibles en ligne, par exemple sur le hub de ZeroCostDL4Mic, il serait intéressant de pouvoir construire des pipelines entièrement basés sur le Deep Learning où les images sont envoyées à un premier réseau de rehaussement du signal (par exemple DeepSpot ou SaliencyNet) et dont la sortie serait directement reliée à l'entrée d'un réseau de détection ou de segmentation (par exemple deepBlink pour la détection de spots, ou StarDist pour la segmentation de noyaux).

Bien entendu, l'utilisation de tels pipelines reste à la discrétion de l'utilisateur. En effet, il n'est pas toujours utile de rehausser le signal dans les images de microscopie. Néanmoins, dans les cas d'usages que nous avons explorés (faible SNR, faible luminosité), cela faciliterait la tâche des biologistes et des bioanalystes. Cela donnerait, par ailleurs, davantage de visibilité aux tâches de prétraitement des images, conduisant à l'apparition de nouvelles données et donc de nouvelles méthodes et la possibilité de développer des modèles de plus en plus génériques et ainsi estomper la frontière entre le traitement et l'analyse d'images.

Le Deep Learning est porteur de nombreuses promesses pour la microscopie de fluorescence. À l'heure actuelle, les entrées et les sorties sont des images, mais des études permettent déjà d'extraire des informations sur les patterns de localisation des molécules en utilisant du Deep Learning, il se peut qu'à l'avenir des modèles soient capables de transformer les images directement en connaissances.

Annexe A

DeepMeta : un réseau de neurone pour la détection de métastases pulmonaires dans les images de poumons de souris



OPEN ACCESS

EDITED BY
Andrew Cohen,
Drexel University, United States

REVIEWED BY
Rachel Sparks,
King's College London, United Kingdom
Yong-Zi Chen,
Tianjin Medical University Cancer
Institute and Hospital, China

*CORRESPONDENCE
Edgar Lefevre,
edgar.lefevre@u-bordeaux.fr
Macha Nikolski,
macha.nikolski@u-bordeaux.fr

†These authors share last authorship

SPECIALTY SECTION
This article was submitted to
Computational Biomedicine,
a section of the journal
Frontiers in Bioinformatics

RECEIVED 21 July 2022
ACCEPTED 26 September 2022
PUBLISHED 12 October 2022

CITATION
Lefevre E, Bouilhol E, Chauvière A,
Souleyreau W, Derieppe M-A, Trotier AJ,
Miraux S, Bikfalvi A, Ribot EJ and
Nikolski M (2022), Deep learning model
for automatic segmentation of lungs
and pulmonary metastasis in small
animal MR images.
Front. Bioinform. 2:999700.
doi: 10.3389/fbinf.2022.999700

COPYRIGHT
© 2022 Lefevre, Bouilhol, Chauvière,
Souleyreau, Derieppe, Trotier, Miraux,
Bikfalvi, Ribot and Nikolski. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

Deep learning model for automatic segmentation of lungs and pulmonary metastasis in small animal MR images

Edgar Lefevre^{1*}, Emmanuel Bouilhol^{1,2}, Antoine Chauvière¹, Wilfried Souleyreau³, Marie-Alix Derieppe⁴, Aurélien J. Trotier⁵, Sylvain Miraux⁵, Andreas Bikfalvi³, Emeline J. Ribot^{5†} and Macha Nikolski^{1,2*†}

¹Bordeaux Bioinformatics Center, University of Bordeaux, Bordeaux, France, ²IBGC, CNRS, University of Bordeaux, Bordeaux, France, ³BRIC, INSERM, U1312, University of Bordeaux, Pessac, France, ⁴Service Commun des Animaleries, University of Bordeaux, Bordeaux, France, ⁵Centre de Résonance Magnétique des Systèmes Biologiques, CNRS, University of Bordeaux, Bordeaux, France

Lungs are the most frequent site of metastases growth. The amount and size of pulmonary metastases acquired from MRI imaging data are the important criteria to assess the efficacy of new drugs in preclinical models. While efficient solutions both for MR imaging and the downstream automatic segmentation have been proposed for human patients, both MRI lung imaging and segmentation in preclinical animal models remains challenging due to the physiological motion (respiratory and cardiac movements), to the low amount of protons in this organ and to the particular challenge of precise segmentation of metastases. As a consequence post-mortem analysis is currently required to obtain information on metastatic volume. In this work, we have developed a complete methodological pipeline for automated analysis of lungs and metastases in mice, consisting of an MR sequence for image acquisition and a deep learning method for automatic segmentation of both lungs and metastases. On one hand, we optimized an MR sequence for mouse lung imaging with high contrast for high detection sensitivity. On the other hand we developed DeepMeta, a multiclass U-Net 3+ deep learning model to automatically segment the images. To assess if the proposed deep learning pipeline is able to provide an accurate segmentation of both lungs and pulmonary metastases, we have longitudinally imaged mice with fast- and slow-growing metastasis. Fifty-five balb/c mice were injected with two different derivatives of renal carcinoma cells. Mice were imaged with a SG-bSSFP (self-gated balanced steady state free precession) sequence at different time points after the injection of cancer cells. Both lung and metastases segmentations were manually performed by experts. DeepMeta was trained to perform lung and metastases segmentation based on the resulting ground truth annotations. Volumes of lungs and of pulmonary metastases as well as the number of metastases per mouse were measured on a separate test dataset of MR images. Thanks to the SG method, the 3D bSSFP images of lungs were artifact-free, enabling the downstream detection and serial follow-up of metastases. Moreover, both lungs and metastases segmentation was

accurately performed by DeepMeta as soon as they reached the volume of $\sim 0.02 \text{ mm}^3$. Thus we were able to distinguish two groups of mice in terms of number and volume of pulmonary metastases as well as in terms of the slow *versus* fast patterns of growth of metastases. We have shown that our methodology combining SG-bSSFP with deep learning, enables processing of the whole animal lungs and is thus a viable alternative to histology alone.

KEYWORDS

small animal MRI, motion-correction, pulmonary metastases, deep learning, automatic segmentation

1 Introduction

There is a plethora of preclinical studies investigating the efficacy of innovative treatments on primary tumors as well as on the subsequent metastases, especially pulmonary ones. Indeed, metastases in lungs are of particular interest since they are a sanctuary of many cancer cells (Leong et al., 2006) probably due to the high oxygenated environment and the dense capillary network. Moreover, there is evidence that the number (Cho et al., 2015) and size (Javed et al., 2014) of the pulmonary metastases are related to the survival prognosis. To obtain such quantitative information, most preclinical studies perform post-mortem lung extraction to determine the amount of metastases and their average area or volume. Histology is frequently used due to its high sensitivity of detection, however this technique only allows the analysis of a few lung slices (Shimada et al., 2018; Pein et al., 2020). The number of metastases per slice and their average size (mm^2) are then measured and interpreted as markers of a treatment efficiency.

The usual imaging technique for detecting lung tumors is through a CT scan. However, the X-rays doses limit patient's follow-up, repeated and close sessions, preventing early detection of the disease and patient monitoring during treatment. For example, the French Nuclear Safety Authority (ASN) has published an action plan (n°2011-DL-0019, 2011) aiming to limit the doses delivered to patients, and to favor examinations by Magnetic Resonance Imaging (MRI). MRI is a non-invasive and non-traumatic technique used to perform serial follow-ups to detect lesions. It is a method of choice in oncology since it does not involve ionizing radiation, thus enabling repeated and close sessions to monitor tumor growth. In small animal models, preclinical MRI makes it possible to assess the efficiency of cancer treatments before translation to human studies. However, lung imaging is challenging due to physiological motion (respiratory and cardiac movements) and to the low amount of protons in this organ. We intended to exploit this latter property to efficiently detect pulmonary metastases. For this purpose, the balanced Steady State Free Precession (bSSFP) sequence was chosen as it has been previously shown that high tumor contrast can be obtained in the brain and in the liver (Miraux et al., 2008; Ribot et al., 2011, 2015). When combined with the Self-Gating (SG)

method, motion-induced echoes can be canceled, resulting in images of the abdomen without motion artifacts (Ribot et al., 2015). This sequence is also of high interest due to its high SNR and short scan time to obtain 3D stacks necessary to cover a whole organ with high spatial resolution. Nevertheless, to our knowledge, it has never been applied to detect cancer lesions in lungs.

Deep learning associated with MRI has gained a lot of interest in recent years for different image quantification tasks. Applications range from image acquisition and image retrieval to segmentation and disease prediction (Lundervold and Lundervold, 2019). This is particularly true for human brain diagnostics. For example, recent studies have been conducted to automatically segment tumors or metastases in the human brain, using deep learning, such as the DeepMedic Neural Network (Liu et al., 2017; Charron et al., 2018; Grovik et al., 2020). In the case of segmentation of metastases within lungs of human patients, (Wang et al., 2019), have proposed a deep learning patient-specific method to segment metastases in expert annotated VOIs around them. The model is trained specifically for each patient on early timepoints and can be used in follow-up MRI scans of the same patient, thus lacking the generalization capacity.

However, to date few studies are conducted on small animals such as rats or mice (see e.g. for (Tan et al., 2018) the segmentation of the left ventricle) and, up to our knowledge, no studies has been conducted on lungs, as a result of very low standardization of preclinical protocols, of large variety of mouse lineages, tumor models, as well as of MR sequences and MR instruments (mainly reception coils). Moreover, the low number of animals used in preclinical studies in order to comply as much as possible with the 3R regulations (Replace, Refine, Reduce) guidelines for animal experimentation, represents an additional challenge for downstream automation of image quantification.

In the case of cancer-related studies, the lack of methods for tumor detection and segmentation specific for small-animal preclinical studies, results in manual and time-consuming segmentation by experts. In the case of lung tumors, the development and validation of automated solutions is further hampered by the fact that there are no public databases of mouse or rat MR lung images. Even for human lung images, to our knowledge, there are only two public databases (Lung Image Database Consortium image collection LIDC-IDRI and ELCAP

Public Lung Image Database), both based on CT scans, which limits the development of AI approaches.

In this study we developed a complete methodological pipeline for the automatised analysis of lung and metastases in mice, consisting of an MR sequence for imaging acquisition and a deep learning method for lung and metastases segmentation. This deep learning method enables the measurement of metastases volume in lungs at a given time and the assessment of their longitudinal growth. To ensure reproducibility and foster the use of our method by a large scientific community, all relevant software resources and MR image data are made publicly available. As such, our full pipeline consisting of an MR sequence and downstream automatic image quantification with DeepMeta, constitutes the first step toward standardization of lung and metastases imaging and segmentation in mice, while our MR image database constitutes the first publicly available data resource to foster further methods development.

2 Materials and methods

2.1 Animals

The murine renal cancer RENCA cell lines were maintained in RPMI-1640 (Eurobio) supplemented with 10% (v/v) FBS and 1% (v/v) penicillin-streptomycin, and incubated at 37°C with 5% CO₂. To acquire time-series images of metastases growth showing different growth patterns, 2 cell lines were used: (i) interleukin-34 (il34) knock-out by Crispr/Cas9 method using the 5'-GACCTTACAGGCTACCTTCGGGG-3' targeted sequence for slow-growing metastases pattern and (ii) targeting the LacZ gene using the 5'-TGCGAATACGCCACGCGAT-3' targeted sequence for fast-growing metastases pattern. These 2 cell lines were further injected either intravenously or orthotopically under the renal capsule. For sub-capsular implantations, 1 × 10⁵ RENCA cells were injected under the left kidney capsule of 6–8 weeks old female BALB/c ByJ mice (Charles River Laboratories), whilst for intravenous injections 5 × 10⁶ cells were injected into the caudal vein.

Additional fourteen BALB/c ByJ mice (8 weeks old female, Charles River Laboratories) were injected in the mammary fat pad with 2500 4T1 murine breast cancer cells.

All animal experiments were approved by the “Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI)” (authorization numbers 2016072015478042 and 2015110618597936), and were carried out in accordance with the approved protocols.

2.2 MRI system

Experiments were performed on a 7T Bruker BioSpec system equipped with a gradient coil of 660 mT/m maximum strength and 110 μs rise time. A volume resonator operating in quadrature

mode was used for RF transmit (75.4 mm inner diameter, 70 mm active length) and a proton phased array (RAPID Biomedical GmbH) containing four elements of 30 mm long around an elliptic cylinder (housing: 19 × 25.5 mm³) was used for signal reception.

2.3 MRI acquisitions

A total of 55 mice were imaged after RENCA injection, 27 mice with slow-growing metastases and 28 mice with fast-growing metastases. Animals were imaged every week: from day 6 to day 32 post-implantation for fast-growing metastases mice group and from day 8 until their condition deteriorated (up to day 141 at most) for slow-growing metastases mice group. Also, the 14 mice bearing 4T1 pulmonary metastases were imaged once between day 21 and day 30 after the primary tumor implantation. Two additional healthy mice were scanned three times, two times without repositioning and one time after waking them up in order to evaluate the reproducibility of lung segmentations.

Before imaging, mice were anesthetized with isoflurane (1.5% in air) and placed in the supine position with the lungs in the center of the NMR coil. The breathing rate was monitored using an air balloon placed on top of the lungs (SA Instruments, Inc., NY). The respiration rates between mice were similar for every experiment. The 3D bSSFP images were acquired with the following parameters: TE/TR = 2/4 ms; flip angle (FA) = 30°; FOV: 25 × 20 × 20 mm³; matrix: 128 × 128 × 128; resolution after reconstruction: 195 × 156 × 156 μm; reception bandwidth: 100 kHz; anterior–posterior read direction, according to Ribot et al. (2015). Four different phase offsets were used (180°; 0°; 90°; 270°) to generate four bSSFP images. Each image was acquired with four repetitions per offset. The Analog to Digital Converter (ADC) was turned on immediately after the excitation RF pulse. The corresponding acquired signal, a Free Induction Decay (FID), was recorded at each TR in addition to the echo used for imaging.

2.4 Image reconstruction

For image reconstruction we followed the procedure described in (Ribot et al., 2015). Briefly, the amplitude of the FID signal was characteristic of the animal's respiration and made it possible to identify stable phases and peaks that reflected non-corrupted and motion-corrupted data, respectively. Peaks were identified using the “peak detection” function in MatLab and were then used to delete approximately 30% of the echoes distributed around these peaks. Stacks were then retrospectively reconstructed by deleting the respiration-corrupted k-space lines (or echoes) corresponding to respiration peaks. The different k-spaces obtained for each repetition were then averaged so that missing lines from one k-space could be filled by the k-space lines from another repetition. This results in a single k-space corresponding to one phase offset. All these steps were

TABLE 1 Summary of the dataset obtained from 186 mice. The number of annotated slices is reported alongside the final number of annotated images. Notice that the number of multiclass masks is lower than the lungs one, it is because some metastases are not annotated inside the lungs, so these slices cannot be transformed as a multiclass mask. We split this dataset into a training set and validation set with a ratio of 80/20.

—	Lungs	Metastases	Multiclass masks	Total after augmentation
Slices Number	8156	1296	5762	46096

performed for each phase offset k-space to produce four complete k-spaces. FFT was then applied to generate four SG-bSSFP images containing banding artifacts. Finally, the four images obtained from the acquisition of the four different offsets were summed using the square root of the sum of square (SOS) method to produce the final SG-bSSFP image. Images were exported and processed in tiff format.

2.5 Image annotation

The size of acquired 3D images was $128 \times 128 \times 128$ pixels. Each 3D image was split into 128 2D slices (see Table 1). Each 2D slice was considered as an independent image for downstream analysis. To obtain the corresponding ground truth, each slice was annotated using Fiji Schindelin et al. (2012): (i) masks were manually drawn around the lungs on every slice; (ii) masks were manually drawn around the pulmonary metastases on stacks containing metastases. Both tasks were performed by two different investigators. These annotations enabled us to select slices containing lungs or metastases. Lung and metastases masks were further concatenated with a different value for each area of interest (background, lungs, metastases).

2.6 Dataset

2.6.1 Acquired 3D images

Three datasets have been acquired: (1) RENCA dataset with mice having two different metastases growth patterns, (2) 4T1 dataset and (3) healthy mice dataset. Image acquisitions have resulted in a total of 186 3D stacks for RENCA dataset; the 4T1 dataset contained 14 3D stacks; and the healthy dataset contained 6 separate 3D stacks.

2.6.2 Annotation

RENCA dataset yielded a total of 24576 slices with 128×128 pixel size. Metastases were visible on 62 of the 186 3D stacks. Annotation of the RENCA dataset has resulted in a total of 8156 slices for the lungs segmentation, 1296 slices for the metastases segmentation and a total number of 5762 images with associated multiclass masks (see Table 1). Due to the imaging of the same mice at different timepoints during the development of metastases, large variation in their volume was observed, ranging from 0.0188 mm^3 to more than 200 mm^3 (see Figure 2 for an example of slice annotation).

2.6.3 Data augmentation

Notice that the resulting number of slices is quite small for training a deep neural network. Consequently, we created a data augmentation pipeline, composed of rotations and elastic transformations, which applied a factor 8 to the dataset volume (these numbers are recapitulated in the column “#Total after augmentation”, in Table 1). Rotations (90° , 180° , 270°) ensured that the network was exposed to mouse slices in every position and the elastic transformation ensured a better robustness of the network by providing a training dataset that contains slices subjected to small deformations. Together these augmentation steps helped to reduce overfitting and improved the quality of the segmentations.

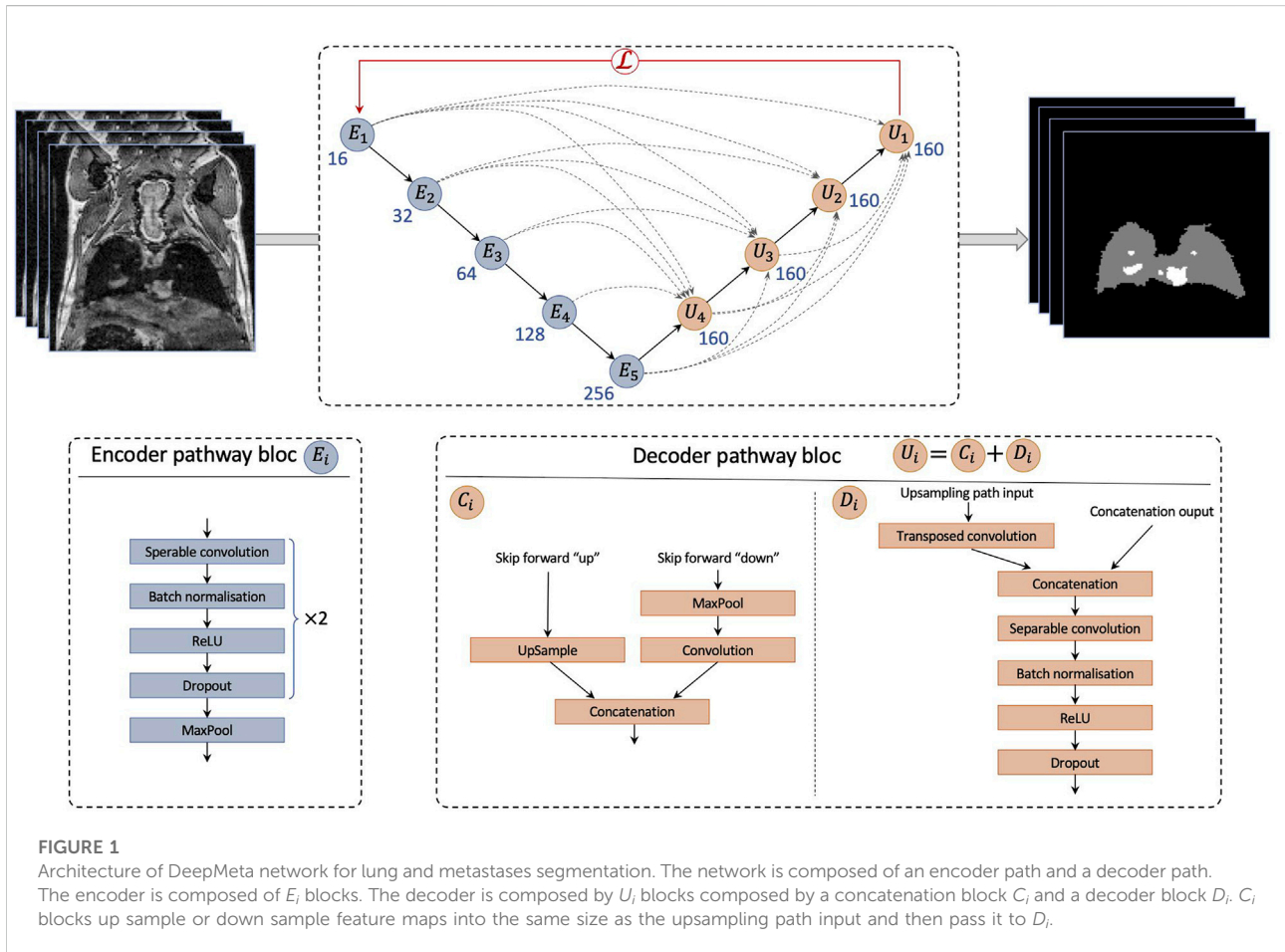
2.6.4 Test dataset

To create the test dataset, we selected four 3D stacks representative of the complete RENCA dataset (representing a total number of 512 slices), sampled from our dataset of 3D image stacks. It is composed of one control 3D stack without metastasis, two 3D stacks with small metastases and one 3D stack with large metastases, defined as under or above 0.4 mm^3 , respectively. This test dataset was annotated by two different annotators in order to obtain a precise inter-observer variation and to quantify whether and to what degree the possible associated bias impacts the training.

2.7 Deep learning processing pipeline

DeepMeta network architecture was designed to perform multiclass segmentation and was based on the U-Net general network structure (see Figure 1). Specifically, we designed our network as a tailored implementation of the U-Net 3+ (Huang et al., 2020). The U-Net 3+ architecture uses full-scale skip connections which combine small and large feature maps from the encoder. We postulated that basing our approach on U-Net 3+ would be suitable to capture both small details and semantic features within images, and as a consequence to accurately segment both lungs and metastases. Specifically, inter- and intra-connection between the encoder E_i down-sampling and decoder U_j up-sampling pathways enabled the U-Net 3+ to account for both fine and coarse level details: in the case of DeepMeta, low level details contain the spatial and boundary information of both lungs and metastases, while high-level details encode their location.

Modifications we made to the classical U-Net 3+ architecture are the following.



- Convolution blocks of the contracting path are composed of one depthwise separable convolution (Chollet, 2017), which speeds up the training process without losing accuracy, followed by a batch norm layer, a ReLU activation and a dropout, repeated twice.
- The expanding path takes its' usual input and concatenates it with feature maps from skip connections (C_i) as in U-Net 3+. However, convolution blocks D_i are composed of a 2D transposed convolution layer, followed by the same architecture as the encoder convolutional blocks.
- The concatenated feature maps have 160 filters. The encoder starts with 16 filters for E_1 and doubles this number for each consecutive block.

2.8 Loss

To train the network, we defined a custom loss function using a loss combination of a cross entropy, a Lovasz-Softmax loss Berman et al. (2018) and a focal loss Lin et al. (2017).

$$\mathcal{L} = \alpha \times \mathcal{L}_{CE} + \beta \times \mathcal{L}_L + \gamma \times \mathcal{L}_F \quad (1)$$

In our data the number of background pixels greatly exceeded that of foreground (lungs and metastases) pixels and thus one of the goals in the definition of our loss function was to solve this class imbalance. To achieve this, we chose to define \mathcal{L}_{CE} as a weighted version of the cross-entropy loss function. Indeed, the classical cross-entropy loss would have been close to 0 due to the high prevalence of True Negatives (background pixels). Moreover, for each class we defined specific weights, one for the background, five for the lungs and 15 for the metastases. The third term of the loss function \mathcal{L}_F , the focal loss, further tweaks the classical cross entropy for solving two issues of the classical cross entropy: the class imbalance problem and learning hard examples. The second term \mathcal{L}_L , Lovasz loss, is based on the IoU and helps to obtain a better segmentation, by minimizing errors that penalize IoU the most. We have set $\alpha = 0.7$, $\beta = 0.4$ and $\gamma = 0.2$.

2.9 Training

During training, the cosine annealing schedule (Loshchilov and Hutter, 2016) was used, which prevents the network from

getting stuck in local minima. Experimental procedure that imaged mice at different time points allowed us to trace the metastases development but also to train the network regardless of the size of the metastasis, as we had access to a wide range of sizes. We split the slices from the training dataset using a ratio of 80% for training and 20% for validation. Network was trained on Nvidia T4 GPUs, for 100 epochs with a learning rate of 0.001 on a dedicated cluster with two CPU Intel Xeon Silver 4114 and 128Go RAM. DeepMeta was implemented in Python 3.10 and Pytorch 1.11, for both neural network creation and training.

2.10 Post-processing pipeline

DeepMeta network generates segmentations in the form of masks, with values of 0 for the background, one for the lungs and two for metastases (see Figure 2). However, the output of the network required post-processing steps to improve the resulting segmentation quality. We defined a post-processing pipeline consisting of four steps.

- 1) The first step is to remove slices that do not contain mouse tissue. For this purpose, we applied a Laplacian of Gaussian (LoG) (Sotak Jr and Boyer, 1989) filter ($\sigma = 7$) to each slice. The Gaussian filter smoothes noise and the Laplacian operator detects edges at 0 crossing while giving a zero response in homogeneous regions. Thus, if a slice contains only noise, the Laplacian of Gaussian operator will not detect any edge and for each pixel of the slice the value yielded by the LoG filter will be close to zero. Consequently, the mean intensity of a filtered slice containing only noise will be close to zero. For this reason, we defined a threshold value of one to separate slices containing mouse tissue (≥ 1), from slices that do not contain any tissue (< 1). This processing step starts from the first slice and is stopped when the first tissues are detected. The same method is applied starting from the last slice. Once the slices that do not contain any tissue are identified, their output masks are defined to contain only 0 (background only).
- 2) In the second step, the mask is split into two binary masks, one mask for lungs and one mask for metastases.
- 3) In the third step, small blobs are removed from each of the remaining slices, more specifically blobs that are smaller than 10 pixels for lungs and three pixels for metastases (blobs are removed in the network's output and in the ground truth in order to not bias the statistics). Additionally, a closing operation (dilation followed by an erosion) with a 3×3 kernel is applied to close small gaps and connect contiguous components. This results in binary masks per slice, for both the lungs and the metastases.
- 4) In the final step, the resulting two binary masks are concatenated to reconstitute the multiclass mask.

Finally, to obtain each volume, we counted the number of pixels for each class in each mask of the stack and multiplied this number by the volume of one voxel, here 0.0047 mm^3 , corresponding to the spatial resolution of the MR images.

In addition, to be able to compute the volume per individual metastasis (and not only the total volume), an additional step was defined. We performed the connected component analyses (Rosenfeld and Pfaltz, 1966) of the 3D stacks of metastases masks using an 18-connected neighborhood to find the instance of each metastasis and thus be able to compute its volume.

2.11 Evaluation

To evaluate the performance of our DeepMeta model, we used two metrics: (i) Intersection over Union (IoU), which provides a coefficient indicating how well the masks overlap, (ii) F1-score, which represents the precision and recall balance. The latter metric is particularly suited for problems with class imbalance, as it is the case in our data.

Moreover, an inter-observer comparison was performed to evaluate the dataset consistency. For a given slice three metrics were computed to compare annotations by the two experts: (1) IoU, (2) positive pixel difference and (3) a ratio of surface difference between the two masks. The IoU metric indicates how well the annotations are overlapping, i.e. whether the experts have annotated the same objects. The positive pixel difference measures the difference in mask sizes. And the ratio shows how much the surface differs in number of pixels. The ratio is calculated with the following formula: $1 - \frac{\min(\sum_i p_i, \sum_j p_j)}{\max(\sum_i p_i, \sum_j p_j)}$, where p_i are pixels from the first annotation and p_j are pixels from the second annotation.

