A dictionary-based denoising method toward a robust segmentation of noisy and densely packed nuclei in 3D biological microscopy images

Lamees Nasser Khalafallah Mahmoud

SORBONNE UNIVERSITÉ

Sorbonne Université

Ecole doctorale : Informatique, Télécommunications et électronique de Paris

Laboratoire : Image and Pervasive Access Lab., CNRS UMI 2955, Singapour

Titre de la thèse

Une méthode de débruitage basée sur la méthode des dictionnaires pour une segmentation robuste de noyaux bruités et denses dans des images biologique 3D de microscopie

Présentée par

Lamees NASSER KHALAFALLAH MAHMOUD

Thèse de doctorat de Sorbonne Université

Dirigée par Thomas BOUDIER

Présentée et soutenue publiquement le 8 Juillet 2019

Devant un jury composé de :

Dr. Xavier Descombes, INRIA, France.	Rapporteur
Dr. Désiré Sidibé, Université de Bourgogne, France.	Rapporteur
Dr. Joëlle Sobczak-Thépot, Sorbonne Université, France.	Examinateur
Dr. Valeriu Vrabie, Université de Reims Champagne-Ardenne, France.	Examinateur
Dr. Sylvie Chambon, Université de Toulouse, France.	Examinateur
Dr. Thomas Boudier, Sorbonne Université, France.	Directeur de thèse

SORBONNE UNIVERSITÉ

Sorbonne University

Doctoral School: Computer Science, Telecommunications and Electronics of Paris

Laboratory: Image and Pervasive Access Lab., CNRS UMI 2955, Singapore

Thesis title

A dictionary-based denoising method toward a robust segmentation of noisy and densely packed nuclei in 3D biological microscopy images

Presented by

Lamees NASSER KHALAFALLAH MAHMOUD

Doctoral thesis Sorbonne University

Directed by Thomas BOUDIER

Defended on 8 July 2019

In front of a jury composed of:

Dr. Xavier Descombes, INRIA, France.	Reviewer
Dr. Désiré Sidibé, University of Burgundy, France.	Reviewer
Dr. Joëlle Sobczak-Thépot, Sorbonne University, France	Examiner
Dr. Valeriu Vrabie, University of Reims Champagne-Ardenne, France.	Examiner
Dr. Sylvie Chambon, University of Toulouse, France.	Examiner
Dr. Thomas Boudier, Sorbonne University, France.	Thesis director

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my thesis director, Professor. Thomas Boudier for the continuous support of my Ph.D. study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

My Ph.D. was partially supported by A*STAR Graduate Academy (AGA). I would also like to give great thanks to Professor. Chiam Keng Hwee who accepted to be my co-supervisor and for his follow up and support throughout my research.

Special thanks go to Professor. Mounir Mokhtari, the Image, and Pervasive Access Lab's (IPAL) director, for giving me the opportunity to work at IPAL and for his support throughout my Ph.D study.

I am thankful to my beloved husband Ibrahim who has always supported me and also helped me by proofreading of this thesis.

Thanks to my beloved daughter Lujain and my beloved son Anas who made me stronger, better and more fulfilled than I could have ever imagined.

Last but not the least important, very special thanks go to my parents and sibling, who supported and encouraged me during stressful moments.

Abstract

Cells are the basic building blocks of all living organisms. All living organisms share life processes such as growth and development, movement, nutrition, excretion, reproduction, respiration and response to the environment.

In cell biology research, understanding cells structure and function is essential for developing and testing new drugs. In addition, cell biology research provides a powerful tools to study embryo development. Furthermore, it helps the scientific research community to understand the effects of mutations and various diseases.

Time-Lapse Fluorescence Microscopy (TLFM) is one of the most appreciated imaging techniques which can be used in live-cell imaging experiments to quantify various characteristics of cellular processes, i.e., cell survival, proliferation, migration, and differentiation.

In TLFM imaging, not only spatial information is acquired, but also temporal information obtained by repeating imaging of a labeled sample at specific time points, as well as spectral information, that produces up to *five-dimensional* (X, Y, Z + Time + Channel) images. Typically, the generated datasets consist of several (hundreds or thousands) images, each containing hundreds to thousands of objects to be analyzed.

To perform high-throughput quantification of cellular processes, nuclei segmentation and tracking should be performed in an automated manner. Nevertheless, nuclei segmentation and tracking are challenging tasks due to embedded noise, intensity inhomogeneity, shape variation as well as a weak boundary of nuclei. Although several nuclei segmentation approaches have been reported in the literature, dealing with embedded noise remains the most challenging part of any segmentation algorithm.

We propose a novel 3D denoising algorithm, based on unsupervised dictionary learning and sparse representation, that can both enhance very faint and noisy nuclei, in addition, it simultaneously detects nuclei position accurately. Furthermore, our method is based on a limited number of parameters, with only one being critical, which is the approximate size of the objects of interest.

The framework of the proposed method comprises image denoising, nuclei detection, and segmentation. In the denoising step, an initial dictionary is constructed by selecting random patches from the raw image then an iterative technique is implemented to update the dictionary and obtain the final one which is less noisy. Next, a detection map, based on the dictionary coefficients used to denoise the image, is used to detect marker points. Afterward, a thresholding-based approach is proposed to get the segmentation mask. Finally, a marker-controlled watershed approach is used to get the final nuclei segmentation result.

We generate 3D synthetic images to study the effect of the few parameters of our method on cell nuclei detection and segmentation, and to understand the overall mechanism for selecting and tuning the significant parameters of the several datasets. These synthetic images have low contrast and low signal to noise ratio. Furthermore, they include touching spheres where these conditions simulate the same characteristics exist in the real datasets.

The proposed framework shows that integrating our denoising method along with classical segmentation method works properly in the context of the most challenging cases. To evaluate the performance of the proposed method, two datasets from the cell tracking challenge are extensively tested. Across all datasets, the proposed method achieved very promising results with 96.96% recall for the *C.elegans* dataset. Besides, in the *Drosophila* dataset, our method achieved very high recall (99.3%).

Keywords:

Unsupervised dictionary learning; Sparse representation; Denoising; Nuclei segmentation; Embryo development; 3D time-lapse fluorescence microscopy.

Résumé

Les cellules sont les éléments constitutifs de base de tout organisme vivant. Tous les organismes vivants partagent des processus vitaux tels que croissance, développement, mouvement, nutrition, excrétion, reproduction, respiration et réaction à l'environnement.

En biologie cellulaire, comprendre la structure et fonction des cellules est essentielle pour développer et tester de nouveaux médicaments. Par ailleurs, cela aide aussi à l'étude du développement embryonnaire. Enfin, cela permet aux chercheurs de mieux comprendre les effets des mutations et de diverses maladies.

La vidéo-microscopie (Time Lapse Fluorescence Microscopie) est l'une des techniques d'imagerie les plus utilisées afin de quantifier diverses caractéristiques des processus cellulaires, à savoir la survie, la prolifération, la migration ou la différenciation cellulaire.

En vidéo-microscopie, non seulement les informations spatiales sont disponibles, mais aussi les informations temporelles en réitérant l'acquisition de l'échantillon, et enfin les informations spectrales, ce qui génère des données dites « cinq dimensions » (X, Y, Z + temps + canal). En règle générale, les jeux de données générés consistent en plusieurs (centaines ou milliers) d'images, chacune contenant des centaines ou milliers d'objets à analyser.

Pour effectuer une quantification précise et à haut débit des processus cellulaires, les étapes de segmentation et de suivi des noyaux cellulaires doivent être effectuées de manière automatisée. Cependant, la segmentation et le suivi des noyaux sont des tâches difficiles dû notamment au bruit intrinsèque dans les images, à l'inhomogénéité de l'intensité, au changement de forme des noyaux ainsi qu'à un faible contraste des noyaux. Bien que plusieurs approches de segmenta-

tion des noyaux aient été rapportées dans la littérature, le fait de traiter le bruit intrinsèque reste la partie la plus difficile de tout algorithme de segmentation.

Nous proposons un nouvel algorithme de débruitage 3D, basé sur l'apprentissage d'un dictionnaire non supervisé et une représentation parcimonieuse, qui à la fois améliore la visualisation des noyaux très peu contrastés et bruités, mais aussi détecte simultanément la position de ces noyaux avec précision. De plus, notre méthode possède un nombre limité de paramètres, un seul étant critique, à savoir la taille approximative des objets à traiter.

Le cadre de la méthode proposée comprend le débruitage d'images, la détection des noyaux et leur segmentation. Dans l'étape de débruitage, un dictionnaire initial est construit en sélectionnant des régions (patches) aléatoires dans l'image originale, puis une technique itérative est implémentée pour mettre à jour ce dictionnaire afin d'obtenir un dictionnaire dont les éléments mis à jour présentent un meilleur contraste. Ensuite, une carte de détection, basée sur le calcul des coefficients du dictionnaire utilisés pour débruiter l'image, est utilisée pour détecter le centre approximatif des noyaux qui serviront de marqueurs pour la segmentation. Ensuite, une approche basée sur le seuillage est proposée pour obtenir le masque de segmentation des noyaux. Finalement, une approche de segmentation par partage des eaux contrôlée par les marqueurs est utilisée pour obtenir le résultat final de segmentation des noyaux.

Nous avons créé des images synthétiques 3D afin d'étudier l'effet des paramètres de notre méthode sur la détection et la segmentation des noyaux, et pour comprendre le mécanisme global de sélection et de réglage de ces paramètres significatifs sur différents jeux de données. Ces images synthétiques ont un très faible contraste et un rapport signal/bruit aussi très faible. De plus, ces images contiennent des sphères touchantes qui simulent les caractéristiques des jeux de données réels.

Le cadre proposé montre que l'intégration de notre méthode de débruitage ainsi que la méthode de segmentation fonctionne correctement même pour les cas les plus difficiles. Pour évaluer les performances de la méthode proposée, deux jeux de données du défi de suivi des cellules sont testés. Dans l'ensemble, la méthode proposée a obtenu des résultats très prometteurs avec un rappel de 96.96% pour le jeu de données de C. elegans. De même, pour le jeu de données

Drosophile, notre méthode a obtenu un rappel très élevé de (99.3%) %.

Mots-clés:

Apprentissage de dictionnaire non supervisé; Représentation clairsemée; Débruitage; Segmentation de noyaux; Développement embryonnaire; microscopie à fluorescence 3D.

Publications

1. **Nasser, L.**, Boudier, T., 2019. "A novel generic dictionary-based denoising method for improving noisy and densely packed nuclei segmentation in 3D time-lapse fluorescence microscopy images." Scientific reports 9 (2019): 5654. [Impact factor 4.122].
   doi:10.1038/s41598-019-41683-3

2. **Nasser, L.**, Coronado, P., David, E., Boudier, T., 2017. "A dictionary-based approach to reduce noise in fluorescent microscopy images." 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP). IEEE.
   doi: 10.1109/SIPROCESS.2017.8124522

3. Liu, Z., Hu,Y., Xu, H., **Nasser, L.** ; Coquet, P., Boudier, T., Yu, H., 2017. "NucleiNet: a convolutional encoder-decoder network for bio-image denoising." 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE.
   doi: 10.1109/EMBC.2017.8037240

Abbreviations

API Application programming interface

AWGN additive white Gaussian noise

BM3D Block-matching 3D

BP Basis pursuit

CARE content-aware image restoration

C.elegans Caenorhabditis Elegans

C-LSTM Convolutional long short term memory

CNN Convolutional neural network

CTC Cell Tracking Challenge

CT Computed tomography

DAPI $4'$, $6-$diamidino$-2-$phenylindole

DCNNs Deep convolutional neural networks

DFB the directional filter bank

DNA Deoxyribonucleic Acid

FCN Fully convolutional network

FLS Fast level set

FN False negative

FP False positive

GAN Generative adversarial network

GC Graph cut

GFP Green fluorescent protein

GMM Gaussian mixture model

GM-PHD Gaussian mixture probability hypothesis density

HMM Hidden Markov model

ILP Integer linear programming

IMM (interacting multiplemodel

ISBI IEEE International Symposium on Biomedical Imaging

K-SVD K-means with singular value decomposition

LP Laplacian pyramid

LSCM Laser scanning confocal microscopy

LSFM Light sheet fluorescence microscopy

MAP Maximum a posteriori

M Mitotic

MP Matching pursuit

MRI Magnetic resonance imaging

MSE mean square error

MTL multi-task learning

NIH National institutes of health

NLM non-local means

OMP Orthogonal matching pursuit

PDE (Partial differential equation)

PMT photomultiplier tube

PURE-LET Poisson Unbiased Risk Estimation Linear Expansion of Thresholds

SDCM Spinning disk confocal microscopy

SNR Signal to noise ratio

SRF stochastically connected random field

SVD Singular value decomposition

TEM Transmission Electron Microscopy

TLFM Time-Lapse Fluorescence Microscopy

TP True positive

VTL Viterbi track linking

WCHT Wavelet-based circular hough transform

ZIB Zuse institute Berlin

3D Three-dimensional

2D Two-dimensional

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1 Biological Background

Cells are the basic building blocks of all living organisms. All living organisms share life processes such as growth and development, movement, nutrition, excretion, reproduction, respiration and response to the environment. These life processes become the criteria for scientists to differentiate between the living and the non-living things in nature.

Roughly speaking, cells have three main parts (as shown in Figure 1.1), each with a different function. The first part, the membrane is the outermost layer in the animal cell and is found inside the cell wall in the plant cell. The second part is the nucleus that contains hereditary genetic information (DNA) and controls all cell activities. The third part, the cytoplasm, which consists of the complete contents of a biological cell i.e., enzymes, and various organic molecules, excluding the nucleus.

In cell biology research, understanding the structure and the function of the cells is essential for developing and testing new drugs. In addition, it provides a powerful tool to study embryo development. Furthermore, it helps the scientific research community to understand the effects of mutations and various diseases. Indeed, as emphasized long ago by the pioneering cell biologist E.B Wilson [Wilson 1900] "*The key to every biological problem must finally be sought in the cell, for every living organism is, or at some time has been, a cell*".

Nowadays, advances in microscopy techniques have enabled scientists to get a better idea of how the cells are structured and how the new morphological characteristics of living cells are visualized. As these techniques allow high-throughput, high-resolution imaging of a wide variety of samples in t*hree-dimensional* space (X, Y, Z) which can raise up to *five-dimensional* space (X, Y, Z Time Channel) [Meijering 2008].

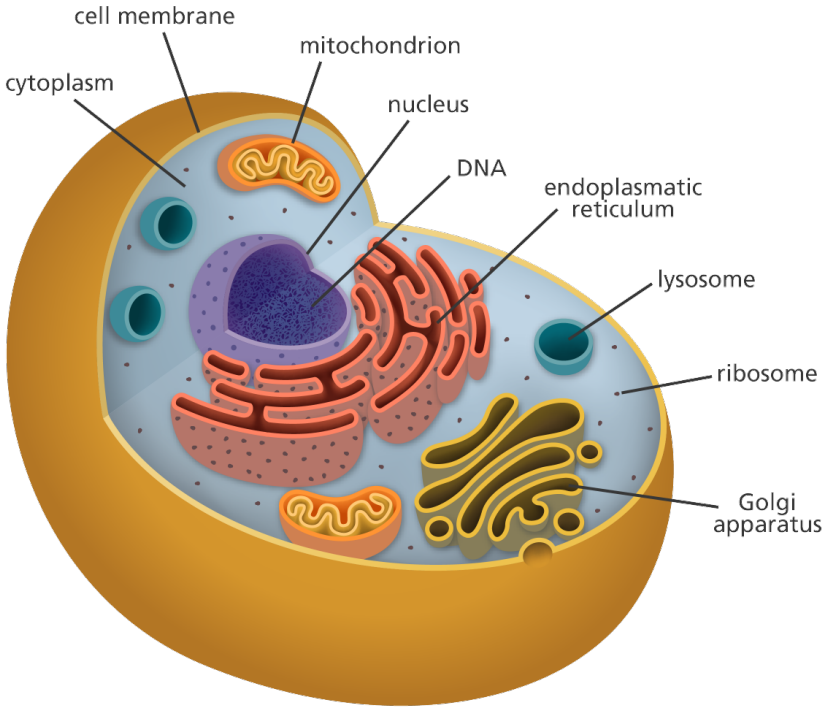

Figure 1.1: The structure of an animal cell. Retrieved from yourgenome: https://www.yourgenome.org/facts/what-is-a-cell

1.1.1 Cell cycle

The cell cycle, as known as the cell-division cycle [Mitchison 1971], is the series of events by which cells grow and divide to produce two daughter cells (as depicted in Figure 1.2). In order for a cell to divide, many important tasks must be completed. To explain, it should grow, copy its genetic information (DNA), and physically split to produce two identical daughter cells. The cell cycle consists of two main phases, namely an interphase, and a mitotic (M) phase. In the interphase, the cell grows, replicates its genetic material and produces proteins. Similarly, during the mitotic (M) phase, the cell divides in two identical daughter cells. Each of these phases includes sub-phases that correspond to certain cellular events. At any given time, a cell is either in an interphase or a mitosis.

Furthermore, the cell-division cycle is a vital process by which a single-cell fertilized egg develops into a mature organism, as well as, the process by which all organs and tissues are renewed. In the following sections, we discuss some of the common model organisms used for studying the cell cycle.

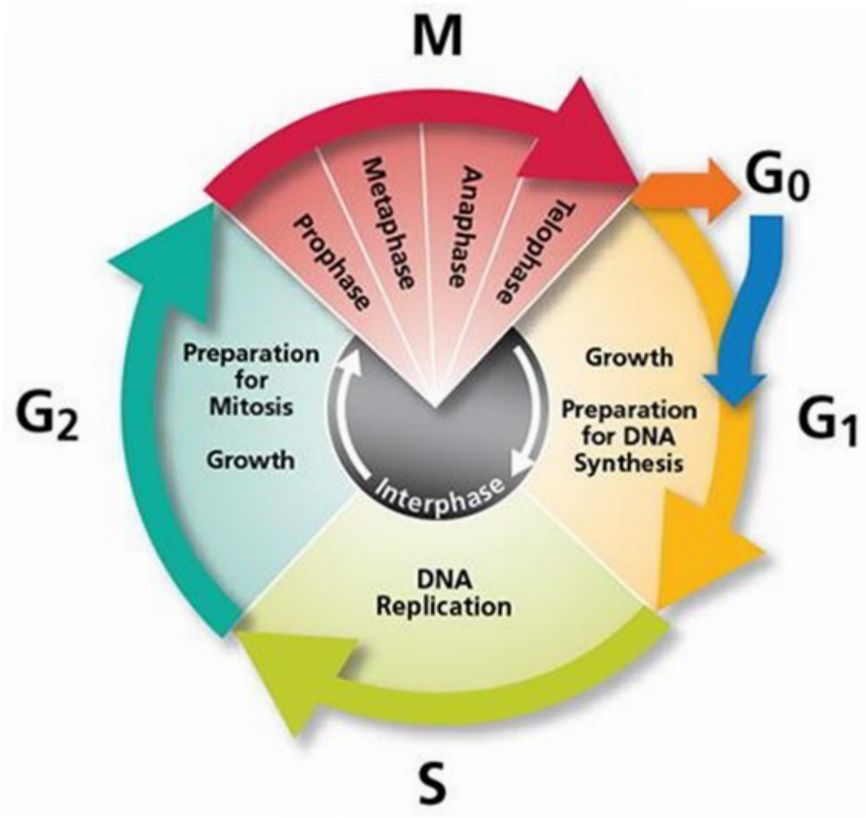

Figure 1.2: The two major phases of the cell cycle include mitosis (M), and interphase, as well as their sub-phases. . Retrieved from Earth's Lab website: https://www.earthslab.com/physiology/the-cell-cycle/

1.1.2 Model organisms for studying the cell cycle.

Model organisms are used to investigate the basic mechanisms common to all living organisms and to understand the biological processes that may be difficult or unethical to experiment in humans. The model organisms are usually chosen

for reasons that make the study easier. Some of these reasons are the transparent bodies of the organisms, the ability to grow and reproduce quickly in a small space and the prominent cell structure of interest or the ability to closely model some aspect of human biology. Most model organisms combine many if not all of these characteristics. In the following section, we present the two model organisms relevant to our experiments.

Caenorhabditis elegans (C.elegans)

The nematode *C. elegans* [Brenner 1974] has emerged as an important animal model in different fields such as neurobiology, developmental biology, and genetics. The characteristics of this animal model that have contributed to its success are being easy to culture; quick reproduction with a short generation time enabling large-scale production of organisms; small size, which allows organisms to grow in a single well of a 96-well plate; transparency throughout its life, which enables the use of fluorescent markers to study biological processes *in vivo*; as well as cellular complexity because *C. elegans* is a multicellular organism which has multiple organs and tissues. Moreover, *C. elegans* investigations have already provided a better understanding of the underlying mechanism for several diseases such as neurological disorders, congenital heart disease, kidney disease, and other diseases.

Drosophila melanogaster

Similar to the *C.elegans* model, the fly Drosophila melanogaster [Adams 2000] is one of the most extensively studied organism in biology and it serves as a model system for the investigation of many developmental and cellular processes.

Several features of *Drosophila* make it a powerful tool in developmental biology. These features include having a close relationship with human genes (in particular the sequences of recently discovered human genes including disease genes can be matched with equivalent genes in the fly); short and simple reproduction cycle that is usually about $8-14$ days, based on the environmental temperature; small size that allows scientists to keep millions of them in the laboratory at a time; in addition providing a simple means of creating genetically modified animals that

express certain proteins such as the green fluorescent protein (GFP) [Chalfie 1994] of jellyfish for live-cell imaging experiments.

1.2 Microscopy imaging techniques

Time-lapse fluorescence microscopy (TLFM) is one of the most appreciated imaging techniques which can be used in live-cell imaging experiments to quantify various characteristics of cellular processes, such as cell survival [Payne 2018], proliferation [Rapoport 2011], migration [Bise 2011a], and differentiation [Zhang 2018]. Time-lapse imaging is a technique by which a series of images are acquired at regular time points to capture the dynamics of what is being observed.

In TLFM imaging, not only spatial information is acquired, but also temporal information, as well as spectral information, that produces up to *five-dimensional* (X, Y, Z + Time + Channel) images. Typically, the generated datasets consist of several (hundreds or thousands) images, each containing hundreds to thousands of objects to be analyzed [Meijering 2008].

The basic principle of fluorescence microscopy [Lichtman 2005] is to label the sample of interest with a fluorescent marker known as a fluorophore and then illuminate the sample with light, which is absorbed by fluorophores to emit light with a wavelength different from illumination wavelength. In order to visualize subcellular structures and processes, different markers have been proposed. The most commonly used markers are Green Fluorescent protein (GFP) [Chalfie 1994] and $4'$, 6−diamidino−2−phenylindole (DAPI) [Kapuscinski 1995].

To start with, GFP refers to the protein derived from the jellyfish Aequorea victoria. GFP has been utilized in research over a wide of biological disciplines. Scientists use GFP for a vast number of functions including examining gene expression, studying protein-protein interactions, and much more. Then, DAPI is a popular marker which is used to stain DNA and allow easy visualization of the nucleus in interphase cells and chromosomes in mitotic cells.

The following section presents a brief introduction of microscopy techniques that are relevant to this thesis. Furthermore, a comparison of different fluorescence imaging modalities is shown in Figure 1.3.

Table 1.1: **Performance of different fluorescence imaging modalities. Speed, sensitivity, and photo-toxicity are rated +, ++, and +++ from worst to best, respectively. Table adapted from [Thorn 2016]**

Type of microscopy	Maximum sample thickness	Speed	Sensitivity	Resolution
Laser scanning confocal	$100 - 200\mu m$	+	+++	+
Spinning disk confocal	$30 - 50\mu m$	+++	++	++
Light sheet	>1 mm	+++	+	+++

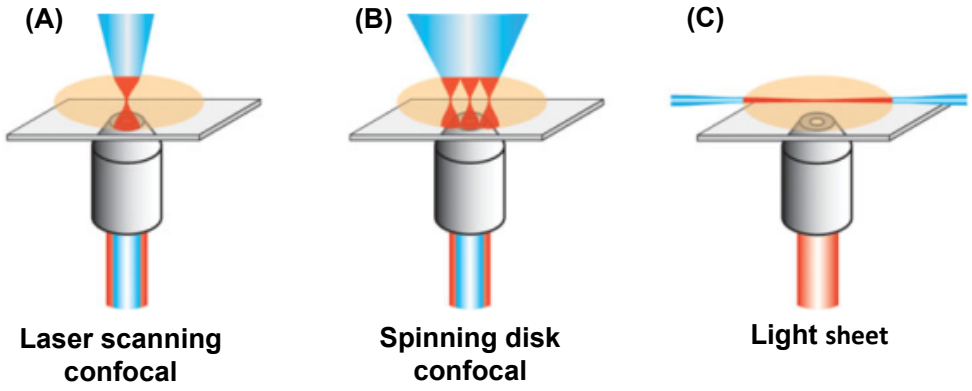

Figure 1.3: **Comparison of different fluorescence imaging modalities. (A) Laser scanning confocal where a pinhole is used to obscure the light from out-of-focus planes. (B) Spinning disk confocal where a disk of pinholes is used to speed up the scan. (C) Light sheet fluorescence microscopy (LSFM) where an excitation laser is focused into the sample from an orthogonal direction and fluorescence is collected by a separate imaging objective. Image adapted from [Adams 2015]**

1.2.1 Laser scanning confocal microscopy (LSCM)

In LSCM (Figure 1.4), a laser beam is used to illuminate a single point in the sample focal plane. Light from this point is passing through the pinhole toward a detector, and so it ensures that only light emitted from the sample focal plane is recorded on the detector. On the other hand, the light from out-

of-focus planes is obscured by the pinhole. LSCM records an image point by point by (a raster pattern) using a point detector rather than using cameras, and as a result, the system becomes less sensitive than cameras. To overcome this limitation, systems that scan multiple points across the sample simultaneously (illuminates the sample with multiple pinholes simultaneously) and image the resulting emission on a camera have been designed and it is defined as spinning disk confocal microscopy [Stephens 2003, Adams 2015, Thorn 2016].

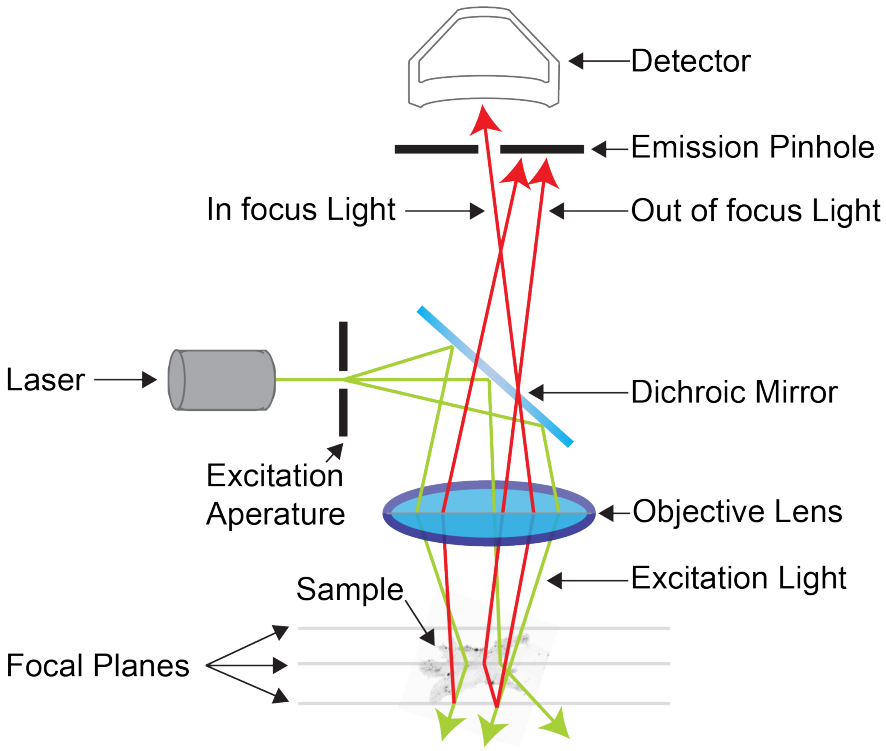

Figure 1.4: Principle of the laser scanning confocal microscopy. Retrieved from the Institute for Molecular Bioscience, University of Queensland: https://imb.uq.edu.au/facilities/microscopy/hardware-software/confocal-microscopes

1.2.2 Spinning disk confocal microscopy (SDCM)

SDCM (Figure 1.5) has been proposed to overcome the limitations of LSCM by exploiting the multiplex principle. This principle relied on using a disk of pinholes that sweep across every point in the sample. Consequently, a rotation of the disk scans over every point in the sample during a single exposure (Toomre and Pawley, 2006). SDCM allows images to be acquired rapidly (up to hundreds of frames per second), with high sensitivity, so SDCM has become widely used in cell biology [Stephens 2003, Adams 2015, Thorn 2016].

A major problem with all confocal microscopes, i.e., LSCM and SDCM is compounded by thicker samples, which usually exhibit such a high degree of fluorescence emission that results in loss of fine detail and color saturation. To tackle this problem light sheet fluorescence microscopy has emerged as a powerful imaging tool to study thick biological samples.

1.2.3 Light sheet fluorescence microscopy (LSFM)

In the last decade, LSFM (Figure 1.6) has become a well-established imaging tool for developmental biology, and more generally for the investigation of thick 3D biostructures (thick biological samples).

The fundamental principle of LSFM is to exploit two decoupled optical paths for illumination and detection; the first path is an orthogonally oriented illumination that is responsible for restricting the illumination within a thin planar region in the focal region, the second one is a widefield based detection path which uses for fast imaging.

LSFM system is able to provide intrinsic optical sectioning and three-dimensional (3D) imaging capabilities with significant background rejection. Additionally, it is particularly convenient for imaging of thick biological samples such as embryos and whole brain, to avoid the illumination of the entire sample volume, because the imaging performances will be useful in reducing the photobleaching as well as the phototoxicity. Although LSFM can provide high-quality data of living organisms, it is not yet widely available [Adams 2015, Thorn 2016].

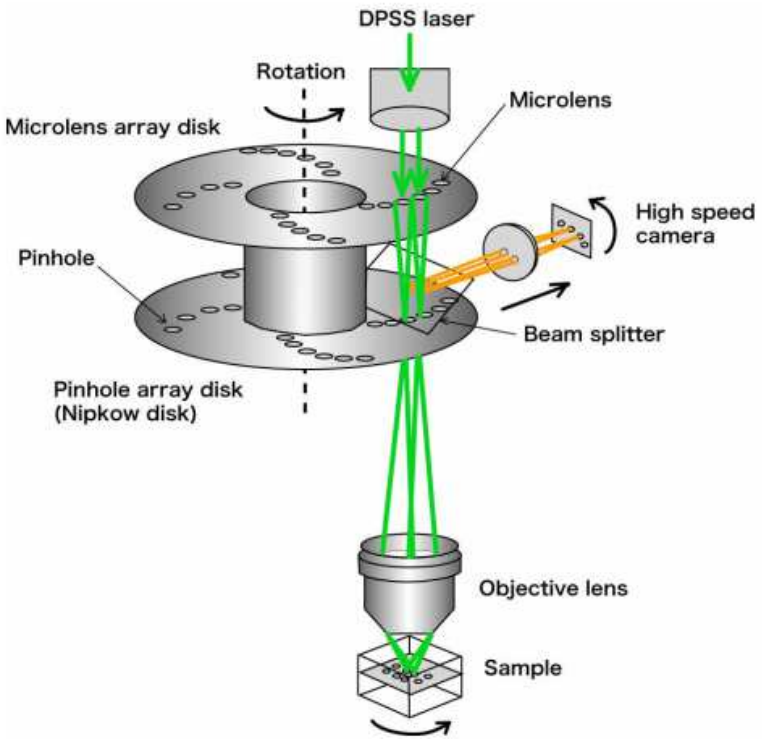

Figure 1.5: Principle of the spinning disk confocal microscopy. Image adapted from [Lima 2006]

1.3 Noise in fluorescence microscopy images

In this section, we briefly explain the main sources of noise existing in the different microscopy imaging techniques discussed in Section 1.2 which include photon noise, dark noise, read noise, and sample noise.

1.3.1 Photon noise

With the advancement of electronics used in detectors, photon noise or shot noise, is the most significant source of uncertainty in most datasets, in particular, fluorescence microscopy images [Luisier 2010a, Meiniel 2018]. Photon noise is generated from the statistical fluctuations in the number of photons detected at a given exposure level. This inherent stochastic variation in the emission rate of

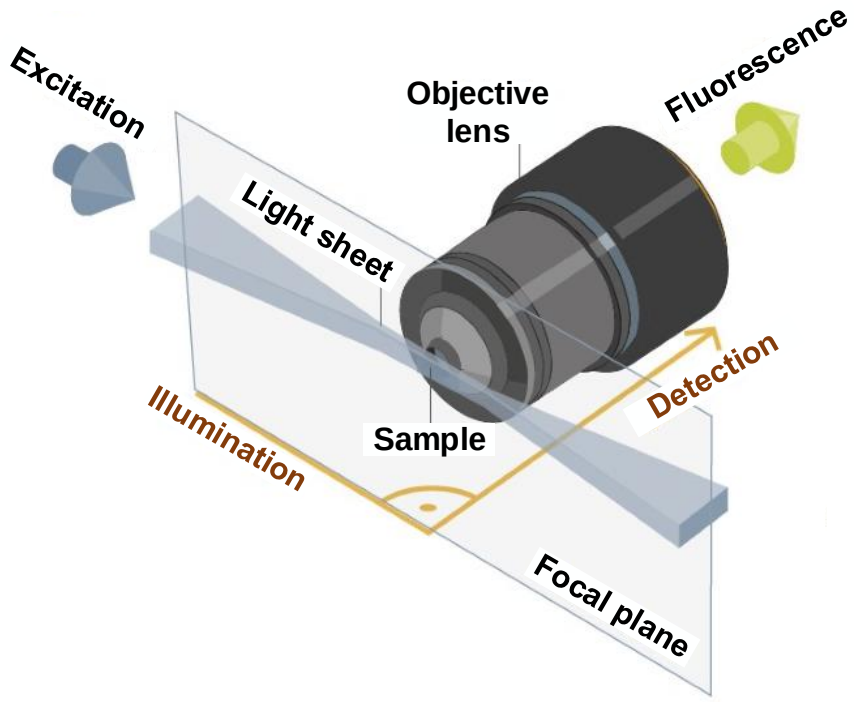

Figure 1.6: Principle of the light sheet fluorescence microscopy. Retrieved from ZEISS website: https://blogs.zeiss.com/microscopy/news/en/light-sheet-microscopy-with-zeiss-lightsheet-z-1

photons is well-described by a Poisson distribution $P(N = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$, where N is the number of photons measured over a time interval t, λ is the expected number of photons per unit time interval.