3 Results

3.1 SG-bSSFP images of the mouse thorax

The SG signal could be detected even though the lungs generate low signal at this echo time (see Figure 3). The 3D SG-bSSFP images were mostly free of any motion and banding artifacts, especially in the thoracic area. As expected, lungs appeared with low signal on the images, with hyperintense areas representing blood vessels. Also, the myocardium appeared with more signal than the lungs, and the blood inside the ventricles was dark. In mice bearing pulmonary metastases, the lesions appeared in hyper-intense signals, similar to those of the blood vessels. This contrast enabled us to detect pulmonary metastases as small as four voxels, and to assess their longitudinal growth.

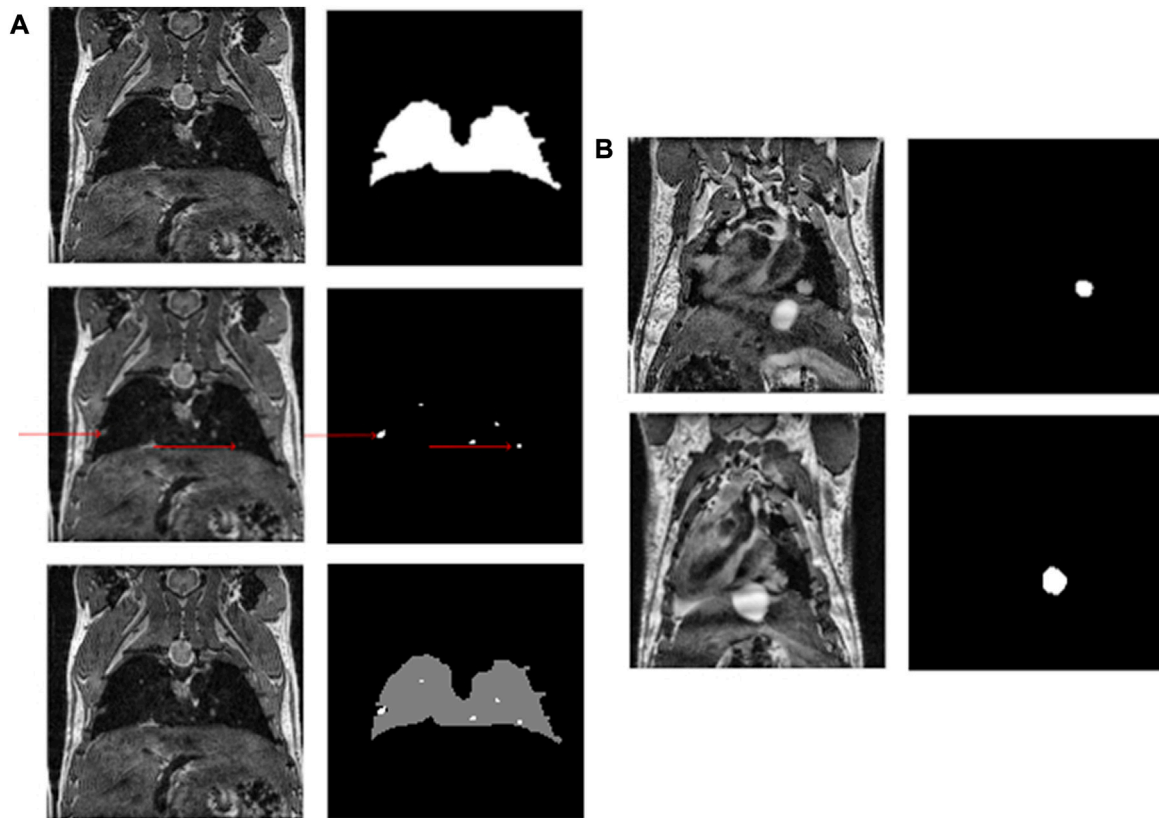


FIGURE 2

(A) An example of a mouse slice, the lung and metastasis masks (red arrows indicate metastases) and the multiclass resulting mask. (B) Shows the growth of a metastasis over time.

3.2 Model performance

3.2.1 Assessment of our model on lungs and metastases segmentation

The performance of our DeepMeta model is shown in Table 2, considering different mice from the test set: those without metastasis (1 mouse) and those with small (2 mice) and large metastases (1 mouse). To highlight the contribution of both the U-Net 3+ architecture and of our loss function, the results are compared with both a vanilla U-Net and a vanilla U-Net 3+.

When interpreting these values, it should be noted that manual segmentation resulted in high inter-observer variability that could decrease to an IoU of 0.71 and 0.82 for small and large metastases, respectively. Moreover, 29% and 12% differences were measured between the two expert segmentations, highlighting the difficulty of manual metastases segmentation. Considering these constraints, the performance of

the DeepMeta model for metastases segmentation is within the range of human experts.

3.2.1.1 Lung segmentation

The mean volume of the mouse lungs was $463.5 \mu\text{L} \pm 72 \mu\text{L}$ in the test set. The IoU index between the manual and the DeepMeta lung segmentation was 0.86 (see Table 2), improving on both vanilla U-Net and vanilla U-Net 3+. The inter-observer variability of lung segmentation was low, with a mean IoU of 0.87, leading to 8% difference (see Table 3).

Moreover, the reproducibility of lung segmentation was assessed on the healthy dataset that was obtained by imaging two healthy mice 3 times each. The resulting lung volume for the first mouse was 484.5 mm^3 , 474.9 mm^3 and 494 mm^3 ; for the second mouse: 490.4 mm^3 , 512.1 mm^3 and 486.2 mm^3 . The coefficients of variance were 1.9% for the first mouse, 2.8% for the second one and 2.3% globally, showing high reproducibility for lung segmentation. The mean lung volume of the RENCA

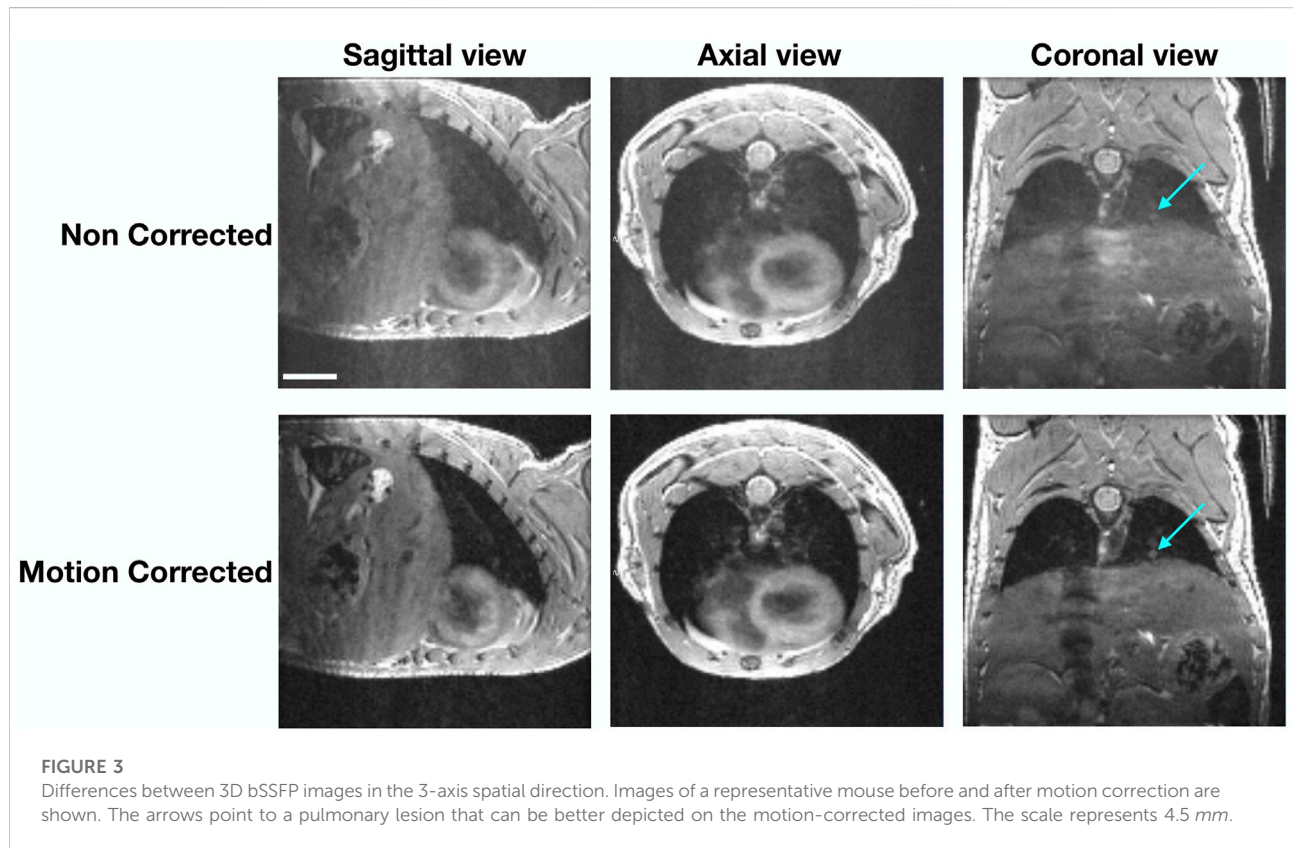


TABLE 2 Models performance are measured by IoU and F1-score for both lungs and metastases segmentation in our representative test set for DeepMeta, vanilla U-Net and vanilla U-Net 3+ architectures. This set includes mice without metastases and having either small or large metastases. The last column shows the mean values across the entire test set.

Object	No metastases		Small metastases		Big metastases		Total	
	IoU	F1	IoU	F1	IoU	F1	IoU	F1
<i>Vanilla U-Net</i>								
Lungs	0.86	0.93	0.86	0.88	0.76	0.84	0.84	0.88
	±0.04	±0.04	±0.07	±0.07	±0.06	±0.06	±0.05	±0.05
Metastases	NA	NA	0.69	0.74	0.70	0.64	0.70	0.69
			±0.19	±0.18	±0.17	±0.15	±0.15	±0.13
<i>Vanilla U-Net 3+</i>								
Lungs	0.87	0.91	0.86	0.90	0.80	0.87	0.84	0.89
	±0.04	±0.04	±0.07	±0.07	±0.04	±0.03	±0.05	±0.04
Metastases	NA	NA	0.71	0.75	0.62	0.67	0.68	0.71
			±0.17	±0.17	±0.16	±0.15	±0.14	±0.12
<i>DeepMeta</i>								
Lungs	0.89	0.93	0.87	0.95	0.82	0.88	0.86	0.92
	±0.03	±0.02	±0.06	±0.06	±0.05	±0.04	±0.04	±0.04
Metastases	NA	NA	0.72	0.76	0.67	0.72	0.70	0.74
			±0.19	±0.2	±0.15	±0.14	±0.14	±0.13

TABLE 3 Inter-observer differences. Measured Inter-observer differences in manual annotation of the test set. Reproducibility between two annotations is measured by the IoU; annotation difference is measured by the count of differing pixels and the surface ratio, with a total pixels per slice of 16384.

Metric	Lungs	Small metastases	Large metastases
IoU	0.87	0.71	0.91
Mean difference in number of pixels	94.48	12.18	33.74
Mean surface ratio	0.08	0.29	0.12

dataset is within 10% range from the mean volume of the healthy dataset, considered as reference.

3.2.1.2 Metastases segmentation

The metastases were detected and automatically segmented for the test dataset. See Table 3, panel B, for an example of segmentation of a slice from the same mouse imaged at day 85 and day 92. The total metastasis volume for this mouse was 30.8 mm^3 and 47.3 mm^3 at these time points. Moreover, the model has shown high sensitivity as it could detect metastases as small as four voxels (representing here 0.0188 mm^3) in a given slice. The IoU index of 0.72 was obtained between the manual and the DeepMeta segmentation for small metastases (see Table 2) defined as metastases with an area size of less than 85 pixels (approximately 0.4 mm^3). This index decreased to 0.67 when metastases were larger than 85 pixels. Notice the systematic improvement relative to the vanilla architectures.

3.4 Capacity to distinguish different growth patterns of metastases

In the RENCA dataset metastases were detected and segmented by our model in the fast-growing metastases mice group as early as 19 days post-injection (see Figure 4A). The metastases growing subsequently after the injection of the il34 cells became detectable at 22 days after injection. Also, the model was able to measure the lung volumes occupied by the pulmonary metastases over time. This enabled us to differentiate the two groups of mice. The DeepMeta model was also able to measure the total volume of metastases at different time points. This made it possible to count the number of metastases per volume range (see Figure 4B). Figure 4B highlights that the metastases in the fast-growing group are more homogeneous in size than the slow-growing group. Also, the slow-growing group had fewer metastases compared with the fast-growing one.

3.5 DeepMeta's use case

To assess whether our DeepMeta model can be used to process pulmonary metastases from another cell line we

applied the complete procedure from imaging to lung and metastases segmentation on the 4T1 dataset. The mean lung volume found by DeepMeta was $547 \text{ mm}^3 \pm 61.2$ which is in the 10% range of lung volume from the healthy dataset. The network managed to also segment the pulmonary metastases (see Figure 5 and Table 4) and found an average number of 4.5 metastases per mouse with a volume range from 0.01 to 6.9 mm^3 . These results show that the DeepMeta model generalizes well to other data and in particular that it can process metastases from different cell lines. As such, this indicates that the use of DeepMeta can streamline the analysis of lungs and metastases in MR images in preclinical studies on small animals.

4 Discussion

This study involves the optimization and development of both the image acquisition method and the processing to automatically segment mouse lungs and subsequently, the pulmonary metastases at different stages of their growth.

The MR sequence used here is the bSSFP sequence, due to its high contrasts between the lung parenchyma and pulmonary metastases and also because the 3D images can be obtained relatively fast. This sequence is commonly used for human cardiac imaging, although a 2D version is used in this case. The combination of the bSSFP sequence with a Self-Gating module makes it possible to combine high spatial resolution in the three directions and robustness to motion to detect early-growing metastases throughout the whole lung. Indeed, the bSSFP sequence can not be combined with respiration-gating in order to suppress motion artifact, so as to not perturb the steady-state of the signal. To perform lung imaging, Ultra-short Echo Time (UTE) sequences can be used to enhance the signal from the lung parenchyma. Additionally, these sequences are less sensitive to motion than Cartesian sequences because of the radial encoding. Nevertheless, due to the proton-density weighting, the contrasts between the different tissues are more homogeneous than with a bSSFP sequence, which may have decreased the detection sensitivity of the metastases and make the annotation and thus the automatic segmentation less reliable. One of the drawbacks of the radial encoding is the lower sharpness of the images. Yet, this characteristic is highly

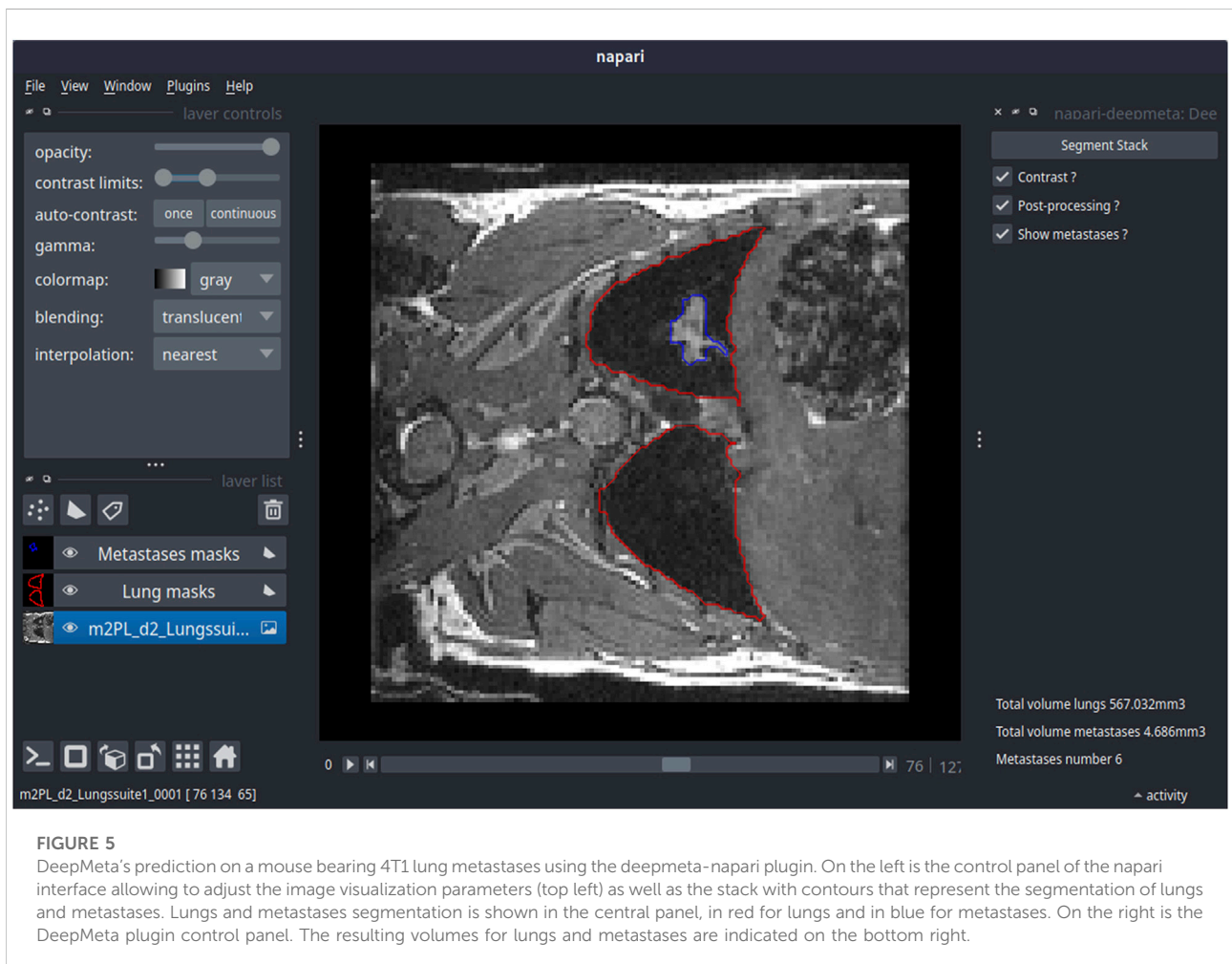
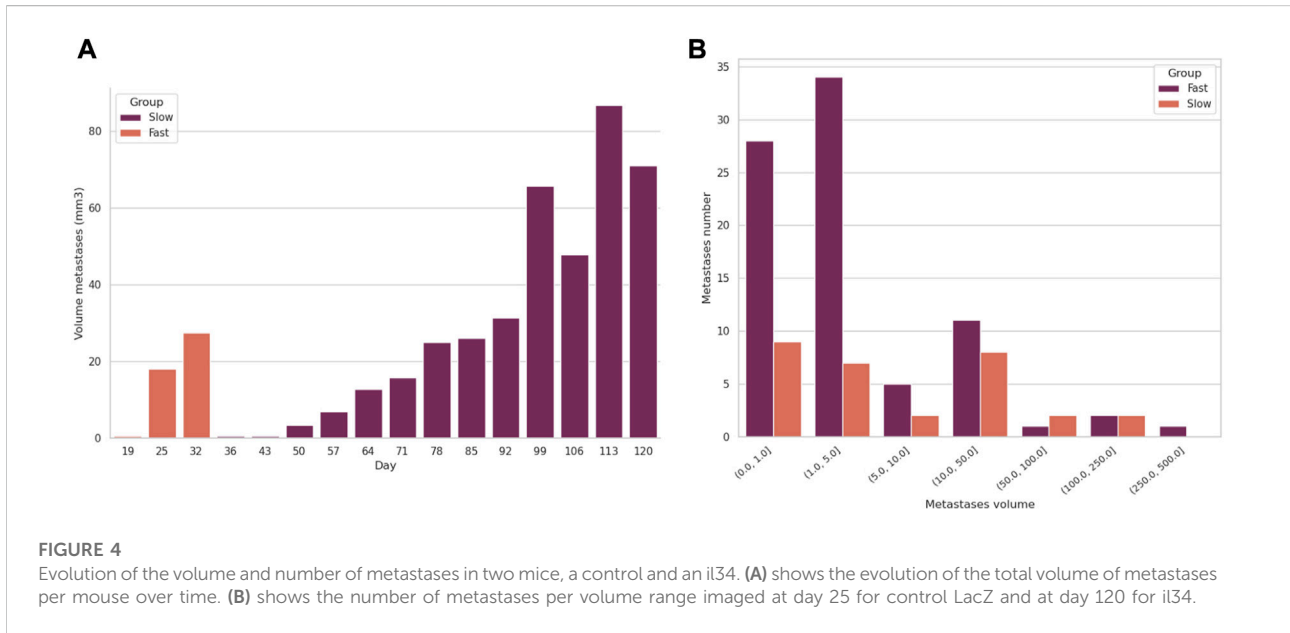


TABLE 4 Results of DeepMeta's prediction on mice bearing 4T1 lung metastases.

Name	Lungs volume	Metastases total volume (mm ³)	Metastases number
4T1_1	561.3 mm ³	1.4 ³	4
4T1_2	522.1 mm ³	1.6 ³	15
4T1_3	586.6 mm ³	0 ³	0
4T1_4	627.8 mm ³	3 ³	6
4T1_5	579.8 mm ³	0 ³	0
4T1_6	550.5 mm ³	6.9 ³	22
4T1_7	467.5 mm ³	0.02 ³	1
4T1_8	434 mm ³	0 ³	0
4T1_9	471.4 mm ³	0.7 ³	4
4T1_10	511.5 mm ³	0.3 ³	4
4T1_11	599.6 mm ³	0 ³	0
4T1_12	627.8 mm ³	3 ³	6
4T1_13	519.8 mm ³	0.02 ³	1
4T1_14	597.9 mm ³	0.01 ³	1

needed when measuring the volume of small structures such as pulmonary metastases.

Lungs and tumor segmentation using deep learning techniques is an active research field. In particular, CNN-based networks are popular and efficient both for 2D and 3D x-ray images (human and mouse) (Tang et al., 2019; Nishio et al., 2021; Osadebey et al., 2021; van de Worp et al., 2021). Models used for these tasks rely on a particular CNN architecture, the U-Net. As for MRI images, U-Net models are also used for brain and tissues segmentation in mice (2D and 3D) (Holbrook et al., 2020; Zhang et al., 2021). However, up to our knowledge, no model has been developed before this paper to segment lungs and metastases in small animals using MRI images.

Although conventional image processing techniques exist for lung segmentation (Egger et al., 2014), these approaches typically require human supervision to achieve high precision. Moreover, previous deep learning applications to MR images, such as En-DeepMedic (Grøvik et al., 2020), had to perform manual preprocessing steps (skull stripping before brain metastases segmentation), while DeepMeta pipeline implements what can be considered as a fully automated lung stripping and the automated metastases segmentation, both based on deep learning and automatic post-processing.

As lung slices are more susceptible to contain artifacts due to respiratory movements, we used the U-Net 3+ that helps propagate semantic and spatial information along the network reducing the network's sensitivity to motion blur.

The U-Net 3+ architecture uses full-scale skip connection blocks that allow the network to retrieve semantic information in the reconstruction path and also prevent small objects from disappearing. In our case, lungs are not especially small objects compared to the image size (128 × 128), but metastases are

relatively small and can be considered as small objects. A previous study (Zhang et al., 2021) shows that classical architectures do not perform well on small objects. Thus, this U-Net 3+ based network architecture is particularly suited for the task of small object detection.

DeepMeta U-Net 3+ model enables the reconstruction of lung volumes consistent with previous studies (Heverhagen et al., 2004; Soutiere and Mitzner, 2004; Ribot et al., 2011), thus ensuring the viability of this approach. Note that the chosen supervised learning approach relies on manual data annotation for both lungs and metastases. For example, lung segmentation alone was performed on 8156 slices, which is both time consuming and potentially bias-prone. A possible avenue to augment the dataset with known ground truth could be by data augmentation based on function decomposition from a template (Tustison et al., 2019). This technique is promising for lung data augmentation, but not for metastases. Indeed, these lesions have their shape and volume that grow over time, and their location varies among mice.

The high variability between annotators, especially for metastases, might come from partial volume effects that can greatly modify the volume of small metastases. For larger metastases, especially those growing close to the heart or the thoracic muscles, the main issue comes from the proper delineation of the metastases, as all these structures show similar signals on the bSSFP images.

The method presented here is the first stone to further developments to make deep learning models more robust to variability. First, imaging with different reception coils would enrich the training dataset. Indeed, our study was conducted using a prototype coil that is not widely used in imaging laboratories. It will thus be relevant to train the DeepMeta

models with images generated by more commonly used coils, such as surface or volume coils. In addition, different models of pulmonary metastases could be evaluated. Indeed, their shape and growth pattern usually depend on the cancer cells from which they originate. Some metastases grow along blood vessels (co-option), while others develop neo-angiogenesis; the RENCA cells used here grow like spheres, whereas others have a very invasive phenotype, *etc.*

4.1 Limitations

Deep learning networks do not carry object semantics, which results in the network segmenting any object that has a similar structure to the metastases (e.g., blood vessels). The performance of the network depends on the architecture and on the training and thus on the training dataset. In this work, the images were acquired using the bSSFP MRI sequence, which implies that the pre-trained DeepMeta models are expected to perform well on similarly acquired images, but that the performance might decrease if the images to be segmented are acquired with a different sequence. Consequently, retraining might be necessary by including images acquired with other MRI sequences in the training dataset to achieve a higher level of generalization.

The bSSFP sequence has several advantages for mice imaging. Nevertheless, the inherent banding artifacts severely affect the quality of the images. It is thus necessary to acquire images at multiple phase offsets, which lengthens the acquisition time. Our MR reconstruction pipeline also required manual selection of the most sensitive coil to the motion self-gated signal variation and the rejecting window size. Nevertheless, the whole reconstruction pipeline can be fully automated through the use of an advanced signal processing pipeline such as SSA-FARY (Rosenzweig et al., 2020).

In conclusion, we have developed a freely available and highly reproducible deep learning model that allows segmentation of lungs and metastases as well as the measurement of the corresponding volumes over time in small animal MR images without need for human intervention.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://zenodo.org/record/7014776>.

Ethics statement

The animal study was reviewed and approved by Ministère de l'Enseignement Supérieur, de la Recherche

et de l'Innovation (MESRI) authorization numbers 2016072015478042 and 2015110618597936.

Author contributions

MN and AB conceptualized the approach. MN and ER designed the study. MN, SM and ER supervised the research. MN, EL, EB and AC contributed to the computational experimental design. EL implemented DeepMeta. ER, AT and SM performed image acquisition. WS and M-AD perform animals experiments. ER, AC and EL performed data annotation. All the authors contributed to writing the manuscript.

Funding

This work was supported by the funding from Plan Cancer 2017 project SystemsRCC (C18005GS), by the funding from INCa SIRIC 2018 project (BRIO/Commucan) and by the Inserm Cancer PCSI funding N°20CP093-00.