In the context of confocal microscopy, the measurement process for every scan position, i.e., pixel/voxel in the image is performed in the same way. Among the photosensitive devices in use today, the photomultiplier tube (PMT) is commonly used to measure the visible light photons. A typical PMT consists of a photoemissive cathode (photocathode) followed by focusing electrodes, an electron multiplier (Dynodes) and an electron collector (anode) in a vacuum tube, as shown in Figure 1.7.

To explain, the measurement process can be described in the following steps. Firstly, when the photons arrive at the PMT and enter the photocathode, the photocathode emits photoelectrons into the vacuum. Afterward, these photoelectrons are directed by the focusing electrode voltages towards the electron multiplier where electrons are multiplied by the process of secondary emission. Finally, the multiplied electrons are collected by the anode as an output signal. The recorded signal is essentially proportional to the number of photoelectrons. Consequently, the amplification factor may require considerable adjustment to account for the noise.

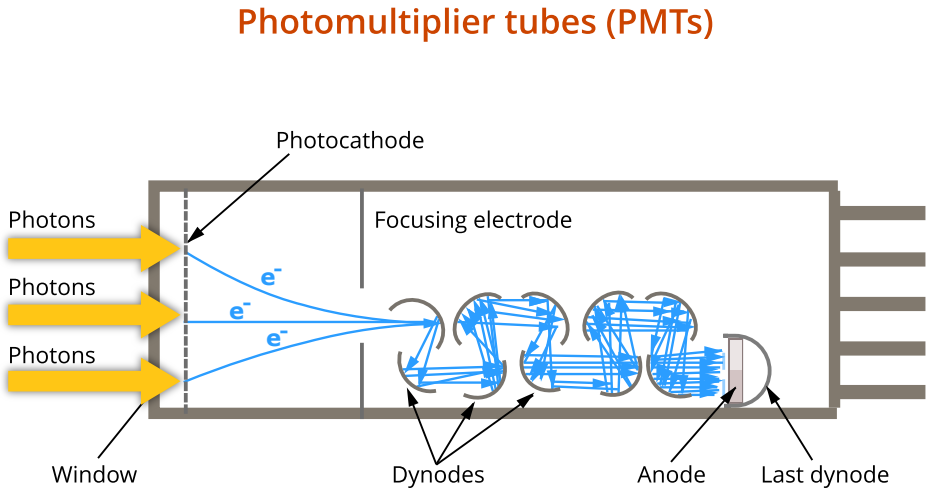

Figure 1.7: Schematic diagram of a photomultiplier tube. Retrieved from Myscope, Microscopy Australia website: https://myscope.training/index.html#/LMlevel_3_3

1.3.2 Dark noise

Besides photon counting noise, a signal independent noise that results from statistical fluctuation in the number of electrons thermally is liberated from the silicon atoms of the camera sensor's substrate [Luisier 2010a, Meiniel 2018]. This noise depends strongly on the device temperature where the generation rate of

thermal electrons at a certain temperature is known as dark current.

Similar to photon noise, dark noise follows a Poisson distribution to dark current and equals the square root of the number of thermal electrons liberated within the image exposure time. Since this noise can be reduced by cooling the sensors, the high performance cameras are usually cooled to a temperature at which dark current is negligible over a typical exposure interval.

1.3.3 Read noise

Read noise is a combination of several system noise components generated by the process of converting the charge carriers into a voltage signal and an analog signal to digital [Meiniel 2018]. In contrast to photon and dark noise, it is usually described as an additive white Gaussian noise (AWGN) with a normal distribution.

1.3.4 Sample noise

The sample itself also includes two significant factors of image quality degradation [Luisier 2010a, Meiniel 2018]. The first arises from comes from the reduction in the emission efficiency of these proteins along the time that decreases the effect on the image intensity. This phenomenon called photobleaching (Figure 1.9), that can be reduced by limiting the exposure time of the fluorophores, lowering power to the light source and increasing the fluorophores concentration. Nevertheless, these techniques also reduce the number of detected photons and thus decrease the SNR.

The second comes from the intrinsic fluorescence properties of the sample. To explain, several unlabeled molecules can emit fluorescence radiation as well. This may result in the interfere between labeled and unlabeled molecules where this phenomenon called autofluorescence (Figure 1.8). The autofluorescence affects the extraction of the interest signal if the emission wavelengths of the unlabeled molecules overlap with the labeled ones.

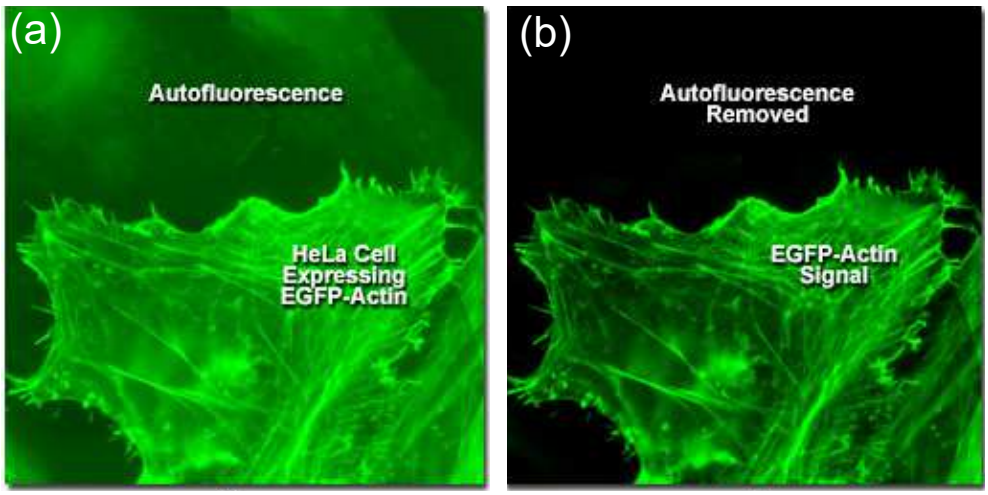

Figure 1.8: Example of autofluorescence phenomena. Retrieved from ZEISS website: http://zeiss-campus.magnet.fsu.edu/articles/spectralimaging/introduction.html

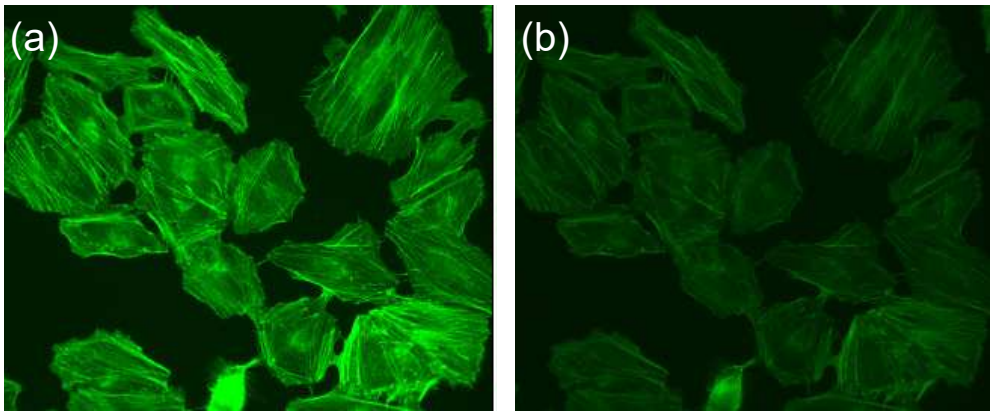

Figure 1.9: Example of a photobleaching in HeLa cells. (a) The initial intensity of the fluorophore. (b) The photobleaching that occurs after 36 seconds of constant illumination. Retrieved from Thermo Fisher Scientific website: https://www.thermofisher.com/sg/en/home/life-science/cell-analysis/cell-analysis-learning-center/molecular-probes-school-of-fluorescence/imaging-basics/protocols-troubleshooting/troubleshooting/photobleaching.html

1.4 Image analysis

Automated analysis of time-lapse microscopy images is an active topic in the field of image processing. Due to the development of high-resolution microscopes, the generated datasets consist of several (hundreds or thousands) images, each containing hundreds to thousands of objects to be analyzed [Meijering 2008]. Thus, these large volumes of data cannot easily be parsed and processed, via visual inspection or manual processing within any reasonable time.

Nowadays, there is a growing consensus that automated analysis methods are necessary to manage the time issue and provide a level of reliability and validity. Accordingly, the implementation of automated high-throughput techniques may be able to improve the clinical diagnosis, predict the treatment outcome, and help to enhance therapy planning.

In this section, we introduce the general workflow for microscopy image analysis (as shown in Fig 1.10.) ranging from sample imaging to quantitative analysis.

To start with, the sample is prepared, placed under the microscopy, and the image acquired at different time points to record the dynamic cell behavior. Afterward, a pre-processing step is applied to the captured image to enhance the signal to noise ratio (SNR) by reducing variability without losing essential information. This step varies depending on the experimental setup and microscopy modalities. Then, the processing step which includes cell detection (i.e., locate specific features of interest), segmentation (i.e., determine precise boundaries of cells) and tracking (identifying the location of cells in consecutive frames) is employed to locate cells in individual frames and then associates them in consecutive frames. Finally, cell detection, segmentation masks, and cells trajectories are used to compute some biologically relevant information (like cell velocity, displacement, phenotype, etc.)

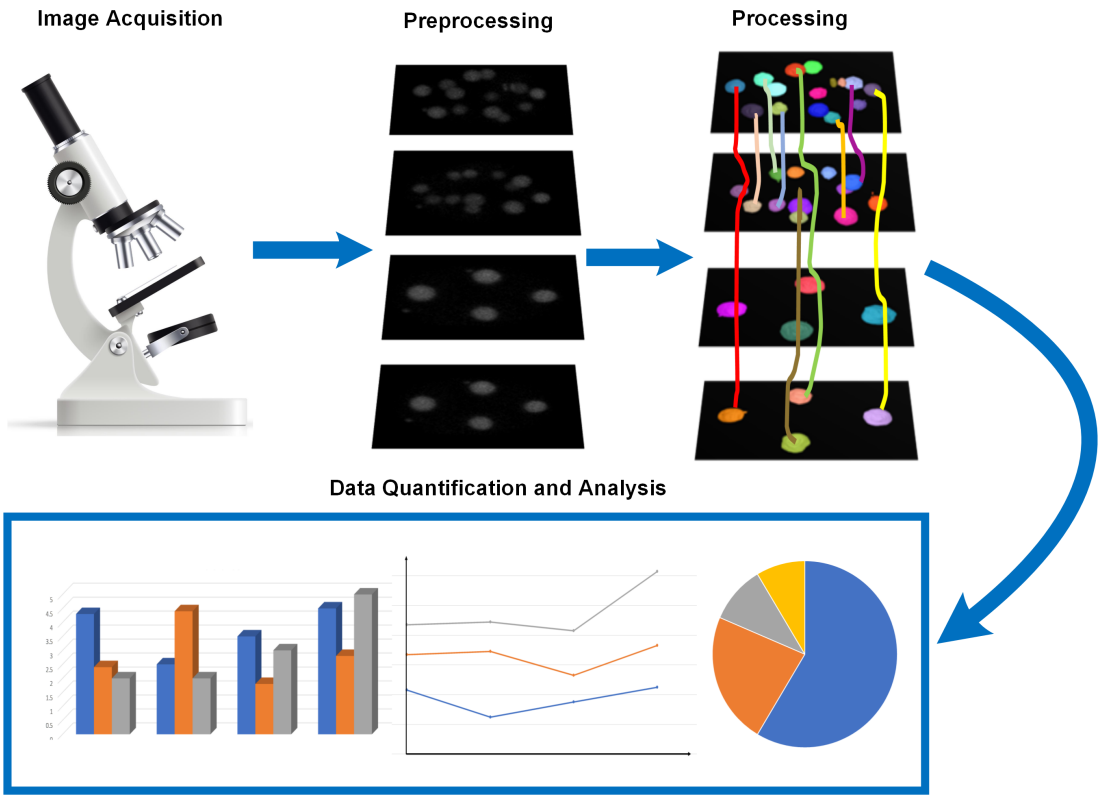

Figure 1.10: Typical framework for microscopy image analysis. After sample preparation, automatically acquired microscopy images are processed for subsequent data manipulation and analysis. First: image preprocessing includes denoising, smoothing, and contrast enhancement. Second: image processing, i.e., cell/nuclei detection, segmentation, and tracking. Finally: data quantification and analysis (like cell velocity, displacement, and phenotype) .

1.5 Cell tracking challenge

In 2012, Cell Tracking Challenge (CTC) [1] [Ulman 2017, Maška 2014] was launched under the auspices of the IEEE International Symposium on Biomedical Imaging (ISBI). The goal of this challenge is to provide a benchmark for a thorough evaluation and fair comparison of available cells and nuclei segmentation and tracking methods under the same criteria. The CTC includes both real 2D and 3D time-lapse microscopy videos of cells and nuclei, tandem with computer generated 2D and 3D video sequences simulating nuclei moving in realistic environments. Since most researchers focus their interest on the segmentation process, a new time-lapse cell segmentation benchmark on the same datasets was launched in 2018. In this research, we mainly focus on testing two very noisy 3D datasets for *C. elegans* and *Drosophila melanogaster* embryonic cells, namely *Fluo-N3DH-CE* (Section 4.1.2) and *Fluo-N3DL-DRO* (Section 4.1.2).

1.6 Problem statement

Automatic segmentation and tracking of biological structures such as cells or nuclei in 3D time-lapse microscopy images are important tasks in scientific research ranging from cell development and growth to disease discovery. However, segmentation and tracking of biological cells encounter several problems. These problems can be summarized as depicted below:

- Noise embedded in the images such as photon noise, dark noise, and read noise;
- Non-uniform background illumination because of the fluorescence in cytoplasm and mounting medium;
- Low contrast and weak boundaries of non-obvious nuclei;
- The degradation of image intensity over time due to photo-bleaching of fluorophores [Meijering 2008].

[1] http://celltrackingchallenge.net/

Following these problems, the acquired images become very noisy (as shown in Figure 1.11) and difficult to interpret, which might lead to false detection, segmentation as well as false tracking results.

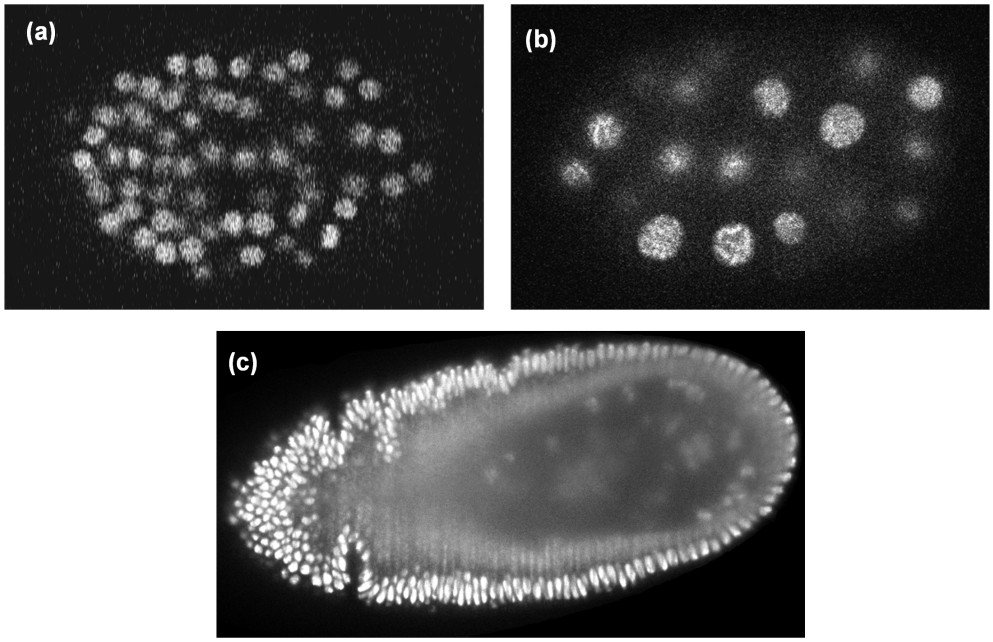

Figure 1.11: Representative samples of a selected slice from 3D datasets used in this thesis. (a) ***Fluo-N3DH-CE*** [Ulman 2017]. (b) ***CE-UPMC*** [Gul-Mohammed 2014a]. (c) ***Fluo-N3DL-DRO*** [Ulman 2017].

1.7 Research objectives

The ultimate objective of this research is to develop a robust approach for denoising 3D microscopy images toward a better segmentation of noisy and densely packed nuclei. So, a full statement of the aims and objectives would be as follows.

- First, designing and developing a generic approach for denoising 3D microscopy images with as few parameters as possible.
- Second, validating the proposed approach and comparing it with the top-ranked approach from the IEEE International Symposium on Biomedical

Imaging (ISBI) Cell Tracking Challenge (discussed in Section 1.5).

- Third, applying the proposed approach to various 3D biological datasets, for example, the human breast adenocarcinoma cells.

1.8 Structure of the thesis

Following this introductory chapter, the thesis is organized as detailed below:

1. **Chapter 2:** gives a comprehensive literature review of available denoising, cell segmentation and tracking techniques in microscopy images. In addition, it includes the most widely used software tools for cell detection, segmentation, and tracking. These tools involve commercial and open-source platforms.

2. **Chapter 3:** discusses in detail the proposed method for simultaneous denoising and detecting of cell nuclei in 3D time-lapse fluorescence microscopy images which is based on an unsupervised dictionary learning and sparse representation approach. Moreover, it discusses the integration of the classical segmentation methods such as adaptive thresholding and 3D marker-controlled watershed together with our denoising method.

3. **Chapter 4:** provides the results and discussions of the proposed methodology that extensively tested on three real datasets for embryonic cells and one dataset of synthetic images that have different values for the signal to noise ratio (SNR) and the object size. Besides, it presents an experimental setup and suitable parameters selection, where the approximate size of the objects of interest is the only one being critical. Additionally, a comparison between our proposed method and various denoising and segmentation algorithms is provided. Finally, additional experiments on real datasets coming from various tissues such as human breast adenocarcinoma cells have been conducted to emphasize the genericity of the proposed method.

4. **Chapter 5:** presents the conclusions, strengths, and limitation of the proposed methodology. Furthermore, it recommends some future directions to enhance the proposed methodology in term of computational time.

Chapter 2

Literature Review

Contents

2.1 Denoising of microscopy images

Despite recent advances in microscopy industry, image denoising is still an essential step in various image processing and computer vision tasks, such as object

segmentation and tracking. Over the last few years, several methods have been described in existing literature for filtering and denoising of microscopy images and a recent comprehensive review can be found in [Meiniel 2018, Roels 2018]. Such methods can be divided into four groups: patch-based, wavelet-based, median-based, and CNN-based methods.

In the first group, the original non-local means (NLM) algorithm introduced by Buades *et al.* [Buades 2005] was considered as a very popular and powerful family of denoising methods. This method relied on the aggregation of patches from the entire image rather than just filtering in a local manner. For each noisy patch, other patches existed somewhere in the image and having the same structure (geometry or texture) can be added together to remove such noise. The main limitation existed in the NLM algorithm is that remarkable denoising results had been obtained at a high expense of computational cost due to the enormous amount of weight computations.

Inspired by the original NLM, Darbon *et al.* [Darbon 2008] developed a fast non-local means approach that was capable of reducing the computational cost of calculating an approximate measure regarding the similarity of neighborhood windows. This algorithm demonstrated a good performance in enhancing particles' contrast and reducing the noise in electron cryomicroscopy images.

Indeed, the widely used NLM and fast NLM filters were not the optimal methods employed for noisy biological images containing small features of interest. The reason was that the noise existed in the image prevented a precise determination of the weights used for the averaging process. As a result, this led to over-smoothing and other image artifacts.

To circumvent the aforementioned problem, Yang *et al.* [Yang 2010] presented an adaptive non-local means filter for improving feature recovery and particle detection in live cell fluorescence microscopy images. This method started by constructing a particle feature probability image, which relied on Haar-like feature extraction. Then, the particle probability image was used to enhance the estimation of the correct weights for the averaging process.

An extension of the NLM was described by Deledalle *et al.* [Deledalle 2010] for denoising images corrupted by Poisson noise. The authors used the probabilistic

similarities proposed by Deledalle *et al.* [Deledalle 2009] to compare patches infected by noise and patches of a pre-estimated image. Additionally, a risk estimator for NLM that was adopted from Van De Ville *et al.* [Van De Ville 2009] was utilized in the optimization problem to automatically find the more appropriate filtering parameters in a few iterations.

In the same way, Danielyan *et al.* [Danielyan 2014] developed a method for denoising two-photon fluorescence images using Block-matching 3D (BM3D) filtering. This method was implemented in two steps. First, the image was divided into overlapping blocks and then similar blocks were collected mutually into groups. Second, collaborative filtering was employed to reduce noise effectively from all similar blocks in each group. This filtering process was implemented in a 3D transform domain, in which 3D transforms modeled both the content of the blocks and their mutual similarity or difference as well. Such joint models were able to enhance sparse representation of the signal which can then be efficiently separated from noise by thresholding the transform coefficients.

On the contrary to the aforementioned patch-based approaches, Boulanger *et al.* [Boulanger 2010] presented a spatiotemporal patch-based adaptive statistical method for denoising of 3D+t fluorescence microscopy image sequence. The merits of this approach are twofold. Firstly, a variance stabilization step was applied to the data to find the independence between the mean and the variance. Second, the spatiotemporal neighborhoods were considered to restore the series of 3D images as already proposed for 2D image sequences in [Boulanger 2007].

Haider et al. [Haider 2016] introduced a noise reduction problem that was formulated as a maximum a posteriori (MAP) estimation problem. In this method, a novel random field model called stochastically connected random field (SRF), that coupled random graph and field theory was used to solve the estimation problem. In the SRF model, each node represented a pixel in the image, and the random variables u_i and u_j were the noise-free image intensities of the ith and jth node. The stochastic edges between nodes aimed to establish weights within a local neighborhood surrounding a pixel region in the random field under an assumption of local spatial-feature smoothness for noise reduction. One of the main advantages of the SRF approach was that it was able to achieve strong

performance in terms of high signal to noise ratio and contrast to noise ratio in the real fluorescence microscopy data results. Other advantages included its flexibility to maintain cell structure and fine details while reducing background and intracellular noise.

In the second group, although signal modeling in the wavelet-domain methods provided effective image denoising while preserving edges, it could not capture the smoothness along contours in a compact manner. As a result, new algorithms such as curvelet and contourlet transforms have been developed to improve such shortcomings.

An example of contourlet-based denoising method was proposed by Yang and Lee [Yang 2015]. This method utilized the hidden Markov model (HMM) in the Contourlet transformed domain for mixed Poisson-Gaussian noise in fluorescence microscopy images. The Contourlet transforms consisted of two distinct filters: (1) the Laplacian pyramid (LP) filter and (2) the directional filter bank (DFB). The LP filter was used in the existing wavelet filter as a low pass filter to separate high and low frequencies. The DFB provided information on image direction components. In this study, the HMM algorithm reformulated an independent mixture model to match the non-Gaussian nature of the Contourlet coefficients. Furthermore, this HMM was adapted to characterize the key dependencies between the contourlet coefficients.

Luisier *et al.* [Luisier 2010b] suggested a *Poisson Unbiased Risk Estimation-Linear Expansion of Thresholds (PURE-LET)* technique for denoising images corrupted with Poisson noise. The method was based on three criteria: 1) minimizing of an unbiased estimate of mean square error (MSE) for Poisson noise, 2) linear parameterization of the denoising process 3) preserving of Poisson statistics across scales. This algorithm was particularly promising for large datasets as well as images having a low signal to noise ratio. In addition, it had limited system requirements.

To complement, some biological studies employed conventional denoising methods like bandpass filter, and Wiener filter.

To explain, Sim *et al.* [Sim 2016] and Aguirre [Aguirre 2012] introduced an adaptive Wiener filter that was adapted from the classical Wiener filter for

reducing noise in electron microscopy images. The main idea of the classical filter was to implement the linear estimation of a target signal sequence from another related sequence (corrupted signal). Such process was performed by minimizing the mean square error (MSE) between the estimated random process and the desired process. As opposed to the classical filter, the adaptive Wiener filters had been considered the noise variance and so it led to enhancing filter effectiveness.

Magnusson [Magnusson 2016] introduced a Gaussian band-pass filter to reduce the noise in fluorescence and phase contrast microscopy images. This filter was based on the difference of Gaussians algorithm that discarded the noise by removing high spatial frequencies and non-uniform background illumination by removing low spatial frequencies.

In the third group, Kushwaha *et al.* [Kushwaha 2012] implemented a median filter for removing noise presented in Transmission Electron Microscopy (TEM) images. The median filter was based on the idea of considering the noisy intensities as outliers. Thus, the median of neighborhood window of a reference pixel provided an estimate of the denoised reference pixel. This filter was usually used to remove speckle (salt and pepper) noise without blurring the picture. Nevertheless, in most practical situations, the median of the neighborhood window might not guarantee the true intensity and this led to performing less well than expected.

A similar studies also were proposed by Akram *et al.* [Akram 2014], Gul-Mohammed *et al.* [Gul-Mohammed 2014a], Gul-Mohammed and Boudier [Gul-Mohammed 2014b] and Cascio *et al.* [Cascio 2019] to employ median filter as a noise removal tool.

In the fourth group, Weigert *et al.* [Weigert 2017a] proposed a method to enhance the axial resolution in 3D microscopy images by reconstructing isotropic 3D data from non-isotropic acquisitions using a convolutional neural network. This method relied on restoring isotropy by deconvolving the image under the trained and sample specific image prior. Training with two different learning strategies, i.e., IsoNet-1 and IsoNet-2, was performing unsupervised and end-to-end, on the same anisotropic image data.

Another approach had also been proposed by Weigert *et al.* [Weigert 2017b] that presented a content-aware image restoration (CARE) networks method to

denoise fluorescence microscopy data. This method was based on the U-Net architecture together with a per-pixel similarity loss (such as the absolute error) which introduced a solution to the problem of missing training data for deep learning in fluorescence microscopy by generating training data without the need for laborious manual annotations. With the CARE networks, high-quality restored mages were obtained even if the SNR of the images was very low. Table 2.1 summarizes the denoising approaches used for biological microscopy images.

Table 2.1: Summary of microscopy images denoising approaches.

References	Methodology	Image dimensions	Modality
[Darbon 2008]	Fast non-local means	2D	Cryomicroscopy
[Yang 2010]	Adaptive non-local means	2D	Fluorescence microscopy
[Deledalle 2010]	Poisson non-local means	2D	Fluorescence microscopy
[Boulanger 2010]	Spatiotemporal patch-based adaptive statistical method	3D	Fluorescence microscopy
[Luisier 2010b]	Poisson Unbiased Risk Estimation-Linear Expansion of Thresholds (PURE-LET)	2D/3D	Fluorescence microscopy
[Aguirre 2012]	Adaptive wiener filter	2D	Electron microscopy
[Gul-Mohammed 2014a, Gul-Mohammed 2014b]	3D median filter	3D	Fluorescence microscopy
[Akram 2014]	2D/3D median filter	2D/3D	Fluorescence microscopy
[Danielyan 2014]	Block-matching 3D (BM3D) filtering	2D	Two-photon fluorescence microscopy
[Yang 2015]	Contourlet hidden Markov model (HMM)	2D	Fluorescence microscopy
[Haider 2016]	Stochastically connected random field model	2D	Fluorescence microscopy
[Magnusson 2016]	Gaussian band-pass filter	2D/3D	Fluorescence and phase contrast microscopy
[Sim 2016]	Adaptive wiener filter	2D	Scanning electron microscopy
[Weigert 2017a]	A convolutional neural network	3D	Fluorescence microscopy
[Weigert 2017b]	Deep convolutional neural network	3D	Fluorescence microscopy
[Cascio 2019]	Median filter	2D	Fluorescence microscopy

2.2 Cells/Nuclei segmentation

Cell segmentation arises as the process of labeling every pixel in an image and therefore pixels with the same label share certain characteristics. The ultimate goal of this process is to extract interesting regions, i.e., cells or nuclei from an image. In image analysis, segmentation is considered as a fundamental step since its impossible to get good tracking results with poor segmentation even with using powerful tracking techniques. Many scientists contributed with a lot of efforts into developing various algorithms for automated segmentation of cells. However, there are still no systems that can handle the extensive variability during time-lapse experiments of cells. Additionally, several cell segmentation methods use a preprocessing step to reduce the noise and enhance the image quality as mentioned in (Section 2.1) toward better segmentation results.

The cell nuclei segmentation approaches are broadly classified into three categories: simple approaches such as thresholding method [Bise 2015, Liao 2016, Gul-Mohammed 2014a], edge detection [Wählby 2004] and shape matching [Cicconet 2013, Türetken 2017, Türetken 2015], and more sophisticated approaches like region growing [Cliffe 2017, Liu 2014, Tonti 2015, Gul-Mohammed 2014b] and energy minimization [Tarnawski 2013] and machine learning based approaches [Ronneberger 2015, Akram 2016b, Liu 2017, Sadanandan 2017].

In the first category, Liao *et al.* [Liao 2016] applied a classical thresholding approach to segment cell nuclei. This approach assumed that the cell was usually brighter than its surrounding areas and there often existed an optimal threshold where individual cells can be segmented as separate objects. This assumption was not applicable for challenging regions, because it was impossible to find a suitable threshold to separate all touching cells.

Bise and Sato [Bise 2015] adapted a multiple-level thresholding technique to identify and segment all cell nuclei. In this method, a set of level thresholds that were equally spaced had been used to segment cell nuclei at a particular level of intensity with the fact that the intensities inside the cell were slightly higher than those at its boundaries among touching cells.

A similar approach was proposed in [Gul-Mohammed 2014a], where authors

used an adaptive iterative thresholding algorithm to find the best threshold for each cell and eventually to obtain all segmented cell nuclei. This algorithm was able to correctly segment nuclei even when they were in touch with each other. In addition, this algorithm became effective in the presence of temporal and spatial intensity variations.

An edge detection approach was implemented by Wählby *et al.* [Wählby 2004] for segmentation of 2D and 3D cell nuclei in images. In such approach, an edge filter was applied to the image and therefore pixels were classified as edge or non-edge. These edges were usually detected by the first or the second order derivative method. Nevertheless, this method failed to detect the non-obvious cell's boundary. Thus, the authors combined the intensity information, the gradient information (edge strength) as well as the shape of the nuclei to improve segmentation result.

The shape matching-based approach was dependent on the assumption that cells, in particular nuclei have round shapes. To give an example, a shape matching technique using wavelet-based circular hough transform (WCHT) was presented by Cicconet *et al.* [Cicconet 2013]. The original idea of the Hough transform was to produce circle candidates by "voting" each point on the image in the Hough parameter space and then select the local maxima. Due to the mathematical properties of the transform, this "voting" allowed to locate prominent circles in the image. However, the original implementation, at which votes were considered pixel-wise, was very sensitive to noise. Thus, the votes in WCHT were weighted by wavelet kernels and a fine-tuning stage based on dynamic programming.

Another approach based on ellipses fitting, described in [Türetken 2015, Türetken 2017]. This method was designed to cluster cell nuclei in a hierarchical manner and then fit ellipses to each resulting cluster using a non iterative least squares approach.

In the second category, region-based segmentation approaches can be used rather than defining the border between an object and a background via a threshold in the image intensity. The basic idea of these approaches is to combine the neighbouring pixels of initial seed points which have similar properties to form individual cells. The most used region-based approaches for nuclei segmentation were watershed and region growing methods. These methods also worked properly

in the context of crowded nuclear population. However, the main drawback of these methods was over-segmentation that can be reduced by fusing touching regions a posteriori using a cost function.

For instance, Gul-Mohammed and Boudier [Gul-Mohammed 2014b] employed a classified region growing for 3D segmentation of packed cell nuclei. This method was able to overcome the over-segmentation problem and it can be explained as follows. First, the seed points were obtained by applying a local maxima filter to the raw or filtered image. Then, intensity threshold with decreasing levels were applied to the image, and hence the regions around local maxima grow and may eventually touch. At each threshold level, every set of touching regions was updated by the combinations of individual regions that leads to a predefined object. If the combination of individual regions not led to one of the predefined objects, individual regions were not merged, and the procedure continues.

Another potential solution to account for over-segmentation problem was introduced by Magnusson [Magnusson 2016]. The author proposed to use the marker-controlled watershed along with h-maxima transforms to find marker points. By using the h-maxima transform, all undesired regional maxima whose values less than h-values were suppressed.

Tonti *et al.* [Tonti 2015] employed an adaptive marker-controlled watershed technique for the segmentation of HEp-2 cells. The authors applied a marker-controlled watershed on the gradient of the normalized image. Typically, the watershed algorithm considers the gradient image as a topography surface and it starts a flooding from its regional minima. In this method, the gradient image was modified to deal with the over-segmentation problem, so that the regional minima occur at the locations only specified by the internal and external markers.

Tarnawski *et al.* [Tarnawski 2013] utilized the active contour approach for both segmentation and tracking of clustered cells in time-lapse fluorescent microscopy. As a first step, the least squares method was used to identify the ellipse that best characterizes the generic cell. Then, active contours model, starting from an elliptic curve, was employed to evolve the contour towards the cellular edge. In the same way, Cascio *et al.* [Cascio 2019] adapted the active contour approach to segment HEp-2 cells. However, the authors used a randomized Hough transform

for ellipse rather than least squares fitting.

Cliffe *et al.* [Cliffe 2017], introduced two methods for segmentation nuclei and cytoplasm in fluorescence microscopy images. The author was used a Gaussian mixture model (GMM) for segmentation of nuclei segmentation and 3D marker watershed to segment cytoplasm.

Some studies had combined multiple segmentation methods for improving the segmentation accuracy as introduced by Chen *et al.* [Chen 2006]. In such method, the Otsu's thresholding method was use to segment nuclei from the background and then a watershed technique was deployed to further separate touching nuclei.

In the same vein, Akram *et al.* [Akram 2014] proposed a multi-stage approach for segmentation of cells in spinning disk confocal images. The first stage was to used edge detector to compute edge probability at each pixel. Then, cells were separated from background using graph cuts, in which the intensity in local neighborhood of a pixel was used to set terminal edge weights. The third stage was to find the seeds for individual cells by sing another graph cuts stage, in which edge probability map was used to adjust terminal edge weights. Finally, marker-controlled watershed was employed to find boundaries of individual cells.

In the third category, much of the current literature on cell segmentation pays particular attention to machine learning and deep learning methods. To give some examples, Ronneberger *et al.* [Ronneberger 2015] proposed U-Net convolutional networks, a type of fully convolutional network (FCN) [Long 2015], to segment cells by labeling every pixel in the image. Unlike other deep neural networks, the U-Net can be trained end-to-end from very few images by applying elastic deformations to the available training images. This allowed the network to learn invariant to such deformations, without the need to see these transformations in the annotated image corpus. Furthermore, it was able to deal with the touching objects of the same class by using a weighted loss, where the separating background labels between touching cells obtain a large weight in the loss function. U-Net is still a very successful and widely used model for semantic segmentation of various medical images.

Liu *et al.* [Liu 2017] designed a deep convolutional encoder-decoder network architecture known as NucleiNet for segmenting cell nuclei in noisy fluorescence

microscopy images. The idea of NucleiNet was to learn the stochastic characteristics of noise as well as the shape of nuclei for the learning step and then regenerate the clean nuclei image based on the learned prior knowledge. Finally, the segmented cell nuclei can be obtained by applying the thresholding- based method on the clean nuclei image.

A two-stage convolutional neural network (CNN) method presented by Akram *et al.* [Akram 2016b] to precisely segment the cells. In the first stage,a FCN was used to propose cell bounding boxes along with their scores, i.e., the probability of being a cell. In the second stage, another CNN was employed to segment cells within the regressed bounding box.

Sadanandan *et al* [Sadanandan 2017] advised to use the deep convolutional neural networks (DCNNs) to segment cells in fluorescence microscopy images. Inspired by the U-Net [Ronneberger 2015], both the data and their labels were subjected to data augmentation to create the final dataset for training the DCNN.

A fully convolutional network (FCN) was employed by Hernandez *et al.* [Hernandez 2018] to segment cells in phase and fluorescence microscopy images. In contrast to classical convolutional networks, the FCN learns a mapping from pixels to pixels, without extracting the region proposals. Furthermore, it only had convolutional and pooling layers. As a result, the output was a spatial map, which can be used to perform pixel-wise segmentation on arbitrary-sized inputs. This method had a relatively low processing time, as manually labeling around $100 - 200$ cells was enough to train a network which can then segment a large number of new images with high accuracy

A more effort has been put into the development of deep learning approaches algorithms to use as few as possible number of labels data. A recent study by Arbelle and Raviv [Arbelle 2018a] involved a novel method for cell segmentation in microscopy images that was adapted from the original generative adversarial network (GAN) [Goodfellow 2014]. The GAN framework is based on two models, a generative model to capture the data distribution, and a discriminative model to estimate the probability that a sample came from the training data rather than generative model. One of the main advantages of this method was that it did not require a formulation of a loss function for the optimization process. Other

advantages included accurate segmentation in the presence of a low number of training dataset, that resulted in dramatically reducing the annotation effort.

Although recent approaches had consistently shown promising segmentation results using convolutional encoder-decoders such as the U-Net [Ronneberger 2015] and a two-stage convolutional neural network (CNN) [Ronneberger 2015], these methods are still unable to incorporate temporal information that can facilitate segmentation of individual touching cells or of cells that are partially visible.

Another study was proposed by Arbelle and Raviv [Arbelle 2018b] to exploit cell dynamics by coupling of convolutional long short term memory (C-LSTM) and the well established network architecture (U-Net) to allow compact spatio-temporal representations in multiple scales. In addition, the authors planned to incorporating adversarial loss as well to weaken the dependency on training set size. Tables 2.2 and 2.3 summarize the cell/nuclei segmentation approaches used for biological microscopy images.

Table 2.2: Summary of cell /nuclei segmentation approaches using deep learning.

References	Methodology	Image dimensions
[Ronneberger 2015]	U-Net Convolutional Networks	2D
[Akram 2016b]	Two-stage convolutional neural network (CNN)	2D
[Liu 2017]	A deep convolutional encoder-decoder network	3D
[Arbelle 2018a]	Generative Adversarial Neural Network (GAN)	2D
[Arbelle 2018b]	Coupling of Convolutional long short term memory (C-LSTM) and the well established network architecture (U-Net)	2D
[Hernandez 2018]	A Fully Convolutional Network (FCN)	2D

Table 2.3: Summary of cell/nuclei segmentation approaches. NA: not available.

References	Preprocessing	Methodology	Image dimensions
[Chen 2006]	NA	A classical thresholding approach followed by a watershed .	2D
[Wählby 2004]	A classical Gaussian filter	An edge detection based approach	2D/3D
[Cicconet 2013]	NA	Wavelet-based Circular Hough Transform	2D
[Tarnawski 2013]	NA	Active contour with least squares fitting	2D
[Gul-Mohammed 2014b]	A classical 3D median filter	A classified region growing	3D
[Akram 2014]	A classical median filter	A multi-stage approach including edge detector, graph cut and watershed	2D/3D
[Gul-Mohammed 2014a]	A classical 3D median filter	An adaptive iterative thresholding algorithm	3D
[Bise 2015]	NA	A multiple-level thresholding technique	2D/3D
[Tonti 2015]	A classical Gaussian filter	Adaptive marker-controlled watershed	2D
[Magnusson 2016]	A band-pass filter	Adaptive marker-controlled watershed	2D/3D
[Türetken 2017]	NA	Ellipses fitting	2D
[Cliffe 2017]	A classical Gaussian filter	3D marker-controlled watershed and Gaussian mixture model (GMM)	3D
[Cascio 2019]	A classical median filter	Active contour with a randomized Hough transform	2D

2.3 Cells/Nuclei tracking

Cell tracking has received great attention over the past few decades therefore many approaches have been proposed to perform this task. Reviews of existing cell tracking approaches can be found in [Rohr 2010, Kanade 2011, Meijering 2012]. Besides, the general description of the approaches implemented in cell tracking challenges have been reviewed in [Maška 2014, Chenouard 2014, Ulman 2017]. Broadly speaking, cell tracking approaches can be categorized into two groups: (1) model-based contour evolution approaches and (2) detection-based association approaches.

The model-based contour evolution approaches are getting started by segmenting the cells contours in the first frame of a video sequence and then evolving these contours dynamically to fit the cells in successive frames. Thus, both segmentation and tracking tasks can be solved simultaneously as a one process, with the hypothesis that obvious spatiotemporal overlapping exist between the corresponding cell regions. These approaches represent cells contours using either a parametric or an implicit model.

In the former model, active contours [Dufour 2005], active meshes [Dufour 2011], or Gaussian mixture [Amat 2014] are employed to represent the cells, which require more steps to handle key events such as cell division, cells touching and cells entering/leaving the field of view.

In the latter model, the contours of the cells are represented in an implicit manner through the level-set method. Such an approach can handle changes of topology and provide enough support to the estimation of the geometric properties, but it is computationally expensive. The main advantage of contour-based evolution approaches is that they can provide robust segmentation of cell boundary even in a noisy environment. Moreover, such techniques are very useful in tracking various cell topological changes that occur during the cell cycle and growth.

On the other hand, the main idea of detection-based association approaches is to segment (identify) the cells first in all frames of a video sequence independently and later associate (link) them among frames. The key benefit of detection-based

association approaches is the decoupling of detection and association process in a two-step procedure which enables easy tracking of new cells entering the field of view. Since the segmentation is usually more application dependent than the track association, the same track association algorithm is commonly used, along with multiple segmentation algorithms related to a specific application.

Few approaches in the existing literature dedicated for cells tracking are based on contour evolution models. This is due to a burning need for tracking the whole cell boundary instead of just locating the cell, which slows the processing speed. In addition, they are sensitive to parameter settings.

As for the first category, Dufour *et al.* [Dufour 2005] presented a classical active contours model [Chan 2001] along with a volume preservation constraint and several optimization problem to track cells in 3D+t microscopy data. In this model, each cell was described by its own level-set function to address the problem of overlapping cells. Furthermore, this model proves its ability to process data with low signal to noise ratios (SNRs) and handles the tracking of the cells that tend to touch, divide and enter/leave the field of view. However, a limited performance was inherited to this model especially when applied to real biological images that have very low SNRs and also contain multiple cells in close contact with each other.

Dzyubachyk *et al.* [Dzyubachyk 2007] proposed an advanced level-set based model that depends on the fusion of the variational model described in [Rousson 2002] with the classical Chan and Vese model [Chan 2001] for segmenting and tracking cell movement and deformation in time-lapse fluorescence microscopy images. The added value of this model was that its results became less sensitive to different combinations of parameter values.

In order to overcome the shortcomings of the early study introduced by Dufour *et al.* [Dufour 2005], Dzyubachyk *et al.* [Dzyubachyk 2008, Dzyubachyk 2010] presented an improved level-set based method to simultaneously segment and track multiple cells in fluorescence microscopy image sequences. The proposed method had implemented significant modifications that noticeably improve its performance in terms of accuracy, robustness, and computational cost. More specifically, the energy functional had been modified to be minimized, and this

resulted in reducing the number of energy weights.

In another study, Dzyubachyk *et al.* [Dzyubachyk 2009] introduced a novel multi-object tracking framework using graph cuts based active contours. The graph cuts method uses the information about the state of an object at a previous time point as initialization for the detection and segmentation of the same object in the next time point. This greatly handled the segmentation and tracking of clustered nuclei simultaneously. In contrast to the framework of active contours, graph cuts-based energy minimization had a fast performance because it did not need any a priori global shape model, which makes it useful for tracking objects with deformable shapes and appearances.

Li et al [Li 2008] employed a two-level image analysis system by integrating multiple collaborative modules. This modular design provided a fast-geometric active contour tracker together with a biologically relevant interacting multiple model (IMM) motion filter, and spatiotemporal trajectory optimization. The system allowed automatic quantification of cell migration, proliferation, apoptosis, and construction of cell lineage maps. As a result, it facilitated the analysis of massive biological datasets. However, this approach can be applied only to 2D images.

Padfield et al. [Padfield 2009] applied a model-based shape/size constraint to control the level set spatio-temporal surface evolution ,and then exploited the fast-marching method to establish links between detected cells from frame to frame based on the shortest distance.

In a different study, Dufour *et al.* [Dufour 2011] presented a novel deformable model framework by reformulating the conventional variational approaches [Dufour 2005, Dzyubachyk 2008, Dzyubachyk 2010] in the discrete active mesh framework to reduce the memory and time costs drastically. Nonetheless, this framework only works in 3D and cannot straightforwardly be used to analyze 2D sequences.

Maška *et al.* [Maška 2012] designed the whole cell tracking by a model evolution approach to make the algorithm as fast as possible. Unlike prior efforts [Dufour 2005, Dzyubachyk 2008, Dzyubachyk 2010, Dufour 2011], the cells were detected by minimizing the original Chan–Vese model [Chan 2001] in the fast level set-like (FLS) frameworks [Maška 2010] without solving any partial differential equation

(PDE). Afterward, the multiple separate models were replaced with a topological prior that uses the object indication function [Maška 2011] to enable simultaneous tracking of multiple cells over time. An extended version of this work was presented in [Maška 2013]. The extended framework integrated the original Chan–Vese model with graph cut (GC) [El Zehiry 2007] and the FLS framework. In addition, this framework was expanded into higher dimensions that can be applied directly to 2D as well as 3D timelapse series.

A sequential Bayesian approach in conjunction with Gaussian mixture models (GMMs) was introduced by Amat *et al.* [Amat 2014] to perform segmentation and tracking simultaneously using a parametric contour evolution. The idea of this approach was to propagate the information obtained from a previous time point to the next time point instead of fitting a GMM at each time point independently. Due to the temporal coherence between consecutive time points, the authors hypothesized, as a *priori* knowledge (Bayesian approach), that the location, shape, and appearance of cell nuclei cannot variate suddenly between time points.

As for the second category, most cell tracking approaches belong to such category i.e, detection-based association tracking and summarized in Table 2.5. As a starting point, Li *et al.* [Li 2010] improved the tracking accuracy by introducing a new dissimilarity measure which considered the morphological appearance, migration distance, neighboring relationship, and intensity information. An integer linear programming and a sophisticated matching strategy, in conjunction with the novel dissimilarity measure were employed to match and track cell nuclei in the time-lapse fluorescence microscopy images.

An integrated study was introduced by Bise *et al.* [Bise 2011b] for cell tracking based on global spatiotemporal data association which considers hypotheses of initialization, termination, translation, division and false positive detections. At first, reliable tracklets (i.e., short trajectories) were generated by linking detection responses based on frame by frame association approach. Afterward, these tracklets were globally associated over time to obtain final cell trajectories and lineage trees. The global association for cell trajectories was formulated as a maximum a posteriori (MAP) problem and then solved by a linear programming strategy to match the nuclei of two consecutive frames.

Magnusson and Jaldén [Magnusson 2012] developed a batch algorithm for track linking, in which the information from the whole sequence were employed to establish local decisions about cell tracks. The algorithm integrated image-based likelihoods of cell division and cell death into the tracking, without using heuristic post-processing algorithms or separate detection algorithms that make hard detection decisions over time. Hence, it presented more robust decisions over algorithms that process the data sequentially. This algorithm included a scoring function to rank different tracking hypotheses and an iterative algorithm with low complexity which looks forward to a set of tracks that maximizes the scoring function in a computationally efficient way using the Viterbi algorithm

Another study, as an extension to the previous algorithm [Magnusson 2012], was designed by Magnusson *et al.* [Magnusson 2015a] to manage false detections, missed detections, detections containing multiple cells, mitosis, apoptosis, and cells migrating in and out of the field of view. Moreover, the extended algorithm with reduced complexity was developed, to enable processing of the image sequences with thousands of cells affecting its performance.

In the previous studies [Magnusson 2012, Magnusson 2015a], the Viterbi track linking (VTL) algorithm was only used with Brownian motion models as it did not work well for particles with non-Brownian motion. This is because, a Markov property, which states that the linking scores should not depend on how the tracks were linked in prior images. In order to tackle this problem, Magnusson and Jaldén [Magnusson 2015b] presented a global track linking algorithm based on the Viterbi algorithm, with a Gaussian mixture probability hypothesis density (GM-PHD) filter. GMPHD filter was used as the first step for detection of particle locations. Consequently, a target distribution for each image along with Gaussian components representing possible locations and velocities of particles can be produced. Then, the Gaussian components were linked into tracks using VTL. This allowed distinguishing closely spaced particles based on their velocities, without losing the required Markov property.

A similar scoring function in conjunction with a track linking algorithm relied on an integer linear programming was later proposed by Schiegg *et al.* [Schiegg 2013]. This method included a probabilistic graphical model that can

handle detection errors in the tracking graph and even correct the false positives, false negatives, and under-segmentation errors. Since the detection and tracking were performed separately, i.e., did not incorporate any time information, the tracking result became highly dependent on the detection quality. As a result, the overall quality was limited due to the lack of cooperation between detection and assignment decisions.

A follow-up study, avoiding segmentation errors by jointly optimizing segmentation and tracking using one probabilistic graphical model was presented by Schiegg *et al.* [Schiegg 2014]. In this study, global spatial and temporal information were coupled together to select the hypotheses that best fit the overall tracking. Thus, the proposed model enabled to solve the problem especially if groups of cells appeared clumped together in some frames. Although this model had high computational complexity, the optimization problem still can be solved within a reasonable time.

Another joint segmentation and tracking method was proposed by Türetken *et al.* [Türetken 2017] to automatically detect and track cells in time-lapse images. The first step is to segment cells and generate over-complete segmentation hypotheses based on fitting ellipses hierarchically, then construct a spatiotemporal graph from all segmentation hypotheses and eventually formulate an integer linear programming (ILP) to select the consistent cells and find the optimal trajectories in this graph. Nevertheless, the use of ILP may be computationally expensive and may not be feasible for very dense and long sequences.

Similarly, Akram*et al.* [Akram 2016a] presented a novel proposal based joint cell segmentation and tracking method. This method employed cell segmentation proposals to represent potential cell hypotheses and construct a directed acyclic graph. Next, an iterative process was used to find the shortest path in this graph, which provided segmentation, tracks, and events for individual cells. In this method, a greedy iterative shortest path algorithm for performing inference was used rather than ILP to overcome the issue raised in the previous research [Türetken 2017].

Lou et al. [Lou 2014] introduced learning-based tracking which combines the use of two supervised learning algorithms to obtain better predictive performance.

The first was active learning algorithm that uses to retrieve informative training samples. The second was regularized max-margin learning approach which uses to obtain a regularized and globally optimized model. Similar to proposal-based tracking approaches, this method started with identify segmentation proposals in all frames and then use the learning algorithms along with an ILP to select the consistent cells and find the optimal trajectories.

A general, consistent, and extensible tracking approach that explicitly models cell behaviors in a graph-theoretic framework was proposed by Padfield *et al.* [Padfield 2011]. With this general approach, no specific models, parameter tuning, and training are needed and therefore not limited to tracking cells of a particular type or stained in a particular way. In this framework, a tracking approach based on extending the minimum cost flow algorithm has been employed to establish the association costs in a weighted directed graph. In addition, an efficient edge coupling approach has been applied for handling mitosis (splitting) and merging events.

Nowadays, deep learning-based methods have been proposed for cell tracking. For instance, He *et al.* [He 2017] designed a cell tracking method based on multi-task learning (MTL) and convolutional neural networks (CNNs). It consists of three models: particle filter motion model, a MTL observation model, and an optimized model update strategy. In the first, a cell position had been identified in the first-time frame. Next, the particle filter motion model was used to produce a set of candidates bounding boxes in the following frames. Finally, the MTL observation model was enabled to select the best candidate as the final target. CNNs act as part of MTL observation model which enables to learn robust cell feature whereas the optimized model update strategy employed to update the observation model to improve the tracking performance. A similar study was presented by Wang et al. [Wang 2017]. However, it employed the CNNs for feature extraction and two output layers for simultaneous detection and tracking. The two outputs shared the same features from convolutional layers that individually handles tracking or mitosis detection tasks. The tracking output selected the correct target in each frame, while the mitosis detection output detected whether the target cell is splitting (undergoing mitosis). A key

limitation of the convolutional neural networks(CNNs) based methods is that a single cell tracking was only considered and therefore the methods had not yet been widely adopted to multi object cell tracking. Table 2.4 summarizes the cell/nuclei tracking by model-based contour evolution approaches used for biological microscopy images. Table 2.5 summarizes the cell/nuclei tracking by detection-based association approaches for biological microscopy images.

Table 2.4: Summary of tracking by model-based contour evolution approaches.

References	Methodology	Image dimensions
[Dufour 2005]	Classical active contour model	3D
[Dzyubachyk 2007]	Level-set based variational model	2D
[Dzyubachyk 2008]	Multi-phase level set model	3D
[Li 2008]	Fast active contour model in conjunction with adaptive interacting multiple models (IMM) motion filtering and spatiotemporal trajectory optimization.	2D
[Padfield 2009]	Level set model with the fast-marching method and spatio-temporal constrains	2D
[Dzyubachyk 2009]	Graph cuts based active contours	3D
[Dzyubachyk 2010]	Multi-phase level set model	2D/3D
[Dufour 2011]	Adapted variational level set in the discrete active mesh framework	3D
[Maška 2012]	Integrated the original Chan–Vese model with the fast level set-like (FLS) framework	2D
[Maška 2013]	Integrated the original Chan–Vese model with the fast level set-like (FLS) and graph cut (GC) framework	3D
[Amat 2014]	A sequential Bayesian approach with Gaussian mixture models	3D

Table 2.5: Summary of tracking by detection-based association approaches.

References	Methodology	Image dimensions
[Li 2010]	Global nearest neighbor based on similarity measures i.e., appearance, distance, relationship, and intensity	2D
[Bise 2011b]	A spatio-temporal association with maximum a posteriori probability (MAP)	2D
[Padfield 2011]	Graph-based minimum-cost flow global optimization	2D/3D
[Magnusson 2012, Magnusson 2015a]	Graph-based shortest-path global optimization using Viterbi algorithm	2D/3D
[Schiegg 2013, Schiegg 2014]	Probability graph-based global optimization	2D/3D
[Lou 2014]	Global optimization with a supervised machine learning-based technique	2D
[Akram 2016a]	Multiple hypothesis and graph-based shortest-path global optimization	2D
[Türetken 2017]	Multiple hypothesis and graph-based tracking	2D
[He 2017, Wang 2017]	Similarity-based deep learning approach	2D

2.4 Existing software

In recent years, several software tools are made available for analysis of microscopy images. These tools involve commercial and open source platforms. A wide range of software tools include techniques for cell detection, segmentation, and tracking. Typically, techniques distributed with software tools are based on standard methods such as threshold-based methods for segmentation, watershed-based methods for splitting cell clusters and distance or size based association for cell tracking. Additionally, more sophisticated techniques are made available, but they are not always used. This is because they require tuning multiple parameters.

In this work, we highlight the more prominent and widely used tools and we present a comparison between them in Table 2.6. Furthurmore, we introduce in Table 2.7 the denoising and segmentation algorithms implemented with the existing software tools. An overview of various microscopy images analysis tools is presented in [Hilsenbeck 2016, Wiesmann 2015, Meijering 2012].

2.4.1 ImageJ/Fiji

ImageJ is developed throughout a grant from the National Institutes of Health (NIH) by Schneider *et al.* [Schneider 2012]. Still, it is the most widely used software for a variety of research areas in medicine and biology applications. Furthermore, it is a public-domain software and a completely open-source tool that allows users to modify the Java scripts and create new tools. This tool has a huge and fast-growing community, with many people developing their own plug-ins.

More tools, for instance, CellProfiler and Icy, have taken the initiative to support and employ ImageJ plug-ins. Since ImageJ has a flexible interface, a huge number of plug-in development projects have been presented which results in implementing several image processing algorithms and convenient utilities for everyday tasks.

Due to a rapid increase in the development of plug-ins, the dependencies between those plug-ins became complicated, and often hard to untangle. To overcome this issue, a distribution of ImageJ bundled with several plug-ins called

Fiji is emerged in 2007 and has been actively maintained by a large group of developers [Schindelin 2012]. The main specification of Fiji is maintaining an automatic plug-in updating function. In addition, Fiji includes more than 900 commands compared with ImageJ that involves about 500 commands.

2.4.2 Icy

Icy is a free, open-source platform developed at the Quantitative Image Analysis Unit, Institute Pasteur [De Chaumont 2012]. It offered a set of image analysis tools to be used in biological applications, i.e., image enhancement and filtering, active contours, colocalization, registration, particle detection, cell segmentation, and tracking. Additionally, it has enabled the development of a variety of image processing algorithms by integrating new plug-ins. Icy has provided powerful visualization capabilities (2D, montage and 3D volume rendering with VTK). Furthermore, it has been designed to interactively handle the multidimensional images (3D, time and channels).

2.4.3 CellProfiler

Likewise, CellProfiler is a free open-source software, designed for measuring and analyzing cell in biological images.

The first version is introduced by Anne Carpenter's team at the Broad Institute, MIT and Harvard [Carpenter 2006]. It is highly appreciated worldwide, in which it enables biologists to quantitatively measure phenotypes robustly from whether a few or thousands of images.

The second version of CellProfiler derived in Python from its original MATLAB implementation [Kamentsky 2011]. Furthermore, it has supported ImageJ plug-ins. This version has included methods for segmenting and tracking moving cells in video sequences and measuring neurons, worms, and tissue samples. In the latest version, i.e., CellProfiler 3.0 release [McQuin 2018], new methods are introduced for analyzing both whole-volume and plane-wise of three-dimensional (3D) images, using deep learning architectures and cloud computing resources. Moreover, the team has improved the usability and capabilities features.

2.4.4 ilastik

ilastik is developed by the ilastik team in Anna Kreshuk's lab at the European Molecular Biology Laboratory [Sommer 2011]. Similarly, ilastik is a user-friendly free open-source software for automated segmentation, classification, tracking, and analysis of the 2D+t and 3D+t sequences. It allows biologists to even manually annotate a small set of training samples. These annotation samples along with a set of image features are then used in machine learning-based techniques for classifying and segmenting cells from more challenging images.

ilastik projects can be imported to different software tools such as CellProfiler, FIJI to perform further post-processing or analysis tasks. The plug-in functionality takes advantage of a rich application programming interface (API) that enables advanced users to develop their own specific algorithms.

2.4.5 Imaris

Imaris [1] is a commercial software tool that is being for studying functionality, visualization, segmentation, and interpretation of 3D and 4D microscopy datasets. The user-friendly interface along with powerful rendering engines allow the users to interactively reconstruct and explore big data. Also, Imaris provides eight unique, tightly integrated modules that provide the opportunity to build a package based on a researcher's specific needs. These modules include automatic image processing of multiple 2D/3D + time images; quantitative image analysis; 3D and 4D object tracking; visualizing and measuring; isolating, visualizing and quantifying colocalized regions and much more.

For flexible use, Imaris can be customized with plug-ins (called Xtensions) via the ImarisXT API. Commonly, Imaris XTensions scripts are written in MATLAB. Recently, it added an easy integration for Python scripts and their plug-ins.

[1] http://www.bitplane.com

2.4.6 Amira

Amira is a commercial software platform for advanced 3D/4D data manipulation [Stalling 2005], developed by Thermo Fisher Scientific [2] in collaboration with the Zuse Institute Berlin (ZIB)[3]. It allows visualizing, processing, and understanding life sciences data coming from computed tomography (CT), microscopy, magnetic resonance imaging (MRI), and several other imaging modalities. In particular, it gives researchers the opportunity to do intra- and intercellular image segmentation and employing powerful automated object tracking solution.

Amira can be controlled via the TCL scripting language and interfaced with Matlab. Moreover, various modules can be created using C++ programming language. Recently, an alternative shell-script based on Python language becomes available, and so it gives Amira software more strength.

[2] https://www.thermofisher.com
[3] http://www.zib.de/

Table 2.6: Comparison of available software tools for microscopy image analysis.

Software name	Operating system bundle	Scripting language	Primary function	Image dimensions	Available
ImageJ/Fiji [1]	All	ImageJ Macro, Javascript, Python, JRuby, BeanShell,Groovy, Clojure	Image analysis	2D/3D	Free
Icy [2]	All	ImageJ Macro, Javascript, Python, Protocol (Graphical Programming)	Data visualization, annotation and analysis	2D/3D	Free
CellProfiler [3]	All	Python	Image processing and analysis	2D/3D	Free
ilastik [4]	All	Python	Image processing and machine learning	2D/3D	Free
Amira [5]	All	TCL,Python	Data visualization, processing, and analysis	3D/4D	Commercial
Imaris [6]	Win, OSX	Matlab	Data visualization and analysis	3D/4D	Commercial

[1] https://imagej.net/ImageJ
[2] http://icy.bioimageanalysis.org/
[3] https://cellprofiler.org/
[4] https://www.ilastik.org/
[5] https://www.fei.com/software/
[6] http://www.bitplane.com

Table 2.7: Denoising and segmentation algorithms implemented with the existing software tools. NA: not available.

	Implemented algorithms	
Software name	**Denoising**	**Segmentation**
ImageJ/Fiji[1]	Median filter Gaussian filter Wiener filter Band-pass filter Fast non-local means PURE-LET Deep convolutional neural network	Classical thresholding Multiple-level thresholding Adaptive iterative thresholding Edge detection Ellipses fitting Hough transform circle detection Active contour Region growing Marker controller watershed Machine learning U-Net convolutional networks
Icy[2]	Median filter Gaussian filter Band-pass filter	Iterative thresholding Classical and adaptive thresholding Edge detection Ellipses fitting Active contour Marker controller watershed
CellProfiler[3]	Median filter Gaussian filter Non-local means	Classical and adaptive thresholding Ellipses fitting Edge Detection Active contour Marker controller watershed Machine learning U-Net convolutional networks
ilastik[4]	NA	Edge Detection Classical watershed Machine learning U-Net convolutional networks
Amira[5]	Median filter Gaussian filter Non-local means	Classical thresholding Region growing Marker controller Watershed
Imaris[6]	Median filter Gaussian filter	Classical thresholding Edge Detection Ellipses fitting Region growing Marker controller Watershed

[1] https://imagej.net/ImageJ
[2] http://icy.bioimageanalysis.org/
[3] https://cellprofiler.org/
[4] https://www.ilastik.org/
[5] https://www.fei.com/software/
[6] http://www.bitplane.com

2.5 Summary

This chapter reviewed existing approaches for denoising, segmentation, and tracking of cell/nuclei in microscopy images. In addition, it discussed available software tools for analyzing microscopy images. Among all image processing tasks, denoising remains an essential step before performing higher level tasks, as the noise has a significant effect on extraction of information.

In denoising techniques, PURE-LET showed very promising denoising results for 3D and time-lapse fluorescence microscopy data. On the other hand, median filter is a simple and widely used method in most of microscopy image analysis. Moreover, for cell/nuclei segmentation, we noticed that thresholding and watershed-based methods are the most used ones. In the same manner, most researchers focused on using detection-based association approaches for tracking of cell/nuclei.

Although, existing software tools as well as the aforementioned denoising, segmentation, and tracking methods have achieved good results in microscopy images, most of them require tuning multiple set of parameters. Moreover, they could be used effectively only for specific applications. As a result, implementing a proper and robust approach for various datasets regarding high variation in cell nuclei volume, shape, and stain distribution as well as high cell density has become a major challenge in analysis of microscopy images.

From this perspective, we proposed a generic framework (explain in Chapter 3) for denoising 3D microscopy images that only require few parameters. This method is based on unsupervised dictionary learning and sparse representation technique. Next, two segmentation methods based on thresholding and watershed are employed to get segmentation mask and subsequently to obtain the final cell nuclei segmentation results.

Chapter 3

Methodology

Contents

This section introduces a novel method for denoising and detection of cell nuclei in 3D TLFM images based on a sparse representation approach [Aharon 2006, Elad 2006]. The use of sparse signal representation is becoming popular in several fields such as face recognition [Adamo 2015], image denoising [Nasser 2017, Qiu 2012, Boulanger 2010] and inpainting [Ogawa 2013], and image classification [Alegro 2017, Theodorakopoulos 2014]. Indeed, natural images represent very sparse data, especially in biology where numerous instances of the same structure,

i.e. cell or nucleus, are present in the image. Moreover, a dictionary-based approach is usually linked to unsupervised learning since the data itself can be used to learn the basis vectors to build a sparse representation matrix.

The sparse representation method (shown in Figure 3.1) is implemented as described by M. Elad and M. Aharon [Elad 2006], we have only changed the construction of the initial dictionary as depicted in the following steps. Firstly, the patches are extracted by moving a window with a step size of one pixel over the raw image. For each extracted patch, pixels' intensities are averaged. Then, the mean intensity over all patches is calculated. Secondly, an initial dictionary is constructed by selecting random patches from extracted patches among those having intensities greater than the obtained average intensity. By doing that, we are ensuring the presence of cell nuclei patches in the initial dictionary. Thirdly, a technique based on K-means with singular value decomposition (K-SVD) [Aharon 2006] is implemented to update and obtain the final dictionary. Fourthly, the updated dictionary is used to reconstruct the denoised image as well as the detection map that will be used for detection of cell nuclei.

In the cell nuclei segmentation stage, the maximum response image, which is obtained by multiplying the denoised image with the detection map is used to detect the potential location of cell nuclei. Then, a thresholding-based approach [Bradley 2007, Otsu 1979] is proposed to get the segmentation mask. Finally, a marker-controlled watershed approach [Parvati 2008] is used to obtain the final individual cell nuclei segmentation result.