Acknowledgments

For the lentiviral Crispr/Cas9 construct, lentiCRISPRv2 plasmid containing spCas9 cassette and expressing guide RNA targeted LacZ or mouse il34 exon one were kindly provided by Michel Tremblay (McGill University). The authors thank Pr. Slim Karkar for his valuable suggestions for this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Berman, M., Triki, A. R., and Blaschko, M. B. (2018). "The lovasz-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 4413–4421. doi:10.1109/CVPR.2018.00464
- Charron, O., Lallement, A., Jarnet, D., Noblet, V., Clavier, J. B., and Meyer, P. (2018). Automatic detection and segmentation of brain metastases on multimodal mr images with a deep convolutional neural network. *Comput. Biol. Med.* 95, 43–54. doi:10.1016/j.combiomed.2018.02.004
- Cho, J. H., Seok, K., Namgung, M., Choi, Y. S., Kim, H. K., Zo, J. I., et al. (2015). The prognostic importance of the number of metastases in pulmonary metastasectomy of colorectal cancer. *World J. Surg. Oncol.* 13, 222–228. doi:10.1186/s12957-015-0621-7
- Chollet, F. (2017). "Xception: Deep learning with depthwise separable convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 1251–1258. doi:10.1109/CVPR.2017.195
- Egger, C., Gérard, C., Vidotto, N., Accart, N., Cannet, C., Dunbar, A., et al. (2014). Lung volume quantified by mri reflects extracellular-matrix deposition and altered pulmonary function in bleomycin models of fibrosis: Effects of som230. *Am. J. Physiology-Lung Cell. Mol. Physiology* 306 12, L1064–L1077. doi:10.1152/ajplung.00027.2014
- Grovik, E., Darvin, Y., Iv, M., Tong, E., Rubin, D., and Zaharchuk, G. (2020). Deep learning enables automatic detection and segmentation of brain metastases on multisquence mri. *J. Magn. Reson. Imaging* 51, 175–182. doi:10.1002/jmri.26766
- Heverhagen, J. T., Hahn, H. K., Wegmann, M., Herz, U., Whitaker, C. D. S., Matschl, V., et al. (2004). Volumetric analysis of mice lungs in a clinical magnetic resonance imaging scanner. *MAGMA* 17, 80–85. doi:10.1007/s10334-004-0053-9
- Holbrook, M. D., Blocker, S., Mowery, Y. M., Badea, A., Qi, Y., Xu, E. S., et al. (2020). Mri-based deep learning segmentation and radiomics of sarcoma in mice. *Tomography* 6, 23–33. doi:10.18383/j.tom.2019.00021
- Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., et al. (2020). Unet 3+: A full-scale connected unet for medical image segmentation. *Speech Signal Process. (ICASSP). International Conference on Acoustics*, 1055–1059. doi:10.1109/ICASSP40776.2020.9053405
- Javed, M. A., Sheel, A. R. G., Sheikh, A. A., Page, R. D., and Rooney, P. S. (2014). Size of metastatic deposits affects prognosis in patients undergoing pulmonary metastasectomy for colorectal cancer. *annals* 96, 32–36. doi:10.1308/003588414X13824511650371
- Leong, S. P. L., Cady, B., Jablons, D. M., Garcia-Aguilar, J., Reintgen, D., Jakub, J., et al. (2006). Clinical patterns of metastasis. *Cancer Metastasis Rev.* 25, 221–232. doi:10.1007/s10555-006-8502-8
- Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. *Proc. IEEE Int. Conf. Comput. Vis.*, 2980–2988. doi:10.1109/ICCV.2017.324
- Liu, Y., Stojadinovic, S., Hryckushko, B., Wardak, Z., Lau, S., Lu, W., et al. (2017). A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery. *PloS one* 12, e0185844. doi:10.1371/journal.pone.0185844
- Loshchilov, I., and Hutter, F. (2016). Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*.
- Lundervold, A. S., and Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on mri. *Z. für Med. Phys.* 29, 102–127. doi:10.1016/j.zemedi.2018.11.002
- Miriaux, S., Philippe, P. M., Ribot, E. J., Franconi, J. M., and Thiaudiere, E. (2008). 3d truefisp imaging of mouse brain at 4.7t and 9.4t. *J. Magn. Reson. Imaging* 28, 497–503. doi:10.1002/jmri.21449
- Nishio, M., Fujimoto, K., and Togashi, K. (2021). Lung segmentation on chest x-ray images in patients with severe abnormal findings using deep learning. *Int. J. Imaging Syst. Technol.* 31, 1002–1008. doi:10.1002/ima.22528
- Osadebey, M., Andersen, H. K., Waaler, D., Fossaa, K., Martinsen, A. C. T., and Pedersen, M. (2021). Three-stage segmentation of lung region from ct images using deep neural networks. *BMC Med. Imaging* 21, 112–119. doi:10.1186/s12880-021-00640-1
- Pein, M., Insua-Rodriguez, J., Hongu, T., Riedel, A., Meier, J., Wiedmann, L., et al. (2020). Metastasis-initiating cells induce and exploit a fibroblast niche to fuel malignant colonization of the lungs. *Nat. Commun.* 11, 1494–1518. doi:10.1038/s41467-020-15188-x
- Ribot, E. J., Duriez, T. J., Aurélien, A. J. T., Thiaudiere, E., Franconi, J. M., and Miriaux, S. (2015). Self-gated bssfp sequences to detect iron-labeled cancer cells and/or metastases *in vivo* in mouse liver at 7 tesla. *J. Magn. Reson. Imaging* 41, 1413–1421. doi:10.1002/jmri.24688
- Ribot, E. J., Martinez-Santesteban, F. M., Simeanea, C., Steeg, P. S., Chambers, A. F., Rutt, B. K., et al. (2011). *In vivo* single scan detection of both iron-labeled cells and breast cancer metastases in the mouse brain using balanced steady-state free precession imaging at 1.5 t. *J. Magn. Reson. Imaging* 34, 231–238. doi:10.1002/jmri.22593
- Rosenfeld, A., and Pfaltz, J. L. (1966). Sequential operations in digital picture processing. *J. ACM* 13, 471–494. doi:10.1145/321356.321357
- Rosenzweig, S., Scholand, N., Holme, H., Christian, M., and Uecker, M. (2020). Cardiac and respiratory self-gating in radial mri using an adapted singular spectrum analysis (ssa-fary). *IEEE Trans. Med. Imaging* 39, 3029–3041. doi:10.1109/TMI.2020.2985994
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi:10.1038/nmeth.2019
- Shimada, K., Ushijima, K., Suzuki, C., Horiguchi, M., Ando, H., Akita, T., et al. (2018). Pulmonary administration of curcumin inhibits b16f10 melanoma lung metastasis and invasion in mice. *Cancer Chemother. Pharmacol.* 82, 265–273. doi:10.1007/s00280-018-3616-6
- Sotak Jr, G. E., and Boyer, K. L. (1989). The laplacian-of-Gaussian kernel: A formal analysis and design procedure for fast, accurate convolution and full-frame output. *Comput. Vis. Graph. Image Process.* 48, 147–189. doi:10.1016/S0734-189X(89)80036-2
- Soutiere, S. E., and Mitzner, W. (2004). On defining total lung capacity in the mouse. *J. Appl. Physiol.* 96, 1658–1664. doi:10.1152/jappphysiol.01098.2003
- Tan, L. K., McLaughlin, R. A., Lim, E., Abdul Aziz, Y. F., and Liew, Y. M. (2018). Fully automated segmentation of the left ventricle in cine cardiac mri using neural network regression. *J. Magn. Reson. Imaging* 48, 140–152. doi:10.1002/jmri.25932
- Tang, H., Zhang, C., and Xie, X. (2019). "Automatic pulmonary lobe segmentation using deep learning," in 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019) (Venice, Italy: ISBI), 1225–1228. doi:10.1109/ISBI.2019.8759468
- Tustison, N. J., Avants, B. B., Lin, Z., Feng, X., Cullen, N., Mata, J. F., et al. (2019). Convolutional neural networks with template-based data augmentation for functional lung image quantification. *Acad. Radiol.* 26, 412–423. doi:10.1016/j.acra.2018.08.003
- van de Worp, W. R. P. H., van der Heyden, B., Lappas, G., van Helvoort, A., Theys, J., Schols, A. M., et al. (2021). Deep learning based automated orthotopic lung tumor segmentation in whole-body mouse ct-scans. *Cancers* 13, 4585. doi:10.3390/cancers13184585
- Wang, C., Tyagi, N., Rimner, A., Hu, Y.-C., Veeraraghavan, H., Li, G., et al. (2019). Segmenting lung tumors on longitudinal imaging studies via a patient-specific adaptive convolutional neural network. *Radiotherapy Oncol.* 131, 101–107. doi:10.1016/j.radonc.2018.10.037
- Zhang, F., Breger, A., Cho, K. I. K., Ning, L., Westin, C. F., O'Donnell, L. J., et al. (2021). Deep learning based segmentation of brain tissue from diffusion mri. *NeuroImage* 233, 117934. doi:10.1016/j.neuroimage.2021.117934

Annexe B

FluoSIM : Simulateur de diffusion de molécules fluorescente de type SMLM



OPEN

FluoSim: simulator of single molecule dynamics for fluorescence live-cell and super-resolution imaging of membrane proteins

Matthieu Lagardère¹, Ingrid Chamma¹, Emmanuel Bouilhol², Macha Nikolski² & Olivier Thoumine¹✉

Fluorescence live-cell and super-resolution microscopy methods have considerably advanced our understanding of the dynamics and mesoscale organization of macro-molecular complexes that drive cellular functions. However, different imaging techniques can provide quite disparate information about protein motion and organization, owing to their respective experimental ranges and limitations. To address these issues, we present here a robust computer program, called FluoSim, which is an interactive simulator of membrane protein dynamics for live-cell imaging methods including SPT, FRAP, PAF, and FCS, and super-resolution imaging techniques such as PALM, dSTORM, and uPAINT. FluoSim integrates diffusion coefficients, binding rates, and fluorophore photo-physics to calculate in real time the localization and intensity of thousands of independent molecules in 2D cellular geometries, providing simulated data directly comparable to actual experiments. FluoSim was thoroughly validated against experimental data obtained on the canonical neurexin-neurologin adhesion complex at cell–cell contacts. This unified software allows one to model and predict membrane protein dynamics and localization at the ensemble and single molecule level, so as to reconcile imaging paradigms and quantitatively characterize protein behavior in complex cellular environments.

Critical cellular functions such as membrane adhesion, receptor-mediated signaling, or synaptic transmission, involve the diffusional trapping of specific molecules in sub-cellular compartments^{1,2}. To quantitatively describe such molecular dynamics in living cells, several fluorescence imaging techniques are currently available^{3,4}: (1) single particle tracking (SPT) which resolves the motion of individual proteins at camera frame rate; (2) photo-activation and photo-bleaching methods, namely PhotoActivation of Fluorescence (PAF) and Fluorescence Recovery After Photobleaching (FRAP) which infer protein turnover at the population level; and (3) Fluorescence Correlation Spectroscopy (FCS), which analyzes molecular dynamics by correlating intensity fluctuations. More recent approaches based on single molecule localization such as PhotoActivated Localization Microscopy (PALM)^{5,6}, direct Stochastic Optical Reconstruction Microscopy (dSTORM)⁷, and Point Accumulation In Nanoscopic Topography (PAINT)⁸, yield images of protein distribution at improved resolution (below 50 nm), giving unprecedented information about the nanoscale organization of biological structures⁹.

Despite such progress in imaging power, many experimental parameters remain difficult to estimate or control, including protein expression levels, probe labeling density, potential fixation artifacts, spatial and temporal sampling of the recordings, and protein motion below the system resolution, such that results from different experimental paradigms are often difficult to reconcile¹⁰. Thus, there is a pressing need for computer simulators that could unify those different imaging modes in a unique framework, estimate their respective biases, and serve as a predictive tool for experimenters, with the aim to quantitatively decipher protein organization and dynamics in living cells. Several particle-based packages relying on Monte Carlo simulations already exist to predict random motion and multi-state reactions of biological molecules, but either they do not integrate fluorescence properties or are limited to a specific type of imaging mode, and are usually not performing real-time visualization^{11–18}.

In this context, we provide here fast, robust, and user-friendly software (*FluoSim*) that allows real time simulation of membrane protein dynamics in live-cell imaging (SPT, FRAP, PAF, and FCS) and super-resolution (PALM, dSTORM, uPAINT) modalities. We also show that FluoSim can be further used to produce large virtual data sets

¹CNRS, Interdisciplinary Institute for Neuroscience, IINS, UMR 5297, Univ. Bordeaux, 33000 Bordeaux, France. ²CNRS, IBGC, UMR 5095, Univ. Bordeaux, 33000 Bordeaux, France. ✉email: olivier.thoumine@u-bordeaux.fr

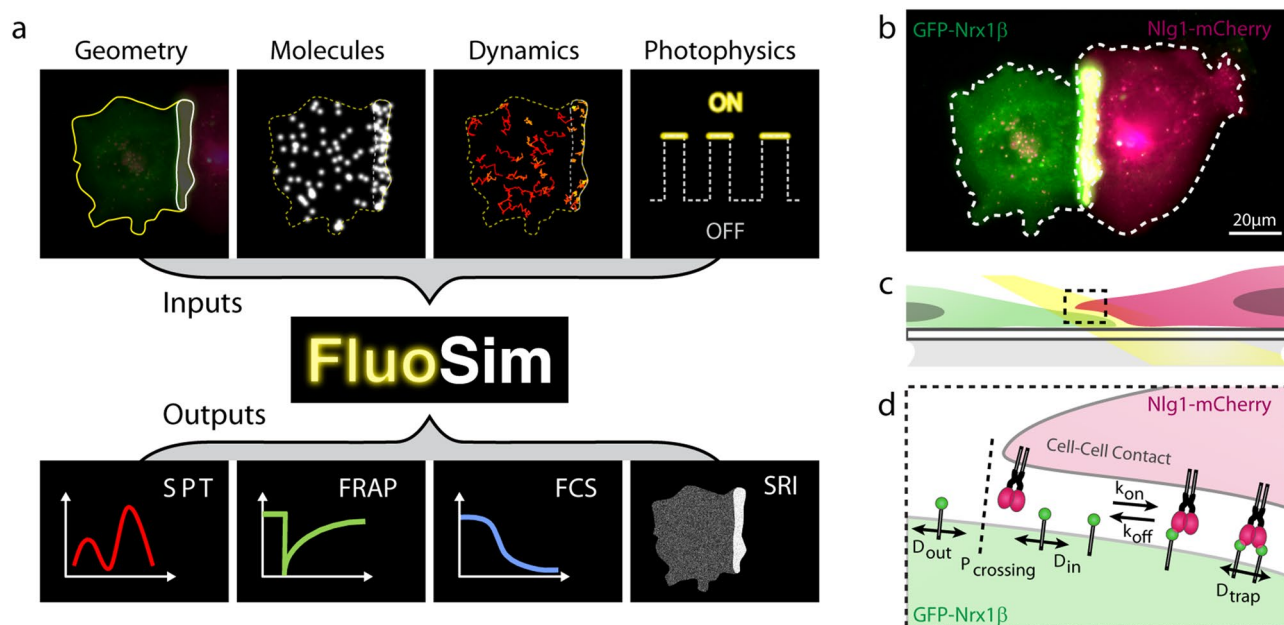


Figure 1. Schematics of the simulator and experimental system. (a) General principle of FluoSim. (b) Contact between two COS cells, one expressing GFP-Nrx1 β (green) and the other expressing Nlg1-mCherry (magenta), resulting in molecule accumulation through adhesive interactions (yellow zone). (c) Diagram showing a zoomed section of the cell–cell interface at the coverslip; the yellow beam represents oblique laser illumination. (d) Cartoon of Nrx1 β diffusional trapping by Nlg1 in a cell–cell contact.

for training deep neural networks for image reconstruction¹⁹. This software should thus be of great interest to a wide community specialized in imaging methods applied to cell biology and neuroscience, with the common aim to better understand membrane dynamics and organization in cells.

Results

General principle of FluoSim. The FluoSim interface looks like performing a real experiment: the user imports a 2D cellular geometry from a microscopy image, and populates it with a realistic number of molecules (a parameter which depends on protein expression level, cell surface area, and labeling density) (Fig. 1a, Supplemental software and accompanying user manual). Kinetic parameters characterizing the diffusion and trapping of molecules in cellular regions of interest (ROI), are entered as inputs (Table 1). At each time increment (typically ~ 1 – 100 ms, adjusted to sensor acquisition rates), the algorithm updates the instantaneous positions of all molecules based on random number generation. Photo-switching rates of organic fluorophores or fluorescent proteins attached to the proteins of interest, further determine the fluorescence intensity associated to each molecule over time. The algorithm is optimized to visualize in real time the cellular system and provide post-processing information including molecule trajectories, image stacks, and output graphs (i.e. histograms of diffusion coefficients, FRAP, PAF, and FCS curves). Examples of simulated data sets for a realistic range of parameter values are given in Fig. S1. Moreover, since the positions of molecules are known with near-infinite accuracy, the program can generate super-resolved images comparable to those obtained with PALM or dSTORM, after introducing additional parameters describing protein labeling density, fluorophore duty cycle, and localization precision of the system. Finally, various levels of noise were introduced in the simulator, including a Poisson shot noise characterizing intensity fluctuations of individual fluorescent molecules, and a Gaussian readout noise applied around the camera offset^{20,21}. This approach led to the simulation of very realistic images that closely matched those obtained by Total Internal Reflection Fluorescence (TIRF) imaging of single GFP-tagged molecules immobilized on a glass substrate (Fig. S2). A complete view of the parameters used in each experimental paradigm is available from the individual examples provided in the software menu.

Experimental system to validate FluoSim. To thoroughly validate FluoSim, we performed SPT, dSTORM, FRAP, and FCS experiments essentially on the canonical neurexin-neuroligin complex that mediates trans-synaptic adhesion in neurons²². We used COS-7 cells as a model expression system because they form large and flat lamellipodia that can be approximated as 2D environments for membrane diffusion. Cells were separately electroporated with recombinant GFP-tagged neurexin-1 β (GFP-Nrx1 β) and mCherry-tagged neuroligin-1 (Nlg1-mCherry), then cultured together for 24 h (Fig. 1b–d). Both molecules reached the cell membrane and accumulated at cell–cell contacts (GFP-Nrx1 β enrichment = 3.9 ± 0.5 , $n = 20$ cells), revealing adhesive interactions.

Simulations of SPT experiments. First, we experimentally tracked single GFP-Nrx1 β molecules labeled with Atto647N-conjugated anti-GFP nanobody at 50 Hz using uPAINT^{8,23,24}. Nrx1 β exhibited fast free diffusion

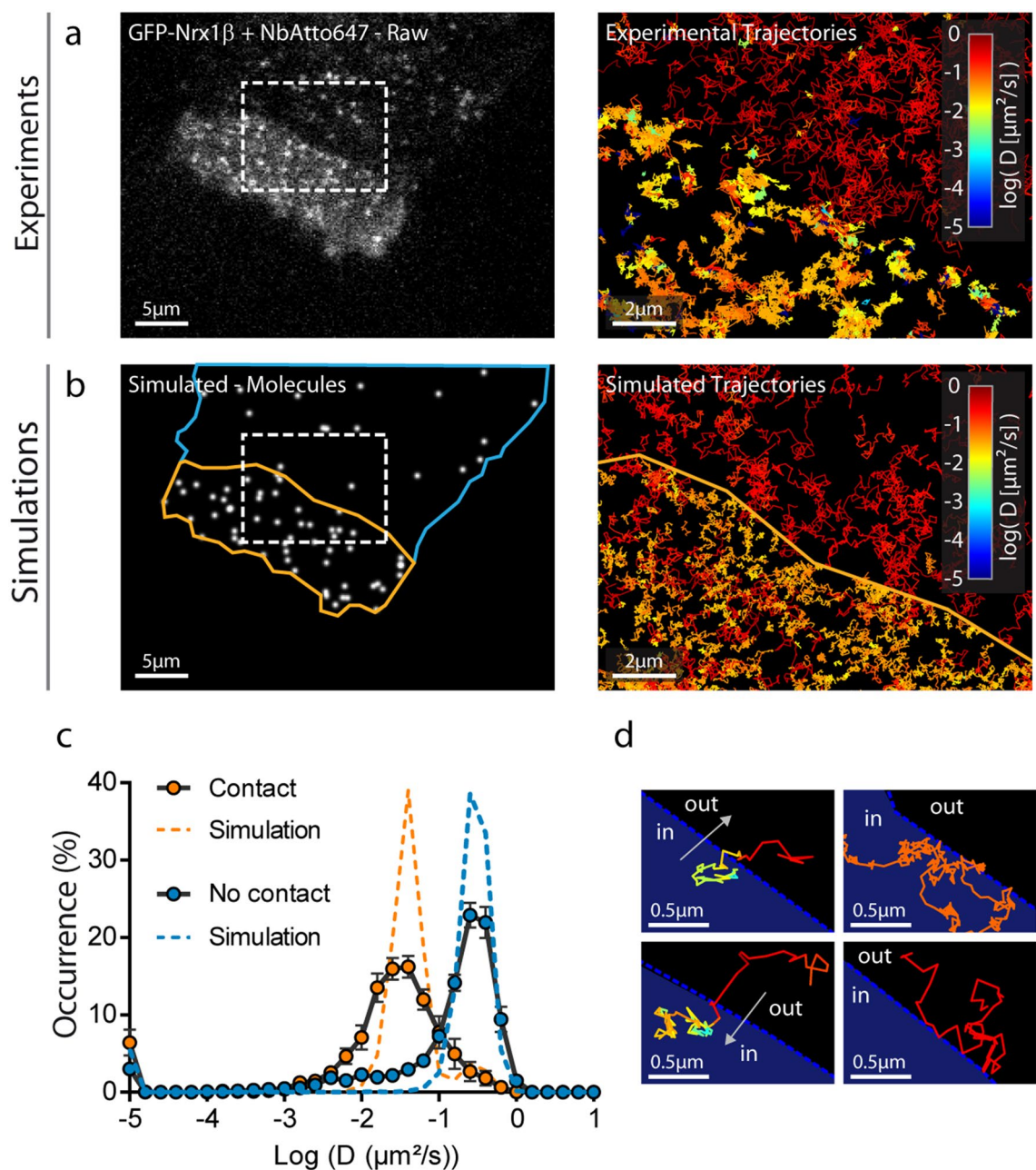


Figure 2. Fitting SPT experiments. (a) Raw image of a COS-7 cell expressing GFP-Nrx1 β sparsely labeled with Atto647N-conjugated GFP nanobody (white signal). (b) Image of simulated molecules in the same geometry. Rectangles highlight the interface between the cellular region with freely diffusing GFP-Nrx1 β molecules (blue outline), and the contact region with a cell expressing Nlg1-mcherry (yellow outline). On the right, a zoom on this ROI shows single molecule trajectories for both experiments and simulations. The diffusion coefficient expressed in log scale is color coded. (c) Distributions of GFP-Nrx1 β diffusion coefficients for experiments (circles, average \pm sem of 3 cells, 4145 trajectories in contact and 4867 outside contacts) and simulations (dashed lines, 5 repetitions, 8757 trajectories in contact and 2645 trajectories outside contacts) on a semi-log plot. The Spearman correlation coefficient comparing experiment and simulation was $r = 0.85$ ($P < 0.001$, $n = 36$ bins) for contact regions, and $r = 0.81$ ($P < 0.001$, $n = 36$ bins) for outside regions. (d) Representative examples of single molecule trajectories at the interface: (left) a molecule escapes the contact and diffuses out freely, or enters the contact and gets trapped; (right) a molecule stays stuck in the contact with low diffusion, or bounces on the contact without entering it.

outside the adhesive contact ($D_{\text{out}} = 0.3 \mu\text{m}^2/\text{s}$), and was slowed down by \sim tenfold in the contact, reflecting the formation of Nrx1 β -Nlg1 bonds between apposed membranes ($D_{\text{trap}} = 0.04 \mu\text{m}^2/\text{s}$) (Fig. 2a,c). Nrx1 β molecules often bounced at the contact border, revealing steric hindrance to penetrate the narrow cell-cell junction²⁵, an effect which was described by a crossing probability ($P_{\text{crossing}} < 1$) (Fig. 2d). Next, we simulated Nrx1 β diffusion

using FluoSim (Fig. 2b). We used the diffusion coefficients obtained experimentally and defined a sparse number of molecules (250) corresponding to the number of experimental detections per frame, together with two kinetic rates describing the Nr x 1 β -Nlg1 interaction taken from the literature ($k_{\text{on}}=0.15\text{ s}^{-1}$ and $k_{\text{off}}=0.015\text{ s}^{-1}$)^{26,27}. We also chose realistic photophysics parameters ($k_{\text{on}}^{\text{Fluo}}=1\text{ s}^{-1}$ and $k_{\text{off}}^{\text{Fluo}}=3\text{ s}^{-1}$) to match experimental trajectory number and duration, and single molecule rendering parameters for Atto647N ($\sigma=0.22\text{ }\mu\text{m}$, FWHM=0.53 μm). Using these parameters, FluoSim generated diffusion coefficient distribution curves inside and outside the contact that aligned well on experimental data (Fig. 2c). Experimental distributions were somewhat more spread than theoretical ones, most likely because of local membrane heterogeneities which are not accounted for in the model.

Simulations of dSTORM experiments. To simulate dSTORM experiments that were experimentally performed on GFP-Nr x 1 β using saturating labeling with Alexa647-conjugated GFP nanobody (Fig. 3a), we introduced a relatively large number of molecules in the imported geometry (70,000, corresponding to the sum of experimental detections per frame integrated over the cell surface area), and set all diffusion coefficients to zero to mimic cell fixation. To simulate stochastic fluorescence emission of Alexa647²⁸, we calculated the switch-on ($k_{\text{on}}^{\text{Fluo}}=0.006\text{ s}^{-1}$) and switch-off ($k_{\text{off}}^{\text{Fluo}}=9.3\text{ s}^{-1}$) rates of Alexa647 fluorophores from isolated Alexa647-conjugated nanobodies in dSTORM imaging conditions (Fig. 3c). We then simulated the accumulation of single molecule localizations for 80,000 frames, including a realistic localization precision ($\sigma=25\text{ nm}$, FWHM=58 nm), to mimic the experimental super-resolved maps of Nr x 1 β distribution (Fig. 3b). Simulated images had intensity values very similar to experimental ones, but as expected from an idealistic model, could not capture specific membrane features (e.g. ruffles) seen experimentally (also see Fig. S4 for a more homogeneous cell-cell contact).

Training of deep neural networks for image reconstruction. Deep learning is becoming increasingly popular for image reconstruction in fluorescence microscopy^{19,29,30}. Convolutional Neural Networks (CNNs) are especially relevant for image treatment and have to be trained using large exemplary data sets obtained either experimentally, or from simulations. In this context, we tested the ability of FluoSim to train deep CNNs based on simulated data. We generated large image sets of randomly distributed single molecules represented as Gaussian intensity profiles plus Poisson noise, together with their localization maps as ground truth, and trained the previously described CNN called Deep-STORM¹⁹. FluoSim-trained Deep-STORM network was able to faithfully reconstruct super-resolved maps of both simulated and experimentally-observed microtubules, from single molecule localization image stacks (Figs. S3-S4). Strikingly, the CNN worked well at both low and high single molecule density, thus offering a considerable gain in acquisition time ($\times 20$) for an equivalent resolution²⁹. FluoSim-trained Deep-STORM network also allowed the reconstruction of images of Nr x 1 β -Nlg1 contacts that were comparable to images obtained with PALM-Tracer³¹, or to images directly simulated by FluoSim (Fig. S5). FluoSim can therefore be used to train deep CNNs.