3.1 Denoising of 3D cell nuclei images.

3.1.1 An introduction to sparse representation

The idea of sparse representation is to obtain an efficient representation of a signal as a linear combination of few atoms chosen from a dictionary. Given a dictionary $D \in R^{n \times K}$ that contains K atoms, i.e., a unit vector of length n, as column vectors $d_j \in R^n, j = 1, 2, \ldots, K$, where n is a patch size. The sparse representation problem of a signal $y \in R^n$ can be described as finding the sparsest

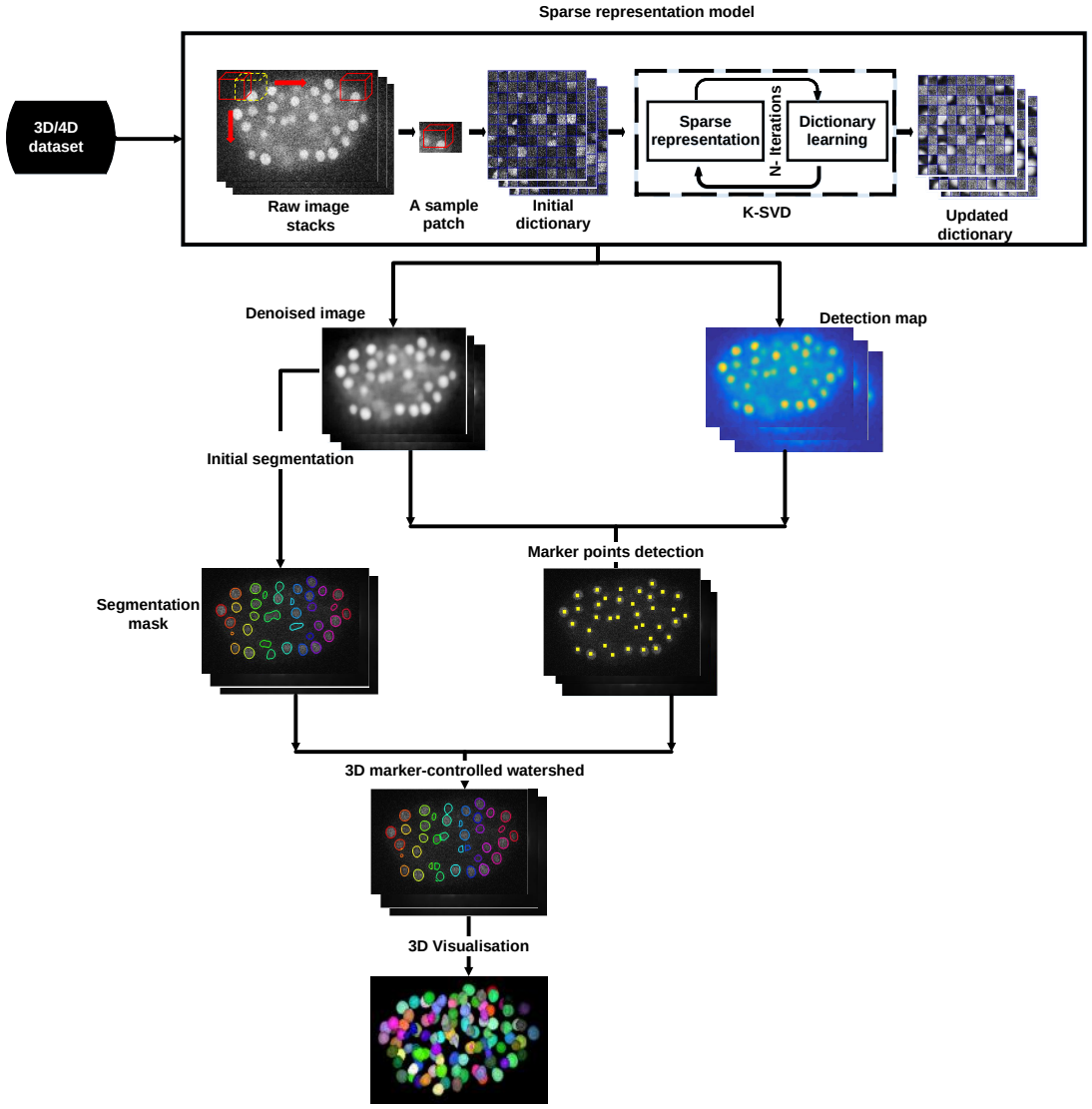

Figure 3.1: General representation of the proposed framework for denoising and segmentation of cell nuclei in 3D time-lapse fluorescence microscopy images. The proposed pipeline consists of data preprocessing, initial cell nuclei segmentation, cell nuclei detection, final segmentation as well as 3D visualization. In the preprocessing step, an initial dictionary is constructed by selecting random patches from the raw image as well as a K-SVD technique is implemented to update the dictionary and obtain the final one. Then, the maximum response image which is obtained by multiplying the denoised image with the detection map is used to detect marker points. Furthermore, a thresholding-based approach is proposed to get the segmentation mask. Finally, a marker-controlled watershed approach is used to get the final cell nuclei segmentation result and hence cell nuclei are displayed in a 3D view.

vector $a \in R^K$ where $y \simeq Da$. The problem can be formulated as an energy optimization problem as follows:

$$\hat{a} = \underset{a}{\operatorname{argmin}} \|y - Da\|_2^2 \quad s.t. \quad \|a\|_0 \leq L \tag{3.1}$$

where y is the signal, a denotes the sparse representation coefficients, $\|a\|_0$ is the L_0 pseudo-norm that counts the number of non-zeros of a and L is a predetermined sparsity threshold.

Solving the previous optimization problem is NP-hard and numerically intractable, thus several methods have been developed to get an approximate solution for this particular problem. The first type of methods uses $\mathrm{L}_0 -$ norm minimization, such as matching pursuit (MP) [Mallat 1993b] or orthogonal matching pursuit (OMP) [Mallat 1993a]. The second type of method uses $\mathrm{L}_1 -$ norm for optimization. The objective of $\mathrm{L}_1 -$ norm is to make the optimization problem convex, which can be addressed efficiently using basis pursuit (BP) [Chen 1998].

Among these methods, we discuss the orthogonal matching pursuit (OMP) algorithm that has been widely used because it is easy to implement and it provides a satisfactory stable results. Moreover, it has a low computational complexity. The orthogonal matching pursuit (OMP) is a greedy algorithm which aims to find a sparse representation of a signal of interest, given a predefined dictionary. The main idea of this algorithm is to iteratively search for the best atoms that minimize a residual. To explain, initially, the atom $\hat{k} = \underset{k}{\operatorname{argmax}} |d_k^T r|$ with the maximal correlation to the residual r is selected. Then, the current signal y is orthogonally projected onto the span of the selected atoms by computing $a_I = (D_I)^+ y$, where I denotes a set containing indices of selected atoms, D_I are the corresponding columns of D, $(D_I)^+$ represents pseudoinverse of D_I and a_I are the corresponding columns of A. Finally, the new residual is computed by $r = y - D_I a_I$ and the process is repeated until the stopping criterion is satisfied (see Algorithm 1). This criterion depends on sparse constraint so that this process is terminated when a specific number of distinct atoms have been selected.

Algorithm 1 Orthogonal matching pursuit

Inputs:
Dictionary D, signal y, sparsity threshold L
Output:
Sparse representation a such that $y \simeq Da$
Initialize:
set $I := ()$, $r := y$, $a := 0$
while stopping criterion not met **do**
  $\hat{k} := \underset{k}{\operatorname{argmax}} |d_k^T r|$
  $I := (I, \hat{k})$
  $a_I := (D_I)^+ y$
  $r := y - D_I a_I$
end while

Dictionary selection

The crucial issue for any practical applications is to select the dictionary D. Basically, dictionaries are of two types: (1) fixed dictionaries and (2) adaptive dictionaries. The fixed dictionary as is the case for curvelet, discrete cosine, wavelet, ridgelet, or bandlet [Rubinstein 2010] which use pre-defined and fixed atoms. This dictionary might not ensure a well-defined representation of all given signals. As a result, it is more appealing to use an adaptive dictionary approach to learn the dictionary directly from the data itself.

Learning the dictionary requires two steps, the first step is to compute an initial dictionary. It is usually computed by taking random patches directly from the raw image. These patches are overlapped with a step size of one pixel. To ensure the presence of patches containing nuclei in the initial dictionary beside background patches, we select patches among those having intensity greater than the average intensity of all patches extracted from the image. The second step is to update the initial dictionary using the K-SVD algorithm [Aharon 2006].

The K-SVD algorithm (Algorithm 2) is a standard unsupervised adaptive dictionary learning algorithm that generalizes the well-known K-means clustering algorithm [Hartigan 1979]. It jointly learns a dictionary $D = [d_1, d_2, \ldots, d_K]$, $d_j \in R^n, j = 1, 2, \ldots, K$ and a related sparse representation matrix $A = [a_1, a_2, \ldots, a_m], a_i \in$

R^K, from a set of training signals $Y = [y_1, y_2, \ldots, y_m]$, $y_i \in R^n$, $i = 1, 2, \ldots, m$ by solving the following problem:

$$\underset{D,A}{\operatorname{argmin}} \|Y - DA\|_2^2 \quad s.t. \quad \|a_i\|_0 \leq L \tag{3.2}$$

This technique solves the optimization problem by alternating between finding the sparse representation coefficients A and the dictionary D using an iterative approach. Assuming that D is known, the best sparse representation matrix is constructed by solving Equation (3.2) using an orthogonal matching pursuit algorithm (OMP).

Following the sparse representation stage, we search for the optimal dictionary D toward better sparse representation of an image. Since finding the whole dictionary at the same time is impractical, the dictionary is updated atom by atom while assuming the sparse representation vectors (A) are fixed. The update of the $k-$th column of D is done by reformulating the optimization problem as follow:

$$\|Y - DA\|_F^2 = \|Y - \sum_{j=1}^{K} d_j a_L^j\|_F^2 = \|(Y - \sum_{j \neq k} d_j a_L^j) - d_k a_L^k\|_F^2 = \|E_k - d_k a_L^k\|_F^2 \tag{3.3}$$

where, a_L^k denotes the $k - th$ row of A and the matrix E_k stands for the error for all the signals when the $k - th$ atom is removed.

In this problem, the multiplication term DA is first decomposed into a summation of K rank-1 matrices, whereas the other $K - 1$ terms are assumed to be fixed, and the $k-$th remains unknown. Afterward, the optimization problem can be solved by approximating the E_k term with a rank-1 matrix using singular value decomposition (SVD), then updating d_k. However, such solution can not enforce the sparsity constraint of the resulting matrix A .

To handle sparsity constraint problem, we define a group of indices ω_k representing the signals y_i that use the d_k atom (i.e., those where is nonzero) as depicted below:

$$\omega_k = \{i | 1 \leq K \leq N, a_L^k(i) \neq 0\} \tag{3.4}$$

Then, we define Ω_k as a matrix of size $N \times |\omega_k|$, with ones on the (i, ω_k)-th entries and zeros elsewhere. We obtain a_R^k by multiplying a_T^k and Ω_k, thus the row vector is shrunk by discarding of the zero entries, resulting with the row vector a_R^k of length $|\omega_k|$. Likewise, we obtain Y_k^R by multiplying Y and Ω_k and this creates a matrix of size $N \times |\omega_k|$ that includes a subset of the signals that are currently using the d_k atom. In the same way, $E_k^R = E_k \Omega_k$.

Thus, the optimization problem described in Equation (3.3) can be accomplished practically by reformulating it as:

$$\|E_k \Omega_k - d_k a_L^k \Omega_k\| = \|E_k^R - d_k a_R^k\|_F^2 \tag{3.5}$$

and this optimization problem can be done directly via singular value decomposition (SVD). The main idea of such technique is to compute the singular value decomposition (SVD) of an image and then discard small singular values. Small singular values mainly represent the noise. On the other hand, the large singular values represent the significant features.

To explain, the SVD has several application in image processing. In particular, it has been widely used to restore a corrupted image by distinguishing between significant structure and the noise data. The SVD decomposes the restricted matrix E_k^R to $E_k^R = U \Delta V$. The solution for d_k is defined as the first column of U, and the coefficient vector a_R^k as the first column of V multiplied by $\Delta(1,1)$. After updating the whole dictionary, the process is iterated alternately between solving A and D. Once the best dictionary and sparse representation coefficients are obtained, the denoised image and detection map can be constructed. An example of the final adaptive dictionary that leads to those images is presented in Figure 3.2.

Algorithm 2 K-SVD algorithm

1: **Inputs:**
Noisy image Y, dictionary D, sparsity threshold L, number of iterations N
2: **Output:**
The optimal sparse representation coefficients A and optimal dictionary D
3: **Initialize:**
Dictionay D with patches randomly extracted from Y
4: **for** i<N **do**
5: Sparse representation: compute sparse representation coefficients $\hat{A}$ (Equation (3.1)) using OMP algorithm (Algorithm 1)
6: Dictionary update: for each column $k = 1, 2, \ldots, K$ in D, update it using the following steps:
   - Define the set of signals that use this atom by $\omega_k = \{i | 1 \leq K \leq N, a_L^k(i) \neq 0\}$
   - Compute the representation error matrix by $E_k = (Y - \sum_{j \neq k} d_j a_L^j)$
   - Find E_k^R by choosing only the columns corresponding to ω_k, $E_k^R = E_k \Omega_k$.
   - Apply SVD decomposition $E_k^R = U \Delta V$, choose the updated dictionary atom d_k to be the first column of U. Update the sparse representation coefficients a_R^k to be the first column of V multiplied by $\Delta(1,1)$.
7: **end for**

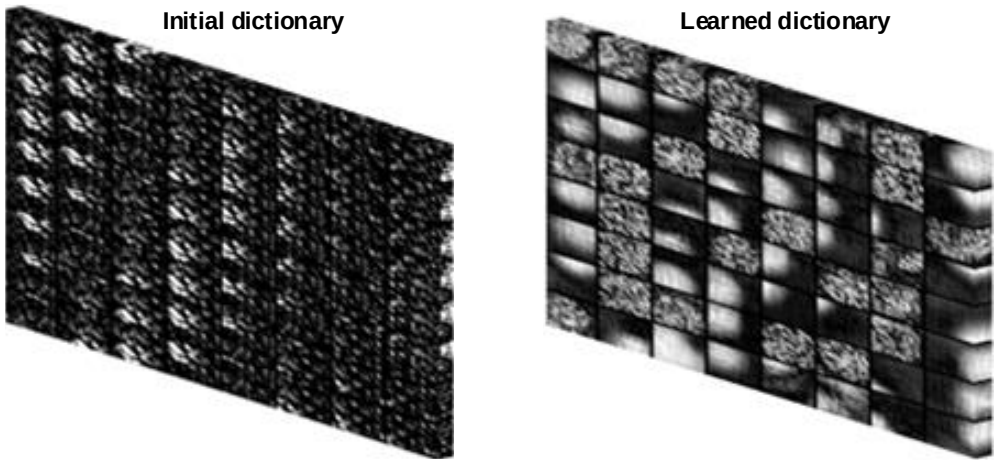

Figure 3.2: Example of the adaptively learned dictionary.

3.1.2 Images with sparse representation.

After obtaining the optimal dictionary and sparse representation coefficients, the denoised image and detection map are reconstructed as explained in Algorithm 3, which will be used later in the detection and segmentation of cell nuclei.

Denoised image reconstruction.

The denoised image with sparse representation is formed by solving the following optimization problem:

$$\hat{X} = \underset{X}{\operatorname{argmin}} \ \lambda \ \|X - Y\|_2^2 + \|X - DA\|_2^2 \tag{3.6}$$

The first term in Equation (3.6) introduces the log-likelihood global force that demands the proximity between X and Y. λ is the regularization parameter that controls this fidelity term. When the noise level in the image is low, Y is close to X, and λ should be larger. In contrast, with the high noise level, Y is getting away from X, and λ should be smaller. The second term is used to makes sure that in the constructed image, every patch in every location has a sparse representation.

Equation (3.6) can be solved as a set of smaller optimization problems (i.e. at patch level) which is defined by:

$$\hat{X} = \underset{X}{\operatorname{argmin}}\ \lambda\ \|X - Y\|_2^2 + \|H_i X - Da_i\|_2^2 \quad i = 1, 2, 3, \ldots, m. \tag{3.7}$$

where H_i is the matrix which selects the $i-$th patch from X (i.e. $x_i = H_i X$), X is the latent clean image, Y is the noisy image and λ is the regularization parameter.

The obtained A and D from Algorithm 2 which are the optimal dictionary and sparse representation coefficients are now fixed and thus we turn to update X .

Returning to Equation (3.7), we need to solve the the following optimization problem:

$$\hat{X} = \underset{X}{\operatorname{argmin}}\ \lambda\ \|X - Y\|_2^2 + \|H_i X - Da_i\|_2^2 \tag{3.8}$$

This is a simple quadratic term that has a closed-form solution of the form:

$$\hat{X} = \frac{(\lambda Y + \sum_i H_i^T D a_i)}{(\lambda I + \sum_i H_i^T H_i)} \tag{3.9}$$

where I is an identity matrix and $\hat{X}$ is the denoised image. The solution to this problem includes averaging of the overlapping patches with a weighted sum of the original noisy patch. As the patches are overlapping, the final value of each voxel is an average of all representations obtained from the sparse representation stage.

Detection map reconstruction.

In point of fact, the denoised image does not have sufficient contrast to completely separate touching nuclei. In order to improve cell nuclei detection, a detection map image that indicates the potential locations of cell nuclei will be built. The construction of this image (Imap) is based on the computation of the sparse

coefficients (a_i) of each image patch. It can be obtained by:

$$p_i = \frac{1}{C} \sum_{k=1}^{K} |a_i^k| \quad i = 1, 2, 3, \ldots, m. \tag{3.10}$$

where, C is a normalization term, p_i is the probability value corresponds to the $i-$th patch and a_i^k denotes the $k-$th element of a_i.

The final detection map can be obtained from the following form :

$$Imap = Op_i \quad i = 1, 2, 3, \ldots, m. \tag{3.11}$$

where O is a vector of all ones with dimensions $n \times 1$, and n represents patch size. Thus, all voxels values i.e., n elements of the $i-$th patch in the detection map have only one value equals to p_i. The value of each voxel in the detection map ranges between 0 and 1, where a value of 0 indicates the background voxel and 1 indicates the nucleus voxel. To give an example, assume all atoms in the dictionary have the same average intensity, to reconstruct a patch with higher (brighter) intensity values, the sparse coefficients of such patch ((as presented in Figure 3.3 (c))) will be higher than sparse coefficients of a patch in the background (as presented in Figure 3.3 (d)). Thus, this detection map is only based on the voxel intensity values of the raw image. Notably, voxels within the center of the nucleus have very high sparse coefficients values, in contrast to voxels far away from the center having low values. Consequently, the p_i value of the patches containing nuclei (as shown in Figure 3.3 (b, c)) tend to be large compared with the p_i value of background patches (as presented in Figure 3.3 (d)). As a result, the dictionary learning technique with sparse representation can capture strong structures of biological images as well as restrain the noise.

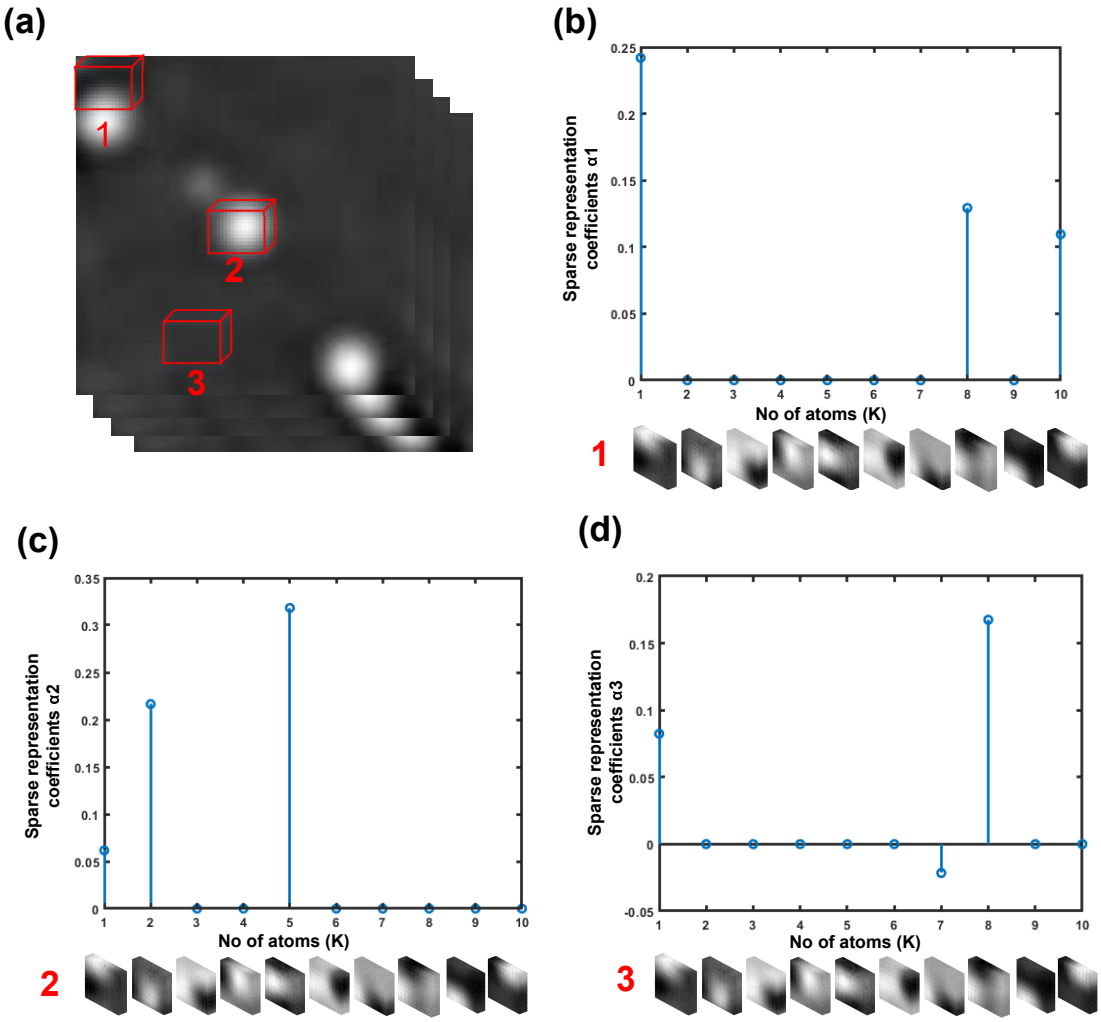

Figure 3.3: Example of patches sparse coefficients (a) sample patches i.e. red rectangles $(1, 2, 3)$ overlaid on the filtered image. (b) sparse coefficients of a sample patch (1) contain part of nucleus. (c) sparse coefficients of sample patch (2) contain nucleus. (d) sparse coefficients of a sample patch (3) from background.

Algorithm 3 Denoised image and detection map

Inputs:
Noisy image Y, Dictionary D, Sparse coefficients matrix A

Output:
Denoised Image $\hat{X}$ and detection map $Imap$

Initialize:
The optimal dictionary D and optimal sparse coefficients matrix A obtained from (Algorithm 2)

Denoised image:

$$\hat{X} = \underset{X}{\operatorname{argmin}}\ \lambda\ \|X - Y\|_2^2 + \|R_i X - Da_i\|_2^2$$

$$\hat{X} = \frac{(\lambda Y + \sum_i R_i^T D a_i)}{(\lambda I + \sum_i R_i^T R_i)}$$

Detection map:

$$p_i = \frac{1}{C}\sum_{k=1}^{K} a_i^k \quad i = 1, 2, 3, \ldots, m.$$

$$Imap = Op_i$$

3.2 Cell nuclei segmentation.

In time-lapse fluorescence microscopy images, nuclei appear as bright objects on a dark background. Consequently, nuclei can be neatly extracted from the background using a simple thresholding method, but such method unable to segment touching nuclei.

In previous research, a watershed-based method was proven to be an effective tool for segmenting such touching nuclei scenarios. However, classical watershed approach usually leads to an over-segmentation problem.

In this section, we present how a thresholding-based approach is used to get initial cell nuclei segmentation. Then, we discuss the generic marker extraction method, which based upon both the denoised image and detection map (Section 3.1.2, Section 3.1.2). Afterward, we briefly review the watershed-based algorithm and the marker-controlled watershed. Algorithm 4 present the main steps of cell nuclei segmentation process.

3.2.1 Initial cell nuclei segmentation.

A local adaptive-thresholding approach [Bradley 2007] is applied to the denoised image, in which the general concept of the algorithm is that for every image's voxel the threshold is determined by the following equation:

$$\mathrm{T}_{\mathrm{local}} = \mathrm{mean}_{\mathrm{local}} * (1 - \mathrm{SensitivityFactor}) \tag{3.12}$$

where, $\mathrm{mean}_{\mathrm{local}}$ is the mean intensity value in the neighbourhood of each voxel and the SensitivityFactor is a scalar value within a range from zero to one which controls sensitivity towards thresholding more voxels as foreground. Accordingly, voxels with intensity values larger than $\mathrm{T}_{\mathrm{local}}$ are set to 1, all others are set to 0. Small regions detected as foreground and smaller than a predefined volume denoted by $MinNucleiVolume$ are discarded. This threshold corresponds to the volume of the smallest cell nucleus and is determined prior to the segmentation step. The resulting image is called the segmentation mask.

Alternatively, the global Otsu's thresholding-method [Otsu 1979] can also

be employed to obtain the segmentation mask, in which the basic idea of Ostu method is to iterate through all the possible threshold values and to choose the optimal one that minimizes the within-class variance. The within-class variance is simply defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(t) = w_b^2(t)\sigma_b^2(t) + w_f^2(t)\sigma_f^2(t) \tag{3.13}$$

where, weights ω_b and ω_f are the probabilities of the two classes separated by a threshold t ,and σ_b^2 and σ_f^2 are variances of these two classes.

Subsequently, for each potential threshold t, first, the voxels are separated into two classes according to the threshold. Then, the mean of each class is calculated. Afterwards, the squared difference between the mean of the two classes is determined as well. Finally, the number of voxels in one class are multiplied by the number of voxels in the other class.

3.2.2 Marker points detection.

For splitting of touching cell nuclei, we employed a marker-controlled watershed technique. The marker points are obtained as follow: first, the denoised image is multiplied by the detection map to provide a maximum response image. Second, The maximum response image is processed to detect the local maxima (Figure 3.4). The obtained local maxima image is multiplied by the segmentation mask to discard local maxima detected in the background. Third, a morphological dilation operator of certain radius denoted by $NucleiSeedDilation$ is employed to avoid detection of multiple local maxima for the same object by merging those maxima that were in close proximity to each other. Finally, the modified image determining the marker points is fed to the subsequent watershed algorithm.

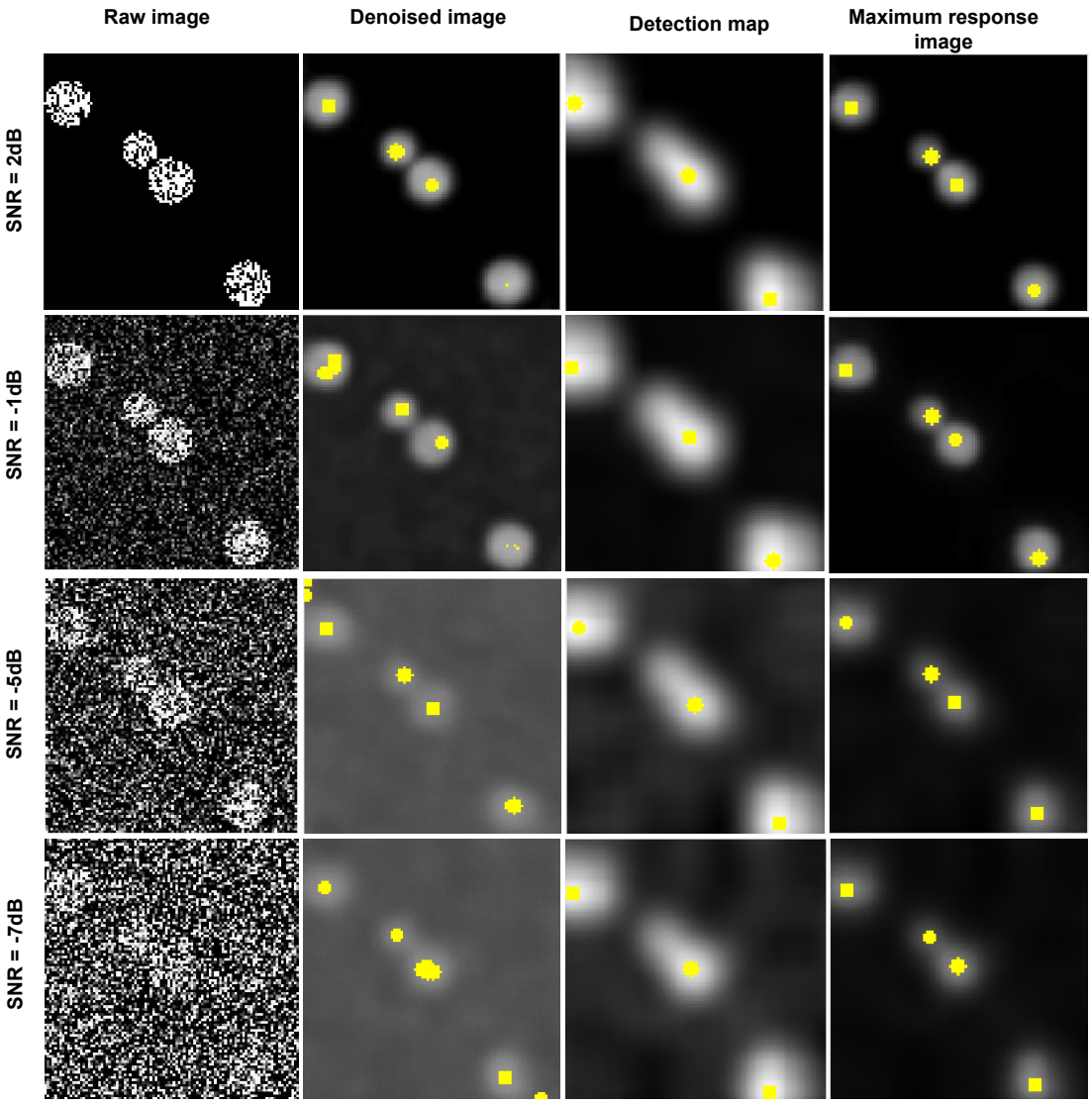

Figure 3.4: A comparison of marker points detection at various levels of noise. First column: representative single plane (Z = 10) of the raw image. Second column: the results of marker points detection from the denoised image. Third column: the result of marker points detection from the detection map. Fourth column: the result of marker points detection from the maximum response image. For all images the marker points depicted by yellow markers. Note that, the marker point detection here is performed in two dimensions for the purpose of explanation and visualisation. However, in the framework it is applied in three dimensions.

3.2.3 Watershed and marker-controlled watershed segmentation.

A classical watershed transform [Gonzalez 2002] can be used to separate clusters into individual cells. This approach is based on the flooding simulation as initially proposed by Pierre Soille and Luc M. Vincent [Soille 1990].

The main concept of the watershed transform is to consider the input image as a topographic surface, where, the process starts with placing a water source within each regional minimum in the image. Then, the entire relief is flooded from the sources at last the dams are built where the different water sources meet.

All points on the surface at a specified minimum represent the catchment basin associated with that minimum. On the other hand, the zones that separate adjacent catchment basins are considered as watersheds as shown in Figure 3.5.

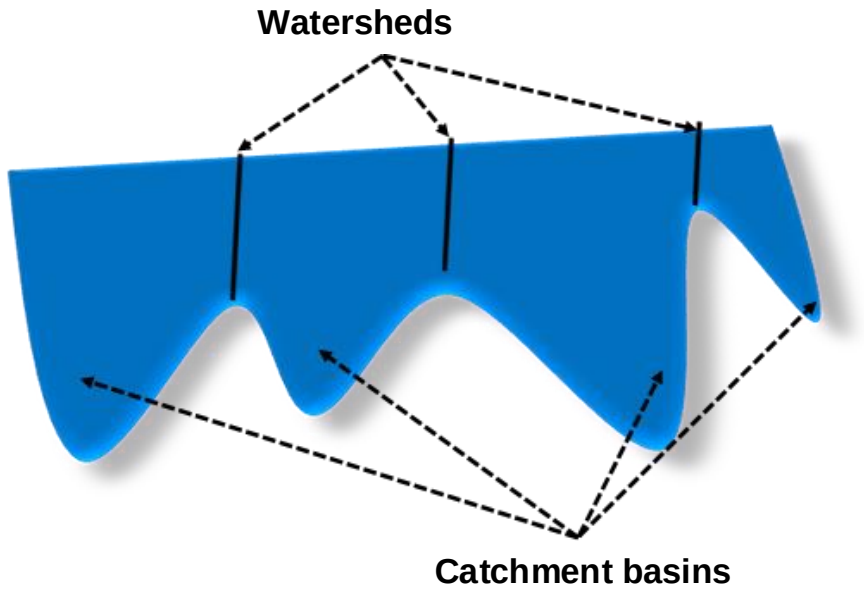

Figure 3.5: The watershed transform: strategy for clustered objects segmentation.

As the homogeneity of the grey values of the objects present in the image is considered as the fundamental principle of the segmentation process, a gradient

image or a distance map is often used in the watershed transformation rather than grayscale images. The gradient magnitude image (shown in Figure 3.6 (c)), which has high pixel values for object edges and lower pixel values everywhere else, can be easily constructed using the first or the second order derivative methods. Likewise, the distance map (presented in Figure 3.6 (b)) is assigned a number for each pixel that is described as the distance from that pixel to the nearest non-zero pixel in the image.

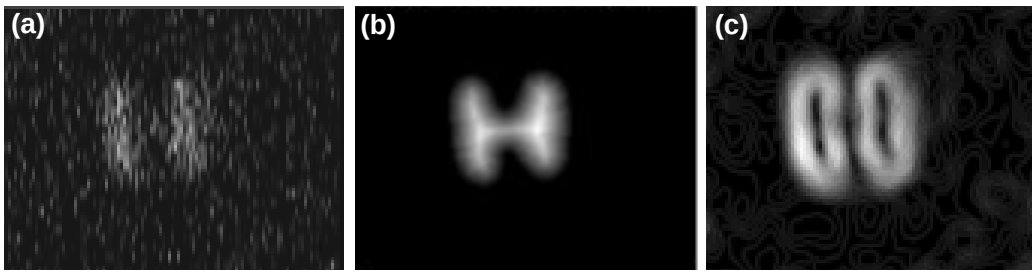

Figure 3.6: Example of the images fed to the watershed transform (a) raw image. (b) distance map. (c) gradient magnitude image.

The main problem of the classical watershed transformation is the over-segmentation in which individual objects might split into several objects as shown in Figure 3.7. This happens mainly owing to noise or other irregularity in the image.

In order to overcome the shortcomings of the classical watershed, various solutions are proposed in existing literature. One solution is to merge adjacent regions according to some criteria after using the watershed, but this approach is a highly time consuming and the design of merging criteria is hard to create.

Another solution, is to use marker-controlled watershed transform. The main idea of marker-controlled watershed is to start the flooding from specific markers (seed points) instead of the regional minima. Each marker has a direct relationship to a specific watershed region, so the number of markers equal to the total number of watershed regions. In such algorithm, proper marker selection has a major impact on the segmentation results.

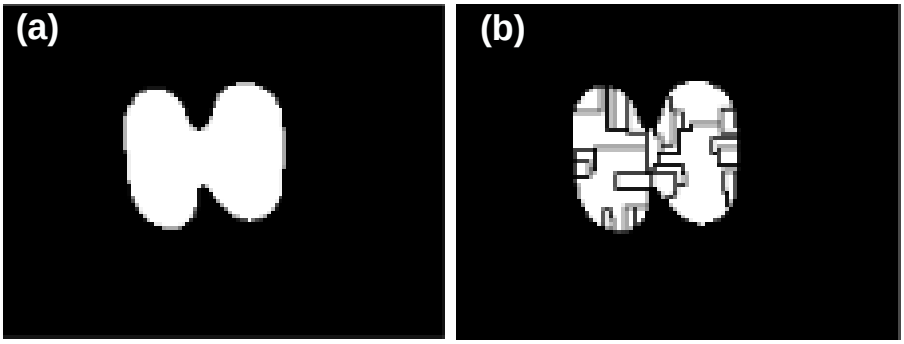

Figure 3.7: Example of over-segmentation problem resulted from classical watershed. (a) Binary image. (b) Segmentation result by classical watershed.

To give an example, Wählby *et al.* [Wählby 2004] and [Magnusson 2016] employed the h-maxima/minima transforms to find marker points. With the h-maxima/minima transform, all undesired regional maxima/minima whose values less than h-values are suppressed.

For this algorithm, the empirical selection of the h-value often makes robust segmentation difficult. This is because that the h-value has a direct influence on the number of segmented regions, and as such the larger the h-value, the fewer the number of the segmented regions. Figure 3.8 shows regional maxima and corresponding watershed segmentation results from a merged nuclei with different h-values adopted.

In this thesis, marker-controlled watershed segmentation is implemented to separate connected cell nuclei clusters. In the begining, initial markers are detected by applying the marker point detection method discussed in Section 3.2.2. Then, the watershed transform is applied to flood the denoised image, which contains merged objects starting at the predefined marker points as sources. Sometimes the flooding process is not stopped at the border of a cell nucleus, therefore the denoised image is multiplied by the segmentation mask prior to the flood. Eventually, watershed dams are built when different sources meet during the flooding process. This approach allows splitting clusters of apparently touching

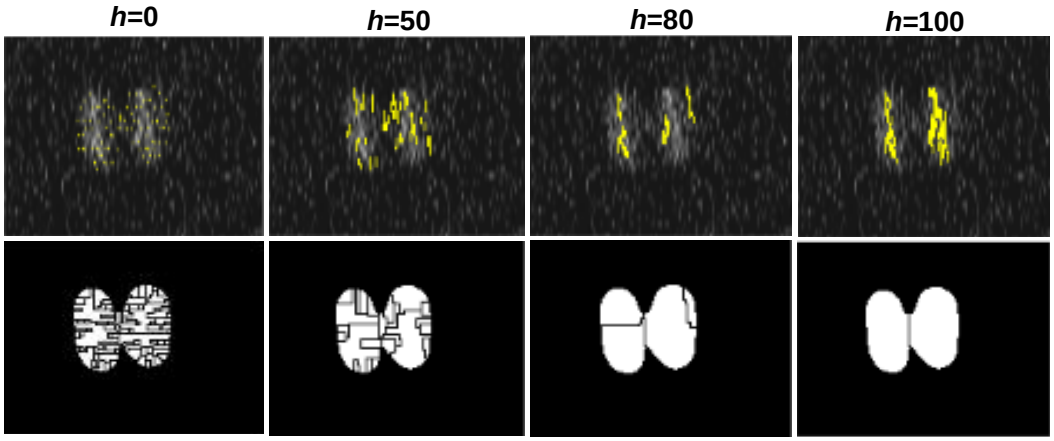

Figure 3.8: **Regional maxima detection and watershed segmentation results with several h-values. First row: regional maxima (marked in yellow) overlaid on the raw image. Second row: watershed segmentation results.**

cell nuclei.

Algorithm 4 Cell nuclei segmentation

1: **Inputs:**
Denoised image, detection map
2: **Output:**
A label image identify the individual cell nuclei
3: **Steps:**

1. Find the segmentation mask by applying a thresholding-based segmentation method to the denoised image.
2. Extract markers by finding the local maxima from the detection map.
3. Apply marker-controlled watershed transform to flood the denoised image starting from the predefined marker points as sources.

3.3 Evaluation method and metrics.

To assess the performance of the proposed algorithm, three metrics are employed. The first two metrics are the *recall* [Powers 2011] and *precision* [Powers 2011] of object detection. The recall is the proportion of the number of relevant detected cell nuclei to the total number of relevant cell nuclei in ground truth. Precision is the proportion of the number of relevant detected cell nuclei to the total number of irrelevant and relevant detected cell nuclei. These parameters are defined as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3.14}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3.15}$$

$$\text{F-measure} = 2 * \frac{\text{Sensitivity} * \text{Precision}}{\text{Sensitivity} + \text{Precision}} \tag{3.16}$$

where, True Positive (TP) represents the total number of correctly detected nuclei, False Negative (FN) represents the number of undetected nuclei and False Positive (FP) represents the number of falsely detected nuclei. To compute these values, we used the following steps: first, we calculated the distance between the centroids of ground truth nuclei and centroids of segmented objects. Second, a weight is assigned to each pair of segmented and ground truth objects, equal to the distance between them. Third, Hungarian algorithm is used to solve this assignment problem. Objects with no match to any other object are considered as FP, objects absent in ground truth, but they appear in the segmentation result are deemed FP, and FN were objects absent in the segmentation result despite these objects appear in ground truth.

The third metric is the Jaccard index [Rohlfing 2012] that measures the segmentation accuracy of the segmented objects. The Jaccard index for each set of segmented (A) and ground truth (B) objects is defined as the intersection between them divided by their union.

$$J(A, B) = \frac{A \cap B}{A \cup B} \tag{3.17}$$

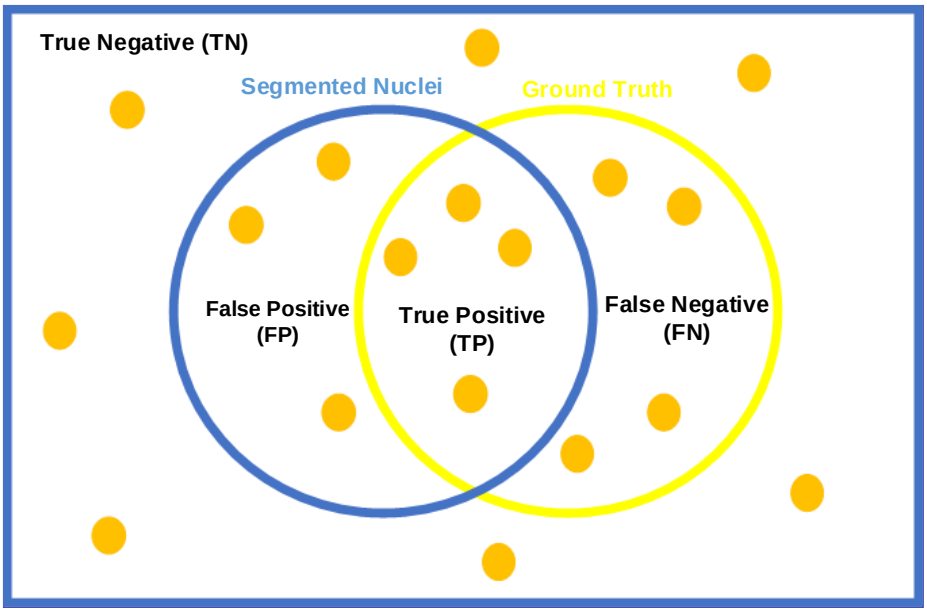

Figure 3.9: The four basic ingredients: TP, FP, TN, and FN for precision and recall measures.

The final measure is then the average of the Jaccard indices of matched pairs.

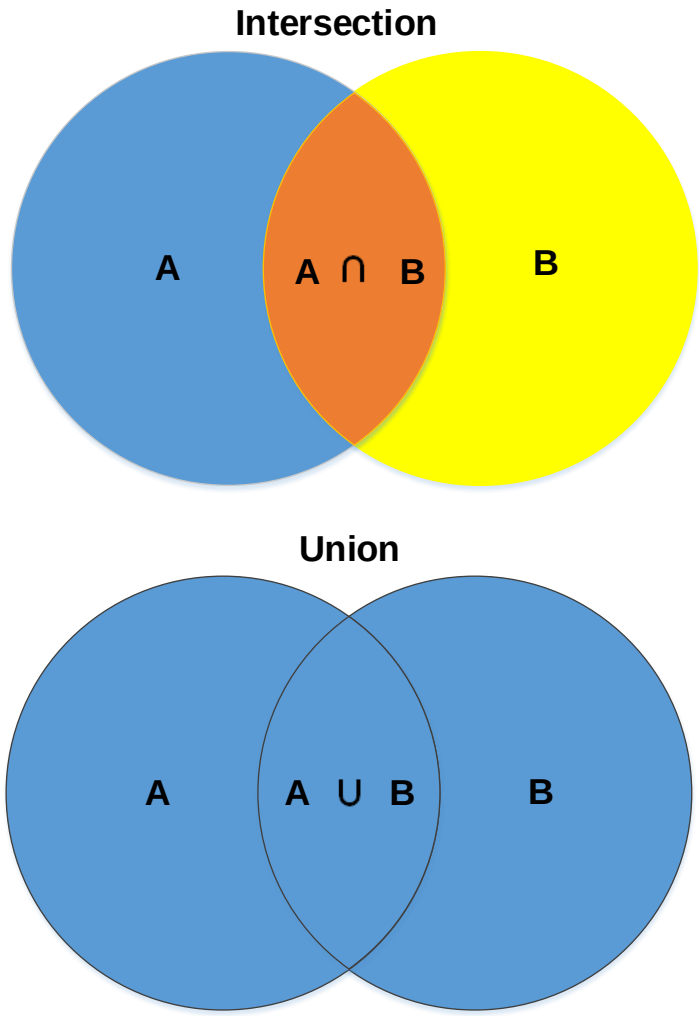

Figure 3.10: A diagram explaining the similarity via Jaccard Index

3.4 Implementation details.

The image analysis framework is developed using MATLAB (R2017b) on a Windows-based computer (Intel Core i7, 3.07 GHz, and 16 GB RAM). Furthermore, the 3D ImageJ viewer plugin [Schmid 2010] along with tools from the 3D ImageJ suite [Ollion 2013] are used for three-dimensional visualisation of the final segmentation result. The source code, as well as datasets, are available upon request. Data processing using the complete framework took 5 mins for the synthetic dataset, 11 mins for the *CE-UPMC* dataset, 35 mins for *Fluo-N3DH-CE* dataset and 48 mins for *Fluo-N3DL-DRO* dataset to process only one time point of 3D image from the complete dataset. Regarding cell tracking challenge, the web site (http://www.codesolorzano.com/celltrackingchallenge) provides access to the datasets with the ground truth. In addition, it provides access to Windows and Linux executable files for the evaluation software as well as an executable program that includes the process description for KTH work.

The next section will discuss in detail the employed datasets and the results achieved by the proposed algorithm.

Chapter 4

Results and Discussion

Contents

Since the main motivation of our work is to automate the detection and segmentation of cell nuclei in time-lapse fluorescence microscopy images, we focused on applying our algorithm to development biology datasets, where only a limited number of existing methods had provided satisfactory results. The proposed framework is extensively tested on three real datasets for embryonic cells

and one dataset of synthetic images with different values for the signal to noise ratio (SNR) and the object size. In addition, we tested our method on another real dataset, which is used to automatically compute the translocation of a transcription factor to the nucleus. SNR is a performance measure for the sensitivity of imaging systems which is defined as the ratio of the average signal level (μ_{signal}) to the standard deviation (σ_{noise}) of the background noise level: $SNR = \mu_{signal}/\sigma_{noise}$ and expressed in logarithmic function as $SNR(dB) = 20 \times \log_{10}(\mu_{signal}/\sigma_{noise})$ [Johnson 2006].

4.1 Datasets description

4.1.1 Synthetic dataset.

In order to measure the robustness of the proposed method, we generated synthetic images of size (XYZ) equal $(100 \times 100 \times 20)$ voxels containing spheres of two radii: 7 and 9 voxels. As it is common in fluorescence microscopy images to have low contrast and low signal to noise ratio (SNR), as a consequence of weak fluorescent staining or microscope properties, the images are distorted with different levels of Poisson-Gaussian noise, resulting in SNRs of 2 dB, -1 dB, -5 dB and -7 dB, respectively. Furthermore, the images include touching spheres where these conditions simulate the same characteristics existing in the real datasets as shown in Figure 4.1.

4.1.2 Real dataset.

The first dataset comes from the work of Gul-Mohammed [Gul-Mohammed 2014a] and it is named as *CE-UPMC*. The other two datasets come from the cell tracking challenge [Ulman 2017, Maška 2014], namely *Fluo-N3DH-CE* and *Fluo-N3DL-DRO*. The last two datasets are proven to be the hardest to be fully segmented automatically [Ulman 2017]. Each dataset from the cell tracking challenge contains 2 sequences. For the cell tracking challenge datasets, all pixels belonging to objects including the centroid are labelled as object by the ground truth. However, for

the other dataset (i.e. *CE-UPMC*), only the centroid of each object is labelled. Another dataset provided by Keiko Kawauchi (Konan University, Japan) is named as *Fluo-MCF7shvec*. The datasets are described as follows:

***CE-UPMC* dataset.**

It involves the *C. elegans* embryonic cells. The size (XYZT) of dataset is $512 \times 512 \times 31 \times 160$. The cells were acquired every 1 minute using a spinning-disk confocal microscopy. This dataset is very challenging, as the intensity of the images is decaying over time due to the labeling technique and acquisition system. Thus, the quality of the acquired images is low.

***Fluo-N3DH-CE* dataset.**

It includes the *C. elegans* embryonic cells. The size (XYZT) of the first sequence is $708 \times 512 \times 35 \times 250$ and of the second sequence is $712 \times 512 \times 31 \times 250$. Both sequences are $8-$bit images with cells imaged every 1.5 minutes. The cells are acquired using a Zeiss LSM 510 Meta Confocal Microscopy. This dataset is challenging as well, since it has a low signal to noise ratio (SNR = 6.74 dB), in addition, the fluorescence can fade when the cells divide. Furthermore, the cells become smaller over time.

***Fluo-N3DL-DRO* dataset.**

It contains the *Drosophila melanogaster* embryonic cells. The size (XYZT) of each sequence is $1272 \times 603 \times 125 \times 49$. Both sequences are $8-$bit images with cells imaged every 30 second. The cells are acquired using a SIMView light-sheet microscopy. This dataset is very challenging as it has a large number of densely packed cells. In addition, it has a low signal to noise (SNR = 2.46 dB).

***Fluo-MCF7shvec* dataset.**

It involves human breast adenocarcinoma cells that are acquired using Confocal Microscopy. This dataset has several image sequences acquired under different

conditions i.e., before and after treatment with the etoposide. The size (XYZ) of dataset is $512 \times 512 \times Z$ where Z varies between each sequence. Each image sequence includes two channels. The first channel identifies the cell nuclei that are stained with DAPI marker. The second channel identifies the protein of interest (i.e., p53) that is stained with mCherry marker.

4.2 Experimental setup and suitable parameters selection.

Synthetic datasets are generated to study the effect of parameters (described in Table 4.1) on cell nuclei detection and segmentation, as well as to understand the overall mechanism for selecting and tuning the significant parameters of various datasets (summarized in Table 4.3).

Our approach is based on the building of a dictionary (small patches of the image) that will be eventually used for denoising and detection of cell nuclei. We investigated the optimal size of the patches and the number of patches (called atoms) in the dictionary. Then, the randomly created dictionary is updated, and so we investigated the number of iteration for the update. Finally, we investigated the sparsity level (i.e., the number of used atoms) for the reconstruction of the denoised image and the detection map. To start with, we tested several values for patch size ($p = 5 \times 5 \times 5$, $10 \times 10 \times 5$, $15 \times 15 \times 5$ and $20 \times 20 \times 5$), dictionary size ($K = 64$, 128, 256 and 512), sparsity level ($L = 3$, 6 and 9) and number of iterations ($N = 5$, 10, 15, 20 , 25 and 30) at different noise levels (SNR $= 2$, -1, -5 and -7 dB) as shown in Figure 4.1. Since the coefficient of variation (CV) is a useful statistical descriptor for comparing the degree of variation from one data series to another one. The CV is defined as the ratio of the standard deviation to the mean [Kesteven 1946]. Thus, we employed the CV to measure the effect of changing the parameters on the result. The average CV from the patch size, dictionary size, sparsity level and the number of iterations over the four noise levels are approximately 15, 2, 2 and 2 % respectively.

In all aforementioned parameters, we observed that the patch size is considered

as a critical parameter where the change in this value has a major impact on the subsequent segmentation results as shown in Figure 4.2 (a). On the contrary, changes in the other parameters i.e. dictionary size, sparsity level and the number of iterations often achieve very close results as shown in Figure 4.2 (b), (c) and (d). Therefore, we fixed all parameters while tuned the patch size according to the object's size present in the images.

In order to confirm the importance of patch size tuning, we conducted more analysis in the term of cell nuclei detection as shown in Figure 4.3. For example, at the first three noise levels 2, -1 and -5 dB all cell nuclei are correctly detected for different patch size values. However, at noise level equal -7 dB, many objects are falsely detected with the patch size equivalent to $p = 5 \times 5 \times 5$ and one nucleus is not detected at patch size equal $20 \times 20 \times 5$. Though, for patch size equal $10 \times 10 \times 5$ and $15 \times 15 \times 5$, all cell nuclei are correctly detected.

For all previously mentioned patch size values, the average of recall, precision, F-measure, and Jaccard index with different noise levels are presented in Figure 4.3 (b), where these measures are high when the patch size values equal $p = 10 \times 10 \times 5$ and $15 \times 15 \times 5$ compared to the measures of the other two values .

Following the above experiment, we observed that, for robust detection and segmentation results, the patch size should not be less than 25% and not more than 100 % of the average cell nuclei volume in images.

Table 4.1: Description of denoising and segmentation parameters

Parameters	Description
Denoising	
Patch size [N N M]	The patch is a small region of an image with size ([N N M]). Patches are extracted by moving a window with a step size of one pixel over the raw image.
Dictionary size (K)	The dictionary is constructed by concatenating the patches to vectors (called atoms). Dictionary size (K) is the total number of atoms.
Number of iterations (N)	A specified number of times to update the dictionary.
Sparsity level (L)	The number of nonzero elements (used atoms from the dictionary) for the sparse representation coefficient
Segmentation	
SensitivityFactor	A scalar value within a range from zero to one. It controls sensitivity towards thresholding more voxels as foreground.
NucleiSeedDilation	Radius of the structuring element for morphological dilation.
MinNucleiVolume	The approximate volume of the smallest cell nucleus in the image .

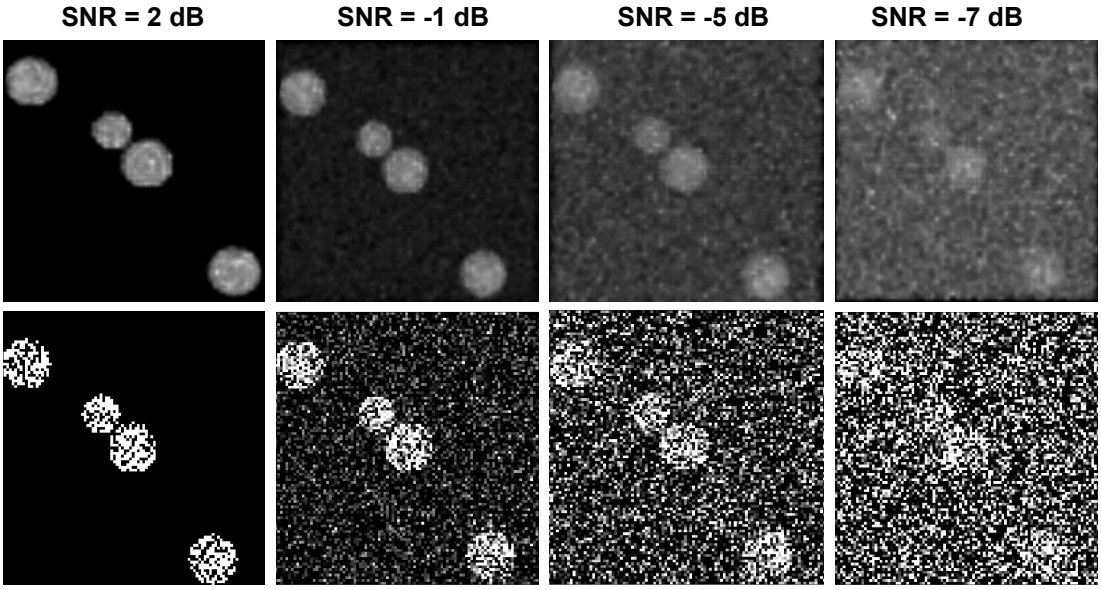

Figure 4.1: Synthetic images with different levels of signal to noise ratio (SNR). Top row: 3D view of the synthetic dataset. Bottom row: single plane (Z = 10) from the synthetic dataset.

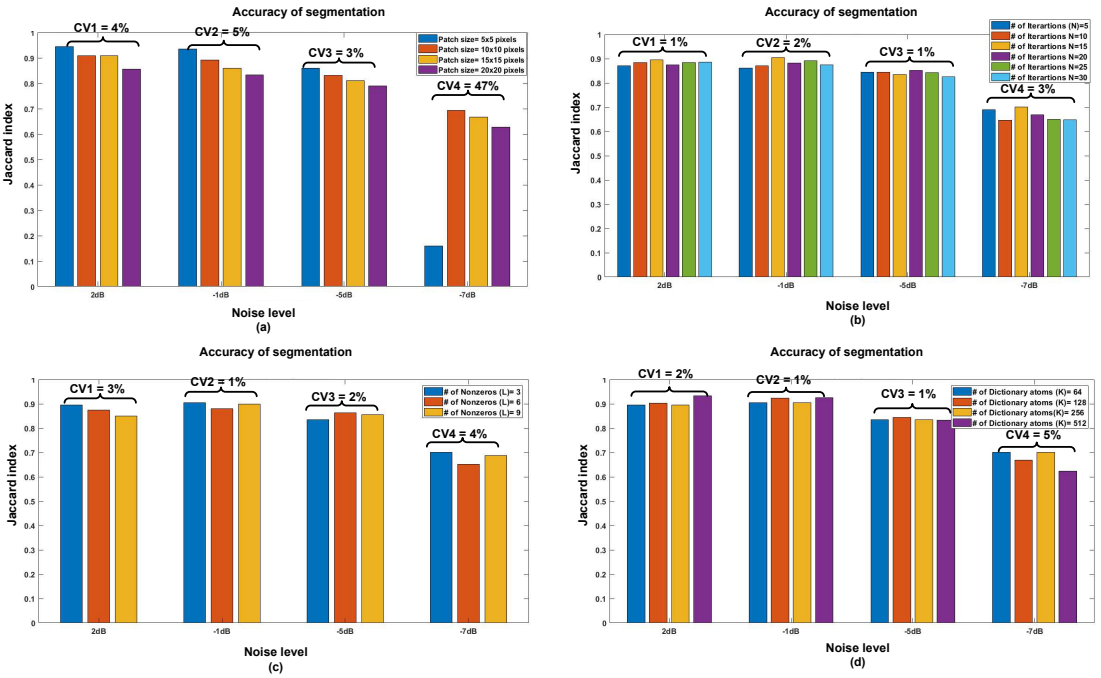

Figure 4.2: Evaluation of the segmentation accuracy with different initialization parameters at different noise levels.

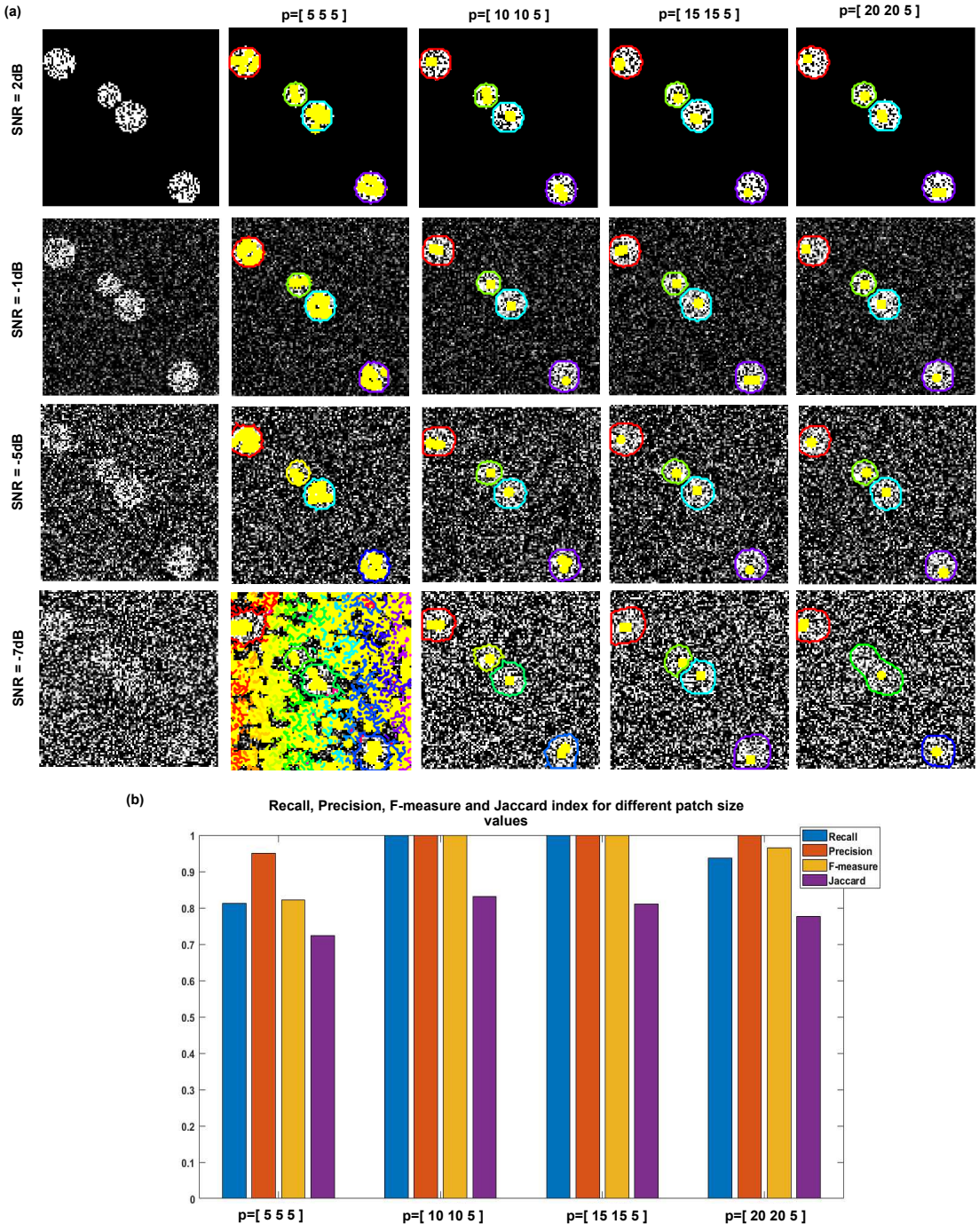

Figure 4.3: Evaluation of patch size values for detection and segmentation results at different noise levels. (a) The results of detection (depicted by yellow dots) and segmentation (delineated by colored contours) overlaid on single plane (Z = 10) from synthetic images for different patch size values p = $5 \times 5 \times 5$, $10 \times 10 \times 5$, $15 \times 15 \times 5$ and $20 \times 20 \times 5$ at various noise levels (SNR = 2, -1, -5 and -7 dB). (b) Average Recall, Precision, F-measure, and Jaccard index values of detection and segmentation results at different noise level as a function of patch size.

4.3 Results of denoising 3D cell nuclei images.

In this thesis, a sparse representation model [Aharon 2006, Elad 2006] is employed to obtain the denoised images. Our method is compared with PURE-LET [Luisier 2010b], which is one of the most efficient, fast and automatic methods for denoising of multi-dimensional fluorescence microscopy images. The main motivation behind the need for cell nuclei denoising is assisting better segmentation of cell nuclei images. Therefore, the comparison between the denoising methods is performed in the context of improving segmentation results.

For instance, the results in Figure 4.4 (first row), Figure 4.5 (a, c) and Figure 4.6 (a, c) show that our method is able to reduce, and almost remove the noise as well as enhance the contrast of cell nuclei. We have also noticed a better contrast than PURE-LET results as shown in Figure 4.4 (third row), Figure 4.5 (b, d) and Figure 4.6 (b, d).

For further assessment, thresholding-based approach is applied to the denoised images to obtain the segmentation mask. It can be noted from Figure 4.4 (second row), Figure 4.5 (e) and Figure 4.6 (e) that our method succeeded to segment all nuclei in comparison with the other method which failed to detect some nuclei as demonstrated in Figure 4.6 (f). Even though, when the PURE-LET method is able to detect all cell nuclei shown in Figure 4.5 (f), the size of segmented nuclei are smaller than their original size. Unfortunately, this method can not detect any cell nuclei at very low signal to noise ratios (-5 dB and -7 dB,) as presented in Figure 4.4 (fourth row)

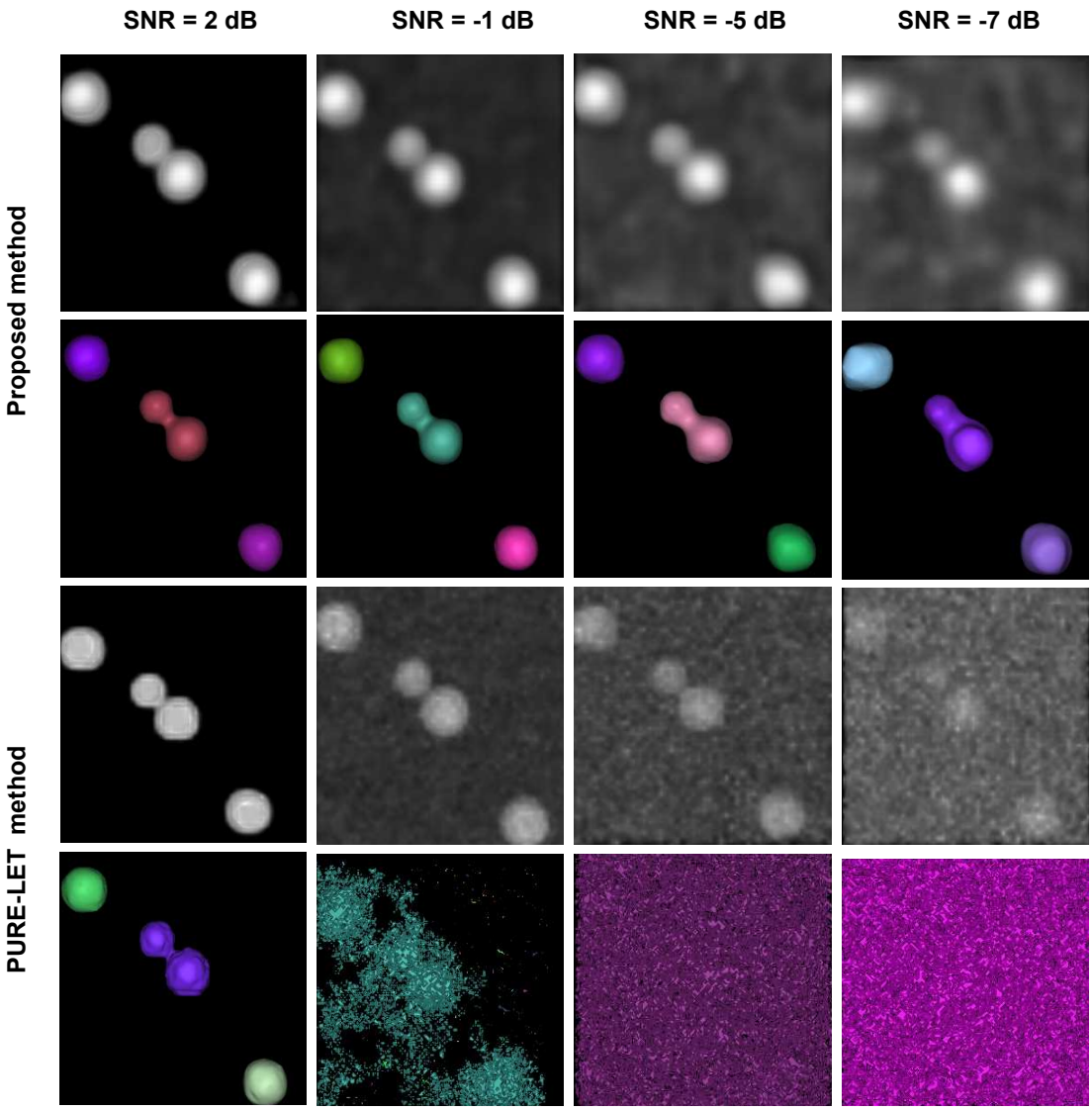

Figure 4.4: A comparison of denoising results on the synthetic dataset (Figure 4.1) using our method and PURE-LET method [Luisier 2010b] at different noise levels. First row: 3D view of the denoised images from the proposed method. Second row: 3D view of the segmentation mask of the denoised images from the proposed method. Third row: 3D view of the denoised images from the PURE-LET method. Fourth row: 3D view of the segmentation mask of the denoised images from the PURE-LET method .