Simulations of FRAP experiments. To challenge the simulator against ensemble measurements, we performed FRAP experiments on GFP-Nr x 1 β expressed in COS-7 cells. GFP-Nr x 1 β accumulates at cell-cell contacts when the opposite cell expresses its molecular partner Nlg1 (Fig. 4a) and shows slower recovery in the adhesive contact as compared to membrane regions not in contact with other cells (Fig. 4b,e). To mimic these experiments, we introduced a large number of molecules in the simulator (up to 150,000) and generated fluorescence-like images by defining a Gaussian intensity profile for each GFP-tagged molecule ($\sigma=0.17\text{ }\mu\text{m}$, FWHM=0.47 μm). Taking kinetic parameters from the literature ($k_{\text{on}}=0.15\text{ s}^{-1}$ and $k_{\text{off}}=0.015\text{ s}^{-1}$) and an intermediate porosity ($P_{\text{crossing}}=0.3$), the simulated images at steady state predicted Nr x 1 β enrichment in the contact area that matched experimental values (Fig. 4c,d). To induce local photo-bleaching, we chose a bleaching rate ($k_{\text{off}}^{\text{bleach}}=4.25\text{ s}^{-1}$) reproducing the initial drop of fluorescence observed experimentally ($\sim 75\%$ in 400 ms). Using those coefficients plus the diffusion parameters determined from SPT, FRAP simulations accurately reproduced experimental data (Fig. 4e).

Simulations of PAF experiments. PAF, which can be viewed as the mirror of a FRAP experiment, is also very popular to estimate membrane molecule dynamics in specific cellular compartments³². We thus confronted FluoSim with actual PAF experiments performed on COS-7 cells expressing Nr x 1 β fused to photoactivatable GFP (PAGFP-Nr x 1 β). Because PAGFP is non-fluorescent in the absence of photo-stimulation, cells were co-electroporated with BFP for detection. Characteristic adhesive contacts were formed between cells expressing PAGFP-Nr x 1 β , and co-cultured cells expressing Nlg1-mcherry (Fig. 5a,c). PAGFP-Nr x 1 β was turned on with a brief pulse of 405 nm laser and the fluorescence decay due to protein diffusion and turnover was recorded for 2 min (Fig. 5b,d). As expected from SPT and FRAP data, the fluorescence decay was slower for PAGFP-Nr x 1 β making contact with Nlg1-mcherry, compared to freely diffusing PAGFP-Nr x 1 β outside contact regions (Fig. 5e). In FluoSim, we introduced similar numbers of molecules, diffusion coefficients and kinetic rates as for FRAP. To match the sudden increase in PAGFP-Nr x 1 β fluorescence intensity at time zero, the photo-activation rate was raised from 0 to 1 s^{-1} for 300 ms. Given this set of parameters, the PAF simulations fit very well the experimental data (Fig. 5e).

Simulations of FCS experiments. To further validate the simulator in conditions of intermediate molecular density, we performed FCS experiments³³ by recording intensity fluctuations over time for Atto647N-nanobody bound to GFP-Nr x 1 β , in a diffraction-limited laser spot (Fig. 6a,b). As expected from slower diffusion, the autocorrelation function obtained for Nr x 1 β in the adhesive contact was shifted to the right compared to the

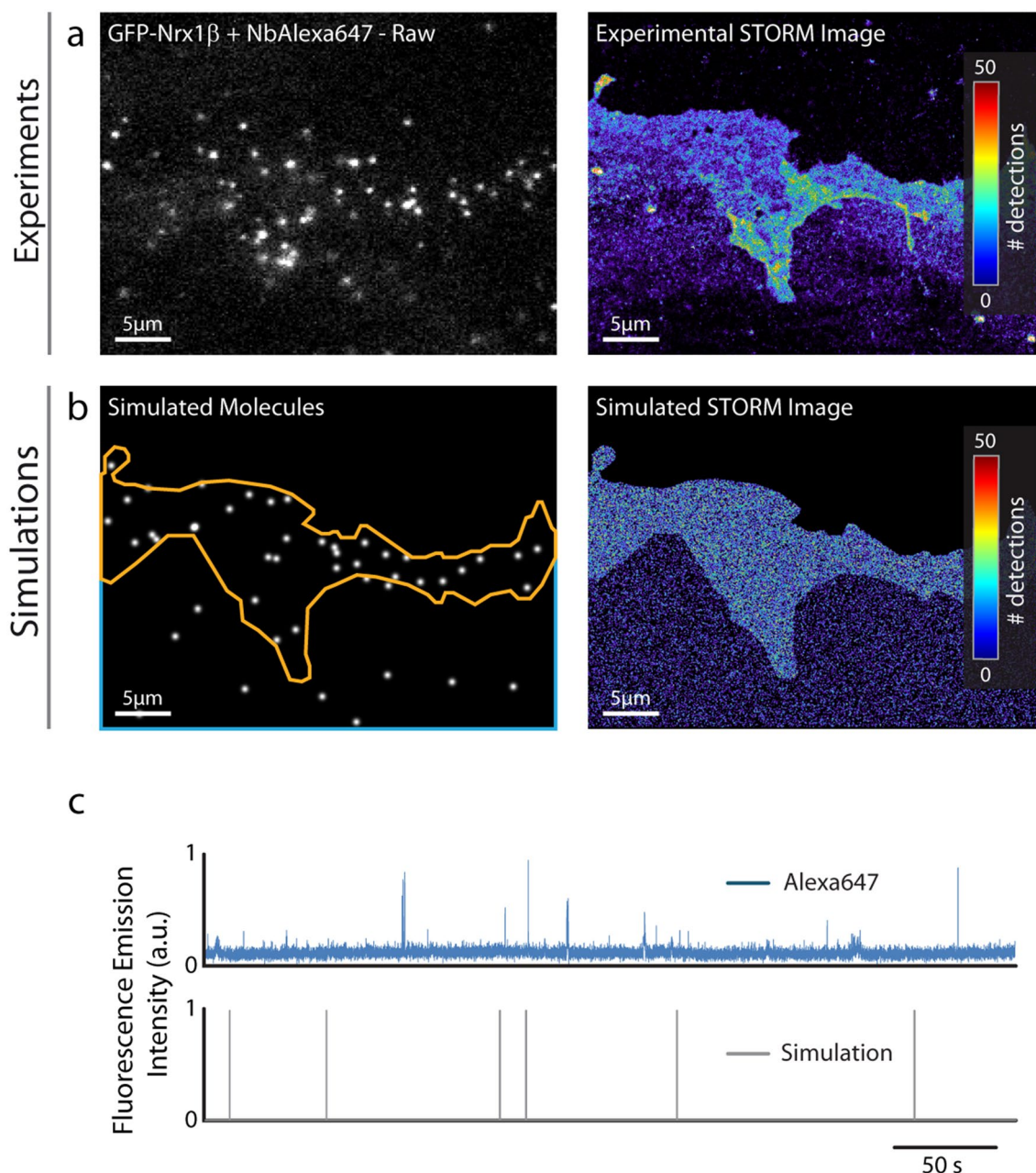


Figure 3. Fitting SRI experiments. **(a)** Representative single frame image of a STORM sequence performed on GFP-Nrx1 β labeled with Alexa647-conjugated GFP nanobody, and corresponding super-resolved image generated from 3.57×10^6 single molecule localizations (pixel size 32 nm, total acquisition time 1600 s). The number of molecules detected per pixel is color coded. **(b)** Simulated image showing single molecule fluorescence emission in the same cell geometry, and corresponding super-resolved map with a localization precision of 58 nm (FWHM). The total number of detections is 3.61×10^6 . **(c)** Fluorescence emission over time from an isolated Alexa647-conjugated GFP nanobody bound to the glass coverslip in a STORM sequence, and simulated emission of fluorescence of an immobile molecule obtained with switch-on and -off rates of 0.006 s^{-1} and 9.3 s^{-1} , respectively.

free region, with a small photo-bleaching bias due to longer residence times (Fig. 6e). Again, the simulations performed with an intermediary number of molecules introduced in FluoSim, matched experimental FCS curves with the same set of kinetic and photophysical parameters used for SPT, FRAP, and PAF (Fig. 6c–e).

Discussion

In summary, FluoSim allows a prediction and comparison of membrane protein dynamics in a wide range of fluorescence cell imaging modes, with a precise control of the relevant kinetic, photo-physical, and acquisition parameters determining measurement outputs. The program is intended to help biologists adjust and interpret their experiments on a variety of cellular systems, and serve as a teaching resource in bio-imaging programs.

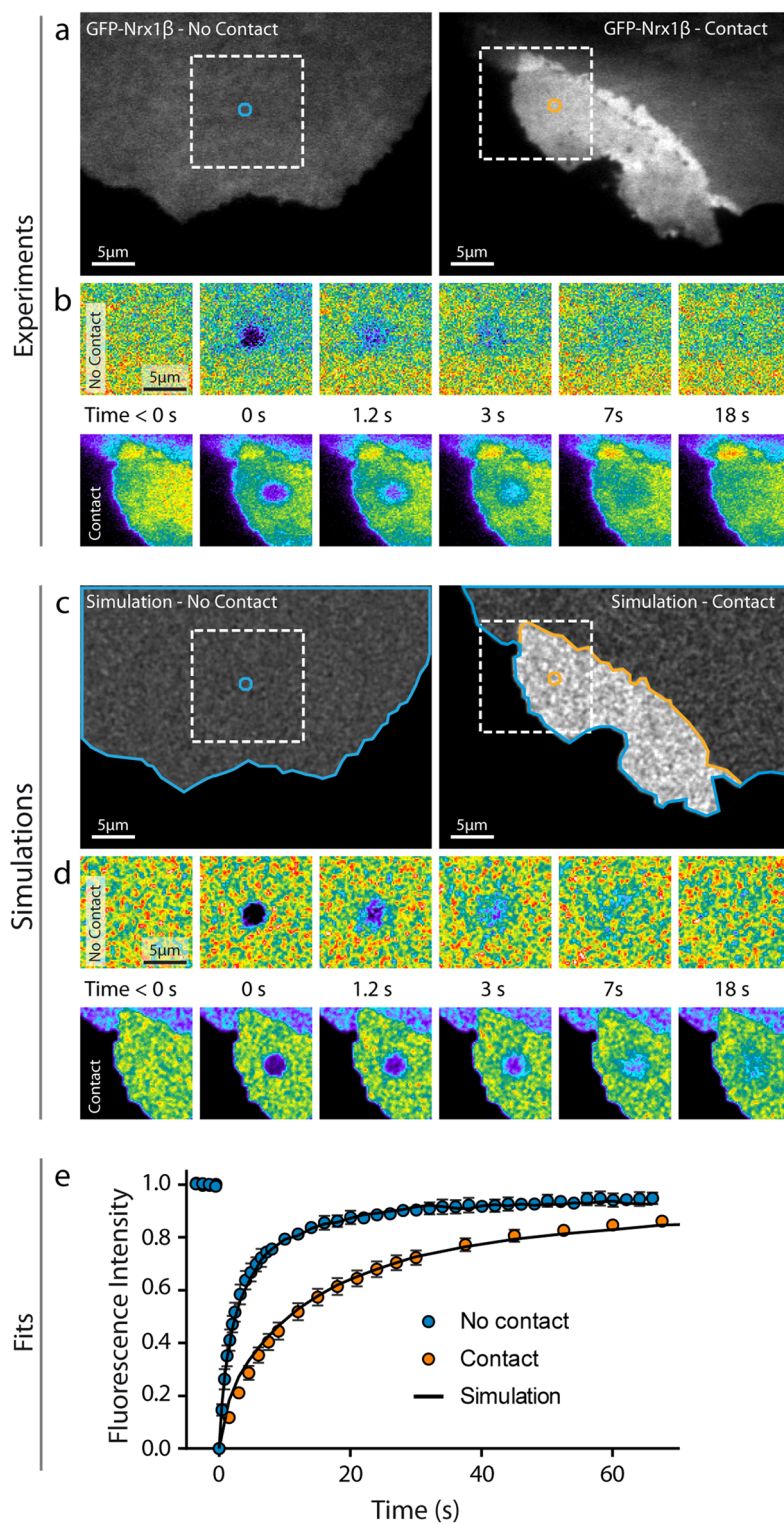


Figure 4. Fitting FRAP experiments. **(a)** Representative images of COS-7 cells expressing GFP-Nrx1 β either not forming contact (left), or forming contact with another cell expressing Nlg1-mCherry (right). **(b)** Corresponding FRAP sequences on the zoomed square areas (bleached circle diameter = 2.8 μ m). Intensity is color coded. **(c,d)** Simulated images of the same cell geometries filled with 150,000 fluorescent molecules, and corresponding FRAP sequences. **(e)** Normalized FRAP curves obtained by experiment (closed circles) outside (mean \pm sem of 6 cells, 44 bleached regions) or inside contact regions (mean \pm sem of 18 cells, 144 bleached regions) and corresponding simulations (solid curves, average of 30 repetitions each, sem < 1% mean, not shown). The Spearman correlation coefficient comparing experiment and simulation was $r = 1.0$ ($P < 0.001$, $n = 21$ time points) for contact regions, and $r = 0.98$ ($P < 0.001$, $n = 43$ time points) for outside regions.

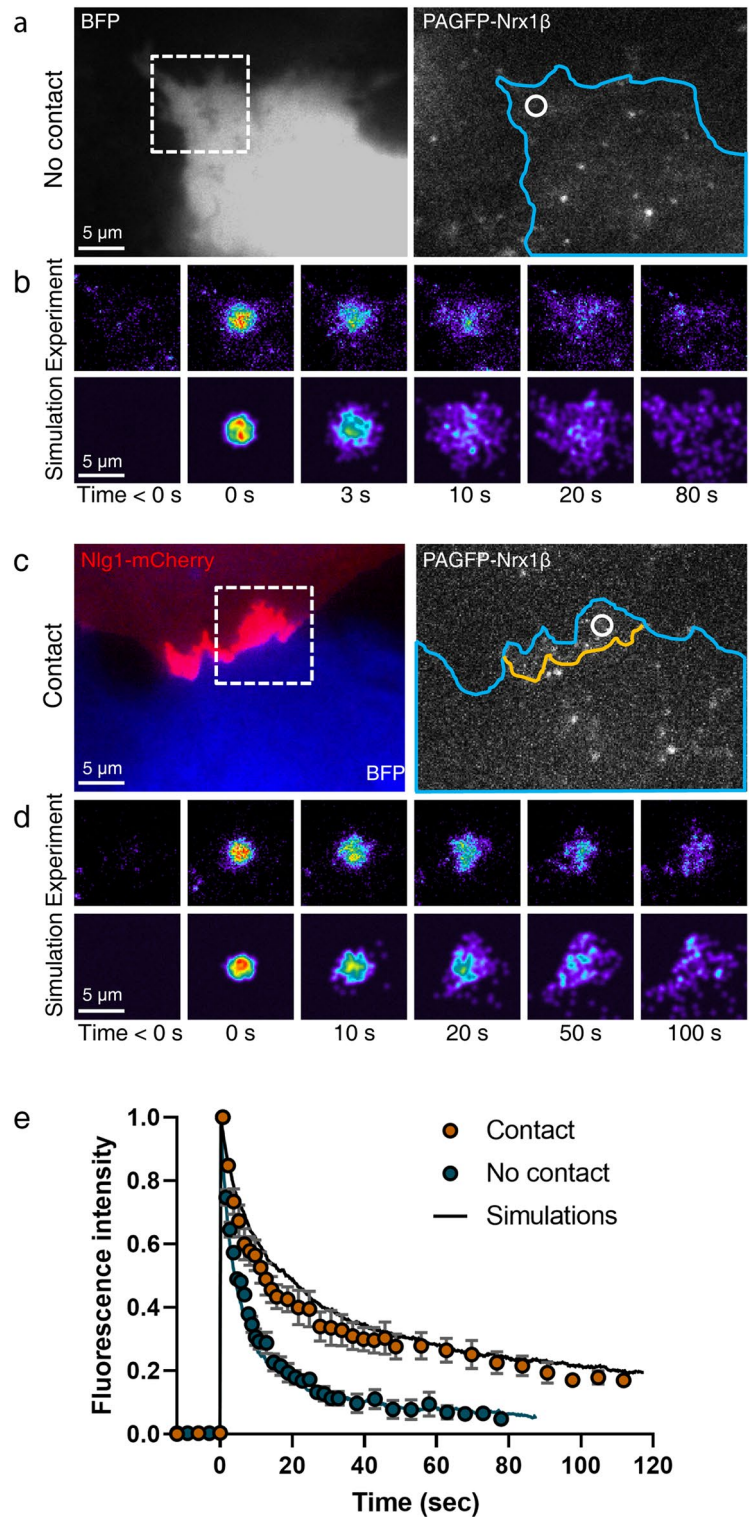


Figure 5. Fitting PAF experiments. (a,c) Representative images of COS-7 cells co-expressing BFP (blue) plus PAGFP-Nrx1β, and either not forming contact, or forming contact with another cell expressing Nlg1-mCherry (red). The cell geometry (blue contour) is entered in the simulator, together with the contact region (yellow line). (b,d) Corresponding PAF sequences on zoomed areas (dashed squares), where the photoactivated spot is indicated in white (circle of diameter = 2.8 μm). The experimental and simulated images are shown on top of each other for the indicated time points. Fluorescence intensity is color coded. (e) Graph showing the fluorescence intensity normalized between 0 and 1 for PAF experiments performed outside (blue circles) or inside (orange circles) contact regions (mean ± sem of 5 cells/1 PAF region per cell, and 3 cells/2 PAF regions per cell, respectively). The corresponding simulations (solid curves) represent the average of 20 repetitions (sem < 2% mean, not shown). The Spearman correlation coefficients comparing experiment and simulation was $r = 0.998$ for contact regions, and $r = 0.995$ for outside regions ($P < 0.0001$, $n = 32$ time points in both conditions).

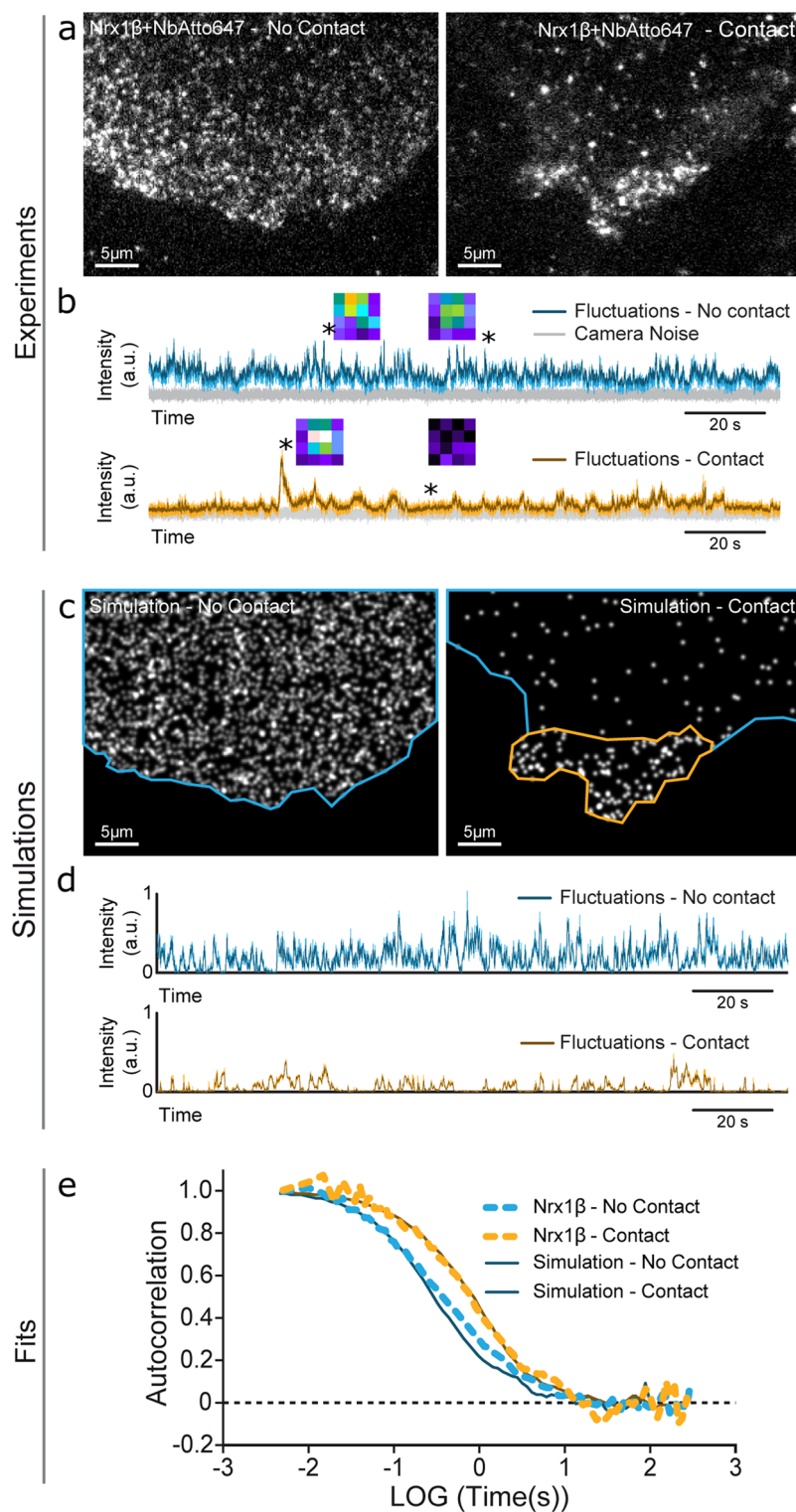


Figure 6. Fitting FCS experiments. **(a)** Images of Atto647N-conjugated GFP nanobody bound to COS cells expressing GFP-Nrx1 β not forming contact (left), or forming contact with a cell expressing Nlg1-mCherry (right). A 642 nm focused laser beam of 0.6 μ m FWHM was parked in contact or no-contact regions, and square images of the illuminated area (1 μ m \times 1 μ m) were collected at 200 or 500 Hz, respectively. **(b)** Intensity fluctuations over time in the two regions (color), above the camera noise (grey). Insets represent images acquired at the indicated times (stars). Intensity is color coded. **(c,d)** Simulated images using the same cell geometries populated with 2500 and 200 molecules, respectively, and corresponding intensity fluctuations. **(e)** Normalized autocorrelation functions for FCS experiments (dashed curves) performed outside (2 cells/10 recordings) and inside contact regions (1 cell/3 recordings), and corresponding simulations (solid curves, average of 10 repetitions each).

Category	Parameter	Notation	Unit/format
Geometry	Reference image		.TIFF, PNG, JPG, GIF
	Image pixel size		$\mu\text{m}/\text{pixel}$
	Region file		.rgn or .roi
	FRAP laser spot		2.8 μm
	FCS laser spot width	σ_{FCS}	0.25 μm
Molecules	Number ^a		1–150,000
	Initial position		Uniform or estimated
Times	Simulation duration ^a		1000–100,000 frames (40–1600 s)
	Time step	Δt	2–100 ms
Diffusion coefficients	Outside contact	D_{out}	0.3 $\mu\text{m}^2/\text{s}$
	Inside contact	D_{in}	0.3 $\mu\text{m}^2/\text{s}$
	Trapped	D_{trap}	0.04 $\mu\text{m}^2/\text{s}$
	Crossing probability	P_{crossing}	25–70%
Kinetics	Binding rate	k_{on}	0.15 s^{-1}
	Unbinding rate	k_{off}	0.015 s^{-1}
Photophysics	Switch-on rate ^a	$k_{\text{on}}^{\text{Fluo}}$	0.005–5 s^{-1}
	Switch-off rate ^a	$k_{\text{off}}^{\text{Fluo}}$	0–10 s^{-1}
	Bleaching rate ^a	$k_{\text{off}}^{\text{bleach}}$	0.4–4 s^{-1}
	Activation rate ^a	$k_{\text{off}}^{\text{activ}}$	1–2 s^{-1}
Export files	SPT or SRI images		.TIFF or multi-TIFF
	SPT trajectories		.trc
	SPT histogram		.txt
	FRAP curves		.txt
	FCS fluctuations and autocorrelation		.txt

Table 1. Simulator parameters. ^aSee the methods above for the specific molecule numbers and photo-physical parameters used in the various imaging modes (SPT, STORM, FRAP, PAE, FCS).

FluoSim reproduced a wide range of experimental results (SPT, FRAP, PAE, FCS, SRI) on the $\text{Nrx1}\beta\text{-Nlg1}$ membrane complex using a unique set of parameters extracted from published in-vitro studies and/or taken from our own measurements^{26,27}, thereby giving strong credit to the correlative approach. The program is very fast and robust, and should be applicable to model a wide range of 2D-like dynamic molecular systems experiencing membrane diffusion and transient confinement, for example integrins at focal contacts in fibroblasts^{34,35}, cadherins at cell–cell contacts^{36–38}, neuronal adhesion proteins and neurotransmitter receptors at synapses^{24,39,40}, and trapping of membrane molecules by lipid rafts or cytoskeletal interactions^{41–43}.

Independently of changes in the fluorescence parameters, FluoSim can also be used to predict changes in the sub-cellular distribution of membrane proteins over time, induced by the modification of diffusion coefficients (D_{out} , D_{in} , D_{trap}) or kinetic rates (k_{on} , k_{off}). For example, the effect of cross-linking surface receptors with antibodies can be simulated by a decrease in diffusion coefficients^{44,45}. Moreover, changes in the parameters k_{on} and k_{off} in FluoSim can mimic the action of pharmacological agents, competing peptides, and chemo- or opto-genetic stimulations that modulate specific protein–protein interaction rates. For instance, cell-permeant peptides competing with glutamate receptor anchorage to scaffolding proteins rapidly decrease receptor density at synapses⁴⁶, a process that is mimicked by changing the binding rates of receptors to their post-synaptic scaffold⁴⁵. Similarly, long term potentiation (LTP) and long term depression (LTD) can be simulated by modulating the AMPA receptor/scaffold affinity, providing insight on the molecular mechanisms underlying synaptic plasticity^{45,47}.

We believe that this quantitative simulation approach will be of great help to optimize experimental design, especially regarding the choice of various parameters such as fluorophores, control of laser powers, acquisition frame rate, and overall timing of the experiment with respect to the internal dynamics of the molecular system, thus replacing lengthy experimental adjustments and saving research time. The software can also be used to test the robustness and predictions of single molecule tracking algorithms that have been implemented those past years, e.g. SR-Tesseler and InferenceMap^{48–50}. Compared to existing software such as PyFRAP, SuReSim, FERNET, or MCell^{13–15}, FluoSim integrates many fluorescence modalities into a single program and achieves real-time display (Supplementary Table 1). In its present version, FluoSim is limited to Brownian motion and first order molecular reactions, but sub- or super-diffusive behaviors as well as more complex multi-state molecular reactions might be implemented on a case-by-case basis, depending on user needs (Table 1).

Finally, with the advent of single-molecule based super-resolution microscopy such as PALM and STORM, many new questions arise regarding the degree of labeling needed, and the number of single molecule localizations to accumulate in order to reconstruct a realistic image of biological structures, with the risk of finding artificial properties in under-sampling conditions¹⁰. By varying the number of molecules, type of fluorophore attached to them, localization precision, and length of the acquisition sequence, the software is able to determine

the conditions for faithful detection of the biological sample. The capacity of FluoSim to mimic molecule dynamics might also be used in the near future to train the next generation of CNNs for the analysis of live imaging experiments⁵¹.