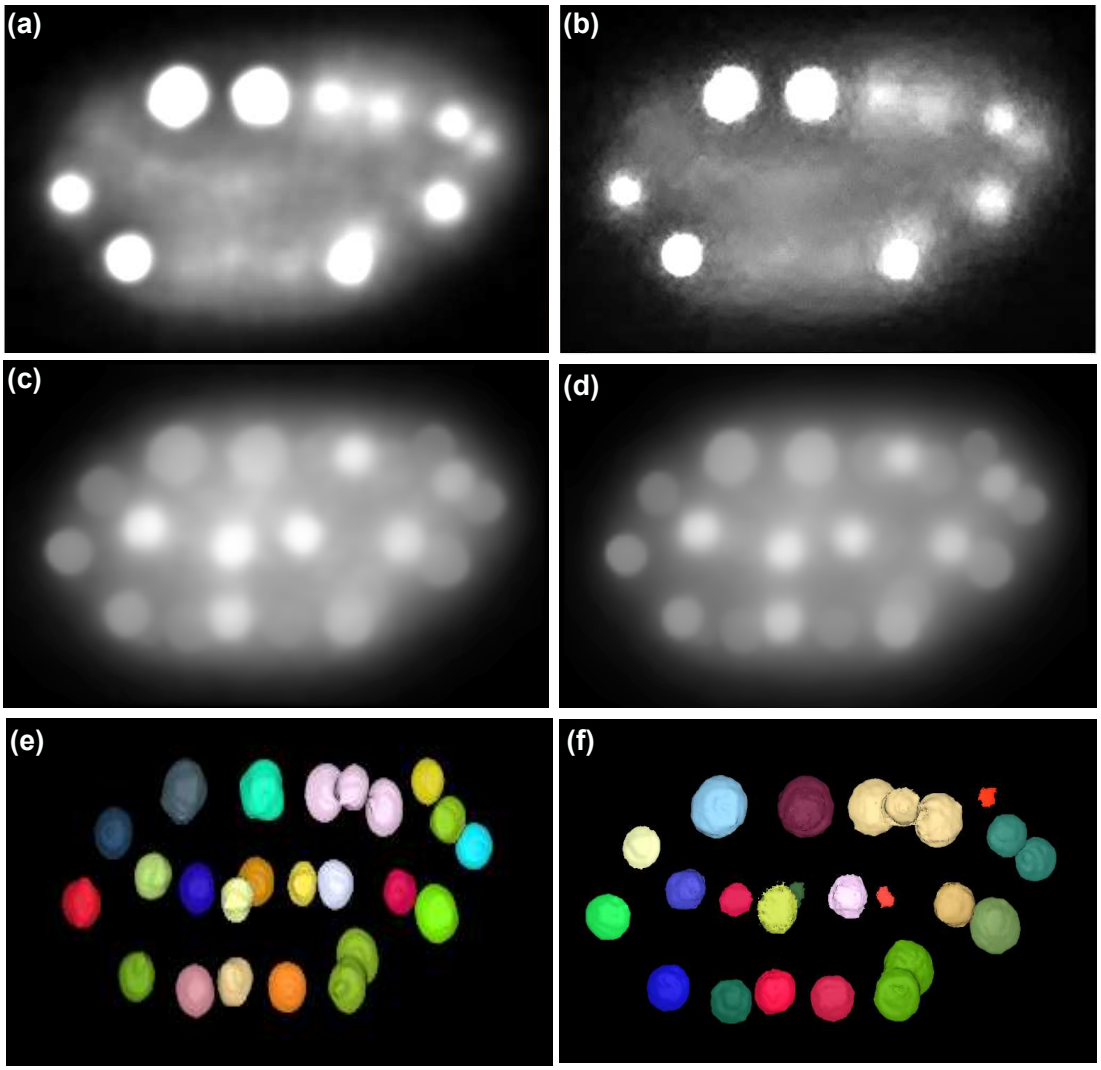

Figure 4.5: A comparison of denoising results on the *CE-UPMC* dataset using our method and PURE-LET method [Luisier 2010b] for the same time points as Figure 4.11 (a). First column: shows the results of the proposed method. Second column: shows the results of the PURE-LET method. (a, b) a single plane (Z = 15) of time point (T = 60) for the denoised images. (c, d) 3D view of the denoised images. (e, f) 3D view of the segmentation mask of the denoised images, colours shown are for illustration purpose only, they are not the final segmentation results.

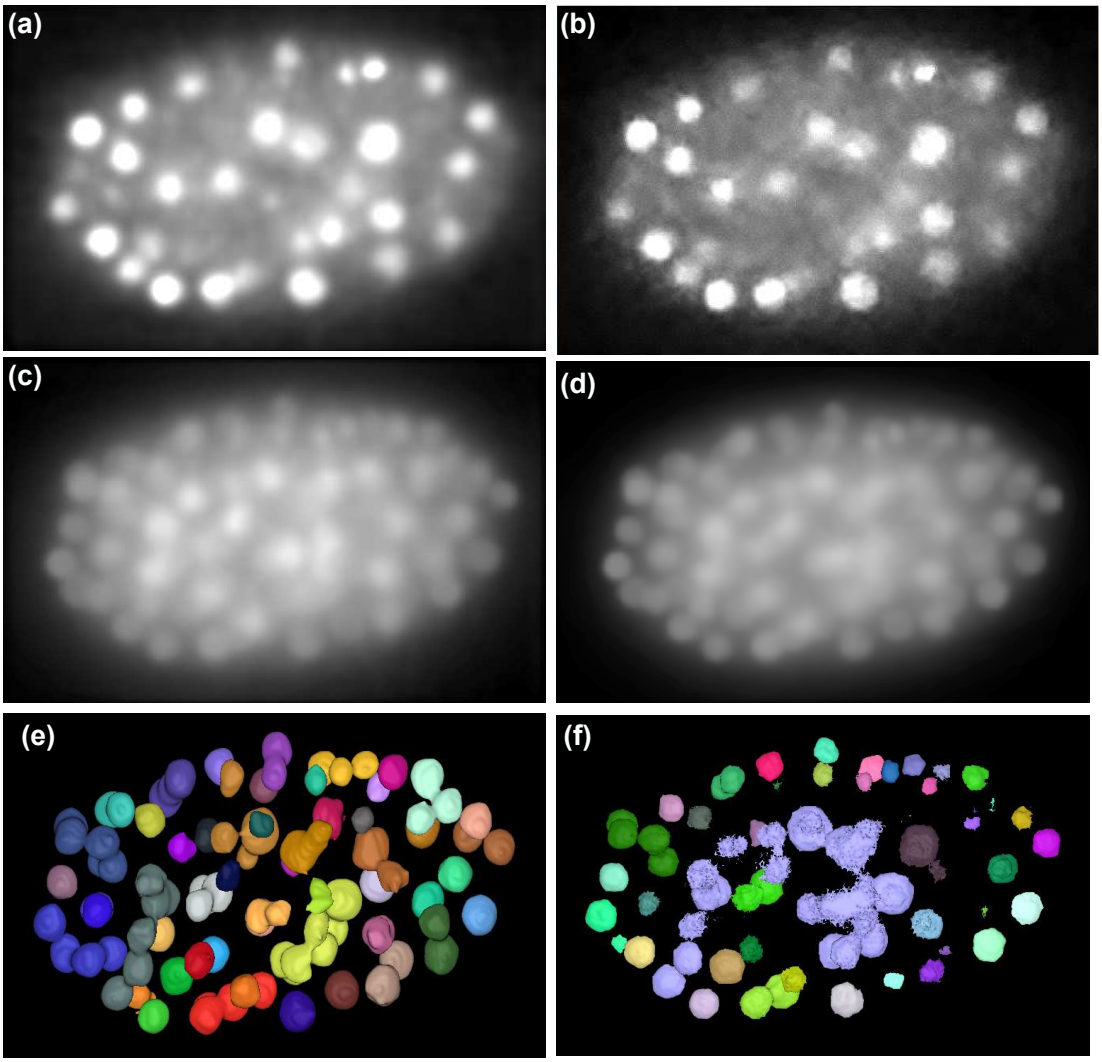

Figure 4.6: A visual comparison of denoising results on the *CE-UPMC* dataset using our method and PURE-LET method [Luisier 2010b] for the same time points as Figure 4.11 (b). First column: shows the results of the proposed method. Second column: shows the results of the PURE-LET method. (a, b) a single plane (Z = 15) of time point (T = 140) for the denoised images. (c, d) 3D view of the denoised images. (e, f) 3D view of the segmentation mask of the denoised images, colours shown are for illustration purpose only, they are not the final segmentation results.

4.4 Results of segmentation 3D cell nuclei images.

Following the denoising step, a local adaptive-thresholding approch [Bradley 2007] is applied to the denoised image to get the segmentation mask of candidates regions. In order to obtain the candidates locations of cell nuclei centres, we used a novel representation called the detection map. Each voxel in this map is computed as the summation of the patch coefficients that are used to reconstruct the denoised image. We then define a maximum response image by multiplying the denoised image with the detection map. This maximum response image is used to detect the local maxima (Figure 4.7). Afterwards, the obtained local maxima are used as an input for a 3D marker-controlled watershed segmentation of the cell nuclei (Figure 4.8).

For the synthetic dataset, great performance is observed on very low signal to noise ratios (2 dB, -1 dB, -5 dB and -7 dB), in which our method is capable of correctly identifying and segmenting all cell nuclei at the various noise levels as presented in Figure 4.9. Furthermore, our method has similar performance as the top-ranked KTH algorithm from cell tracking challenge, [Ulman 2017, Maška 2014] as shown in Figure 4.10. However, at very low SNR (i.e., -7 dB) KTH algorithm is not able to separate touching nuclei.

Regarding the *CE-UPMC* dataset, there is an intensity decay over time owing to the labeling technique and acquisition system. As a result, the acquired image quality is low. Table 4.4 and Figure 4.11, illustrate the results obtained at certain time points. For instance, at early time points (40, 60, 80 and 100) all cell nuclei are correctly detected. In addition, few false positives are also detected (2 objects out of 119 cell nuclei). Even though the image quality at advanced time points (120, 140 and 160) is low, only 11 cell nuclei out of 247 are not detected, also there exists a small number of false positives (3 objects out of 247 cell nuclei which is displayed as yellow arrows in Figure 4.11 (f)). Typically, the reason for missing cell nuclei is the detection of clustered cell nuclei (indicated by red arrows in Figure 4.11 (f)) rather than detecting them separately. Figure 4.12

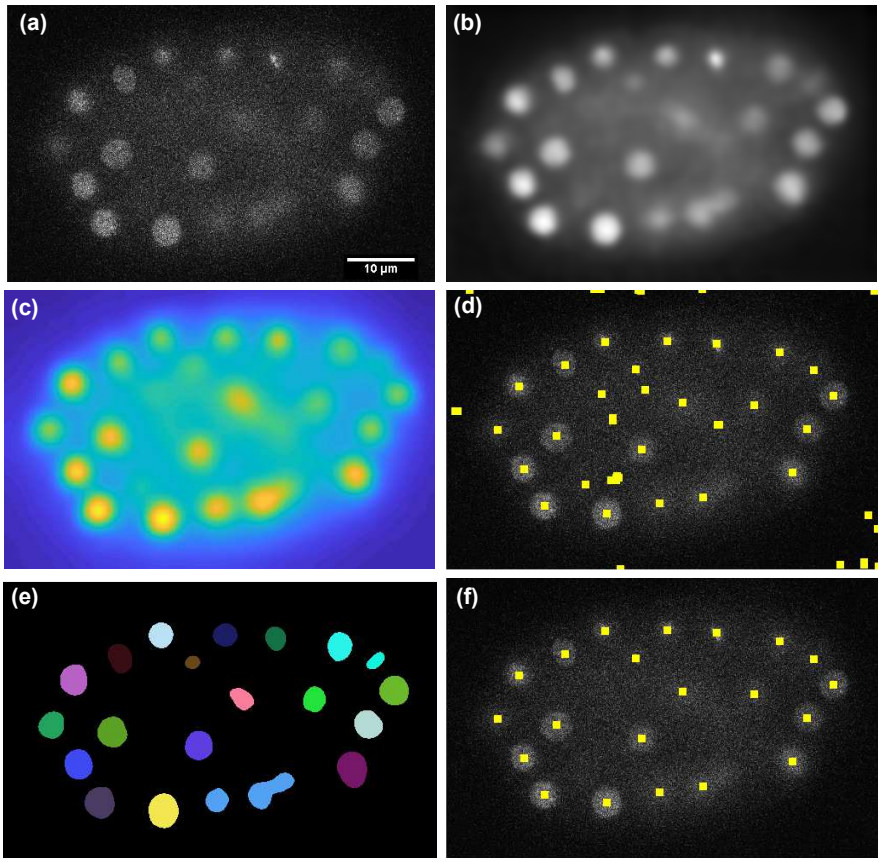

Figure 4.7: Denoising and nuclei detection with the sparse representation model. (a) A single plane (Z = 15) of time point (T = 100) from the *CE-UPMC* dataset. (b) The denoised image obtained by applying the sparse representation model to the image in (a). (c) The detection map obtained from the sparse representation model for image in (a). (d) Marker points detected by applying the local maxima search on the maximum response image, obtained from multiplying image (b) with image (c). Marker points displayed as yellow squares are overlaid on the raw image. (e) Segmentation mask obtained by applying the initial segmentation to the image in (b). (f) Objects detected in the background are discarded by multiplying the detected marker points image (d) with the segmentation mask (e). Note that, the marker point detection here is performed in two dimensions for the purpose of explanation and visualisation, however, in the framework it is applied in three dimensions.

shows the segmentation results of our method and the results from the original paper [Gul-Mohammed 2014a] for the *CE-UPMC* dataset.

4.4.1 Comparison of nuclei segmentation result with the top-ranked approach from ISBI cell tracking challenge.

In the *Fluo-N3DH-CE* dataset, the proposed approach is able to identify and segment correctly more than 96% of total cell nuclei. Furthermore, it detects a small number of false positives (9 objects out of 876 cell nuclei) as well as a small number of false negative (29 cell nuclei). The achieved F-measure of approximately 97.8%, which is comparable to the competing algorithm, i.e. KTH [Ulman 2017, Maška 2014] (Table 4.5, Figure 4.13 and Figure 4.14).

For the *Fluo-N3DL-DRO* dataset, despite our method succeeds to detect more cell nuclei (99%recall), it has low precision (3%) due to the annotated ground truth, which considered only the cell nuclei located in the early nervous system and all other nuclei are deemed as false positives. As a result, the obtained F-measure is low with an approximate value of 6.4% (Table 4.5). Furthermore, our method has a comparable segmentation accuracy with KTH competing approach [Ulman 2017, Maška 2014] (Table 4.5, Figure 4.15 and Figure 4.16).

The results achieved by our method for the two datasets obtained from cell tracking challenge are compared with the top-ranked KTH algorithm [Ulman 2017, Maška 2014]. KTH algorithm is chosen for the reason that, it presented the best overall performance in the challenge. This algorithm is mainly based on adopting the band-pass filter to detect and segment cell nuclei.

Regarding the KTH algorithm, some detected objects are actually noise and some cell nuclei are not detected. This is because the algorithm detected clustered cell nuclei instead of detecting them separately. For example, at time point 28 from *Fluo-N3DH-CE* (seq1) and at time point 106 from *Fluo-N3DH-CE* (seq2), KTH algorithm failed to resolve the fusion of two nuclei (as presented by the red arrows in Figure 4.14 (b). In contrast, our method succeeds to identify and segment each nucleus individually as shown in Figure 4.14 (a) and (c). For the

Fluo-N3DL-DRO dataset, although our method succeeds to detect more cell nuclei than KTH approach (Table 4.5 and Figure 4.16), the evaluation method considered those cell nuclei as false positives, due to the annotation method which considered only the cell nuclei located in the early nervous system.

We have found that, the proposed method is less sensitive to some parameters such as dictionary size (K), sparsity level(L) and number of iterations (N). All these parameters are being fixed for different datasets and experiments with the subsequent values K = 64, L = 3 and N = 15. However, the proposed method is more sensitive to fundamental parameters, i.e., such as patch-size, and in a less critical manner to $SensitivityFactor$ and $NucleiSeedDilation$. As these parameters are easy to understand, this makes them easier to tune-up if needed. We need to stress that all parameters, except patch size, are quite robust, as we only need to use three sets of parameters for all datasets. The set of empirically determined parameter values being applied to the datasets are listed in the Table 4.3.

Concerning the *Fluo-N3DH-CE* dataset, although the average cell nuclei size in sequence (2) is slightly greater than the average size in sequence (1), we have decided to use the same parameter (i.e., patch size) for denoising of both sequences. As a result, we tuned $NucleiSeedDilation$ to avoid detection of multiple local maxima for the same object as explained in section **Marker points detection**.

We have also presented a Table 4.2 to show the detection and segmentation results among various datasets considering the patch size percentage (related to average cell nuclei volume).

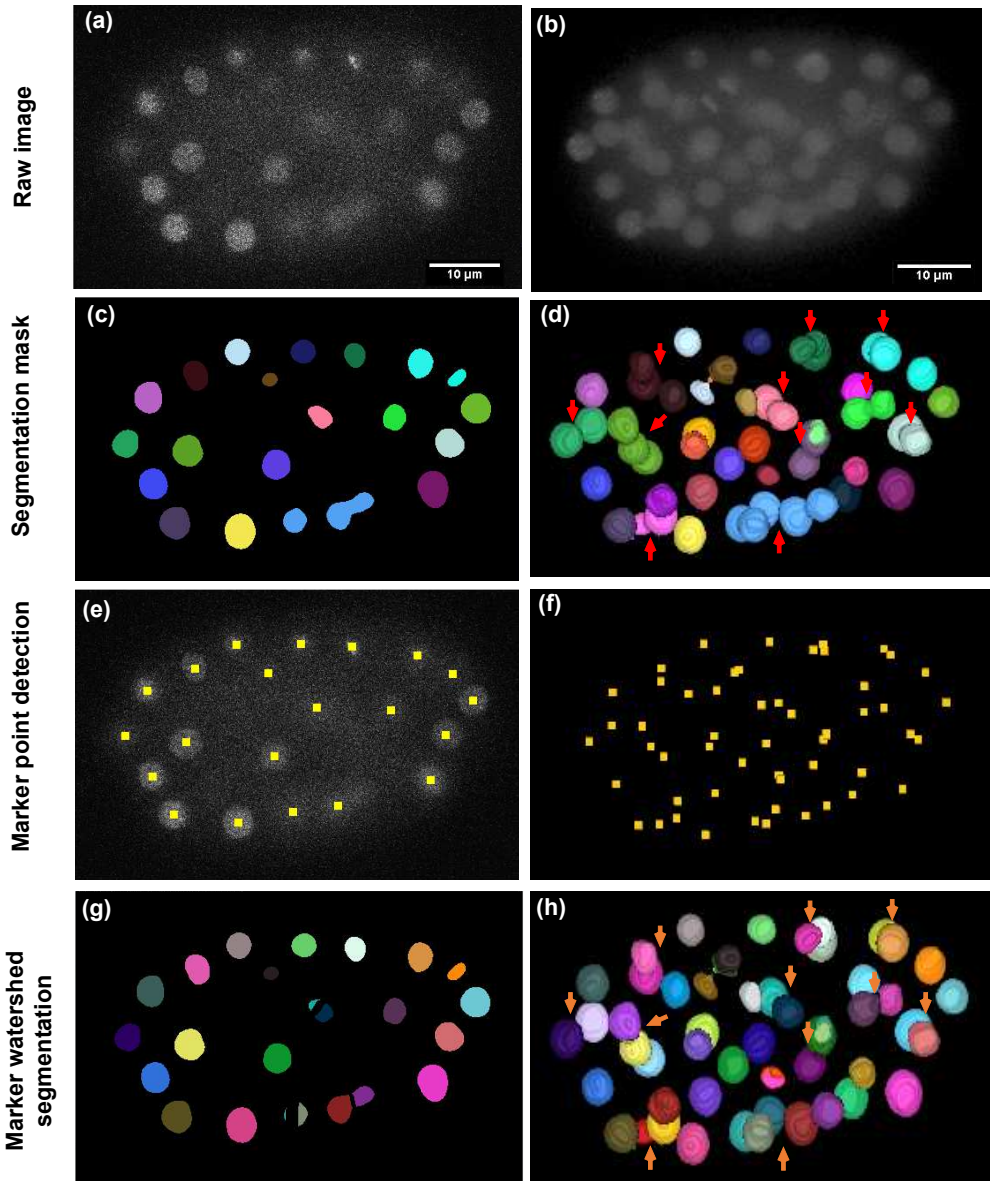

Figure 4.8: An overview of cell nuclei segmentation steps. First column: shows a single plane (Z = 15) of time point (T = 100) from the *CE-UPMC* dataset. Second column: shows a three-dimensional view of the same time point. (a, b) The raw images. (c, d) The segmentation mask, which identifies the cell nuclei (presented as coloured) in the image, but fails to separate apparently touching cell nuclei (shown as red arrows). (e, f) Marker points (indicated by yellow squares) are obtained from the sparse representation model. (g, h) Marker-controller watershed segmentation that succeeds to separate apparently touching cell nuclei (orange arrows). Note that, different colours represent individual components. The marker points detection at (e) is performed in two dimensions for the illustration process. However, in the framework it is applied in three dimensions.

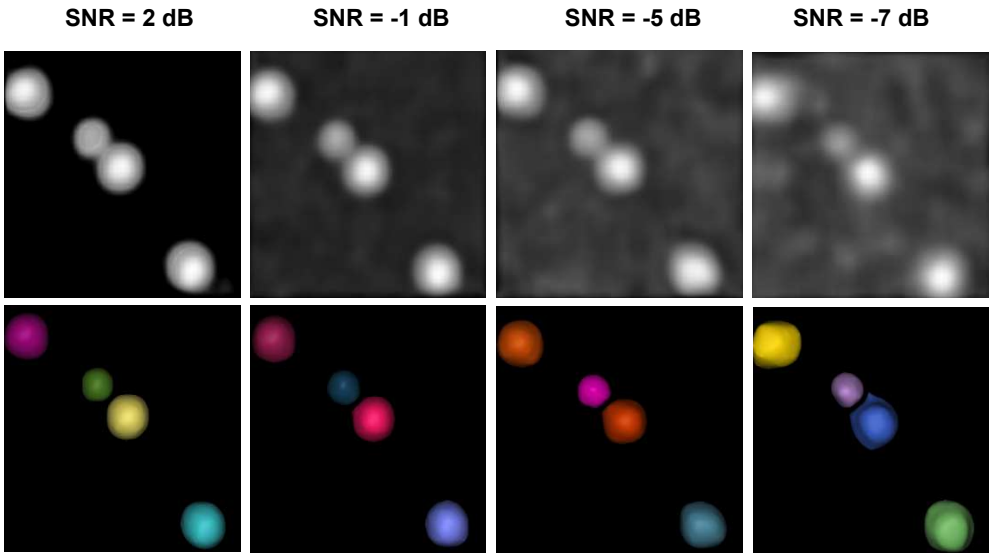

Figure 4.9: Example of denoising and segmentation results on the synthetic dataset (Figure 4.1) at different noise levels. First row: 3D view of denoising result. Second row: 3D view of the segmentation result.

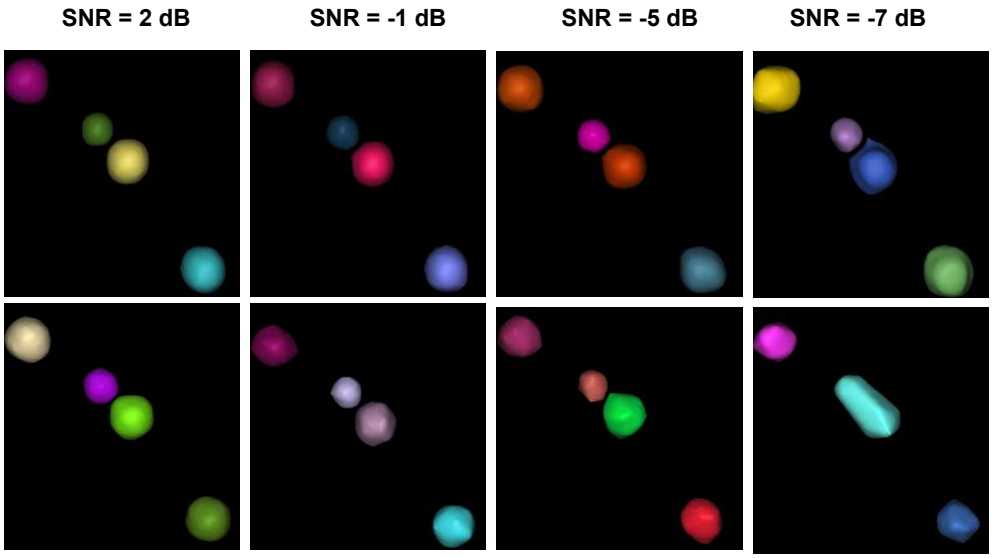

Figure 4.10: A comparison of segmentation results on the synthetic dataset (Figure 4.1) using our method and KTH method [Ulman 2017, Maška 2014] at different noise levels. First row: 3D view of the segmented images from the proposed method. Second row: 3D view of the segmented images from the KTH method.

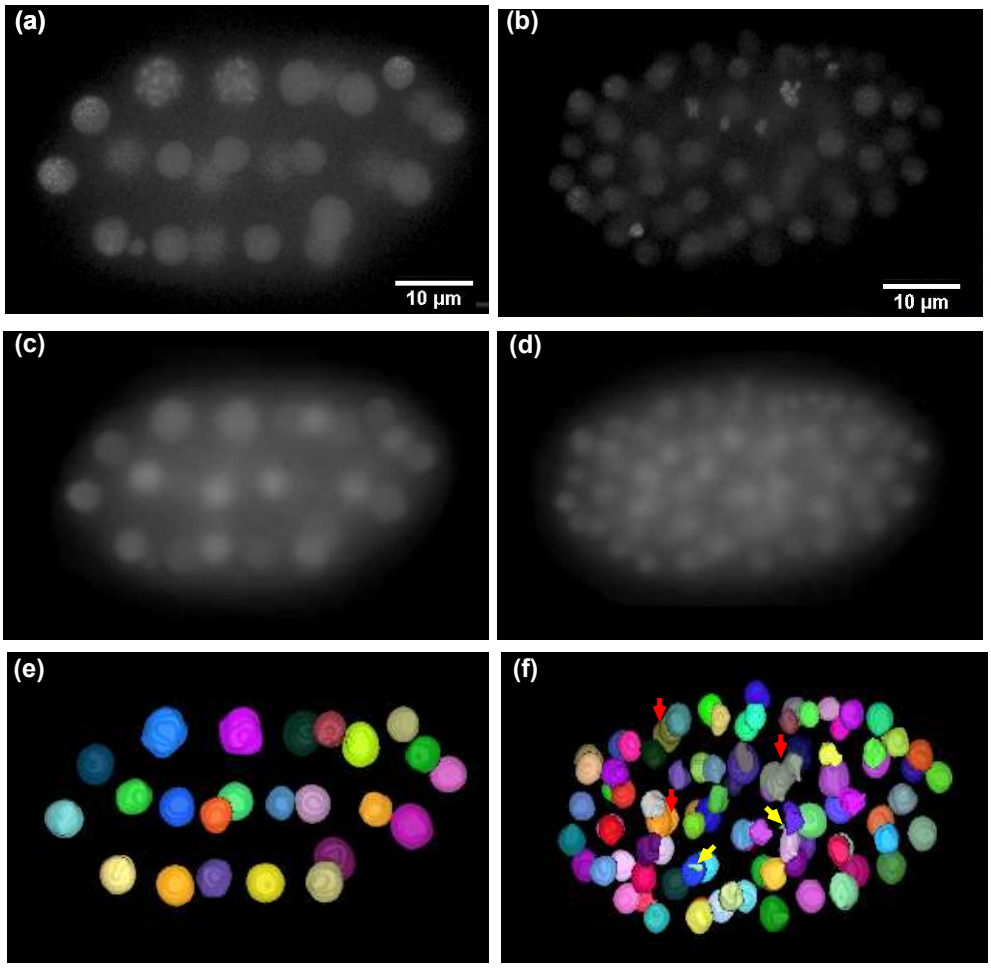

Figure 4.11: Example of denoising and segmentation results on the ***CE-UPMC*** dataset. (a, b) 3D view of the raw data for time points (T = 60 and 140) respectively. (c) 3D view of the denoising result for (a). (d) 3D view of the denoising result for (b). (e) 3D view of the segmentation result for (c). (f) 3D view of the segmentation result for (d). Note that, yellow arrows indicate noisy objects and red arrows indicate merged cell nuclei.

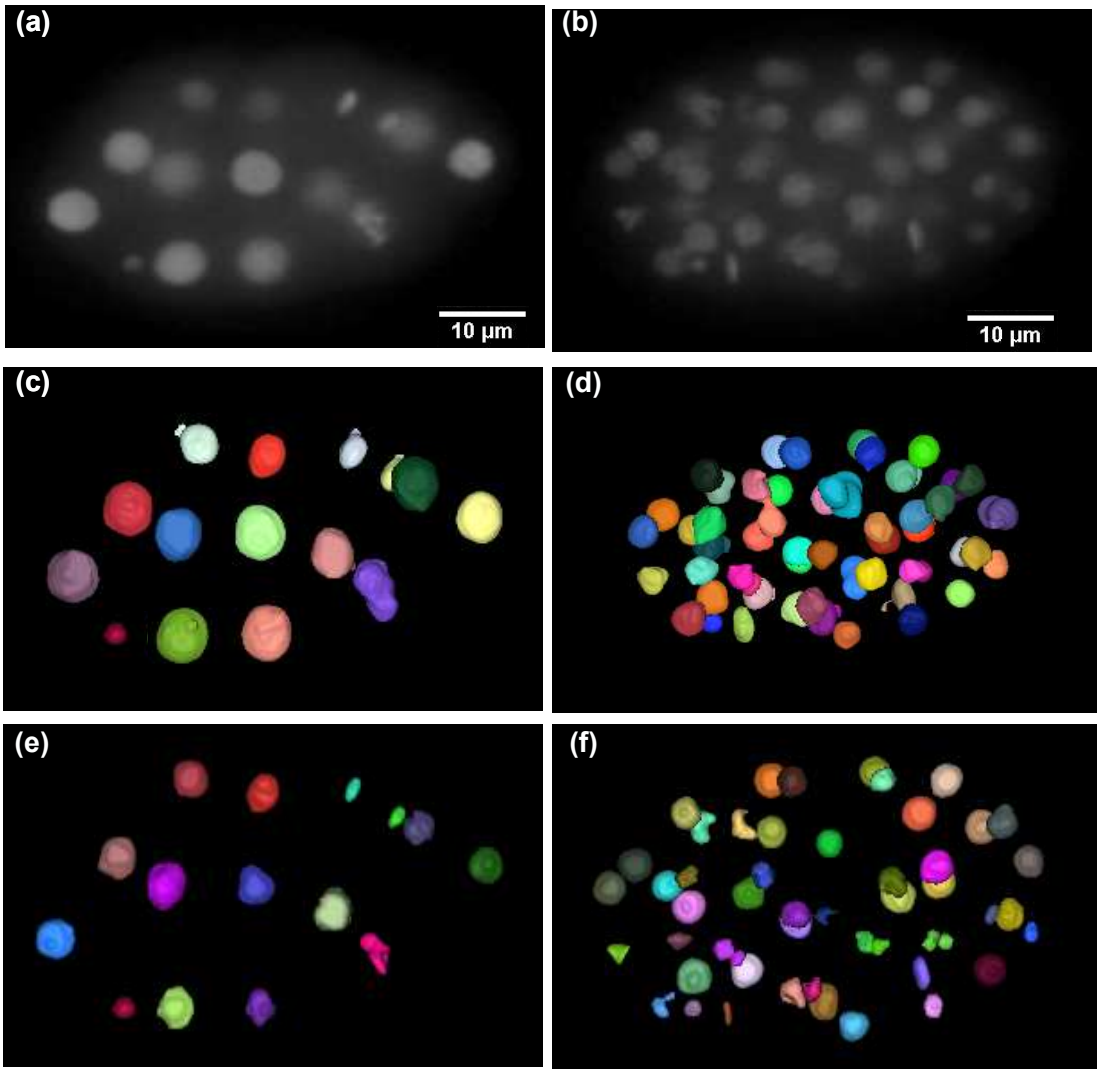

Figure 4.12: An example of the segmentation results for our method and the results from the original paper [Gul-Mohammed 2014a] of ***CE-UPMC*** dataset at time points T = 40 and 120 respectively. First column: shows results of time point (T = 40). Second column: represents results of time point (T = 120). (a, b) 3D view of the raw data. (c, d) 3D view of our segmentation result. (e, f) 3D view of the segmentation result from the original paper [Gul-Mohammed 2014a] of ***CE-UPMC*** dataset.

Table 4.2: Segmentation performance of our method (SRS) among various datasets considering the patch size percentage.

	CE-UPMC	***Fluo-N3DH-CE* (seq1)**	***Fluo-N3DH-CE* (seq2)**	***Fluo-N3DL-DRO***
Patch	[20 20 5]	[25 25 5]	[25 25 5]	[10 10 5]
Approximate average cell nuclei (Voxels)	7000	10,000	12,000	1,300
Percentage of patch size (%)	28	31	26	38
Recall (%)	97.9757	96.03	97.49	99.37
Precision (%)	98.7743	98.9	98.98	3.32
F-measure (%)	98.3476	97.44	98.23	6.43
Jaccard index (%)	-	66	70	66.5

Table 4.3: Denoising and segmentation parameters. When the values of parameters differ between the first and the advanced time points, the value for the advanced time points is given in round brackets.

	Datasets				
Parameters	**Synthetic**	***CE-UPMC***	***Fluo-N3DH-CE* *(seq1)***	***Fluo-N3DH-CE* *(seq2)***	***Fluo-N3DL-DRO* *(seq1 & seq2)***
Patch size (voxels)	[15 15 5]	[20 20 5]	[25 25 5]	[25 25 5]	[10 10 5]
Threshold	global Otsu's [Otsu 1979]	local adaptive [Bradley 2007]	local adaptive [Bradley 2007]	local adaptive [Bradley 2007]	local adaptive [Bradley 2007]
SensitivityFactor	-	0.56 (0.58)	0.5	0.5	0.5
MinNucleiVolume (voxels)	20	5000 (1000)	10,000 (3000)	10,000 (3000)	-
NucleiSeedDilation (voxels)	3	5	10 (5)	20 (5)	5

Table 4.4: **Segmentation performance of our method (SRS) for *CE-UPMC* dataset.**

Time	GT	SN	TP	FN	FP	Recall (%)	Precision (%)	F-measure (%)
40	14	14	14	0	0	100	100	100
60	24	24	24	0	0	100	100	100
80	28	29	28	0	1	100	96.55	98.2447
100	51	52	51	0	1	100	98.08	99.03
120	57	53	53	4	0	92.98	100	96.3623
140	91	93	88	3	2	96.70	97.78	97.2370
160	104	101	100	4	1	96.15	99.01	97.5590

Table 4.5: **Segmentation performance of our method (SRS) and the KTH algorithm [Ulman 2017, Maška 2014], for datasets from cell tracking challenge. The values shown in bold represent the highest performance. GT, number of cell nuclei in ground truth; SN, number of cell nuclei determined by the segmentation; TP, true positives; FN, false negatives; FP, false positives.**

Dataset	Algorithm	GT	SN	TP	FN	FP	Recall	Precision	F-measure	Jaccard
***Fluo-N3DH-CE*_01**	**SRS**	478	464	459	19	5	96.03	98.9	97.44	66
***Fluo-N3DH-CE*_02**		398	392	388	10	4	97.49	98.98	98.23	70
Total		876	856	847	29	9	**96.69**	**98.94**	**97.84**	**68**
***Fluo-N3DH-CE*_01**	**KTH**	478	463	459	19	4	96.03	99.1	97.54	64
***Fluo-N3DH-CE*_02**		398	394	386	12	8	96.98	97.97	97.47	59
Total		876	857	845	31	12	96.5	98.6	97.55	61.5
***Fluo-N3DL-DRO*_01**	**SRS**	3792	122193	3757	35	118436	99.08	3.07	6	60
***Fluo-N3DL-DRO*_02**		4097	120494	4082	15	116412	99.63	3.39	6.56	73
Total		7889	242687	7839	50	234848	**99.37**	3.32	6.43	66.5
***Fluo-N3DL-DRO*_01**	**KTH**	3792	112535	3733	59	108802	98.44	3.32	6.4	62
***Fluo-N3DL-DRO*_02**		4097	104655	4037	60	100618	98.54	3.86	7.4	78
Total		7889	217190	7770	119	209420	98.49	**3.58**	**6.9**	**70**

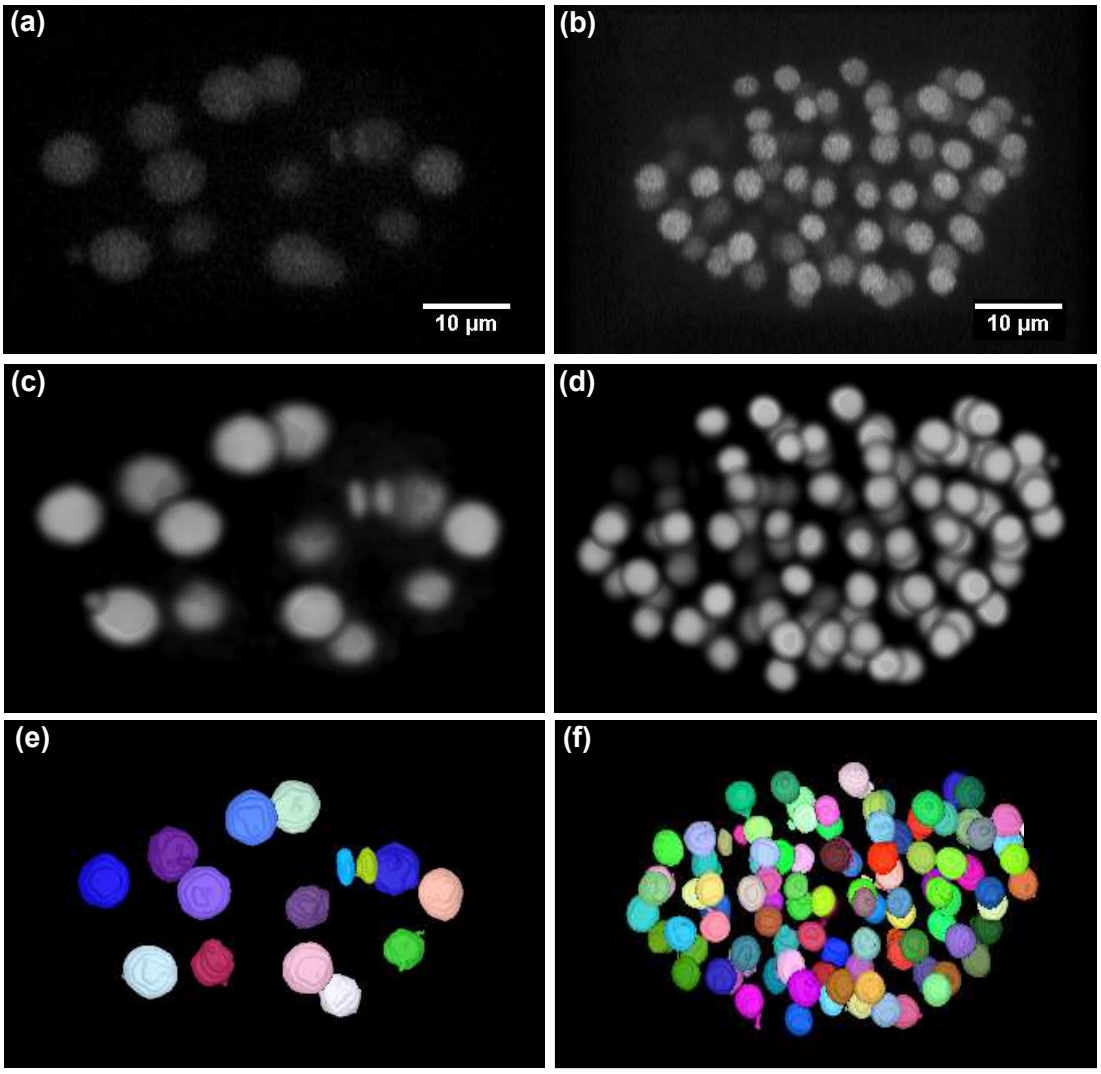

Figure 4.13: Example of denoising and segmentation results on the ***Fluo-N3DH-CE*** dataset. (a) 3D view of the raw data for time point (T = 28) from sequence (1). (b) 3D view of the raw data for time point (T = 106) from sequence (2). (c) 3D view of denoising result for (a). (d) 3D view of denoising result for (b). (e) 3D view of the segmentation result for (c). (f) 3D view of the segmentation result for (d).

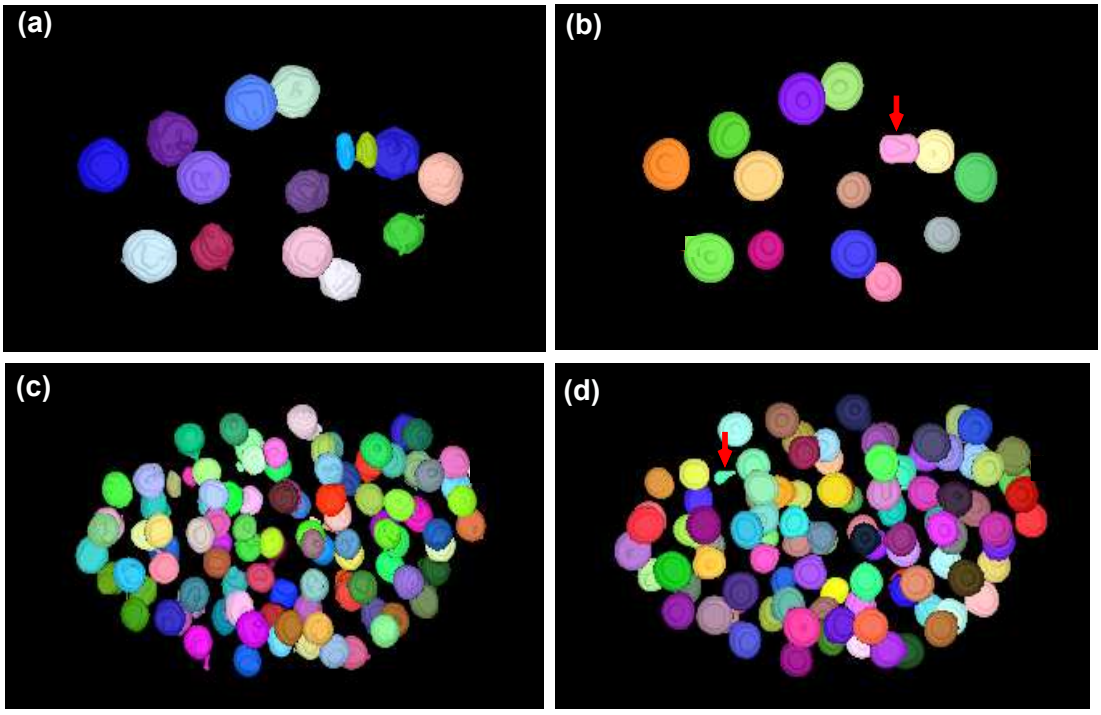

Figure 4.14: A visual comparison of segmentation results over ***Fluo-N3DH-CE*** dataset using our method and KTH method [Ulman 2017, Maška 2014] for the same time points as Figure 4.13). First row: shows results of time point (T = 28). Second row: represents results of time point (T = 106). (a, c) 3D view of our segmentation result. (b, d) 3D view of KTH segmentation result.

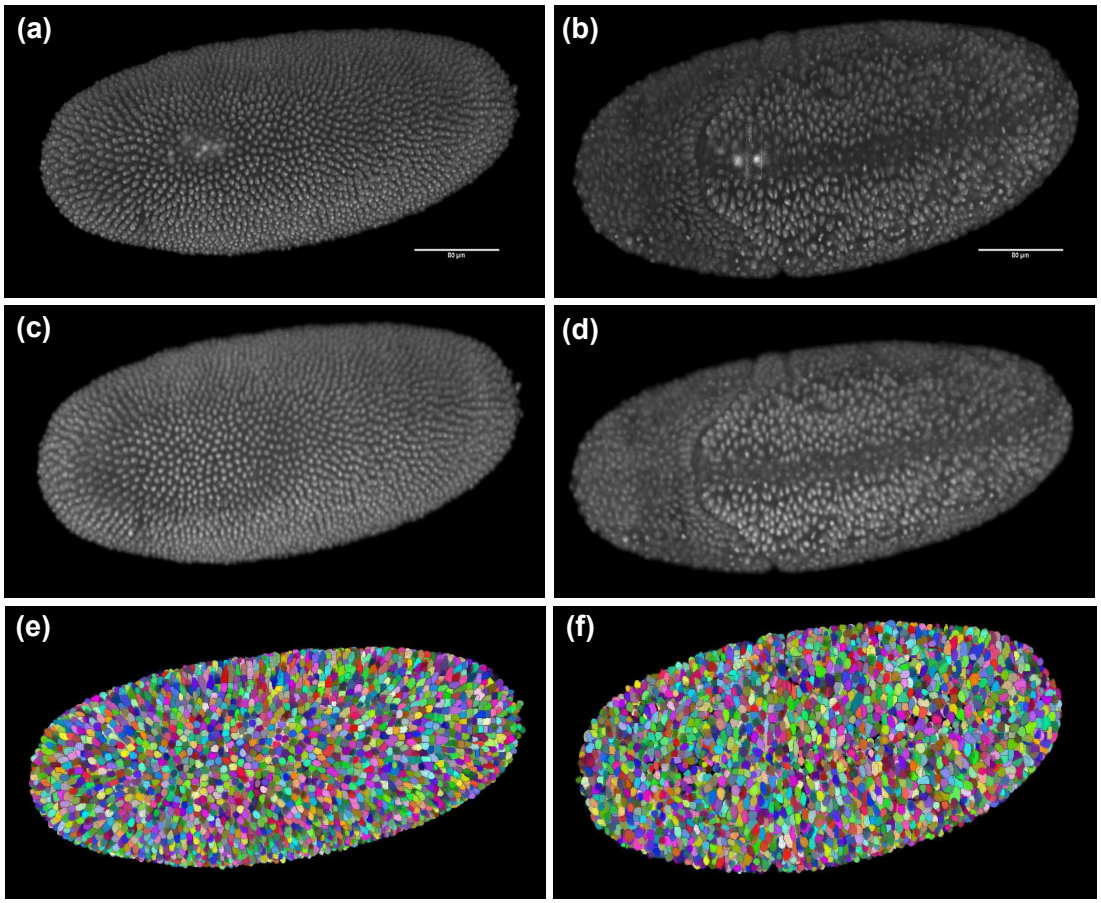

Figure 4.15: Example of denoising and segmentation results on the ***Fluo-N3DH-DRO*** dataset. (a, b) 3D view of the raw data for time point ($\mathrm{T} = 0$) from sequence1 and sequence2, respectively. (c) 3D view of the denoising result for (a). (d) 3D view of the denoising result for (b). (e) 3D view of the segmentation result for (c). (f) 3D view of the segmentation result for (d).

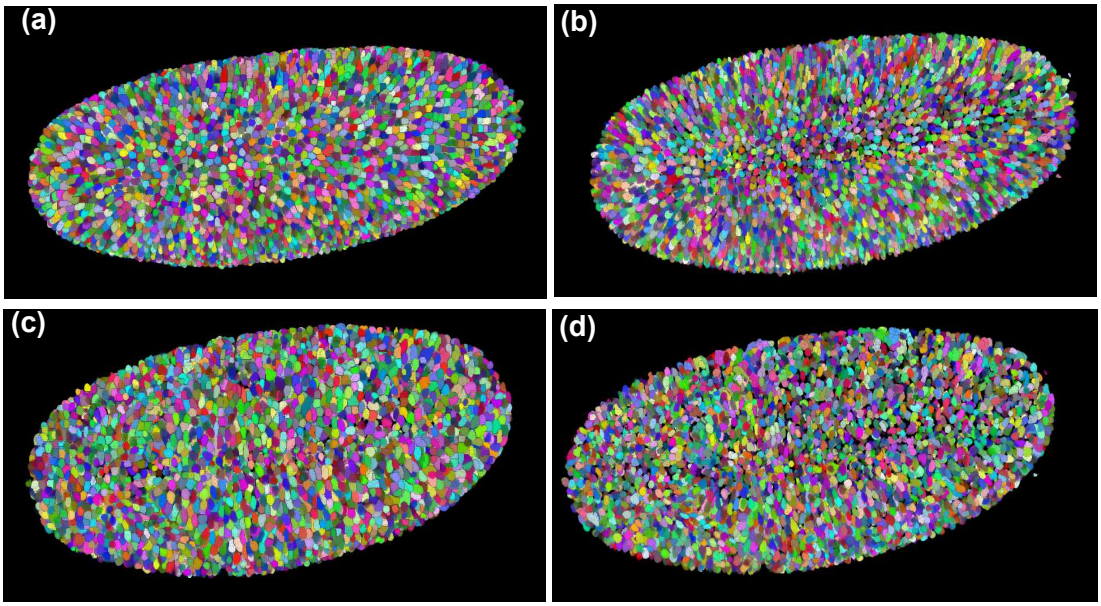

Figure 4.16: A visual comparison of segmentation results over ***Fluo-N3DH-DRO*** dataset using our method and KTH method [Ulman 2017, Maška 2014] for the same time points as Figure 4.15). First row: shows results of time point (T = 0) from sequence1. Second row: represents results of time point (T = 0) from sequence2. (a, c) 3D view of our segmentation result. (b, d) 3D view of KTH segmentation result.

4.4.2 Comparison of nuclei segmentation result with deep learning-based method.

As the deep learning-based approaches have achieved highly promising results in a wide range of applications, here we present some visual results for the proposed method against the NucleiNet deep learning-based method [Liu 2017]. Based on visual observation, we have noticed that the proposed method is able to detect almost all cell nuclei as shown in the second row of Figure 4.17. In contrast, the NucleiNet method did not manage to detect very faint and noisy cell nuclei (as presented in the third row of Figure 4.17). We believe this particular problem occurred due to improper labeled training samples. To explain as the NucleiNet method used the segmentation results obtained from the work of Gul-Mohammed [Gul-Mohammed 2014a] as training samples. Apparently, this is not the optimal result that can be used as labels in the training process. Due to lack of time and technical difficulties, we could not conduct a quantitative comparison between our results and NucleiNet method.

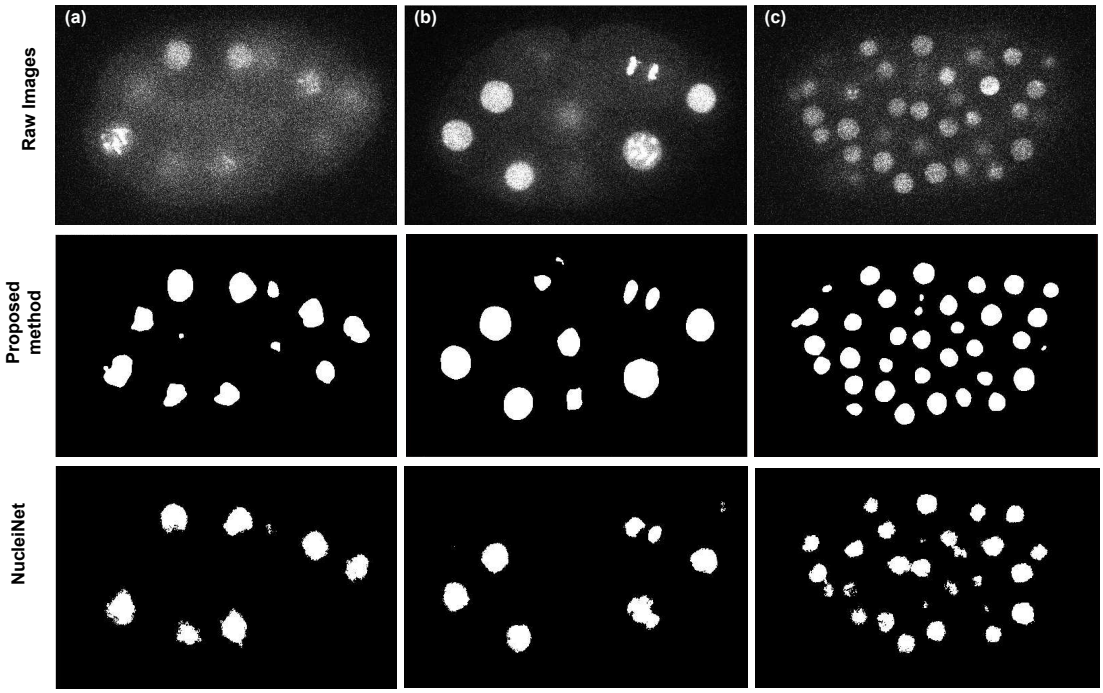

Figure 4.17: **A visual comparison of initial segmentation results over *CE-UPMC* dataset using our method and NucleiNet method [Liu 2017]. (a) A single plane Z = 19 of time point T = 39. (b) A single plane Z = 15 of time point T = 49. (c) A single plane Z = 24 of time point T = 159.**

4.4.3 Segmentation results over different 3D biological datasets.

Regarding *Fluo-MCF7shvec* dataset, the movement (i.e., translocation) of proteins such as transcription factors between the cytoplasm and nucleus has great biological importance in the diagnosis and treatment of cancer disease.

To automatically compute the translocation of a transcription factor to the nucleus using 3D fluorescence microscopy images, we first split each image into two individual channel images (Figure 4.18). The first channel identifies the cell nuclei that are stained with DAPI. The second channel identifies the protein of interest (such as p53) that is stained with mCherry.

For nuclei channel, the proposed denoising and segmentation methods are employed to identify and segment nuclei properly (as presented in Figure 4.19). For the segmentation of cytoplasm (as shown in Figure 4.21), the other channel has a strong cytoplasmic fluorescence in the time-lapse image sequence, and so we just use the proposed denoising method to remove the noise and then we use Otsu's thresholding-based method to obtain the cytoplasm masks. Even though the proposed method has provided good results for denoising the cytoplasm (as shown in Figure 4.20 (b)), it still not able to identify their centroid from the detection map or the maximum response image presented in Figure 4.20 (c, d). This is due to intra-variation of the cytoplasm intensity and also their irregular shape. Consequently, to split merged cytoplasm, the nucleus of each cell is employed as seed points for the watershed algorithm.

Following the segmentation of nuclei and cytoplasm, we calculate the mean of mCherry intensities for cytoplasm ($\bar{I}_C$) and the mean of mCherry intensities for nuclei ($\bar{I}_N$) (as shown in Figure 4.22 (a) and (b)). Finally, 13 cells were counted for the intracellular localization of mCherry fluorescence, that are classified into (1) $\bar{I}_C > \bar{I}_N$ and (2) $\bar{I}_C < \bar{I}_N$ as shown in Figure 4.22 (c).

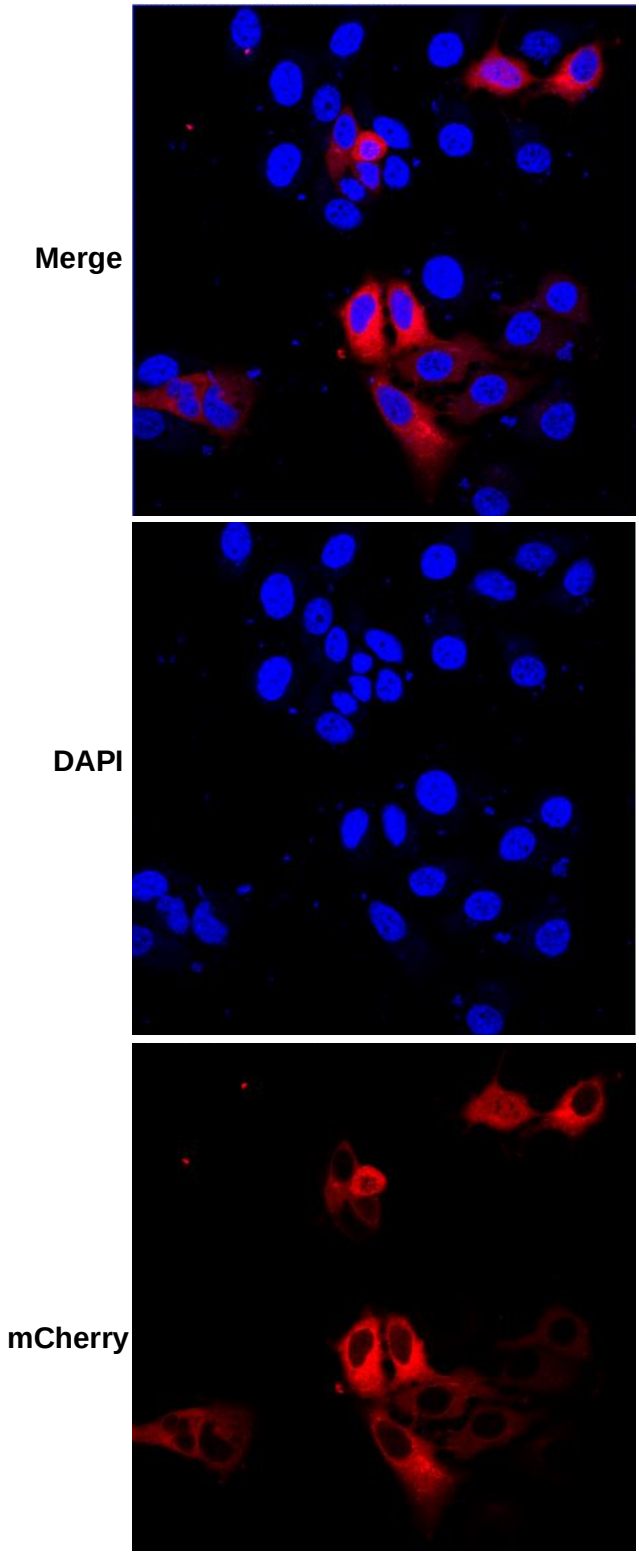

Figure 4.18: Representative image of a selected slice (Z = 3) from *Fluo-MCF7shvec* dataset.

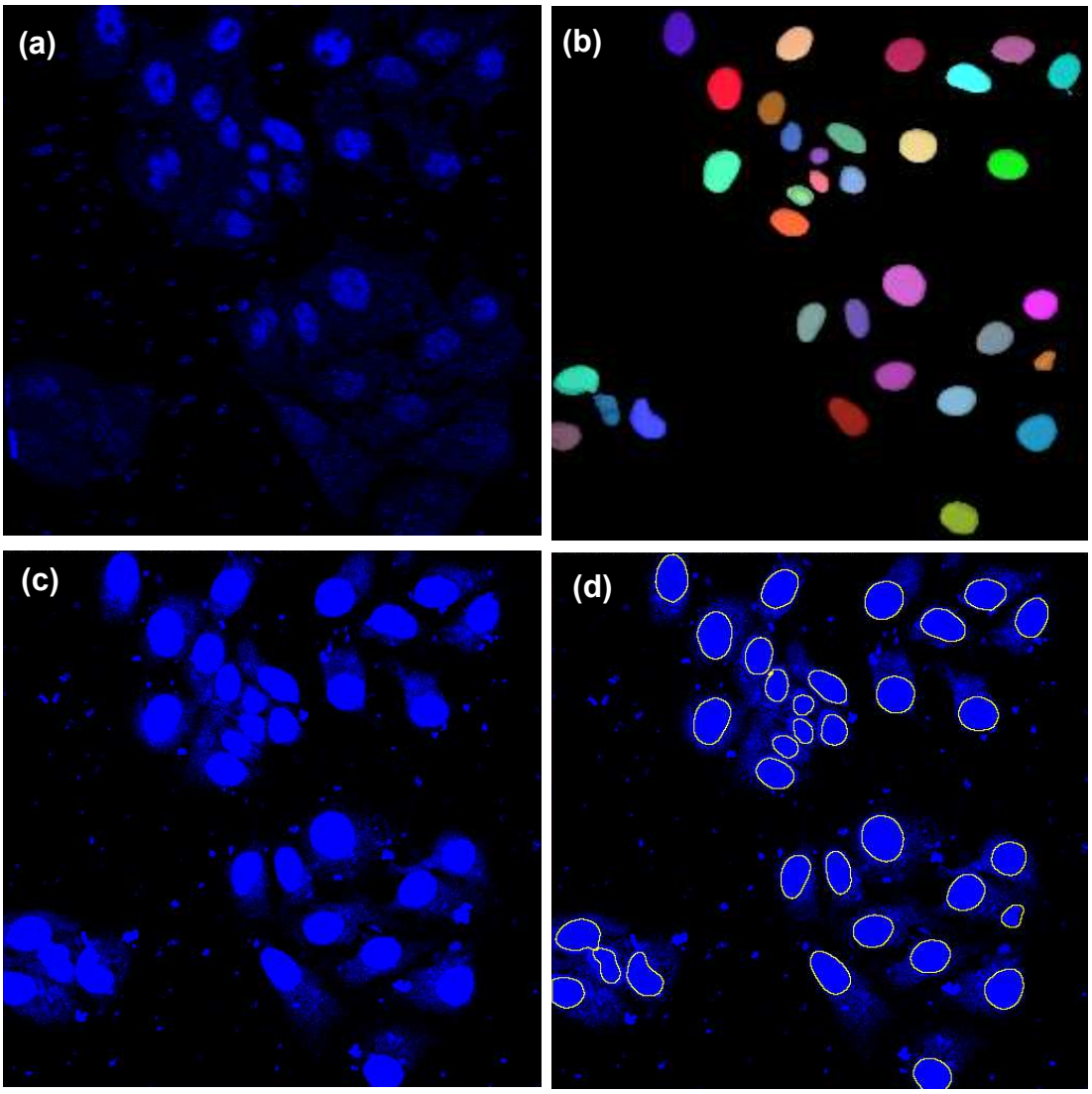

Figure 4.19: Example of nuclei segmentation results on *Fluo-MCF7shvec* dataset. (a) 3D view of the raw image. (b) 3D view of the segmented nuclei. (c) Raw image (a single plane $Z = 3$). (d) Raw image with overlaid nuclei segmentation.

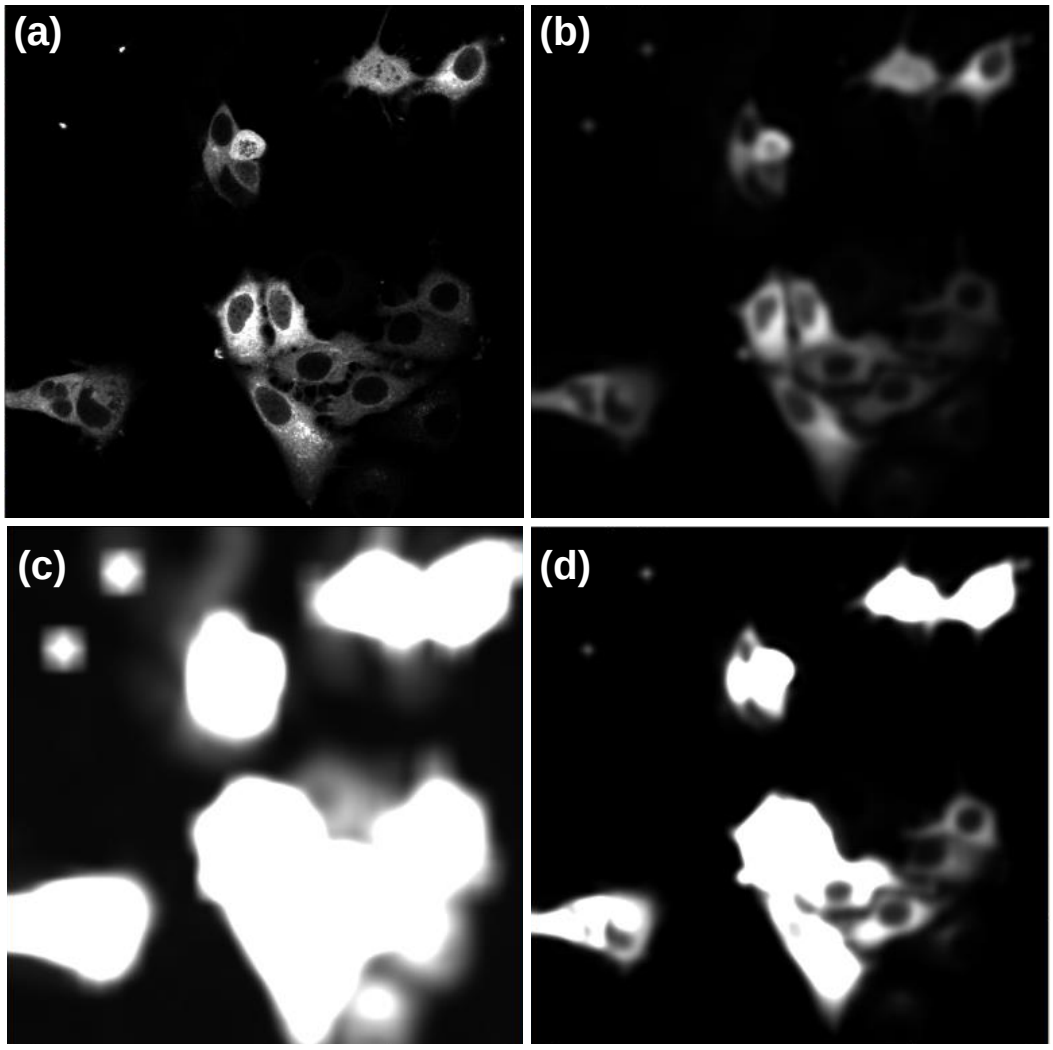

Figure 4.20: Example of the denoising and detection map of the cytoplasm from the mCherry channel of the ***Fluo-MCF7shvec*** dataset. (a) A gray level single plane ($Z = 3$)) from the ***Fluo-MCF7shvec*** dataset. (b) The denoised image obtained by applying the sparse representation model to the image in (a). (c) The detection map obtained from the sparse representation model for image in (a). (d) the maximum response image, obtained from multiplying image (b) with image (c).