Methods

DNA plasmids. GFP-Nrx1 β was a gift from M. Missler (Münster University, Germany). PA-GFP-Nrx1 β was generated by amplifying PA-GFP³² (a gift of J. Lippincott-Schwartz, Janelia Research Campus, USA) with the two primers PAGFP-3F (5'-cgGCTAGCGGAGCAGGAatggtgagcaaggcgagg-3') and PAGFP-4R (5'-cgGCTAGC TCCTGCTCCCTTGACAGCTCGTCCATGC-3') and inserting it into GFP-Nrx1 β in place of GFP at NheI/NheI sites. Proper insert orientation was verified by restriction and sequencing. HA-tagged Nlg1 obtained from P. Scheiffele (Biozentrum, Basel) was used as a backbone to construct Nlg1-mCherry, by inserting mCherry intracellularly at position -21aa before the C-terminus. The HA-Nlg1 sequence was moved from the pNice vector into a pcDNA vector at the HindIII/NotI sites. Two PCRs were performed on Nlg1: the first one from the KpnI site (located inside the Nlg1 sequence) to the insert position of mCherry (AgeI site added), and the other one from the mCherry site insertion (Nhe added) to the end of Nlg1 (NotI). The mCherry gene (AgeI/NheI) was obtained by PCR on pmCherry-N1 (Clontech). A four-fragment ligation was then done to obtain the final construct (HindIII-HA-Nlg1-AgeI-mCherry-NheI-Nlg1Cter-NotI). The plasmid for bacterial expression of the anti-GFP nanobody⁵² was a kind gift from A. Gautreau (Gif-sur-Yvette, France). The bacterial production of anti-GFP nanobody, purification, and conjugation to organic dyes (Atto647N or Alexa647), was described previously²⁴.

Cell culture and electroporation. COS-7 cells (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO/BRL) supplemented with 10% fetal bovine serum (FBS), 100 units mL⁻¹ penicillin, and 100 μ g mL⁻¹ streptomycin, in a 37 °C-5% CO₂ atmosphere. One day before the experiments, cells were rinsed twice in warm PBS, trypsinized for 5 min, mixed in culture medium, and centrifuged for 5 min at 1000 rpm. The cell pellet was suspended in 100 μ L electroporation medium and electroporated for either GFP-Nrx1 β or Nlg1-mCherry plasmids with the Amaxa Nucleofector system (Lonza), using 500,000 cells per cuvette and 3 μ g DNA. Electroporated cells were mixed in culture medium, seeded on 18 mm glass coverslips at a concentration of 50,000–80,000 cells per coverslip, cultured in 12-well plates, and imaged 24–48 h after electroporation.

Single molecule pull-down. Single molecule pull-down of GFP-Nrx1 β molecules was performed essentially as described^{53,54}. Briefly, cleaned glass coverslips (VWR) were treated with N-(2-aminoethyl)-3-aminopropyltrimethoxysilane (United Chemical Technologies, #A0700), then incubated with mPEG-succinimidyl valerate (SVA) containing a 1:100 fraction of biotin PEG-SVA (both from Laysan Bio MW 5000 Da). Coverslips were dried and stored at -20 °C. Just before the experiment, coverslips were incubated with 1 μ M NeutrAvidin for 5 min (Invitrogen, A2666), followed by 10 nM biotinylated anti-GFP antibody for 15 min (ABCAM, 6658), both diluted in T50 buffer (10 mM Tris-HCl, 50 mM NaCl, pH = 7.5). Two days before the experiment, HEK cells were transfected with GFP-Nrx1 β using Lipofectamine (Invitrogen) in optiMEM buffer. Two hours before the experiment, cells were detached in Ca-free PBS (30 min, 37 °C), then centrifuged at 3000 rpm for 2 min. The cell pellet was dissolved in 100 μ L lysis buffer containing 150 mM NaCl, 10 mM Tris-HCl, 1 mM EDTA, 1% Igepal, and protease inhibitor cocktail (Pierce). Samples were rotated for 2 h at 4 °C to solubilize proteins, then centrifuged at 14,000 rpm for 15 min to remove aggregates. The supernatant was kept on ice and diluted 1:200 in observation solution (135 mM NaCl, 5 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂, 10 mM HEPES, 4 mM Trolox, 40 mM D-glucose, 0.03% Igepal, pH = 7.4). 100 μ L of this mix was added to glass substrates for several minutes, then rinsed out with fresh solution when the molecular density reached about 100 per field of view (65 \times 65 μ m). Single GFP-Nrx1 β molecules were observed with a 488 nm laser under TIRF illumination on an inverted Olympus microscope equipped with a 63 \times /1.45 NA objective and an EMCCD camera (Andor Ixon). GFP-Nrx1 β molecules did not stick to a control coverslip containing no anti-GFP antibody, revealing binding specificity to the GFP tag. Acquisition sequences of 500 frames were recorded at 30 Hz, resulting in the photobleaching of nearly all immobilized GFP-Nrx1 β molecules. Images were analyzed using Metamorph by measuring the average fluorescence intensity in small regions drawn around individual molecules.

Single molecule tracking (uPAINT). Universal point accumulation in nanoscale topography (uPAINT) experiments were carried out as reported⁸. Cells were mounted in Tyrode solution (15 mM D-glucose, 108 mM NaCl, 5 mM KCl, 2 mM MgCl₂, 2 mM CaCl₂ and 25 mM HEPES, pH 7.4) containing 1% globulin-free BSA (Sigma A7638) in an open Inox observation chamber (Life Imaging Services, Basel, Switzerland). The chamber was placed on a fully motorized inverted microscope (Nikon Ti-E Eclipse) equipped with perfect focus system, a thermostatic box (Life Imaging Services) providing air at 37 °C, and an APO TIRF 100 \times /1.49 NA oil immersion objective. GFP- and mCherry expressing cells were detected using a mercury lamp (Nikon Xcite) and the following filter sets (SemROCK): EGFP (Excitation: FF01-472/30; Dichroic: FF-495Di02; Emission: FF01-525/30) and mCherry (Excitation: FF01-543/22; Dichroic: FF-562Di02; Emission: FF01-593/40). Cells expressing GFP-Nrx1 β were labeled using a low concentration of Atto647N-conjugated GFP nanobody (1 nM). A four-colour laser bench (405/488/561 nm lines, 100 mW each; Roper Scientific, and 1 W 642 nm line, MPB Communications Inc., Canada) is connected through an optical fiber to the Total Internal Reflection Fluorescence (TIRF) illumination arm of the microscope. Laser power was controlled through an acousto-optical tunable filter (AOTF) driven by the Metamorph software (Molecular Devices). Atto647N was excited with the 642 nm laser line (~2 mW at the objective front lens), through a four-band beam splitter (BS R405/488/561/635, SemRock). Samples were imaged by oblique laser illumination, allowing the excitation of individual Atto-conjugated ligands bound to the cell

surface, without illuminating ligands in solution. Fluorescence was collected on an EMCCD camera with 16 μm pixel size (Evolve, Roper Scientific, Evry, France), using a FF01-676/29 nm emission filter (SemRock). Stacks of 4000 consecutive frames were obtained from each cell with an integration time of 20 ms, using the Nikon perfect focus system to avoid axial drift. Images were analyzed using PALM-Tracer, a program running on Metamorph and based on wavelet segmentation for molecule localization and simulated annealing algorithms for tracking (generously provided by JB Sibarita, Bordeaux)⁵⁵. This program allows the tracking of localized molecules through successive images. Trajectories longer than 20 frames (400 ms) were selected. The diffusion coefficient, D , was calculated for each trajectory, from linear fits of the first 4 points of the mean square displacement (MSD) function versus time. Trajectories with displacement inferior to the pointing accuracy (~ 50 nm in uPAINT conditions) whose MSD function cannot be fitted are arbitrarily taken as $D = 10^{-5} \mu\text{m}^2 \text{s}^{-1}$.

dSTORM experiments. COS-7 cells expressing GFP-Nrx1 β were surface-labeled with a high concentration (100 nM) of Alexa647-conjugated GFP Nanobody in Tyrode solution containing 1% globulin-free BSA (Sigma A7638) for 10 min, rinsed and fixed with 4% PFA-0.2% glutaraldehyde in PBS for 10 min at room temperature, and stored in PBS at 4 °C until imaging. For microtubule staining, COS-7 cells were rinsed twice in PBS, fixed using 4% PFA-20% sucrose for 15 min at room temperature, washed 3 times in PBS, and incubated with 50 mM NH_4Cl in PBS for 10 min. After fixation, cells were washed 3 times in PBS, permeabilized using 0.2% Triton-X 100 for 10 min, washed in PBS, blocked in 1% BSA-PBS for 30 min, and incubated with anti- α -tubulin (ThermoFisher MA1-80017, 1:500) overnight at 4 °C. The next day, cells were washed 3 times in PBS, and incubated 45 min with secondary goat-anti-rat-Alexa647 antibody (ThermoFisher A21247, 1:800) and kept in PBS before dSTORM imaging.

Cells were imaged in Tris-HCl buffer (pH 7.5) containing 10% glycerol, 10% glucose, 0.5 mg/mL glucose oxidase (Sigma), 40 mg/mL catalase (Sigma C100-0.1% w/v), and 50 mM β -mercaptoethylamine (MEA) (Sigma M6500)⁵⁶. The same microscope described for uPAINT was used. Pumping of Alexa647 dyes into their triplet state was performed for several seconds using ~ 60 mW of the 642 nm laser at the objective front lens. Then, a lower power (~ 30 mW) was applied to detect the stochastic emission of single-molecule fluorescence, which was collected using the same optics and detector as described above for uPAINT. 10–20 streams of 3000–4000 frames each were acquired at 50 Hz. To generate images intended to test deep CNN algorithms, a higher density of fluorescent molecules was generated by turning on the 405 nm laser power to 5 mW during acquisition. Multi-color 100-nm fluorescent beads (Tetraspeck, Invitrogen) were used to register long-term acquisitions and correct for lateral drift. The localization precision of our imaging system in STORM conditions is around 60 nm (FWHM). Stacks were analyzed using the PALM-Tracer program, allowing the reconstruction of a unique super-resolved image of 32 nm pixel size (zoom 5 compared to the original images) by summing the intensities of all localized single molecules (1 detection per frame is coded by an intensity value of 1).

FRAP experiments and analysis. COS-7 cells expressing GFP-Nrx1 β in co-culture with cells expressing Nlg1-mCherry were mounted in Tyrode solution, and observed under the same set-up used for uPAINT and STORM. The laser bench has a second optical fiber output connected to an illumination device containing two x/y galvanometric scanning mirrors (ILAS, Roper Instrument) steered by MetaMorph. It allows precise spatial and temporal control of the focused laser beam at any user-selected region of interest within the sample for targeted photo-bleaching. Switching between the two fibers for alternating between imaging and bleaching is performed in the ms time range using an AOTF. Oblique illumination was performed using the 491 nm beam at low power (0.3 mW at the front of the objective) to image GFP-Nrx1 β molecules in the plasma membrane close to the substrate plane. After acquiring a 10 s baseline at 1 Hz frame rate, rapid selective photo-bleaching of 3–9 circular areas of diameter 2.8 μm was achieved at higher laser power (3 mW at the objective front lens), during 400 ms. Fluorescence recovery was then recorded immediately after the bleach sequence for 80 s. The recording period included three phases with decreasing frame rate ranging from 2 to 0.1 Hz. Observational photo-bleaching was kept very low, as assessed by observing control unbleached areas nearby. FRAP curves were obtained by computing the average intensity in the photobleached area, after background subtraction, and normalized between 1 (baseline) and 0 (time zero after photo-bleaching).

PAF experiments and analysis. COS-7 cells co-expressing BFP and PAGFP-Nrx1 β in co-culture with cells expressing Nlg1-mCherry were mounted in Tyrode solution, and observed under the microscope as described for FRAP experiments. Oblique illumination was carried out using the 491 nm beam at low power (0.3 mW at the front of the objective) to image membrane PAGFP-Nrx1 β molecules close to the substrate. After acquiring a 12 s baseline at 0.3 Hz frame rate, photo-activation of 1–2 circular areas of diameter 2.8 μm was achieved with the focused 405 nm laser beam (1.5 mW at the objective front lens), during 300 ms. Fluorescence decay was then recorded for 80–100 s, with decreasing frame rate ranging from 2 to 0.1 Hz. PAF curves were normalized between 0 and 1, by dividing the intensity in the photoactivated area over time by the maximal intensity recorded immediately after the photo-activation pulse, after subtracting the baseline from both values.

FCS experiments and analysis. COS-7 cells expressing GFP-Nrx1 β were surface-labeled at an intermediate concentration (10 nM) of Atto647N-conjugated GFP nanobody for 5 min at 37 °C, then mounted in Tyrode solution on the same microscope. The 642 nm laser beam was parked on a region of interest, either on the free cell membrane, or in the cell–cell contact area. The laser was kept at low power (3% of 100 mW, i.e. ~ 0.2 mW at the objective front lens) to avoid photo-bleaching the organic dye. Using glass coverslips coated with polylysine and higher concentrations of Atto647N-conjugated nanobody (100 nM) to provide a uniform density, we independently measured the Gaussian intensity profile of the laser beam, which gave a standard deviation

$\sigma_{\text{FCS}} = 0.26 \mu\text{m}$ (FWHM = $0.6 \mu\text{m}$). Intensity fluctuations due to Nr $x1\beta$ molecules entering and leaving the laser spot by membrane diffusion, were detected using a second camera (Hamamatsu Orca Flash 4.0) on the opposite port of the microscope. A square area of 16×16 pixels ($1 \mu\text{m} \times 1 \mu\text{m}$) centered on the laser spot was chosen, and streams of 60,000 images at binning 4 were acquired through the Hokawo software (Hamamatsu) at 200 Hz or 500 Hz, for the study of cell–cell contacts or free regions, respectively. Control stacks with the laser off were acquired to record the camera readout noise. The integrated intensity of each image was read from the multi-TIFF stacks, from which the average camera noise was subtracted, and the autocorrelation function was computed using custom-made routines written in C++, and normalized by its first value.

FluoSim simulations. A general description of the FluoSim algorithm and the list of simulations performed to match the experiments, together with the precise set of parameters used in each imaging mode, is available in the Supplementary material.

Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its supplementary information files. The software FluoSim (v1.0) is released under a GNU GPL v3 license as supplementary material accompanying this manuscript, and as an archive that can be downloaded at <https://www.iins.u-bordeaux.fr/SOFTWARE>. The FluoSim (v1.0) release and its source code are available at <https://github.com/mlagardere/FluoSim> under the same license.

Received: 6 February 2020; Accepted: 28 September 2020

Published online: 17 November 2020

References

1. Choquet, D. & Triller, A. The dynamic synapse. *Neuron* **80**, 691–703 (2013).
2. Dustin, M. L. & Choudhuri, K. Signaling and polarized communication across the T cell immunological synapse. *Annu. Rev. Cell Dev. Biol.* **32**, 303–325 (2016).
3. Liu, Z., Lavis, L. D. & Betzig, E. Imaging live-cell dynamics and structure at the single-molecule level. *Mol. Cell* **58**, 644–659 (2015).
4. Lippincott-Schwartz, J., Snapp, E. L. & Phair, R. D. The development and enhancement of FRAP as a key tool for investigating protein dynamics. *Biophys. J.* **115**, 1146–1155 (2018).
5. Manley, S. *et al.* High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nat. Methods* **5**, 155–157 (2008).
6. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
7. Dani, A., Huang, B., Bergan, J., Dulac, C. & Zhuang, X. Superresolution imaging of chemical synapses in the brain. *Neuron* **68**, 843–856 (2010).
8. Giannone, G. *et al.* Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophys. J.* **99**, 1303–1310 (2010).
9. Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: Super-resolution imaging of cells. *Cell* **143**, 1047–1058 (2010).
10. Deschout, H. *et al.* Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods* **11**, 253–266 (2014).
11. Stefan, M. I., Bartol, T. M., Sejnowski, T. J. & Kennedy, M. B. Multi-state Modeling of Biomolecules. *PLoS Comput. Biol.* **10**, (2014).
12. Kerr, R. A. *et al.* Fast Monte Carlo simulation methods for biological reaction-diffusion systems in solution and on surfaces. *SIAM J. Sci. Comput.* **30**, 3126 (2008).
13. Bläßle, A. *et al.* Quantitative diffusion measurements using the open-source software PyFRAP. *Nat. Commun.* **9**, 1–14 (2018).
14. Venkataramani, V., Herrmannsdörfer, F., Heilemann, M. & Kuner, T. SuReSim: simulating localization microscopy experiments from ground truth models. *Nat. Methods* **13**, 319–321 (2016).
15. Angiolini, J., Plachta, N., Mocskos, E. & Levi, V. Exploring the dynamics of cell processes through simulations of fluorescence microscopy experiments. *Biophys. J.* **108**, 2613–2618 (2015).
16. Ibrahim, B. *et al.* Spatial rule-based modeling: a method and its application to the human mitotic kinetochore. *Cells* **2**, 506–544 (2013).
17. Tapia, J.-J. *et al.* MCell-R: a particle-resolution network-free spatial modeling framework. *Methods Mol. Biol.* **1945**, 203–229 (2019).
18. Tolle, D. P. & Le Novère, N. Meredys, a multi-compartment reaction–diffusion simulator using multistate realistic molecular complexes. *BMC Syst. Biol.* **4**, 24 (2010).
19. Nehme, E., Weiss, L., Michaeli, T. & Shechtman, Y. Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica* **18**, 2334–2536 (2018).
20. Mandracchia, B. *et al.* Fast and accurate sCMOS noise correction for fluorescence microscopy. *Nat. Commun.* **11**, 1–12 (2020).
21. Foi, A., Trimeche, M., Katkovnik, V. & Egiazarian, K. Practical Poissonian–Gaussian noise modeling and fitting for single-image raw-data. *IEEE Trans. Image Process.* **17**, 1–18 (2007).
22. Südhof, T. C. Synaptic neuroligin complexes: a molecular code for the logic of neural circuits. *Cell* **171**, 745–769 (2017).
23. Ries, J., Kaplan, C., Platonova, E., Eghlidi, H. & Ewers, H. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. *Nat. Methods* **9**, 582–584 (2012).
24. Chamma, I. *et al.* Mapping the dynamics and nanoscale organization of synaptic adhesion proteins using monomeric streptavidin. *Nat. Commun.* **7**, 10773 (2016).
25. Tanaka, H. *et al.* Higher-order architecture of cell adhesion mediated by polymorphic synaptic adhesion molecules neuroligin and neuroligin. *Cell Rep.* **2**, 101–110 (2012).
26. Comoletti, D. *et al.* Characterization of the Interaction of a recombinant soluble neuroligin-1 with neuroligin-1 β . *J. Biol. Chem.* **278**, 50497–50505 (2003).
27. Saint-Michel, E., Giannone, G., Choquet, D. & Thoumine, O. Neuroligin/neuroligin interaction kinetics characterized by counting single cell-surface attached quantum dots. *Biophys. J.* **97**, 480–489 (2009).
28. Dempsey, G. T., Vaughan, J. C., Chen, K. H., Bates, M. & Zhuang, X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. *Nat. Methods* **8**, 1027–1040 (2011).
29. Ouyang, W., Aristov, A., Lelek, M., Hao, X. & Zimmer, C. Deep learning massively accelerates super-resolution localization microscopy. *Nat. Biotechnol.* **36**, 460–468 (2018).
30. Weigert, M. *et al.* Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).

31. Kechkar, A., Nair, D., Heilemann, M., Choquet, D. & Sibarita, J.-B. Real-time analysis and visualization for single-molecule based super-resolution microscopy. *PLoS ONE* **8**, e62918 (2013).
32. Patterson, G. H. & Lippincott-Schwartz, J. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* **297**, 1873–1877 (2002).
33. Haustein, E. & Schwille, P. Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 151–169 (2007).
34. Rossier, O. *et al.* Integrins β 1 and β 3 exhibit distinct dynamic nanoscale organizations inside focal adhesions. *Nat. Cell Biol.* **14**, 1057–1067 (2012).
35. Ballestrem, C., Hinz, B., Imhof, B. A. & Wehrle-Haller, B. Marching at the front and dragging behind: differential α V β 3-integrin turnover regulates focal adhesion behavior. *J. Cell Biol.* **155**, 1319–1332 (2001).
36. Perez, T. D., Tamada, M., Sheetz, M. P. & Nelson, W. J. Immediate-early signaling induced by E-cadherin engagement and adhesion. *J. Biol. Chem.* **283**, 5014–5022 (2008).
37. Garcia, M. *et al.* Two-tiered coupling between flowing actin and immobilized N-cadherin/catenin complexes in neuronal growth cones. *Proc. Natl. Acad. Sci.* **112**, 6997–7002 (2015).
38. Chazeau, A. *et al.* Mechanical coupling between transsynaptic N-cadherin adhesions and actin flow stabilizes dendritic spines. *Mol. Biol. Cell* **26**, 859–873 (2015).
39. Nair, D. *et al.* Super-resolution imaging reveals that AMPA receptors inside synapses are dynamically organized in nanodomains regulated by PSD95. *J. Neurosci.* **33**, 13204–13224 (2013).
40. Renner, M., Wang, L., Levi, S., Hennekinne, L. & Triller, A. A simple and powerful analysis of lateral subdiffusion using single particle tracking. *Biophys. J.* **113**, 2452–2463 (2017).
41. Kenworthy, A. K. *et al.* Dynamics of putative raft-associated proteins at the cell surface. *J. Cell Biol.* **165**, 735–746 (2004).
42. Lenne, P. F. *et al.* Dynamic molecular confinement in the plasma membrane by microdomains and the cytoskeleton meshwork. *EMBO J.* **25**, 3245–3256 (2006).
43. Heinemann, F., Vogel, S. K. & Schwille, P. Lateral membrane diffusion modulated by a minimal actin cortex. *Biophys. J.* **104**, 1465–1475 (2013).
44. Heine, M. *et al.* Surface mobility of postsynaptic AMPARs tunes synaptic transmission. *Science* **320**, 201–205 (2008).
45. Czöndör, K. *et al.* Unified quantitative model of AMPA receptor trafficking at synapses. *Proc. Natl. Acad. Sci. USA* **109**, 3522–3527 (2012).
46. Sainlos, M. *et al.* Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. *Nat. Chem. Biol.* **7**, 81–91 (2011).
47. Letellier, M., Lagardère, M., Tessier, B., Janovjak, H. & Thoumine, O. Optogenetic control of excitatory post-synaptic differentiation through neuroligin-1 tyrosine phosphorylation. *Elife* **9**, e52027 (2020).
48. Beheiry, M. E., Dahan, M. & Masson, J. B. InferenceMAP: mapping of single-molecule dynamics with Bayesian inference. *Nat. Methods* **12**, 594–595 (2015).
49. Levet, F. *et al.* SR-Tesseler: a method to segment and quantify localization-based super-resolution microscopy data. *Nat. Methods* **12**, 1065–1071 (2015).
50. Briane, V., Salomon, A., Vimond, M. & Kervrann, C. A computational approach for detecting micro-domains and confinement domains in cells: a simulation study. *Phys. Biol.* **17**, ab5e1d (2020).
51. Granik, N. *et al.* Single-particle diffusion characterization by deep learning. *Biophys. J.* **117**, 185–192 (2019).
52. Rothbauer, U. *et al.* Targeting and tracing antigens in live cells with fluorescent nanobodies. *Nat. Methods* **3**, 887–889 (2006).
53. Jain, A. *et al.* Probing cellular protein complexes using single-molecule pull-down. *Nature* **473**, 484–488 (2011).
54. Vafabakhsh, R., Levitz, J. & Isacoff, E. Y. Conformational dynamics of a class C G-protein-coupled receptor. *Nature* **524**, 497–501 (2015).
55. Izeddin, I. *et al.* Wavelet analysis for single molecule localization microscopy. *Opt. Exp.* **20**, 2081–2095 (2012).
56. Heilemann, M. *et al.* Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew. Chem. Int. Ed.* **47**, 6172–6176 (2008).

Acknowledgements

We thank M. Missler, P. Scheiffele, and J. Lippincott-Schwartz for the generous gift of plasmids; B. Tessier, Z. Karatas, and S. Benquet for molecular biology; R. Sterling and the cell culture facility of the Institute; J.B. Sibarita, C. Butler, and A. Kechkar for the kind gift of the PALM-Tracer program, M. Sainlos for the gift of dye-conjugated GFP nanobody, R. Vafabakhsh (Northwestern University, Evanston, USA) for help with single molecule pull-down, E. Isacoff (University of California, Berkeley, USA) for providing generous access to laboratory resources, L. Barousseau (Explora Nova) and F. Levet for programming tips, A. Disher (Hamamatsu Photonics) for help on camera acquisition parameters for FCS, E. Nehme (Technion, Haifa, ISR) for helpful advice on Deep-STORM, O. Rossier and G. Giannone for scientific insight, A. Mehidi and T. Orré for interesting feedback, and A. Florea for help with the FluoSim movie. This work received funding from the Fondation pour la Recherche Médicale (“Equipe FRM” DEQ20160334916 to O.T.), the French Ministry of Research (fellowship to M.L.), the Agence Nationale de la Recherche (Grants «SynAdh» ANR-13-BSV4-0005-01 and «Synthesyn» ANR-17-CE16-0028-01 to O.T.), the national infrastructure France BioImaging (Grant ANR-10INBS-04-01), the French-American Commission (Fulbright grant to O.T.), and the Conseil Régional Aquitaine («SiMoDyn» to O.T.).

Author contributions

M.L. developed FluoSim, performed the SPT and FRAP experiments, carried out the simulations, and made most figures. M.L. and O.T. designed FluoSim, performed the FCS experiments, and wrote the user manual. I.C. and M.L. did proof-of-concept SPT and FRAP experiments. I.C. performed dSTORM acquisition and analysis, and made two figures. E.B. and M.N. installed and run the CNN Deep-STORM. E.B. and M.L. adapted the Deep-STORM MATLAB and Python scripts to train the CNN with FluoSim simulation outputs. O.T. came up with the original idea, supervised the work, performed the PAF experiments, and wrote the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-75814-y>.

Correspondence and requests for materials should be addressed to O.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Annexe C

Prédiction de la localisation

subcellulaire des ARNm :

Apprentissage à partir de sources de
données hétérogènes

Perspective

Prediction of RNA subcellular localization:
Learning from heterogeneous data sourcesAnca Flavia Savulescu,^{1,*} Emmanuel Bouilhol,^{2,3} Nicolas Beaume,⁴ and Macha Nikolski^{2,3,*}

SUMMARY

RNA subcellular localization has recently emerged as a widespread phenomenon, which may apply to the majority of RNAs. The two main sources of data for characterization of RNA localization are sequence features and microscopy images, such as obtained from single-molecule fluorescent in situ hybridization-based techniques. Although such imaging data are ideal for characterization of RNA distribution, these techniques remain costly, time-consuming, and technically challenging. Given these limitations, imaging data exist only for a limited number of RNAs. We argue that the field of RNA localization would greatly benefit from complementary techniques able to characterize location of RNA. Here we discuss the importance of RNA localization and the current methodology in the field, followed by an introduction on prediction of location of molecules. We then suggest a machine learning approach based on the integration between imaging localization data and sequence-based data to assist in characterization of RNA localization on a transcriptome level.

THE IMPORTANCE OF RNA SUBCELLULAR LOCALIZATION

Gene expression in eukaryotes is regulated at various stages of the life cycle of RNA molecules, including transcription initiation, RNA processing, stability, subcellular localization, and translation into protein in the case of mRNAs, stages which are often linked. Subcellular localization of RNA transcripts leads to restriction of translation in a spatial and temporal manner (Bashirullah et al., 1998; Besse and Ephrussi, 2008; Jansen, 2001; Kloc et al., 2002; Martin and Ephrussi, 2009; Mardakheh et al., 2015; Zappulo et al., 2017; Savulescu et al., 2021), as well as serving to avoid toxicity of protein products and to corroborate rapid cellular responses (Shahbaban and Chartrand, 2012). Early studies in the field of RNA subcellular localization investigated the subcellular location of one and up to a few mRNA transcripts in developmental models, such as the *Drosophila* embryo and *Xenopus* oocyte, determining morphogen gradients and cellular fates (Macdonald and Struhl, 1998; Tautz and Pfeifle, 1989), as well as polarized cells such as migrating fibroblasts, budding yeast, and neuronal cells (Mili et al., 2008; Martin and Ephrussi 2009; Batish et al., 2012; Buxbaum et al., 2014; Tzingounis and Nicoll, 2006; Yasuda et al., 2017). However, recent years have seen an increasing number of studies in various organisms/models of RNA subcellular localization, resulting in the consensus that RNA subcellular localization is indeed not limited to a handful of RNAs in a small number of systems but rather a widespread phenomenon, which may apply to the majority of cellular RNAs, including non-coding RNAs (Bouvette et al., 2017; La Manno et al., 2018; Lécuycy et al., 2007; Moor et al., 2017; Sharp et al., 2011; Weis et al., 2013; Cabili et al., 2015; Zappulo et al., 2017).

The subcellular location of the RNA can influence which proteins will bind the transcript, contributing to the RNA's fate, such as directing it toward degradation, increasing/decreasing its rate of translation, and determining molecular interactions. Consequently, the spatial distribution of the RNA can potentially influence the cellular concentration and location of its protein product (Brangwynne et al., 2009; Katz et al., 2012, 2016; Moor et al., 2018; Savulescu et al., 2021), which can, in turn, affect the cell's function and capacity to interact with adjacent cells or react to various environmental cues. Similarly, association of transcripts with specific cellular structures (Hughes and Simmonds, 2019; Khong et al., 2017; Padrón et al., 2019; Savulescu et al., 2021; Suter, 2018; Wilbertz et al., 2019 and others) may indicate a functional correlation between the transcript in the cell where this association occurs and broad cellular processes in the same cells, such as polarization, differentiation, etc. Taken together, this indicates that in addition to expression levels of RNA transcripts, subcellular spatial distribution of these transcripts may contribute to cell state and type (for example, Moor et al., 2018). Using solely single-cell genomics approaches, which are standard in cell

¹Division of Chemical, Systems & Synthetic Biology, Institute for Infectious Disease & Molecular Medicine, Faculty of Health Sciences, University of Cape Town, 7925 Cape Town, South Africa

²Université de Bordeaux, Bordeaux Bioinformatics Center, Bordeaux, France

³Université de Bordeaux, CNRS, IBGC, UMR 5095, Bordeaux, France

⁴Division of Medical Virology, Faculty of Health Sciences, University of Cape Town, 7925 Cape Town, South Africa

*Correspondence: ankasi100@gmail.com (A.F.S.), macha.nikolski@u-bordeaux.fr (M.N.)

<https://doi.org/10.1016/j.isci.2021.103298>



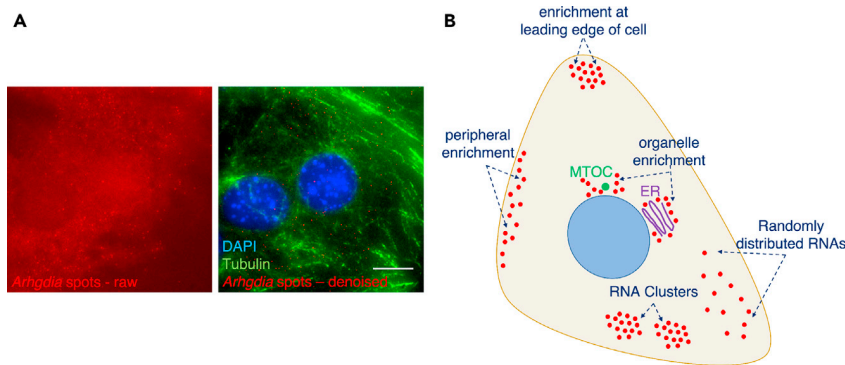


Figure 1. Subcellular RNA localization and visualization

(A) Typical smFISH images. On the left, a raw image showing *Arhgdia* mRNA smFISH spots; on the right, a superimposed and denoised image containing the following stains: DAPI for DNA in blue, anti-tubulin antibody in green, *Arhgdia* mRNA smFISH spots in red. Organelles of interest, such as the microtubule-organizing center (MTOC) and cell contour are extrapolated from z-slices of the anti-tubulin stain. Scale bar 10 μm .

(B) A variety of subcellular spatial distribution features of the RNA can be observed and subsequently analyzed (RNA spots are depicted in red): specific enrichment of RNA in various subcellular locations, random versus clustered distribution of RNA and correlation of RNA with cellular markers, such as the MTOC or ER.

state characterization studies, spatial information of RNA transcripts might be lost. For example, two neighboring cells in a tissue, which possess similar concentrations of the same RNA transcripts, would be classified by single-cell genomics approaches as the same cell type/state; however, the RNAs in these cells may exhibit marked patterns in subcellular dispersion (Savulescu et al., 2020) (Figure 1, Panel B). This again emphasizes the significance of characterizing subcellular distribution alongside expression levels of RNA transcripts for a more thorough classification of cell subtype or/and state.

In addition to contributions in basic research, knowledge of RNA subcellular spatial localization can be beneficial in biomedical applications. For instance, Didiot et al. (2018) show that *htt* mRNA, which encodes the protein responsible for Huntington's disease, is located both in the nucleus and cytoplasm (50/50) in neurons, while being purely cytoplasmic in non-neuronal cells. This has a therapeutic impact as nuclear *htt* mRNA is more stable and more resistant to oligonucleotide therapy than the cytoplasmic *htt*. Consequently, new therapeutic strategies can be envisioned, such as modifying the mRNA encoded by the abnormal allele of huntingtin to redirect it to the cytoplasm where oligonucleotide therapy is more efficient. Taken together, given the functional relevance of subcellular spatial distribution of RNA in both basic and clinical contexts, tools that are able to efficiently detect subcellular localizations at a transcriptome-wide level are urgently needed.

Visualization of RNA using single-molecule fluorescent in situ hybridization (smFISH) or MS2-system-based techniques have been the gold standard method to study RNA subcellular localization. In recent years, smFISH-based techniques and downstream in silico procedures for localization of mRNAs have been developed to cope with a large number of RNA transcripts. However, these methods can be technically challenging, be expensive, and often require sophisticated equipment. Machine learning, which has been successfully applied in various biological fields, including prediction of subcellular locations of proteins and of mRNAs, may be a suitable complementary approach to predict the precise subcellular localization of a large number of RNA species in a variety of systems on a transcriptome-wide level. In the perspective below, we provide a brief overview of techniques to study subcellular RNA localization and then introduce how machine learning can be applied to the study of RNA subcellular localization and discuss how machine learning can be harnessed to predict precise RNA subcellular localization. Finally, we discuss potential synergies between experimental and computational approaches of subcellular localization.

CURRENT METHODS TO STUDY SUBCELLULAR LOCALIZATION OF RNA

To date, the most popular approach to study subcellular localization of RNA is image based. A large number of studies make use of smFISH-based techniques followed by epifluorescent or confocal microscopy to visualize and quantify intracellular mRNAs (typical smFISH images in Figure 1, Panel A). Alternatively, an

MS2 tagging system followed by live cell imaging is applied (for example, [Bertrand et al., 1998](#); [Hocine et al., 2013](#) and others). smFISH makes use of multiple single-stranded DNA oligonucleotides, each labeled with a single fluorophore, which tile a specific RNA target ([Raj et al., 2008](#)). The signal obtained from multiple single fluorophores is sufficiently bright to be seen as a spot on an epifluorescent or confocal microscope and can easily be quantified ([Raj et al., 2008](#)). In recent years, a variety of smFISH-based techniques to increase the capacity of the method have emerged. These include the use of multiple rounds of labeling and imaging of the same sample to label a large number of RNAs. osmFISH is one such smFISH-based technique, in which the cellular organization of the mouse somatosensory cortex was mapped by labeling 33 RNAs over the course of 13 rounds. The number of labeled RNAs can be further increased by applying FISH probe barcoding and sequential labeling approaches, such as multiplexed error-robust FISH (MERFISH) ([Moffitt and Zhuang, 2016](#)) and sequential FISH (seqFISH and seqFISH+) ([Lubeck et al., 2014](#); [Eng et al., 2019](#)). These methods possess the capacity to label hundreds to 1000 RNAs, nearly approaching full-transcriptome imaging. Additionally, [Wang et al. \(2018\)](#), used *in situ* amplification of RNA-specific probe barcode regions, decoding by 3D sequencing within samples converted to a hydrogel matrix, in a method termed spatially resolved transcript amplicon readout mapping (STARmap), to target hundreds of genes in various tissues ([Wang et al., 2018](#)). Branched DNA (bDNA)-based techniques rely on signal amplification using a series of non-isotopic DNA probes hybridized in a sequential manner to detect and quantify RNA ([Player et al., 2001](#); [Wang et al., 2012](#); [Battich et al., 2013](#)).

Following imaging of smFISH samples, data analysis is carried out. The analysis contains several steps, including segmentation, image processing to reduce background noise, readout of barcodes associated to the different transcripts in the case of barcoding-based smFISH methods, as well as precise localization and quantification of RNA FISH spots and in some cases, association of detected spots with cellular landmarks. A wide range of computational methods exist to extract quantified data for RNAs from smFISH images, and a number of widely used microscopy image analysis software programs are available for mRNA spot detection, such as the ICY spot detector ([de Chaumont et al., 2012](#)), ImageJ spot detection, FISH-Quant ([Mueller et al., 2013](#)), its more recent version FISH-quant v2 ([Imbert et al., 2021](#)), and CellProfiler ([McQuin et al., 2018](#)), among others. Many tools also provide cell and nucleus segmentation to apply on nuclear and cellular marker stainings that are often acquired with smFISH images. The extracted vectorized features can take different forms, such as RNA spot coordinates, or statistical features such as RNA counts and/or densities per cell or in different compartments as well as their positioning relative to cellular landmarks. See [Imbert et al. \(2021\)](#); [Samacoits et al., 2018](#) for examples of possible features as well as their representation in relation to cellular landmarks ([Savulescu et al., 2021](#)).

In addition to image-based approaches, subcellular localization of RNA can be studied using methods such as biochemical cellular fractionation or physical separation of cellular compartments, followed by RNA detection using RNA sequencing or microarrays (for example, [Mili et al., 2008](#); [Cajigas et al., 2012](#); [Bigler et al., 2017](#)). Although these image-based and biochemical approaches have provided high-resolution data regarding the subcellular localization of RNA in cells and tissues, they each have limitations both in the spatial resolution of the method, its throughput, as well as in the reliance on effort-heavy computational analysis of the generated images to detect mRNA localizations.

To summarize, given the increasing number of studies that characterize spatial and temporal subcellular localization patterns of RNAs, as well as the functional relevance of variation in spatial distribution of RNA molecules, there is an increasing need to develop appropriate tools and technologies to capture fine-grained variations in spatial distribution of RNAs. These tools would be required to be quantitative, analyze data on a single-cell level, and be accustomed for high-throughput data.

RNA LOCALIZATION PREDICTION

The rise in studies emphasizing the importance of RNA spatial and temporal subcellular localization in biological processes calls for an increased access to such information. Although it is clear that wet-lab experiments are the most straightforward way to characterize subcellular localization of biomolecules, these experiments are typically time- and money consuming. Contrastingly, with sufficient data and powerful *in silico* methods, predictions of subcellular localization of biomolecules would be far less expensive and intrinsically high throughput. We argue that this makes place for computational prediction of RNA localizations.

Prediction of subcellular locations based on sequence

The task of predicting subcellular location of biomolecules is not new. In the previous decade, a significant number of methods have been developed to predict subcellular locations of proteins (Emanuelsson et al., 2007; Imai and Nakai, 2010; Wan and Mak, 2015). The methods are based on certain features extracted from protein sequences, such as specific motifs, and aim to predict a *rough location* in terms of cell regions/departments such as those defined by the UniProt controlled vocabulary for subcellular locations. Often these methods use the guilt by association approach, using the homology with proteins whose location has been experimentally confirmed, the main experimental methods to establish protein localization being mass spectrometry and immunofluorescence.

Some mRNA transcripts contain within their sequences distinct motifs, which on their own, or by forming distinct secondary structures, have been shown to be determinant of the RNA subcellular location (reviewed in Martin and Ephrussi, 2009; Shahbadian and Chartrand, 2012 and others). Hence, several methods for prediction of RNA subcellular location have started to be developed based on mRNA/cDNA sequence composition (Yan et al., 2019; Garg et al., 2020), as well as introducing the predicted secondary structure (Yan et al., 2016). Interestingly, both of these very recent approaches are leveraged by the machine learning technology: deep recurrent networks (convolutional neural network [CNN], long short-term memory network [LSTM], and attention layers) for the former and more classical support vector machine (SVM)-based methodology for the latter. As in the protein world, these methods require training data sets providing mRNA subcellular locations for each annotated human protein-coding gene, such as cytosol, nuclear, endoplasmic reticulum (ER), membrane, etc. However, in contrast to the protein world, where annotation of data that is essential for the development of supervised machine learning typically exists, the RNA data generally lacks annotation. This lack of annotation is reflected in the very recent introduction of such methods for the mRNA, as well as in the fact that there remain to be only a few methods. Consequently, their adoption by the wider scientific community remains to be seen.

Predicting the subcellular location of non-coding RNAs (ncRNA), such as long non-coding RNAs (lncRNAs), microRNAs (miRs), and others is generally a more difficult task than for mRNAs. This is due to several factors, including the complexity of intra-molecular organization that ncRNAs can exhibit (Yan et al., 2016), as well as their typical short lengths and fewer known localization-correlated motifs than mRNAs (Ross et al., 2021; Constanty and Shkumatava, 2021). However, some associations of sequence motifs in lncRNAs with their subcellular localization have been identified (Zhang et al., 2014) and databases containing information on cellular compartments to which the lncRNAs localize have been developed, including lncLocate (Mas-Ponte et al., 2017) and RNALocate (Zhang et al., 2016). Given these recent developments, several methods have been proposed to predict subcellular location of lncRNAs. For example, Su et al. (2018) have combined a PseKNC and SVM to predict subcellular location of lncRNAs to the ribosome and exosome, Gudeñas and Wang (2018) have developed a deep neural network to predict whether lncRNAs are nuclear or cytosolic based on their sequence, and Cao et al., (2018) used k-mer and high-level abstraction features generated by unsupervised deep models to construct four classifiers and predict five subcellular localizations of lncRNAs. Similarly, a number of tools exist to predict the location of miRNAs, such as an approach based on gene ontology (GO)-based functional similarity (Yang et al., 2018) and an SVM-based predictor (Meher et al., 2020), both relying on the miRNA locations from the RNALocate database.

Prediction of subcellular location from microscopy data

More recently, a plethora of tools have been developed to predict the locations of molecules based on microscopy imaging data rather than sequence, in particular, in application to protein subcellular localization. For example, Newberg and Murphy (2008) have developed a framework for image-based protein subcellular location prediction and it has been successfully applied to the Human Protein Atlas database (Newberg and Murphy, 2008). This has paved the way for a whole set of work of microscopy-image-based subcellular localization prediction of proteins, which includes methods relying on k-NN classifiers, SVMs, artificial neural networks, decision trees, and deep CNNs (Xijie Lu and Moses, 2016; Jiang et al., 2019). Recently, a crowd citizen science effort has attracted participants to annotate the subcellular locations of proteins in images and resulted in a novel deep learning method based on transfer learning (Sullivan et al., 2018), capable of predicting distributions of proteins to major organelles. Not only the authors have achieved high accuracy of predictions but also they were able to construct a fully annotated data set (Sullivan et al., 2018). In the next section, we will discuss how sequence data could be combined with smFISH data to improve the predictive accuracy.

PERSPECTIVE ON PREDICTING RNA SUBCELLULAR LOCALIZATION FROM HETEROGENEOUS DATA

As discussed earlier, the two main sources of data for characterization of subcellular localization of RNA are RNA sequence features and smFISH images. In addition, as subcellular localization is considered as a means to restrict translation, among other functions, information regarding the subcellular localization of the encoded protein product, as well as cell-type-specific information may aid in subcellular localization characterization of RNA. However, in some cases, including embryonic development models, subcellular localization of proteins might not be correlated with the subcellular localization of the encoding mRNAs (Knaut et al., 2000; Little et al., 2015; Mardakheh et al., 2015) and, as such, might be misleading if considered as a sole parameter for training the model. Additional information regarding the cellular model should be considered to account for variations in subcellular distribution of an mRNA and its encoded protein.

Sequence features are easy to obtain; however, they do not provide straightforward information regarding subcellular localization. Moreover, current methods of prediction of molecules' location remain limited in their predictive power. For example, precision can widely vary, as reported in the study by Garg et al. (2020), where the area under the curve lies between 0.7 and 0.98 for different compartments. Importantly, no method with good performance for most cellular compartments is available. On the other hand, fluorescent microscopy imaging data are the gold standard to determine the RNA subcellular localization; however, it is typically costly, time-consuming, and complex to perform for all known RNAs. In this part we will discuss how imaging data may be combined with sequence features to accurately predict RNA subcellular location.

Beyond the improvement of both observation techniques and predictive algorithms, we would like to suggest that a synergy between the existing two approaches — sequence and image-based — may assist in deciphering RNA localization (Figure 2). To illustrate our argument, consider the following manual steps to be followed when characterizing the subcellular localization of an RNA transcript that has not yet been visualized using smFISH or MS2-system based techniques.

- (1) The first step would be to extract specific features from the transcript sequence within the RNA, which would aid in characterization of its role and location. These include conserved motifs, regulatory sequences, domains that bind specific RNA binding proteins (RBPs), secondary structures, etc.
- (2) Second, one would have to collect the information on the role and subcellular localization of the protein product in the case of mRNAs and the role of the RNA, if known, in the case of long ncRNAs, as well as additional cell-type- and model-specific localization data.
- (3) Third, identify RNAs with similar sequences and collect the information from databases on RNA localization, protein localization, smFISH and IF images, etc, for these RNAs.

We argue that prediction of the subcellular localization of the RNA could be made by integrating these collected heterogeneous data of the RNA (sequence-based features of the RNA, role, and subcellular localization of the protein product/RNA, etc) as well as imaging data collected from databases of RNAs (Figure 2). A machine learning approach (e.g., using multimodal learning) of this objective would be built of the following steps:

- 1) Gather a database of RNAs for which both sequence features and imaging information are known (Figure 2, Panel A).
- 2) Extract features from sequence and images as detailed in the “RNA localization prediction” section, and vectorize them.
- 3) Train a model based on sequences, represented by their vector of feature. Figure 2 offers an example on how to train a model with heterogeneous data (Figure 2, Panel A).
- 4) Use this model to predict localization of RNA of interest (see the following part for details) (Figure 2, Panel B).

The critical part of this process is the use of a model created from RNAs for which both sequences and imaging data are available, to predict the localization of RNAs for which only sequences are available. Recent developments in machine learning and deep learning (siamese networks, adversarial training) may help to

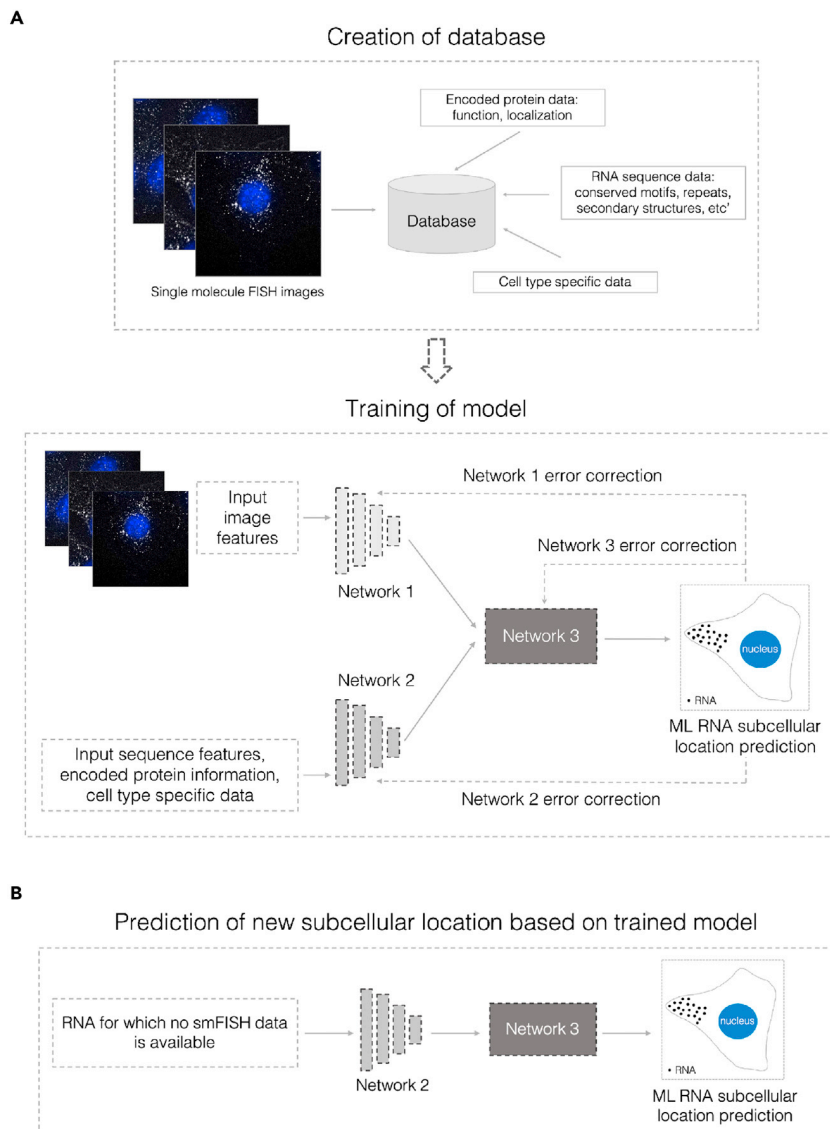


Figure 2. An illustration on how deep learning could use smFISH data and RNA sequence features to build a model, followed by prediction of the subcellular localization of an RNA using the model

(A) Training: RNAs for which both smFISH images and sequence features are available are collected to form a training set. For sequences, the relevant features are extracted (A, top), followed by training network 1 with images and network 2 with sequence features (A, bottom). The output of these networks is used as input to train network 3, which makes the prediction (A, bottom). During the training process, errors are back propagated to improve all networks (A, bottom).

(B) Prediction: the localization of a new RNA, for which only sequence features are available, needs to be predicted. The sequence is processed by network 2, followed by network 3, which outputs the prediction.

address this challenge. For instance, a network from the output of two networks, each based on one data type, could be trained. Once trained, this “aggregate network” could be fed by one type of data and still retain its predicting power. These approaches would create models “boosted” by the availability of smFISH in the training phase but able to predict with sequence features only.

LIMITATIONS IN THE DATA THAT MIGHT CHALLENGE PREDICTION

Prediction greatly depends on the data used to train the model. First, a large amount of data might be required to train a model properly, especially for methods such as deep learning or a combination of classifiers, which might be data greedy. Thus, increasing the number of experimentally observed locations

would be helpful to train better models. Second, it is clear that our knowledge of RNA subcellular localization is not uniform for different RNA species—some classes of RNA, e.g., lncRNA and miRNA, are under-represented in our knowledge map of RNA subcellular localization, potentially leading to bias in the models. Targeted experiments toward the “blind spots” of the RNA localization transcriptome map would help to correct these biases.

One of the avenues to identify the important missing data that has to be acquired is to look at the most distinguishing features, which are features in the model that contribute the most to the prediction. Examining the importance of different features in prediction can be helpful to decipher biological processes behind RNA subcellular localization; however, most importantly, by considering the distinguishing features, it is possible to design experiments to acquire missing data for the subcellular localization of some transcripts. These data in turn increase the impact of distinguishing features and improve prediction accuracy. Among features that are already known to have strong impact on prediction, we can cite k-mer composition, RNA-protein binding motifs, and other genomic sequence features as identified in the study by [Gudenas and Wang \(2018\)](#), who have built a model to classify lncRNAs into cytosolic or nuclear and performed the analysis of feature importance. In this work, the authors have found that the k-mer composition accounted for 90% of the decision. Because it is assumed that such distinguishing features are linked to the biological processes underlying the subcellular localization, guiding acquisition of new data by the principle of enriching the features that contribute to precision accuracy appears to be a promising avenue to increase the impact of machine learning approaches in the field.

An additional limitation concerns the difficulties in segmentation of specialized cells, such as neurons, and the related RNA quantification in relevant compartments. Still, a number of local morphological descriptors such as dendritic tree, radial extension, soma area, and branching complexity have been computed to date ([Shefi et al., 2002](#)). Additionally, invariant measures such as Hu’s moments can be included ([Bhaskar et al., 2019](#)). Such morphological features can be used by a downstream machine learning pipeline. Segmentation, although not always perfect, can be often well performed, and although the annotation of neurons is still not yet fully automated, significant progress has been made on this topic with very promising results ([Li et al., 2019](#); [Lin and Zheng, 2019](#); [Schubert et al., 2019](#)). Normalization of RNA quantities in different compartments can then be done by quantization methods, such as DypFISH ([Savulescu et al., 2021](#)) and others. Naturally, if the identification of certain compartments is imperfect, the corresponding normalized RNA quantities will consequently be skewed, which will inevitably impact predictions that would use these data. In conclusion, any machine learning localization prediction method is only as good as the data that it are built upon.

Finally, regarding features that may involve two-dimensional (2D) or three-dimensional (3D) features such as zip codes, machine learning approaches have been shown to be able to integrate complex features such as 2D and 3D structures from sequences ([Singh et al., 2019](#); [Jumper et al., 2021](#); [Sweeney et al., 2021](#)), and thus, machine learning methods may be able to take the zip code elements into account, at least indirectly.

Addressing the challenge of variability in subcellular locations

One of the major limitations in our understanding of RNA localization is the variability and dynamics of the subcellular localization. In many cases, a given RNA may be addressed to different subcellular compartments, depending on the circumstance (cell type, cell state, environmental conditions, various treatments, as well as temporally). This aspect remains poorly understood for the majority of cellular RNAs. Additionally, a given RNA’s subcellular localization patterns might not be clearly pronounced or their localized enrichment high enough for detection. This typically occurs in developmental systems, such as the early *Drosophila* embryo, where only up to 4% of a particular mRNA localizes to germ granules, while the remaining fraction disperses through the embryo ([Bergsten and Gavis, 1999](#); [Jambor et al., 2015](#); [Trcek et al., 2015](#)). The current practice is to consider only the localization with the highest probability to be the right one. Using highly efficient models, localizations with enough confidence could all be considered as correct, being the sign of a multi-localized RNA. The level of confidence could ultimately give an estimate of the tendency of this RNA to be addressed at different localizations.

Variability in subcellular location of a given RNA can also be addressed by performing smFISH on cells grown on microfabricated patterns ([Savulescu et al., 2021](#)). Micropatterning of cells reduces cell-to-cell variability and allows for a higher resolution, quantitative characterization of subcellular localization of RNAs ([Savulescu et al., 2021](#)). Comparison of the spatial localization of a given RNA in the same cell type on micropatterns under different

conditions or in different micropatterned cell types should allow for a thorough characterization of variability in subcellular localization of the given RNAs in these conditions/cell types. This should, in turn, assist in accounting for variability when predicting subcellular localization of RNAs.

To summarize, variability of RNA location dependent on cell types/conditions/cellular compartments etc is important to account for and should be integrated as parameters in the model. Although current models are limited in this respect, we would like to suggest that with growth of available annotated data, prediction models should gradually improve.

DISCUSSION

High-precision methods for determination and quantification of subcellular localization of RNAs, such as smFISH, are now well established; however, these techniques remain time-consuming and costly. To drive better understanding of cellular processes, there is a need for development of methods to cover the broad landscape of RNA subcellular localization, for a large range of RNAs and conditions, ideally at the whole-transcriptome level. We would like to advance the argument that observation and prediction of RNA subcellular localization are two complementary approaches that can be leveraged together. Although they remain largely disconnected, linking them has the potential to greatly increase knowledge in the field. As mentioned earlier, the quality of the prediction is directly linked both to the quantity and the quality of the available datasets. Thus, building robust models for the prediction of RNA location requires growth of the available and well-annotated data. This increase of relevant data can be driven both by the biological questions, as it is currently the case, the development of relevant data repositories, but also in a complementary fashion by the requirement to fill the gaps in the existing predictive features that are used to populate machine learning models.

As previously discussed, any bias in annotations and/or errors in the upstream analyses are inevitably propagated into predictions. A possible avenue to circumvent these biases would be a non-supervised machine learning approach that would make its own inferences about the structures it finds in the data instead of relying on its vectorized representation. However, unsupervised learning requires even more data than the supervised counterparts. An unsupervised approach would thus be an excellent way to circumvent annotation biases and errors and possibly provide a solution when the field is mature enough and more data are available.

We argue that cost, technical, and time considerations can be alleviated by designing robust predictive methods that take advantage of heterogeneous data, where RNA location prediction is based on both imaging and sequence data. Continued growth of available data sets containing both the data itself and its reliable annotations and covering the diversity of different RNA species in various contexts provides hope that the construction of robust models based on heterogeneous data—both imaging and sequences—for prediction of subcellular RNA localization is realistically feasible in the near future.

ACKNOWLEDGMENTS

We thank Jean-Baptiste Sibarita and Dana M. Savulescu for fruitful discussions and critical reading of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, A.F.S.,M.N., N.B.; writing - original draft, A.F.S.,E.B.,M.N., N.B.; machine learning architecture for location prediction - A.F.S.,E.B.,M.N.; writing - reviewing and editing, A.F.S.,E.B.,M.N, N.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Bashirullah, A., Cooperstock, R.L., and Lipshitz, H.D. (1998). RNA localization in development. *Annu. Rev. Biochem.* *67*, 335–394.
- Bhaskar, D., Lee, D., Knútsdóttir, H., Tan, C., Zhang, M.H., Dean, P., Roskelley, C., and Edelstein-Keshet, L. (2019). A methodology for morphological feature extraction and unsupervised cell classification. *BioRxiv*. <https://doi.org/10.1101/623793>.
- Battich, N., Stoeger, T., and Pelkmans, L. (2013). Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* *10*, 1127–1133.
- Batish, M., van den Bogaarda, P., Kramer, F.R., and Tyagi, S. (2012). Neuronal mRNAs travel singly into dendrites. *Proc. Natl. Acad. Sci. U S A* *109*, 4645–4650.

- Bergsten, S.E., and Gavis, E.R. (1999). Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. *Development* 126, 659–669.
- Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of *ASH1* mRNA particles in living yeast. *Mol.Cell* 2, 437–445.
- Besse, F., and Ephrussi, A. (2008). Translational control of localized mRNAs: restricting protein synthesis in space and time. *Mol. Cell. Biol.* 9, 971.
- Bigler, R.L., Kamande, J.W., Dumitru, R., Niedringhaus, M., and Taylor, A.M. (2017). Messenger RNAs localized to distal projections of human stem cell derived neurons. *Sci. Rep.* 7, 611.
- Bouvet, L.P., Cody, N.A.L., Bergalet, J., Lefebvre, F.A., Diot, C., Wang, X., Blanchette, M., and Lécuyer, E. (2017). CeFra-seq reveals broad asymmetric mRNA and non-coding RNA distribution profiles in *Drosophila* and human cells. *RNA*. <https://doi.org/10.1261/rna.063172.117>.
- Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* 324, 1729–1732.
- Buxbaum, A., Wu, B., and Singer, R.H. (2014). Single β -actin mRNA detection in neurons reveals a mechanism for regulating its translatability. *Science* 343, 419.
- Cabili, M.N., Dunagin, M.C., McClanahan, P.D., Bialesch, A., Padovan-Merhar, O., Regev, A., and Raj, A. (2015). Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* 16, 20.
- Cajigas, I.J., Tushev, G., Will, T.J., tom Dieck, F., Fierst, N., and Schuman, E.M.S. (2012). The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* 74, 453–466.
- Cao, Z., Pan, X., Yang, Y., Huang, Y., and Shen, H.B. (2018). The IncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier. *Bioinformatics* 34, 2185–2194.
- Constanty, F., and Shkumatava, A. (2021). lncRNAs in development and differentiation: from sequence motifs to functional characterization. *Development* 148, dev182741.
- de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., Vannary, M.-Y., Pankajakshan, P., Lecomte, T., Le Montagner, Y., et al. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nat. Methods* 9, 690–696.
- Didiot, M.C., Ferguson, C.M., Ly, S., Coles, A.H., Smith, A.O., Bicknell, A.A., Hall, L.M., Sapp, E., Echeverria, D., Pai, A.A., et al. (2018). Nuclear localization of huntingtin mRNA is specific to cells of neuronal origin. *Cell Rep.* 24, 2553–2560.e5.
- Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat.Protoc.* 2, 953–971.
- Eng, C.L., Lawson, M., Zhu, Q., Dries, R., Koulens, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.C., and Cai, L. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* 568, 235–239.
- Garg, A., Singhal, N., Kumar, R., and Kumar, M. (2020). mRNAloc: a novel machine-learning based in-silico tool to predict mRNA subcellular localization. *Nucleic Acids Res.* 48, W239–W243.
- Gudenas, B.L., and Wang, L. (2018). Prediction of lncRNA subcellular localization with deep learning from sequence features. *Sci. Rep.* 8, 1–10.
- Hocine, S., Raymond, P., Zenklusen, D., Chao, J.A., and Singer, R.H. (2013). Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat. Methods* 10, 119–121.
- Hughes, A.C., and Simmonds, A.J. (2019). *Drosophila* mRNA localization during later development: past, present, and future. *Front. Genet.* 10, 135.
- Imai, K., and Nakai, K. (2010). Prediction of subcellular locations of proteins: where to proceed? *Proteomics* 10, 3970–3983.
- Imbert, A., Ouyang, W., Safieddine, A., Coleno, E., Zimmer, C., Bertrand, E., Walter, T., and Mueller, F. (2021). FISH-quant v2: a scalable and modular analysis tool for smFISH image analysis. *BioRxiv*. <https://doi.org/10.1101/2021.07.20.453024>.
- Jambor, H., Surendranath, V., Kalinka, A.T., Mejstrik, P., Saalfeld, S., and Tomancak, P. (2015). Systematic imaging reveals features and changing localization of mRNAs in *Drosophila* development. *Elife* 4, e05003.
- Jansen, R.P. (2001). mRNA localization: message on the move. *Nat. Rev. Mol. Cell. Biol.* 2, 247–256.
- Jiang, Z., Wang, D., Wu, P., Chen, Y., Shang, H., Wang, L., and Xie, H. (2019). Predicting subcellular localization of multisite proteins using differently weighted multi-label k-nearest neighbors sets. *Technol.Health Care* 27 (Suppl 1), 185–193.
- Katz, Z.B., Wells, A.L., Park, H.Y., Wu, B., Shenoy, S.M., and Singer, R.H. (2012). β -Actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes Dev.* 26, 1885–1890.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.
- Katz, Z.B., English, B.P., Lionnet, T., Yoon, Y.J., Monnier, N., Ovrly, B., Bathe, M., and Singer, R.H. (2016). Mapping translation ‘hot-spots’ in live cells by tracking single molecules of mRNA and ribosomes. *Elife* 5, e10415.
- Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., and Parker, R. (2017). The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol.Cell* 68, 808–820.e5.
- Kloc, M., Zearfoss, N.R., and Etkin, L.D. (2002). Mechanisms of subcellular mRNA localization. *Cell* 108, 533–544.
- Knaut, H., Pelegri, F., Bohmann, K., Schwartz, H., and Nusslein-Volhards, C. (2000). Zebrafish vasa RNA but not its protein is a component of the germ plasm and segregates asymmetrically before germlinespecification. *J.CellBiol.* 149, 875–888.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M.E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. *Nature* 560, 494–498.
- Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174–187.
- Li, R., Zhu, M., Li, J., Bienkowski, M.S., Foster, N.N., Xu, H., Ard, T., Bowman, I., Zhou, C., Veldman, M.B., et al. (2019). Precise segmentation of densely interweaving neuron clusters using G-Cut. *Nat.Comm.* 10, 1549.
- Lin, X., and Zheng, J. (2019). A neuronal morphology classification approach based on locally cumulative connected deep neural networks. *Appl. Sci.* 9, 3876.
- Little, S.C., Sinsimer, K.S., Lee, J.J., Wieschaus, E.F., and Gavis, E.R. (2015). Independent and coordinate trafficking of single *Drosophila* germ plasm mRNAs. *Nat.CellBiol.* 17, 558–568.
- Lubeck, E., Coskun, A.F., Zhyentayev, T., Ahmad, M., and Cai, L. (2014). Single-cell *in situ* RNA profiling by sequential hybridization. *Nat. Methods* 11, 360–361.
- Macdonald, P.M., and Struhl, G. (1998). cis-acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos. *Nature* 336, 595–598.
- Mardakheh, F.K., Paul, A., Kümper, S., Sadok, A., Paterson, H., McCarthy, A., Yuan, Y., and Marshall, C.J. (2015). Global analysis of mRNA, translation, and protein localization: local translation is a key regulator of cell protrusions. *Dev.Cell* 35, 344–357.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719.
- Mas-Ponte, D., Carlevaro-Fita, J., Palumbo, E., Pulido, T.H., Guigo, R., and Johnson, R. (2017). LncAtlas database for subcellular localization of long noncoding RNAs. *RNA* 23, 1080–1087.
- McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., et al. (2018). CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol* 16, e2005970.
- Meher, P.K., Satpathy, S., and Rao, A.R. (2020). miRNALoc: predicting miRNA subcellular localizations based on principal component scores of physico-chemical properties and pseudo compositions of di-nucleotides. *Sci. Rep.* 10, 14557.

- Mili, S., Moissoglu, K., and Macara, I.G. (2008). Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* 453, 115.
- Moffitt, J.R., and Zhuang, X. (2016). RNA imaging with multiplexed error-robust fluorescence in situ hybridization (MERFISH). *Methods Enzymol.* 572, 1–49.
- Moor, A.E., Golan, M., Massasa, E.E., Lemze, D., Weizman, T., Shenhav, R., Baydatch, S., Mizrahi, O., Winkler, R., Golani, O., et al. (2017). Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science* 357, 1299–1303.
- Moor, A.E., Harnik, Y., Ben-Moshe, S., Massasa, E.E., Rozenberg, M., Eilam, R., Bahar Halpern, K., and Itzkovitz, S. (2018). Spatial reconstruction of single enterocytes uncovers broad zonation along the intestinal villus. *Axis. Cell* 175, 1156–1167.e15.
- Mueller, F., Senecal, A., Tantale, K., Marie-Nelly, H., Ly, N., Collin, O., Basyuk, E., Bertrand, E., Darzacq, X., and Zimmer, C. (2013). FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat. Methods.* 10, 277–278.
- Newberg, J.Y., and Murphy, R.F. (2008). A framework for the automated analysis of subcellular patterns in human protein atlas images. *J. Proteome Res.* 7, 2300–2308.
- Padrón, A., Iwasaki, S., and Ingolia, N.T. (2019). Proximity RNA labeling by APEX-seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* 75, 875–887.e5.
- Player, A.N., Shen, L.P., Kenny, D., Antao, V.P., and Kolberg, J.A. (2001). Single-copy gene detection using branched DNA (bdNA) *in situ* hybridization. *J. Histochem. Cytochem.* 49, 603–612.
- Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* 5, 877–879.
- Ross, C.J., Rom, A., Spinrad, A., Gelbard-Solodkin, D., Degani, N., and Ulitsky, I. (2021). Uncovering deeply conserved motif combinations in rapidly evolving noncoding sequences. *Genome Biol.* 22, 29.
- Samacoits, A., Chouaib, R., Safieddine, A., Traboulsi, A.-M., Ouyang, W., Zimmer, C., Peter, M., Bertrand, E., Walter, T., and Mueller, F. (2018). A computational framework to study sub-cellular RNA localization. *Nat. Commun.* 9, 4584.
- Savulescu, A.F., Brackin, R., Bouilhol, E., Dartigues, B., Warrell, J.H., Pimentel, M.R., Beaume, N., Cortunato, I.C., Dallongeville, S., Boule, M., et al. (2021). Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH. *Cell Rep. Methods*, 100068.
- Savulescu, A.F., Jacobs, C., Negishi, Y., Davignon, L., and Mhlanga, M.M. (2020). Pinpointing cell identity in time and space. *Front. Mol. Biosci.* 7, 209.
- Shahbaban, K., and Chartrand, P. (2012). Control of cytoplasmic mRNA localization. *Cell. Mol. Life Sci.* 69, 535–552.
- Sharp, J.A., Plant, J.J., Ohsumi, T.K., Borowsky, M., and Blower, M.D. (2011). Functional analysis of the microtubule-interacting transcriptome. *Mol. Biol. Cell.* 22, 4312–4323.
- Schubert, P.J., Dorkenwald, S., Januszewski, M., Jain, V., and Kornfeld, J. (2019). Learning cellular morphology with neural networks. *Nat. Commun.* 10, 2736.
- Shefi, O., Golding, I., Segev, R., Ben-Jacob, E., and Ayali, A. (2002). Morphological characterization of *in vitro* neuronal networks. *Phys. Rev. E* 66, 021905.
- Singh, J., Hanson, J., Paliwal, K., and Zhou, Y. (2019). RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nat. Commun.* 10, 5407.
- Su, Z.-D., Huang, Y., Zhang, Z.-Y., Zhao, Y.-W., Wang, D., Chen, W., Chou, K.-C., and Lin, H. (2018). iLoc-lncRNA: predict the subcellular location of lncRNAs by incorporating octamer composition into general PseKNC. *Bioinformatics* 34, 4196–4204.
- Sullivan, D.P., Winsnes, C.F., Åkesson, L., Hjelmare, M., Wiking, M., Schutten, R., Campbell, L., Leifsson, H., Rhodes, S., and Nordgren, A. (2018). Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat. Biotechnol.* 36, 820.
- Suter, B. (2018). RNA localization and transport. *Biochim. Biophys. Acta Gene Regul. Mech.* 1861, 938–951.
- Sweeney, B.A., Hoksza, D., Nawrocki, E.P., Ribas, C.E., Madeira, F., Cannone, J.J., Gutell, R., Maddala, A., Meade, C.D., and Williams, L.D. (2021). R2DT is a framework for predicting and visualising RNA secondary structure using templates. *Nat. Commun.* 12, 3494.
- Tautz, D., and Pfeifle, C. (1989). A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma* 98, 81–85.
- Treck, T., Grosch, M., York, A., Shroff, H., Lionnet, T., and Lehmann, R. (2015). *Drosophila* germ granules are structured and contain homotypic mRNA clusters. *Nat. Commun.* 6, 7962.
- Tzingounis, A.V., and Nicoll, R.A. (2006). Arc/Arg3.1: linking gene expression to synaptic plasticity and memory. *Neuron* 52, 403.
- Wan, S., and Mak, M.-W. (2015). Machine Learning for Protein Subcellular Localization Prediction (De Gruyter). <https://doi.org/10.1515/9781501501500>.
- Wang, X., Allen, W.E., Wright, M.A., Sylwestrak, E.L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* 361, eaat5691.
- Wang, F., Flanagan, J., Su, N., Wang, L.-C., Bui, S., Nielson, A., Wu, X., Vo, H.-T., Ma, X.-J., and Luo, Y. (2012). RNAscope: A novel *in situ* RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* 14, 22–29.
- Weis, B.L., Schleiff, E., and Zerges, W. (2013). Protein targeting to subcellular organelles via mRNA localization. *Biochim. Biophys. Acta* 1833, 260–273.
- Wilbertz, J.H., Voigt, F., Horvathova, I., Roth, G., Zhan, Y., and Chao, J.A. (2019). Single-molecule imaging of mRNA localization and regulation during the integrated stress response. *Mol. Cell* 73, 946–958.e7.
- Xijie Lu, A., and Moses, A.M. (2016). An unsupervised kNN method to systematically detect changes in protein localization in high-throughput microscopy images. *PLoS One* 11, e0158712.
- Yan, K., Arfat, Y., Li, D., Zhao, F., Chen, Z., Yin, C., Sun, Y., Hu, L., Yang, T., and Qian, A. (2016). Structure prediction: new insights into decrypting long noncoding RNAs. *Int. J. Mol. Sci.* 17, 132.
- Yan, Z., Lécuyer, E., and Blanchette, M. (2019). Prediction of mRNA subcellular localization using deep recurrent neural networks. *Bioinformatics* 35, i333–i342.
- Yang, Y., Fu, X., Qu, W., Xiao, Y., and Shen, H.B. (2018). MiRGOFs: a GO-based functional similarity measurement for miRNAs, with applications to the prediction of miRNA subcellular localization and miRNA–disease association. *Bioinformatics* 34, 3547–3556.
- Yasuda, Y., Clatterbuck-Soper, S.F., Jackrel, M.E., Shorter, J., and Mili, S. (2017). FUS inclusions disrupt RNA localization by sequestering kinesin-1 and inhibiting microtubule deetyrosination. *J. Cell Biol.* 216, 1015–1034.
- Zhang, B., Gunawardane, L., Niazi, F., Jahanbani, F., Chen, X., and Valadkhan, S. (2014). A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. *Mol. Cell. Biol.* 34, 2318–2329.
- Zhang, T., Tan, P., Wang, L., Jin, N., Li, Y., Zhang, L., Yang, H., Hu, Z., Zhang, L., Hu, C., et al. (2016). Rnalocate: a resource for RNA subcellular localizations. *Nucleic Acids Res.* 45, D135–D138.
- Zappulo, A., van den Bruck, D., CiolliMattioli, C., Franke, V., Imami, K., McShane, E., Moreno-Estelles, M., Calviello, L., Filipchuk, A., Peguero-Sanchez, E., et al. (2017). RNA localization is a key determinant of neurite-enriched proteome. *Nat. Commun.* 8, 583.

Bibliographie

- [1] Mohiuddin AHMED, Raihan SERAJ et Syed Mohammed Shamsul ISLAM. « The k-means algorithm : A comprehensive survey and performance evaluation ». In : *Electronics* 9.8 (2020), p. 1295.
- [2] Laith ALZUBAIDI et al. « Review of deep learning : Concepts, CNN architectures, challenges, applications, future directions ». In : *Journal of big Data* 8.1 (2021), p. 1-74.
- [3] Martin ARJOVSKY, Soumith CHINTALA et Léon BOTTOU. « Wasserstein GAN ». In : (2017). DOI : [10.48550/ARXIV.1701.07875](https://doi.org/10.48550/ARXIV.1701.07875). URL : <https://arxiv.org/abs/1701.07875>.
- [4] Peter BANKHEAD et al. « QuPath : Open source software for digital pathology image analysis ». In : *Scientific reports* 7.1 (2017), p. 1-7.
- [5] Antoine BASSET et al. « Adaptive spot detection with optimal scale selection in fluorescence microscopy images ». In : *IEEE Transactions on Image Processing* 24.11 (2015), p. 4512-4527.
- [6] Mona BATISH et al. « Neuronal mRNAs travel singly into dendrites ». In : *Proceedings of the National Academy of Sciences* 109.12 (2012), p. 4645-4650.
- [7] Nico BATTICH, Thomas STOEGER et Lucas PELKMANS. « Image-based transcriptomics in thousands of single human cells at single-molecule resolution ». In : *Nature methods* 10.11 (2013), p. 1127-1133.
- [8] Anne BEGHIN et al. « Automated high-speed 3D imaging of organoid cultures with multi-scale phenotypic quantification ». In : *Nature Methods* 19.7 (2022), p. 881-892.
- [9] Sean BELL et al. « Inside-outside net : Detecting objects in context with skip pooling and recurrent neural networks ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, p. 2874-2883.
- [10] Chinmay BELTHANGADY et Loic A ROYER. « Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction ». In : *Nature methods* 16.12 (2019), p. 1215-1225.

- [11] J.L. BENTLEY. « Multidimensional binary search trees used for associative searching ». In : *Communications of the ACM* 18.9 (1975), p. 509-517.
- [12] JE BESAG. « Comments on Ripley's paper : Royal Statistical Society ». In : *Journal* 39 (1977), p. 193-195.
- [13] Florence BESSE et Anne EPHRUSSI. « Translational control of localized mRNAs : restricting protein synthesis in space and time ». In : *Nature reviews Molecular cell biology* 9.12 (2008), p. 971-980.
- [14] Eric BETZIG et al. « Imaging intracellular fluorescent proteins at nanometer resolution ». In : *science* 313.5793 (2006), p. 1642-1645.
- [15] David SC BIGGS. « 3D deconvolution microscopy ». In : *Current Protocols in Cytometry* 52.1 (2010), p. 12-19.
- [16] Ryoma BISE et Yoichi SATO. « Cell detection from redundant candidate regions under nonoverlapping constraints ». In : *IEEE Transactions on Medical Imaging* 34.7 (2015), p. 1417-1427.
- [17] Corinna BLASSE et al. « PreMosa : extracting 2D surfaces from 3D microscopy mosaics ». In : *Bioinformatics* 33.16 (2017), p. 2563-2569.
- [18] Susanne BOLTE et Fabrice P CORDELIÈRES. « A guided tour into subcellular colocalization analysis in light microscopy ». In : *Journal of microscopy* 224.3 (2006), p. 213-232.
- [19] William E BRILEY et al. « Quantification and real-time tracking of RNA in live cells using Sticky-flares ». In : *Proceedings of the National Academy of Sciences* 112.31 (2015), p. 9591-9595.
- [20] Claire M BROWN. « Fluorescence microscopy-avoiding the pitfalls ». In : *Journal of cell science* 120.10 (2007), p. 1703-1705.
- [21] Jeremy R BROWN, Chanpasith PHONGTHACHIT et Mikolaj J SULKOWSKI. « Immunofluorescence and image analysis pipeline for Drosophila motor neurons ». In : *Biology Methods and Protocols* 4.1 (2019), bpz010.
- [22] Adina R BUXBAUM, Bin WU et Robert H SINGER. « Single β -actin mRNA detection in neurons reveals a mechanism for regulating its translatability ». In : *Science* 343.6169 (2014), p. 419-422.
- [23] Juan C CAICEDO et al. « Data-analysis strategies for image-based cell profiling ». In : *Nature methods* 14.9 (2017), p. 849.

- [24] Jeffrey A CHAO et Timothée LIONNET. « Imaging the life and death of mRNAs in single cells ». In : *Cold Spring Harbor Perspectives in Biology* 10.12 (2018), a032086.
- [25] KH CHEN et al. *Spatially resolved, highly multiplexed RNA profiling in single cells*. *Science* 348, aaa6090. 2015.
- [26] Racha CHOUAIB et al. « A Dual Protein-mRNA Localization Screen Reveals Compartmentalized Translation and Widespread Co-translational RNA Targeting ». In : *Developmental Cell* 54.6 (2020), 773-791.e5. ISSN : 1534-5807.
- [27] Christian A COMBS. « Fluorescence microscopy : a concise guide to current imaging methods ». In : *Current protocols in neuroscience* 50.1 (2010), p. 2-1.
- [28] Napari CONTRIBUTORS. « napari : a multi-dimensional image viewer for Python ». In : <https://doi.org/10.5281/zenodo.3555620> (2019).
- [29] Christoph CREMER et al. « Super-resolution microscopy approaches to nuclear nanostructure imaging ». In : *Methods* 123 (2017), p. 11-32.
- [30] G CUARTERO et al. « Three dimensional spot detection by multiscale analysis ». In : *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*. T. 1. IEEE. 2001, p. 317-320.
- [31] Pádraig CUNNINGHAM, Matthieu CORD et Sarah Jane DELANY. « Supervised learning ». In : *Machine learning techniques for multimedia*. Springer, 2008, p. 21-49.
- [32] Ankan DASH, Junyi YE et Guiling WANG. « A review of Generative Adversarial Networks (GANs) and its applications in a wide variety of disciplines—From Medical to Remote Sensing ». In : *arXiv preprint arXiv :2110.01442* (2021).
- [33] Fabrice DE CHAUMONT et al. « Icy : an open bioimage informatics platform for extended reproducible research ». In : *Nature methods* 9.7 (2012), p. 690.
- [34] Pietro DELCANALE et al. « Nanoscale mapping functional sites on nanoparticles by points accumulation for imaging in nanoscale topography (PAINT) ». In : *ACS nano* 12.8 (2018), p. 7629-7637.
- [35] Ruben DRIES et al. « Giotto, a toolbox for integrative analysis and visualization of spatial expression data ». In : *bioRxiv* (2020). DOI : [10.1101/701680](https://doi.org/10.1101/701680). URL : <https://doi.org/10.1101/701680>.
- [36] Vincent DUMOULIN et Francesco VISIN. « A guide to convolution arithmetic for deep learning ». In : *ArXiv e-prints* (mars 2016). eprint : [1603.07285](https://arxiv.org/abs/1603.07285).

- [37] Kenneth W DUNN, Malgorzata M KAMOCKA et John H McDONALD. « A practical guide to evaluating colocalization in biological microscopy ». In : *American Journal of Physiology-Cell Physiology* 300.4 (2011), p. C723-C742.
- [38] Sebastian DUNST et Pavel TOMANCAK. « Imaging flies by fluorescence microscopy : principles, technologies, and applications ». In : *Genetics* 211.1 (2019), p. 15-34.
- [39] Bastian Th EICHENBERGER et al. « deepBlink : threshold-independent detection and localization of diffraction-limited spots ». In : *Nucleic Acids Research* 49.13 (2021), p. 7292-7297.
- [40] Amicia D ELLIOTT. « Confocal microscopy : principles and modern practices ». In : *Current protocols in cytometry* 92.1 (2020), e68.
- [41] Seda EMINAGA et al. « Detection of Cell Proliferation Markers by Immunofluorescence Staining and Microscopy Imaging in Paraffin-Embedded Tissue Sections ». In : *Current protocols in molecular biology* 115.1 (2016), p. 14-25.
- [42] Martin ESTER et al. « A density-based algorithm for discovering clusters in large spatial databases with noise. » In : *kdd*. T. 96. 34. 1996, p. 226-231.
- [43] Andreas ETTINGER et Torsten WITTMANN. « Fluorescence live cell imaging ». In : *Methods in cell biology* 123 (2014), p. 77-94.
- [44] Andrea M FEMINO et al. « Visualization of single RNA transcripts in situ ». In : *Science* 280.5363 (1998), p. 585-590.
- [45] Tina FRIIS et al. « Immunocytochemical and immunohistochemical staining with peptide antibodies ». In : *Peptide Antibodies*. Springer, 2015, p. 311-325.
- [46] Kui FU et al. « Intrinsic Relationship Reasoning for Small Object Detection ». In : *arXiv preprint arXiv :2009.00833* (2020).
- [47] Remi GALLAND et al. « 3D high-and super-resolution imaging using single-objective SPIM ». In : *Nature methods* 12.7 (2015), p. 641-644.
- [48] Kaytlyn A GERBIN et al. « Cell states beyond transcriptomics : integrating structural organization and gene expression in hiPSC-derived cardiomyocytes ». In : *Cell Systems* 12.6 (2021), p. 670-687.
- [49] Peter GOLDSBOROUGH et al. « CytoGAN : generative modeling of cell images ». In : *BioRxiv* (2017), p. 227645.
- [50] Edgar R GOMES, Shantanu JANI et Gregg G GUNDERSEN. « Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells ». In : *Cell* 121.3 (2005), p. 451-463.

- [51] Ian GOODFELLOW et al. « Generative adversarial nets ». In : *Advances in neural information processing systems*. 2014, p. 2672-2680.
- [52] Michael GUÉ et al. « Smart 3D-FISH : automation of distance analysis in nuclei of interphase cells by image processing ». In : *Cytometry Part A : The Journal of the International Society for Analytical Cytology* 67.1 (2005), p. 18-26.
- [53] Jaza GUL-MOHAMMED et al. « A generic classification-based method for segmentation of nuclei in 3D images of early embryos ». In : *BMC bioinformatics* 15.1 (2014), p. 1-12.
- [54] Robert HAASE et al. « CLIJ : GPU-accelerated image processing for everyone ». In : *Nature methods* 17.1 (2020), p. 5-6.
- [55] Gal HAIMOVICH et Jeffrey E GERST. « Single-molecule fluorescence in situ hybridization (smFISH) for RNA detection in adherent animal cells ». In : *Bio-protocol* 8.21 (2018), e3070-e3070.
- [56] Christopher M HALE et al. « SMRT analysis of MTOC and nuclear positioning reveals the role of EB1 and LIC1 in single-cell polarization ». In : *Journal of cell science* 124.24 (2011), p. 4267-4285.
- [57] R. HAMAGUCHI et al. « Effective Use of Dilated Convolutions for Segmenting Small Object Instances in Remote Sensing Imagery ». In : *arXiv preprint, arXiv :1709.00179* (2017).
- [58] Ligong HAN, Robert F MURPHY et Deva RAMANAN. « Learning generative models of tissue organization with supervised GANs ». In : *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE. 2018, p. 682-690.
- [59] Kaiming HE et Jian SUN. « Convolutional neural networks at constrained time cost ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, p. 5353-5360.
- [60] Kaiming HE et al. « Deep residual learning for image recognition ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, p. 770-778.
- [61] Stefan W HELL et Jan WICHMANN. « Breaking the diffraction resolution limit by stimulated emission : stimulated-emission-depletion fluorescence microscopy ». In : *Optics letters* 19.11 (1994), p. 780-782.
- [62] Shreyas U HIRWAY et al. « Immunofluorescence image feature analysis and phenotype scoring pipeline for distinguishing epithelial-mesenchymal transition ». In : *Microscopy and Microanalysis* 27.4 (2021), p. 849-859.

- [63] Reka HOLLANDI et al. « nucleAIzer : a parameter-free deep learning framework for nucleus segmentation using image style transfer ». In : *Cell Systems* 10.5 (2020), p. 453-458.
- [64] Reka HOLLANDI et al. « Nucleus segmentation : towards automated solutions ». In : *Trends in Cell Biology* (2022).
- [65] Christine E HOLT et Simon L BULLOCK. « Subcellular mRNA localization in animal cells and why it matters ». In : *Science* 326.5957 (2009), p. 1212-1216.
- [66] Christine E HOLT, Kelsey C MARTIN et Erin M SCHUMAN. « Local translation in neurons : visualization and function ». In : *Nature structural & molecular biology* 26.7 (2019), p. 557-566.
- [67] Peiyun HU et Deva RAMANAN. « Finding tiny faces ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, p. 951-959.
- [68] Jonathan HUANG et al. « Speed/accuracy trade-offs for modern convolutional object detectors ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, p. 7310-7311.
- [69] Jan HUISKEN et Didier YR STAINIER. « Selective plane illumination microscopy techniques in developmental biology ». In : (2009).
- [70] Carole IAMPIETRO et al. « Developmentally regulated elimination of damaged nuclei involves a Chk2-dependent mechanism of mRNA nuclear retention ». In : *Developmental cell* 29.4 (2014), p. 468-481.
- [71] Kyuseok IM et al. « An introduction to performing immunofluorescence staining ». In : *Biobanking* (2019), p. 299-311.
- [72] Arthur IMBERT et al. « FISH-quant v2 : a scalable and modular tool for smFISH image analysis ». In : *RNA* 28.6 (2022), p. 786-795.
- [73] Hellen C ISHIKAWA-ANKERHOLD, Richard ANKERHOLD et Gregor PC DRUMMEN. « Advanced fluorescence microscopy techniques—Frap, Flip, Flap, Fret and flim ». In : *Molecules* 17.4 (2012), p. 4047-4132.
- [74] Phillip ISOLA et al. « Image-to-image translation with conditional adversarial networks ». In : *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, p. 1125-1134.
- [75] Shruti JADON. « A survey of loss functions for semantic segmentation ». In : *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE. 2020, p. 1-7.

- [76] Nick KANOPOULOS, Nagesh VASANTHAVADA et Robert L BAKER. « Design of an image edge detection filter using the Sobel operator ». In : *IEEE Journal of solid-state circuits* 23.2 (1988), p. 358-367.
- [77] Rongqin KE et al. « In situ sequencing for RNA analysis in preserved tissue and cells ». In : *Nature methods* 10.9 (2013), p. 857-860.
- [78] Asifullah KHAN et al. « A survey of the recent architectures of deep convolutional neural networks ». In : *Artificial intelligence review* 53.8 (2020), p. 5455-5516.
- [79] Ismail M KHATER, Ivan Robert NABI et Ghassan HAMARNEH. « A review of super-resolution single-molecule localization microscopy cluster analysis and quantification methods ». In : *Patterns* 1.3 (2020), p. 100038.
- [80] Tae Jin KIM, Silvan TÜRKCAN et Guillem PRATX. « Modular low-light microscope for imaging cellular bioluminescence and radioluminescence ». In : *Nature protocols* 12.5 (2017), p. 1055-1076.
- [81] Naveen KODALI et al. « On convergence and stability of gans ». In : *arXiv preprint arXiv :1705.07215* (2017).
- [82] Hui KONG, Hatice Cinar AKAKIN et Sanjay E SARMA. « A generalized Laplacian of Gaussian filter for blob detection and its applications ». In : *IEEE transactions on cybernetics* 43.6 (2013), p. 1719-1733.
- [83] Alex KRIZHEVSKY, Ilya SUTSKEVER et Geoffrey E HINTON. « Imagenet classification with deep convolutional neural networks ». In : *Communications of the ACM* 60.6 (2017), p. 84-90.
- [84] Thibault LAGACHE et al. « Analysis of the spatial organization of molecules with robust statistics ». In : *PLoS One* 8.12 (2013), e80914.
- [85] Thibault LAGACHE et al. « Mapping molecular assemblies with fluorescence microscopy and object-based spatial statistics ». In : *Nature communications* 9.1 (2018), p. 1-15.
- [86] Lan LAN et al. « Evidence for a transport-trap mode of *Drosophila melanogaster* gurken mRNA localization ». In : *PLoS One* 5.11 (2010), e15448.
- [87] Yann LE CUN et al. « Handwritten digit recognition : Applications of neural network chips and automatic learning ». In : *IEEE Communications Magazine* 27.11 (1989), p. 41-46.
- [88] Yann LECUN, Yoshua BENGIO et Geoffrey HINTON. « Deep learning ». In : *nature* 521.7553 (2015), p. 436-444.

- [89] Yann LECUN et al. « Backpropagation applied to handwritten zip code recognition ». In : *Neural computation* 1.4 (1989), p. 541-551.
- [90] Mickaël LELEK et al. « Single-molecule localization microscopy ». In : *Nature Reviews Methods Primers* 1.1 (2021), p. 1-27.
- [91] Gang LI et al. « 3D cell nuclei segmentation based on gradient flow tracking ». In : *BMC cell biology* 8.1 (2007), p. 1-10.
- [92] Miao LIAO et al. « Automatic segmentation for cell images based on bottleneck detection and ellipse fitting ». In : *Neurocomputing* 173 (2016), p. 615-622.
- [93] Jeff W LICHTMAN et José-Angel CONCHELLO. « Fluorescence microscopy ». In : *Nature methods* 2.12 (2005), p. 910-919.
- [94] Sungjun LIM et al. « CycleGAN with a blur kernel for deconvolution microscopy : Optimal transport geometry ». In : *IEEE Transactions on Computational Imaging* 6 (2020), p. 1127-1138.
- [95] Tony LINDBERG. « Detecting salient blob-like image structures and their scales with a scale-space primal sketch : A method for focus-of-attention ». In : *International Journal of Computer Vision* 11.3 (1993), p. 283-318.
- [96] Howard D LIPSHITZ et Craig A SMIBERT. « Mechanisms of RNA localization and translational regulation ». In : *Current opinion in genetics & development* 10.5 (2000), p. 476-488.
- [97] Xiaolong LIU, Zhidong DENG et Yuhan YANG. « Recent progress in semantic image segmentation ». In : *Artificial Intelligence Review* 52.2 (2019), p. 1089-1106.
- [98] Ziming LIU et al. « HRDNet : high-resolution detection network for small objects ». In : *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, p. 1-6.
- [99] Jun MA. « Segmentation loss odyssey ». In : *arXiv preprint arXiv :2005.13449* (2020).
- [100] Jun MA et al. « Loss odyssey in medical image segmentation ». In : *Medical Image Analysis* 71 (2021), p. 102035.
- [101] Matsilele A MABASO, Daniel J WITHEY et Bhekisipho TWALA. « Spot detection methods in fluorescence microscopy imaging : a review ». In : (2018).
- [102] Paul M MACDONALD et Gary STRUHL. « Cis-acting sequences responsible for anterior localization of bicoid mRNA in *Drosophila* embryos ». In : *Nature* 336.6199 (1988), p. 595-598.

- [103] Klas EG MAGNUSSON. « Segmentation and tracking of cells and particles in time-lapse microscopy ». Thèse de doct. KTH Royal Institute of Technology, 2016.
- [104] Norberto MALPICA et al. « Applying watershed algorithms to the segmentation of clustered nuclei ». In : *Cytometry : The Journal of the International Society for Analytical Cytology* 28.4 (1997), p. 289-297.
- [105] EM MANDERS et al. « Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy ». In : *Journal of cell science* 103.3 (1992), p. 857-862.
- [106] EMM MANDERS, FJ VERBEEK et JA ATEN. « Measurement of co-localization of objects in dual-colour confocal images ». In : *Journal of microscopy* 169.3 (1993), p. 375-382.
- [107] Xudong MAO et al. « Least squares generative adversarial networks ». In : *Proceedings of the IEEE international conference on computer vision*. 2017, p. 2794-2802.
- [108] Faraz K MARDAKHEH et al. « Global analysis of mRNA, translation, and protein localization : local translation is a key regulator of cell protrusions ». In : *Developmental cell* 35.3 (2015), p. 344-357.
- [109] Estibaliz Gómez-de MARISCAL et al. « DeepImageJ : A user-friendly environment to run deep learning models in ImageJ ». In : *Nature Methods* 18.10 (2021), p. 1192-1195.
- [110] Koen JA MARTENS, Bartosz TURKOWYD et Ulrike ENDESFELDER. « Raw data to results : a hands-on introduction and overview of computational analysis for single-molecule localization microscopy ». In : *Frontiers in Bioinformatics* (2022), p. 91.
- [111] Kelsey C MARTIN et Anne EPHRUSSI. « mRNA localization : gene expression in the spatial dimension ». In : *Cell* 136.4 (2009), p. 719-730.
- [112] Vivien MARX. « Method of the Year : spatially resolved transcriptomics ». In : *Nature methods* 18.1 (2021), p. 9-14.
- [113] Warren S MCCULLOCH et Walter PITTS. « A logical calculus of the ideas immanent in nervous activity ». In : *The bulletin of mathematical biophysics* 5.4 (1943), p. 115-133.
- [114] Claire MCQUIN et al. « CellProfiler 3.0 : Next-generation image processing for biology ». In : *PLoS biology* 16.7 (2018), e2005970.
- [115] Erik MEIJERING. « A bird's-eye view of deep learning in bioimage analysis ». In : *Computational and structural biotechnology journal* 18 (2020), p. 2312-2325.

- [116] Stavroula MILI, Konstadinos MOISSOGLU et Ian G MACARA. « Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions ». In : *Nature* 453.7191 (2008), p. 115-119.
- [117] Jeffrey R MOFFITT et al. « High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing ». In : *Proceedings of the National Academy of Sciences* 113.50 (2016), p. 14456-14461.
- [118] Monica MONICI. « Cell and tissue autofluorescence research and diagnostic applications ». In : *Biotechnology annual review* 11 (2005), p. 227-256.
- [119] Andreas E MOOR et al. « Global mRNA polarization regulates translation efficiency in the intestinal epithelium ». In : *Science* 357.6357 (2017), p. 1299-1303.
- [120] Elias NEHME et al. « Deep-STORM : super-resolution single-molecule microscopy by deep learning ». In : *Optica* 5.4 (2018), p. 458-464.
- [121] Hans NETTEN et al. « FISH and chips : automation of fluorescent dot counting in interphase cell nuclei ». In : *Cytometry : The Journal of the International Society for Analytical Cytology* 28.1 (1997), p. 1-10.
- [122] Junhyug NOH et al. « Better to follow, follow to be better : Towards precise supervision of feature super-resolution for small object detection ». In : *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019, p. 9725-9734.
- [123] Adaobi NWANESHIUDU et al. « Introduction to confocal microscopy ». In : *Journal of Investigative Dermatology* 132.12 (2012), p. 1-5.
- [124] Jean-Christophe OLIVO-MARIN. « Extraction of spots in biological images using multiscale products ». In : *Pattern recognition* 35.9 (2002), p. 1989-1996.
- [125] Anton OSOKIN et al. « GANs for biological image synthesis ». In : *Proceedings of the IEEE International Conference on Computer Vision*. 2017, p. 2233-2242.
- [126] Wei OUYANG et al. « Deep learning massively accelerates super-resolution localization microscopy ». In : *Nature biotechnology* 36.5 (2018), p. 460-468.
- [127] Olivia PADOVAN-MERHAR et Arjun RAJ. « Cellular volume is a global controller of mRNA abundance ». In : *Biophysical Journal* 104.2 (2013), 328a.
- [128] Giovanni PALLA et al. « Squidpy : a scalable framework for spatial omics analysis ». In : *Nature methods* 19.2 (2022), p. 171-178.
- [129] Hye Yoon PARK et al. « An unbiased analysis method to quantify mRNA localization reveals its correlation with cell motility ». In : *Cell reports* 1.2 (2012), p. 179-184.

- [130] Taesung PARK et al. « Contrastive learning for unpaired image-to-image translation ». In : *European conference on computer vision*. Springer. 2020, p. 319-345.
- [131] Xue Bin PENG et al. « Deeploco : Dynamic locomotion skills using hierarchical deep reinforcement learning ». In : *ACM Transactions on Graphics (TOG)* 36.4 (2017), p. 1-13.
- [132] Mangal PRAKASH, Alexander KRULL et Florian JUG. « Fully Unsupervised Diversity Denoising with Convolutional Variational Autoencoders ». In : *International Conference on Learning Representations*. 2021. URL : <https://openreview.net/forum?id=agHLC0BM5jP>.
- [133] Bernhard PREIM et Charl P BOTHA. *Visual computing for medicine : theory, algorithms, and applications*. Newnes, 2013.
- [134] Zsuzsanna PÜSPÖKI et al. « Spotcaliper : fast wavelet-based spot detection with accurate size estimation ». In : *Bioinformatics* 32.8 (2016), p. 1278-1280.
- [135] Arjun RAJ et al. « Imaging individual mRNA molecules using multiple singly labeled probes ». In : *Nature methods* 5.10 (2008), p. 877-879.
- [136] Arjun RAJ et al. « Stochastic mRNA synthesis in mammalian cells ». In : *PLoS biology* 4.10 (2006), e309.
- [137] Pauli RÄMÖ et al. « CellClassifier : supervised learning of cellular phenotypes ». In : *Bioinformatics* 25.22 (2009), p. 3028-3030.
- [138] Douglas S RICHARDSON et Jeff W LICHTMAN. « Clarifying tissue clearing ». In : *Cell* 162.2 (2015), p. 246-257.
- [139] Brian D RIPLEY. « Modelling spatial patterns ». In : *Journal of the Royal Statistical Society : Series B (Methodological)* 39.2 (1977), p. 172-192.
- [140] Olaf RONNEBERGER, Philipp FISCHER et Thomas BROX. « U-net : Convolutional networks for biomedical image segmentation ». In : *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, p. 234-241.
- [141] David E RUMELHART, Geoffrey E HINTON et Ronald J WILLIAMS. « Learning representations by back-propagating errors ». In : *nature* 323.6088 (1986), p. 533-536.
- [142] David E RUMELHART et al. « Backpropagation : The basic theory ». In : *Backpropagation : Theory, architectures and applications* (1995), p. 1-34.

- [143] Michael J RUST, Mark BATES et Xiaowei ZHUANG. « Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM) ». In : *Nature methods* 3.10 (2006), p. 793-796.
- [144] Mohammed SABER et al. « Artificial neural networks, support vector machine and energy detection for spectrum sensing based on real signals ». In : *International Journal of Communication Networks and Information Security* 11.1 (2019), p. 52-60.
- [145] Daniel SAGE et al. « Automatic tracking of individual fluorescence particles : application to the study of chromosome dynamics ». In : *IEEE transactions on image processing* 14.9 (2005), p. 1372-1383.
- [146] Aubin SAMACOITS et al. « A computational framework to study sub-cellular RNA localization ». In : *Nature Communications* 9.1 (2018), p. 1-10.
- [147] Gilberto dos SANTOS, Andrew J SIMMONDS et Henry M KRAUSE. « A stem-loop structure in the wingless transcript defines a consensus motif for apical RNA transport ». In : *Development* 135 (2008).
- [148] Anca F SAVULESCU et al. « Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH ». In : *Cell reports methods* 1.5 (2021), p. 100068.
- [149] Anca Flavia SAVULESCU et al. « Prediction of RNA subcellular localization : Learning from heterogeneous data sources ». In : *Isience* 24.11 (2021), p. 103298.
- [150] Manish SAXENA, Gangadhar ELURU et Sai Siva GORTHI. « Structured illumination microscopy ». In : *Advances in Optics and Photonics* 7.2 (2015), p. 241-275.
- [151] Kristine SCHAUER et al. « Probabilistic density maps to study global endomembrane organization ». In : *Nature methods* 7.7 (2010), p. 560-566.
- [152] Johannes SCHINDELIN et al. « Fiji : an open-source platform for biological-image analysis ». In : *Nature methods* 9.7 (2012), p. 676-682.
- [153] Uwe SCHMIDT et al. « Cell detection with star-convex polygons ». In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2018, p. 265-273.
- [154] Uwe SCHMIDT et al. « Cell detection with star-convex polygons ». In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2018, p. 265-273.

- [155] Andreas SCHULZE BÄING. « Spatial Analysis ». In : *Encyclopedia of Quality of Life and Well-Being Research*. Sous la dir. d'Alex C. MICHALOS. Dordrecht : Springer Netherlands, 2014, p. 6259-6261. ISBN : 978-94-007-0753-5. DOI : [10.1007/978-94-007-0753-5_2818](https://doi.org/10.1007/978-94-007-0753-5_2818). URL : https://doi.org/10.1007/978-94-007-0753-5_2818.
- [156] Diane C SHAKES, David M MILLER III et Michael L NONET. « Immunofluorescence microscopy ». In : *Methods in cell biology*. T. 107. Elsevier, 2012, p. 35-66.
- [157] Karen SIMONYAN et Andrew ZISSERMAN. « Very deep convolutional networks for large-scale image recognition ». In : *arXiv preprint arXiv :1409.1556* (2014).
- [158] Ihor SMAL et al. « Quantitative comparison of spot detection methods in fluorescence microscopy ». In : *IEEE transactions on medical imaging* 29.2 (2009), p. 282-301.
- [159] Carolyn L SMITH. « Basic confocal microscopy ». In : *Current protocols in neuroscience* 56.1 (2011), p. 2-2.
- [160] Artur SPEISER et al. « Deep learning enables fast and dense single-molecule localization with high accuracy ». In : *Nature methods* 18.9 (2021), p. 1082-1090.
- [161] Rupesh Kumar SRIVASTAVA, Klaus GREFF et Jürgen SCHMIDHUBER. « Highway networks ». In : *arXiv preprint arXiv :1505.00387* (2015).
- [162] Weston STAUFFER, Huanjie SHENG et Han N LIM. « EzColocalization : An ImageJ plugin for visualizing and measuring colocalization in cells and organisms ». In : *Scientific reports* 8.1 (2018), p. 1-13.
- [163] Ernst HK STELZER. « Light-sheet fluorescence microscopy for quantitative biology ». In : *Nature methods* 12.1 (2015), p. 23-26.
- [164] Ernst HK STELZER et al. « Light sheet fluorescence microscopy ». In : *Nature Reviews Methods Primers* 1.1 (2021), p. 1-25.
- [165] David J STEPHENS et Victoria J ALLAN. « Light microscopy techniques for live cell imaging ». In : *science* 300.5616 (2003), p. 82-86.
- [166] Carsen STRINGER et al. « Cellpose : a generalist algorithm for cellular segmentation ». In : *Nature methods* 18.1 (2021), p. 100-106.
- [167] Michael STUELAND et al. « RDI calculator : an analysis tool to assess RNA distributions in cells ». In : *Scientific reports* 9.1 (2019), p. 1-10.
- [168] Nima TAJBAKHSI et al. « Embracing imperfect datasets : A review of deep learning solutions for medical image segmentation ». In : *Medical Image Analysis* 63 (2020), p. 101693.

- [169] Betty I TARNOWSKI, Francis G SPINALE et James H NICHOLSON. « DAPI as a useful stain for nuclear quantitation ». In : *Biotechnic & histochemistry* 66.6 (1991), p. 296-302.
- [170] Diethard TAUTZ et Christine PFEIFLE. « A non-radioactive in situ hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback ». In : *Chromosoma* 98.2 (1989), p. 81-85.
- [171] Hildegard TEKOTTE et Ilan DAVIS. « Intracellular mRNA localization : motors move messages ». In : *TRENDS in Genetics* 18.12 (2002), p. 636-642.
- [172] Manuel THÉRY et al. « Anisotropy of cell adhesive microenvironment governs cell internal organization and orientation of polarity ». In : *Proceedings of the National Academy of Sciences* 103.52 (2006), p. 19771-19776.
- [173] Jim A THOMAS. « Optical imaging probes for biomolecules : an introductory perspective ». In : *Chemical Society Reviews* 44.14 (2015), p. 4494-4500.
- [174] Marj TONINI et al. « Spatial pattern of landslides in Swiss Rhone Valley ». In : *Natural Hazards* 73.1 (2014), p. 97-110.
- [175] Simone TONTI et al. « An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test ». In : *Computerized Medical Imaging and Graphics* 40 (2015), p. 62-69.
- [176] Tatjana TRCEK et al. « mRNA quantification using single-molecule FISH in *Drosophila* embryos ». In : *nature protocols* 12.7 (2017), p. 1326-1348.
- [177] Maria VERA et al. « Single-cell and single-molecule analysis of gene expression regulation ». In : *Annual review of genetics* 50 (2016), p. 267.
- [178] Sirion VITTAYAKORN et James HAYS. « Quality Assessment for Crowdsourced Object Annotations. » In : *BMVC*. 2011, p. 1-11.
- [179] Duc My VO et Sang-Woong LEE. « Semantic image segmentation using fully convolutional neural networks with multi-scale images and multi-scale dilated convolutions ». In : *Multimedia Tools and Applications* 77.14 (2018), p. 18689-18707.
- [180] Aarno Oskar VUOLA, Saad Ullah AKRAM et Juho KANNALA. « Mask-RCNN and U-net ensembled for nuclei segmentation ». In : *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE. 2019, p. 208-212.
- [181] Carolina WÄHLBY et al. « Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections ». In : *Journal of microscopy* 215.1 (2004), p. 67-76.

- [182] Qi WANG et al. « A comprehensive survey of loss functions in machine learning ». In : *Annals of Data Science* 9.2 (2022), p. 187-212.
- [183] Shulei WANG et al. « Spatially adaptive colocalization analysis in dual-color fluorescence microscopy ». In : *IEEE Transactions on Image Processing* 28.9 (2019), p. 4471-4485.
- [184] Wenjian WANG et al. « Resolution enhancement in microscopic imaging based on generative adversarial network with unpaired data ». In : *Optics Communications* 503 (2022), p. 127454.
- [185] Yueyue WANG et al. « Up and Down Residual Blocks for Convolutional Generative Adversarial Networks ». In : *IEEE Access* 9 (2021), p. 26051-26058. DOI : [10.1109/ACCESS.2021.3056572](https://doi.org/10.1109/ACCESS.2021.3056572).
- [186] Zhou WANG et al. « Image quality assessment : from error visibility to structural similarity ». In : *IEEE transactions on image processing* 13.4 (2004), p. 600-612.
- [187] Jonathan H WARRELL et al. « Generalized statistical tests for mRNA and protein subcellular spatial patterning against complete spatial randomness ». In : *arXiv preprint arXiv :1602.06429* (2016).
- [188] Martin WEIGERT et al. « Content-aware image restoration : pushing the limits of fluorescence microscopy ». In : *Nature methods* 15.12 (2018), p. 1090-1097.
- [189] Cameron G WILLIAMS et al. « An introduction to spatial transcriptomics for biomedical research ». In : *Genome Medicine* 14.1 (2022), p. 1-18.
- [190] Huikai WU et al. « Fastfcn : Rethinking dilated convolution in the backbone for semantic segmentation ». In : *arXiv preprint arXiv :1903.11816* (2019).
- [191] Qiang WU, Fatima MERCHANT et Kenneth CASTLEMAN. *Microscope image processing*. Elsevier, 2010.
- [192] Mai YAMAGISHI, Yoshitaka SHIRASAKI et Takashi FUNATSU. « Size-dependent accumulation of mRNA at the leading edge of chicken embryo fibroblasts ». In : *Biochemical and biophysical research communications* 390.3 (2009), p. 750-754.
- [193] Kyota YASUDA et al. « FUS inclusions disrupt RNA localization by sequestering kinesin-1 and inhibiting microtubule detyrosination ». In : *Journal of Cell Biology* 216.4 (2017), p. 1015-1034.
- [194] Alexander P YOUNG, Daniel J JACKSON et Russell C WYETH. « A technical review and guide to RNA fluorescence in situ hybridization ». In : *PeerJ* 8 (2020), e8806.
- [195] Fisher YU et Vladlen KOLTUN. « Multi-scale context aggregation by dilated convolutions ». In : *arXiv preprint arXiv :1511.07122* (2015).

- [196] Fisher YU et Vladlen KOLTUN. « Multi-Scale Context Aggregation by Dilated Convolutions ». In : *International Conference on Learning Representations (ICLR)*. Mai 2016.
- [197] Rafael YUSTE. « Fluorescence microscopy today ». In : *Nature methods* 2.12 (2005), p. 902-904.
- [198] Alessandra ZAPPULO et al. « RNA localization is a key determinant of neurite-enriched proteome ». In : *Nature communications* 8.1 (2017), p. 1-13.
- [199] Min ZHANG. *Small Blob Detection in Medical Images*. Arizona State University, 2015.
- [200] Xinlei ZHAO et al. « Evaluation of single-cell classifiers for single-cell RNA sequencing data sets ». In : *Briefings in bioinformatics* 21.5 (2020), p. 1581-1595.
- [201] Jun-Yan ZHU et al. « Unpaired image-to-image translation using cycle-consistent adversarial networks ». In : *Proceedings of the IEEE international conference on computer vision*. 2017, p. 2223-2232.