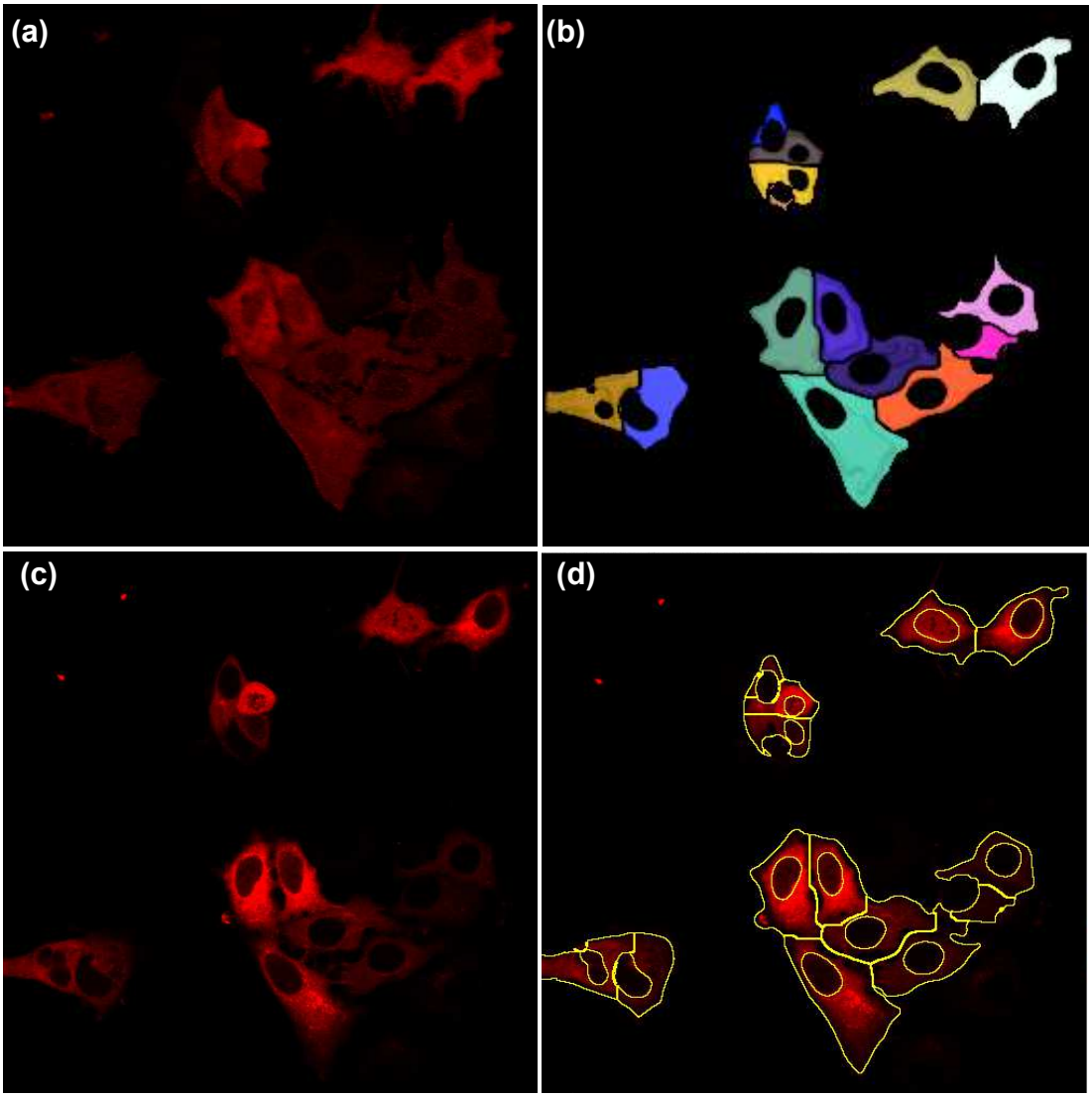

Figure 4.21: Example of cytoplasm segmentation results on *Fluo-MCF7shvec* dataset. (a) 3D view of the raw image. (b) 3D view of the segmented cytoplasm. (c) Raw image (a single plane $Z = 3$). (d) Raw image with overlaid cytoplasm segmentation.

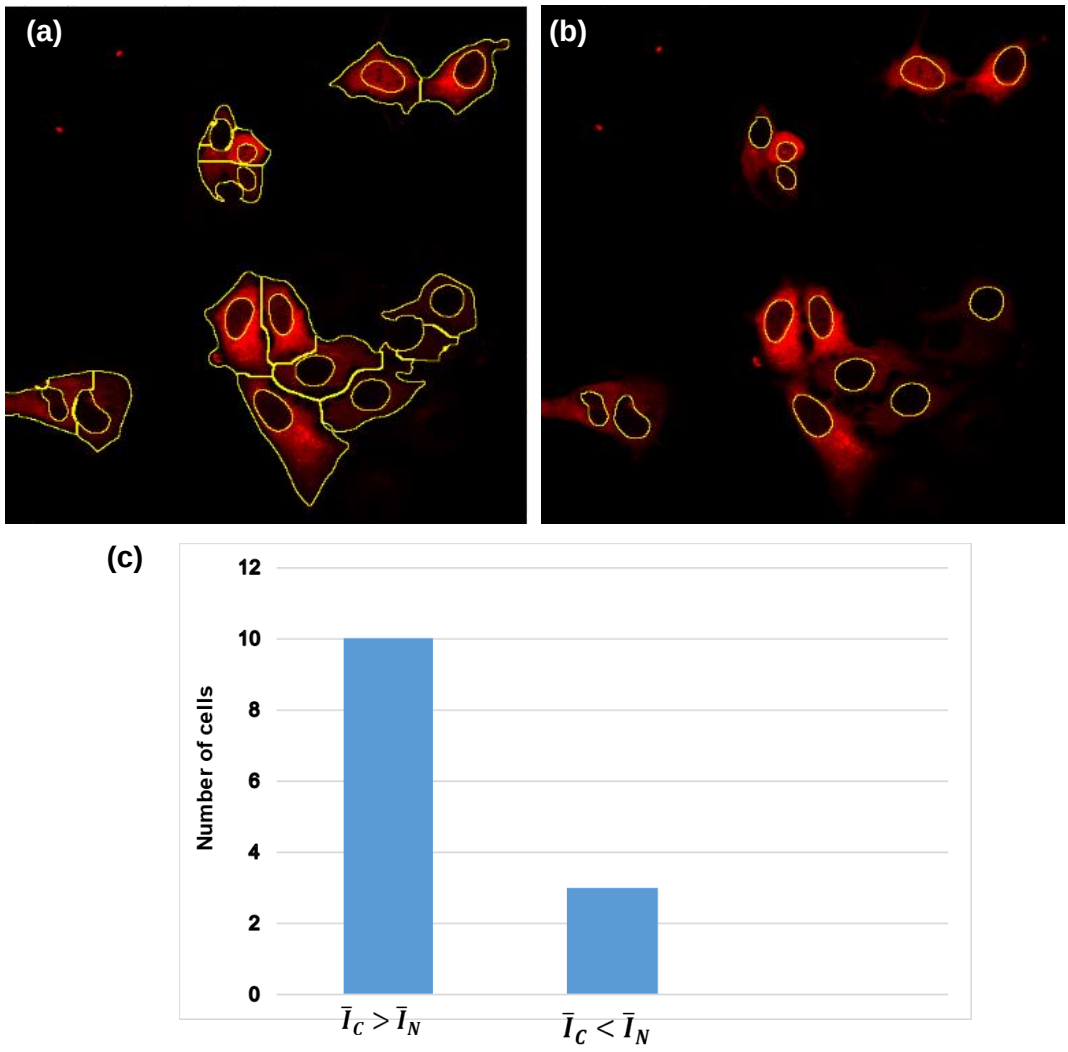

Figure 4.22: Examples of cells classified according to the intracellular localization of mCherry fluorescence. (a) Raw image (mCherry) with overlaid cytoplasm segmentation masks. (b) Raw image (mCherry) with overlaid nuclei segmentation masks. (c) The distribution of the cells according to mCherry fluorescence localization. Note that, the nuclei that don't have cytoplasmic marker are discarded.

In order to test the genericity of the algorithm, we have conducted additional experiments on real datasets coming from various tissues such as thymus tissue (provided by J. Sheridan, Walter and Eliza Hall Institute of Medical Research (WEHI)), lymphoid tissue (provided by JR. Groom, WEHI), and islets of Langerhans tissue (from Tran *et al.* [Nhu 2017]), where robust cell segmentation is still challenging. Despite the noisy and crowded environment, the obtained results from our method are quite encouraging as presented in Figure 4.23.

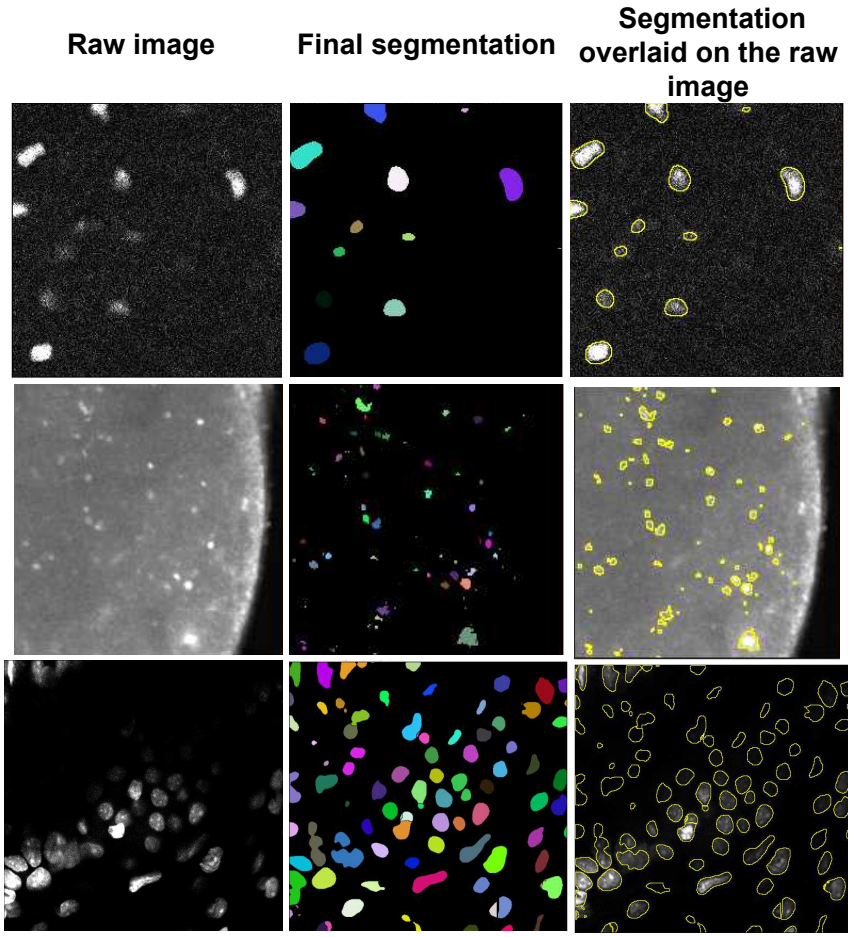

Figure 4.23: A visual example of segmentation results on real datasets coming from various tissues using our method. First raw: thymus tissue (a single plane Z = 106). Second raw: lymphoid tissue (a single plane Z = 204). Third raw: islet of Langerhans tissue [Nhu 2017] (a single plane Z = 100).

Chapter 5

Conclusion and Future Directions

Contents

5.1 Conclusion

In this thesis, we have presented a novel generic method for the denoising and detection of 3D cell nuclei in 3D time-lapse fluorescence microscopy images, based on a sparse representation model. We showed significant improvements among the other denoising methods. This is because the proposed method is a learning-based method and the idea of such method is to learn a dictionary that captures strong and discriminative structures (i.e., cell nuclei) of the images by suppressing unnecessary background information.

We suggested, as to propose a complete workflow for denoising, detection, and segmentation, to pair our denoising algorithm with a rather classical local thresholding method and showed that we obtained similar or better results than state of the art algorithm. We observed that our denoising algorithm is performing extremely well for very noisy data and can hence help to detect very faint or

previously undetectable nuclei. As the strength of our workflow is the denoising part, not so much the segmentation part, we observed (data not shown) little improvements for non-noisy data.

Over the last few years, deep learning approaches have achieved promising results in several domains, including denoising. However, they have some limits, for example, they are implemented to solve a specific problem, i.e., any new dataset will require a new training step. Furthermore, deep learning approaches required ground truth labels. Similarly to machine learning approaches, including deep learning, our method is based on a learning technique, but in dictionary-based methods the learning is completely unsupervised and hence can be performed for any new data without any change.

We also showed that our algorithm can also lead to accurate detection of nuclei centroids, we coupled this detection to a classical segmentation method and showed very good results on challenging datasets. We believe that coupling our algorithm with a more powerful segmentation method may lead to even better results, but this was not the purpose of this article. We focused on a robust and powerful denoising method coupled with classical segmentation to provide an effective workflow with minimal tuning. The fundamental parameters which have a more noticeable impact on the result are patch-size, SensitivityFactor ,and NucleiSeedDilation. All of these parameters are based on the average of cell nuclei volume present in the image. As these fundamental parameters are easy to understand, this makes them easier to tune-up if needed.

The obtained final segmentation results are quite good and stable. In addition, the training step is unsupervised and the dictionary can be directly learned from the image itself. We believe that no similar studies have been reported in existing literature for denoising and simultaneously predicting objects location in images.

The proposed method can handle the most challenging cases involving noisy, densely packed and multiple touching cell nuclei. In addition, it can produce the denoised image and simultaneously the potential locations of cell nuclei. The proposed method is adapted to the segmentation of cell nuclei in 3D time-lapse fluorescence microscopy images, nevertheless, it can be employed to detect and segment the nearly interacting intracellular organelles, including the endosomes,

lysosomes, and lipid droplets.

Additionally, the proposed method can be used for denoising the cytoplasm, but it has a problem of identifying their centroid from the detection map or the maximum response image because of the intra-variation of the cytoplasm intensity and their irregular shape.

Our method is successfully evaluated on two embryo models, the *C. elegans*, and the *Drosophila* datasets. The overall detection and segmentation results are comparable to the existing methods, which is a good starting point for automated cell nuclei tracking process.

5.2 Future Directions

In future work, new directions might be introduced to improve the proposed approach as follows. We will investigate an online learning method to handle long processing time so that we can reduce the time needed for dictionary learning. As an example, to segment the cytoplasm, we have chosen the patch size to be less than 25% of the cytoplasm size and as a result, we were able to avoid timing issue. Moreover, to handle the intra-variation of the cytoplasm intensity issue, we suggest to apply a morphological operation to the raw image such as opening operation.

Following the success in denoising and segmentation of cell nuclei using the proposed method, we would like to conduct further experiments on the test datasets from the cell tracking challenge, as our analysis was performed on the training datasets.

Since the deep learning-based approaches become greatly appreciated and it has been used in different applications, we plan to compare the proposed method with the denoising autoencoder as an unsupervised deep learning-based approach.

Considering the cell tracking process, which aims to identify and link segmented cells across consecutive frames in an image sequence to obtain cell trajectories, the simple method is used to associate each cell in one frame with the spatially nearest cell in a subsequent frame. Unfortunately, such a method may lead to higher mismatches especially when we deal with many cells or rapid cell movements.

Consequently, we plan to implement a method that can account for densely packed nuclei and also can deal with different events such as mitosis, going on and going out the field of view.

Additionally, we are aiming to handle the segmentation errors, i.e., missing detection, over-segmentation (splitting objects) and under-segmentation (merged objects). Basically, we are planning to use the distance between two objects based on the co-localization measurement to create a robust linking for each cell in the image sequence.

Appendix A

Appendix Example

This code is available online on the following link: https://github.com/lameesnasser/3D-Cell-Nuclei-Denoising-and-Segmentation.git

---

```
function [Denoised_Image,Detection_Map,Max_image,Dictionary] =
    ImageswWithKSVD(Image,param);

%   Image Denoising Using Sparse Representation Model and Dictionary
%               Learning on a Noisy Image


% ImageswWithKSVD used for denoising and detecting of 2D/3D cell
    nuclei images
% based on a sparse representation model. A more detailed description
    can
% be found in:
% "A novel generic dictionary-based denoising method for improving
    noisy and
% densely packed nuclei segmentation in 3D time-lapse fluorescence
    microscopy
% images" Scientific Reports9 (2019): 5654 (2019)
% doi:10.1038/s41598-019-41683-3
```

```
% The script is implemented in MATLAB release R2017b by
%  Lamees Nasser
%  April 2019



% INPUT ARGUMENTS : Image            -  The noisy image in gray-level
    scale.
%                  param.K           -  The number of dictionary atoms
    to train.
%                  param.numIteration - The number of K-SVD training
    iterations
%                  param.blockSize   -  The size of the blocks the
    algorithm
%                                     works. Example in 2D imgages p=[N
    N]. 3D images p= [N N M].
%                  param.L           - The number of nonzero elements
    (used atoms from the dictionary)
%                                      for the sparse representation
    coefficient
%                  param.trainnum    -  The number of blocks to train
    on.


% OUTPUT ARGUMENTS : Denoised_Image - The cleaned image.
%
%                   Detection_Map - Image indicates the potential
%                                   locations of cell nuclei.

%                   Dictionary    -  Updated dictionary.


%% Check If 2D or 3D image
[x,y,z]=size(Image);
```

```
if z>1
    p=3; %3D Image
    param.stepsize=[1 1 1]; % Step size of sliding window

else
    p=2; %2D Image
    param.stepsize=[1 1]; % Step size of sliding window

end

%% An initial dictionary is constructed by selecting random patches
    from the patches extracted from the noisy image among those having
    intensities greater than the obtained average intensity.


ids = cell(p,1);
[ids{:}] = reggrid(size(Image)-param.blocksize+1, param.trainnum,
    'eqdist');
blkMatrix = sampgrid(Image,param.blocksize,ids{:});
summation=sum(blkMatrix);
[val ind]=find(summation>mean(summation));
param.initialDictionary = blkMatrix(:,ind(1:param.K ));
%% Waiting for Training the dictionary
counterForWaitBar = param.numIteration+1;
h = waitbar(0,'Trainning Dictionary In Process ...');
param.waitBarHandle = h;
param.counterForWaitBar = counterForWaitBar;



%% Update and obtain the final dictionary by using K-SVD algorithm
    from the paper "The K-SVD: An Algorithm for
% Designing of Overcomplete Dictionaries for Sparse Representation",
```

```
    written by M. Aharon, M. Elad, and A.M. Bruckstein
%and appeared in the IEEE Trans. On Signal Processing, Vol. 54, no.
    11, pp. 4311-4322, November 2006.
% The used K-SVD toolbox can be find in
    https://elad.cs.technion.ac.il/software/

[Dictionary,output] = KSVD(blkMatrix,param);
output.D = Dictionary;
Dictionary(Dictionary>1)=1;
Dictionary(Dictionary<0)=0;
disp('Finished Trainning Dictionary');

%% Reconstruction of the denoised image and detection map

counterForWaitBar = 1
h1 = waitbar(0,'Images Reconstrucion In Process ...')

stepsize=param.stepsize;
blocksize=param.blocksize;
xount=0;
nz=0;
y_clean = zeros(size(Image));
y_Map = zeros(size(Image));
if p==3 %3D Images
    for k = 1:stepsize(3):size(y_clean,3)-blocksize(3)+1
        for j = 1:stepsize(2):size(y_clean,2)-blocksize(2)+1
            xount=xount+1;
            blocks =
                im2colstep(Image(:,j:j+blocksize(2)-1,k:k+blocksize(3)-1)
            ,blocksize,stepsize);
            gamma = full(mexOMP(blocks, Dictionary, param));
            Mapblocks= ones(prod(param.blocksize),param.K)*gamma;
            cleanblocks= Dictionary*gamma;
            nz = nz + nnz(gamma);
```

```
            cleanvol = col2imstep(cleanblocks,[size(y_clean,1)
                blocksize(2:3)],blocksize,stepsize);
            y_clean(:,j:j+blocksize(2)-1,k:k+blocksize(3)-1) =
                y_clean(:,j:j+blocksize(2)-1,k:k+blocksize(3)-1) +
                cleanvol;
            Mapvol = col2imstep(Mapblocks,[size(y_Map,1)
                blocksize(2:3)],blocksize,stepsize);
            y_Map(:,j:j+blocksize(2)-1,k:k+blocksize(3)-1) =
                y_Map(:,j:j+blocksize(2)-1,k:k+blocksize(3)-1) + Mapvol;
        end

        param.waitBarHandle = h1;
        waitbar(k/param.counterForWaitBar);
        param.counterForWaitBar = counterForWaitBar;

    end

elseif p==2 %2D Images

    blocks = im2colstep(Image,blocksize,stepsize);
    gamma = full(mexOMP(blocks, Dictionary, param));
    blocks1 =Dictionary*gamma;
    blocks_map = ones( prod(param.blocksize),param.K)*gamma;
    y_clean = col2imstep(blocks1,[size(Image)],blocksize,stepsize);
    y_Map = col2imstep(blocks_map,[size(Image)],blocksize,stepsize);
      param.waitBarHandle = h1;
        waitbar(param.counterForWaitBar);
        param.counterForWaitBar = counterForWaitBar;
end
lambda=0;
cnt = countcover(size(Image),blocksize,stepsize);
Denoised_Image = (y_clean+lambda*Image)./(1*(cnt + lambda));
Detection_Map= (y_Map+lambda*Image)./(1*(cnt + lambda));
Max_image=Denoised_Image.*Detection_Map;
```

```
disp('Finished Images Reconstrucion')
```

Bibliography

[Adamo 2015] Alessandro Adamo, Giuliano Grossi, Raffaella Lanzarotti and Jianyi Lin. *Robust face recognition using sparse representation in LDA space*. Machine Vision and Applications, vol. 26, no. 6, pages 837–847, 2015. (Cited on page 53.)

[Adams 2000] Mark D Adams, Susan E Celniker, Robert A Holt, Cheryl A Evans, Jeannine D Gocayne, Peter G Amanatides, Steven E Scherer, Peter W Li, Roger A Hoskins, Richard F Galle*et al.* *The genome sequence of Drosophila melanogaster*. Science, vol. 287, no. 5461, pages 2185–2195, 2000. (Cited on page 5.)

[Adams 2015] Michael W Adams, Andrew F Loftus, Sarah E Dunn, Matthew S Joens and James AJ Fitzpatrick. *Light sheet fluorescence microscopy (LSFM)*. Current protocols in cytometry, vol. 71, no. 1, pages 12–37, 2015. (Cited on pages xxiii, 7, 8 and 9.)

[Aguirre 2012] P. Aguirre. *Image processing of microscopic cellular samples*. In 2012 16th IEEE Mediterranean Electrotechnical Conference, pages 366–369, March 2012. (Cited on pages 24 and 27.)

[Aharon 2006] M. Aharon, M. Elad and A. Bruckstein. K -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. IEEE Transactions on Signal Processing, vol. 54, no. 11, pages 4311–4322, Nov 2006. (Cited on pages 53, 54, 57 and 87.)

[Akram 2014] Saad Ullah Akram, Juho Kannala, Mika Kaakinen, Lauri Eklund and Janne Heikkilä. *Segmentation of cells from spinning disk confocal images using a multi-stage approach.* In Asian Conference on Computer Vision, pages 300–314. Springer, 2014. (Cited on pages 25, 27, 31 and 34.)

[Akram 2016a] S. U. Akram, J. Kannala, L. Eklund and J. Heikkilä. *Joint cell segmentation and tracking using cell proposals.* In 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), pages 920–924, April 2016. (Cited on pages 40 and 44.)

[Akram 2016b] Saad Ullah Akram, Juho Kannala, Lauri Eklund and Janne Heikkilä. *Cell Segmentation Proposal Network for Microscopy Image Analysis.* In Gustavo Carneiro, Diana Mateus, Loïc Peter, Andrew Bradley, João Manuel R. S. Tavares, Vasileios Belagiannis, João Paulo Papa, Jacinto C. Nascimento, Marco Loog, Zhi Lu, Jaime S. Cardoso and Julien Cornebise, editors, Deep Learning and Data Labeling for Medical Applications, pages 21–29, Cham, 2016. Springer International Publishing. (Cited on pages 28, 32 and 33.)

[Alegro 2017] Maryana Alegro, Panagiotis Theofilas, Austin Nguy, Patricia A Castruita, William Seeley, Helmut Heinsen, Daniela M Ushizima and Lea T Grinberg. *Automating cell detection and classification in human brain fluorescent microscopy images using dictionary learning and sparse coding.* Journal of neuroscience methods, vol. 282, pages 20–33, 2017. (Cited on page 53.)

[Amat 2014] Fernando Amat, William Lemon, Daniel P Mossing, Katie McDole, Yinan Wan, Kristin Branson, Eugene W Myers and Philipp J Keller. *Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data.* Nature methods, vol. 11, no. 9, page 951, 2014. (Cited on pages 35, 38 and 43.)

[Arbelle 2018a] Assaf Arbelle and Tammy Riklin Raviv. *Microscopy cell segmentation via adversarial neural networks.* In 2018 IEEE 15th International

Symposium on Biomedical Imaging (ISBI 2018), pages 645–648. IEEE, 2018. (Cited on pages 32 and 33.)

[Arbelle 2018b] Assaf Arbelle and Tammy Riklin Raviv. *Microscopy Cell Segmentation via Convolutional LSTM Networks*. arXiv preprint arXiv:1805.11247, 2018. (Cited on page 33.)

[Bise 2011a] Ryoma Bise, Takeo Kanade, Zhaozheng Yin and Seung-il Huh. *Automatic cell tracking applied to analysis of cell migration in wound healing assay*. In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, pages 6174–6179. IEEE, 2011. (Cited on page 6.)

[Bise 2011b] Ryoma Bise, Zhaozheng Yin and Takeo Kanade. *Reliable cell tracking by global data association*. In 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pages 1004–1010. IEEE, 2011. (Cited on pages 38 and 44.)

[Bise 2015] R. Bise and Y. Sato. *Cell Detection From Redundant Candidate Regions Under Nonoverlapping Constraints*. IEEE Transactions on Medical Imaging, vol. 34, no. 7, pages 1417–1427, July 2015. (Cited on pages 28 and 34.)

[Boulanger 2007] Jerome Boulanger, Charles Kervrann and Patrick Bouthemy. *Space-time adaptation for patch-based image sequence restoration*. IEEE transactions on pattern analysis and machine intelligence, vol. 29, no. 6, pages 1096–1102, 2007. (Cited on page 23.)

[Boulanger 2010] Jérôme Boulanger, Charles Kervrann, Patrick Bouthemy, Peter Elbau, Jean-Baptiste Sibarita and Jean Salamero. *Patch-based nonlocal functional for denoising fluorescence microscopy image sequences*. IEEE transactions on medical imaging, vol. 29, no. 2, pages 442–454, 2010. (Cited on pages 23, 27 and 53.)

[Bradley 2007] Derek Bradley and Gerhard Roth. *Adaptive Thresholding using the Integral Image*. Journal of Graphics Tools, vol. 12, no. 2, pages 13–21, 2007. (Cited on pages 54, 66, 91 and 99.)

[Brenner 1974] Sydney Brenner. *The genetics of Caenorhabditis elegans*. Genetics, vol. 77, no. 1, pages 71–94, 1974. (Cited on page 5.)

[Buades 2005] Antoni Buades, Bartomeu Coll and J-M Morel. *A non-local algorithm for image denoising*. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), volume 2, pages 60–65. IEEE, 2005. (Cited on page 22.)

[Carpenter 2006] Anne E Carpenter, Thouis R Jones, Michael R Lamprecht, Colin Clarke, In Han Kang, Ola Friman, David A Guertin, Joo Han Chang, Robert A Lindquist, Jason Moffat*et al.* *CellProfiler: image analysis software for identifying and quantifying cell phenotypes*. Genome biology, vol. 7, no. 10, page R100, 2006. (Cited on page 46.)

[Cascio 2019] Donato Cascio, Vincenzo Taormina and Giuseppe Raso. *An Automatic HEp-2 Specimen Analysis System Based on an Active Contours Model and an SVM Classification*. Applied Sciences, vol. 9, no. 2, page 307, 2019. (Cited on pages 25, 27, 30 and 34.)

[Chalfie 1994] Martin Chalfie, Yuan Tu, Ghia Euskirchen, William W Ward and Douglas C Prasher. *Green fluorescent protein as a marker for gene expression*. Science, vol. 263, no. 5148, pages 802–805, 1994. (Cited on page 6.)

[Chan 2001] Tony F Chan and Luminita A Vese. *Active contours without edges*. IEEE Transactions on image processing, vol. 10, no. 2, pages 266–277, 2001. (Cited on pages 36 and 37.)

[Chen 1998] Scott Shaobing Chen, David L. Donoho and Michael A. Saunders. *Atomic Decomposition by Basis Pursuit*. SIAM Journal on Scientific Computing, vol. 20, no. 1, pages 33–61, 1998. (Cited on page 56.)

[Chen 2006] Xiaowei Chen, Xiaobo Zhou and Stephen TC Wong. *Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy*. IEEE Transactions on Biomedical Engineering, vol. 53, no. 4, pages 762–766, 2006. (Cited on pages 31 and 34.)

[Chenouard 2014] Nicolas Chenouard, Ihor Smal, Fabrice De Chaumont, Martin Maška, Ivo F Sbalzarini, Yuanhao Gong, Janick Cardinale, Craig Carthel, Stefano Coraluppi, Mark Winter*et al.* *Objective comparison of particle tracking methods*. Nature methods, vol. 11, no. 3, page 281, 2014. (Cited on page 35.)

[Cicconet 2013] Marcelo Cicconet, Davi Geiger and Kristin C. Gunsalus. *Wavelet-based Circular Hough Transform and Its Application in Embryo Development Analysis*. In VISAPP, 2013. (Cited on pages 28, 29 and 34.)

[Cliffe 2017] Adam Cliffe, David P Doupé, HsinHo Sung, Isaac Kok Hwee Lim, Kok Haur Ong, Li Cheng and Weimiao Yu. *Quantitative 3D analysis of complex single border cell behaviors in coordinated collective cell migration*. Nature Communications, vol. 8, page 14905, 2017. (Cited on pages 28, 31 and 34.)

[Danielyan 2014] Aram Danielyan, Yu-Wei Wu, Pei-Yu Shih, Yulia Dembitskaya and Alexey Semyanov. *Denoising of two-photon fluorescence images with Block-Matching 3D filtering*. Methods, vol. 68, no. 2, pages 308–316, 2014. (Cited on pages 23 and 27.)

[Darbon 2008] J. Darbon, A. Cunha, T. F. Chan, S. Osher and G. J. Jensen. *Fast nonlocal filtering applied to electron cryomicroscopy*. In 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pages 1331–1334, May 2008. (Cited on pages 22 and 27.)

[De Chaumont 2012] Fabrice De Chaumont, Stéphane Dallongeville, Nicolas Chenouard, Nicolas Hervé, Sorin Pop, Thomas Provoost, Vannary Meas-Yedid, Praveen Pankajakshan, Timothée Lecomte, Yoann Le Montagner*et al.* *Icy: an open bioimage informatics platform for extended repro-*

ducible research. Nature methods, vol. 9, no. 7, page 690, 2012. (Cited on page 46.)

[Deledalle 2009] Charles-Alban Deledalle, Loïc Denis and Florence Tupin. *Iterative weighted maximum likelihood denoising with probabilistic patch-based weights*. IEEE Transactions on Image Processing, vol. 18, no. 12, pages 2661–2672, 2009. (Cited on page 23.)

[Deledalle 2010] Charles-Alban Deledalle, Florence Tupin and Loïc Denis. *Poisson NL means: Unsupervised non local means for Poisson noise*. In 2010 IEEE international conference on image processing, pages 801–804. IEEE, 2010. (Cited on pages 22 and 27.)

[Dufour 2005] Alexandre Dufour, Vasily Shinin, Shahragim Tajbakhsh, Nancy Guillén-Aghion, J-C Olivo-Marin and Christophe Zimmer. *Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces*. IEEE Transactions on Image Processing, vol. 14, no. 9, pages 1396–1410, 2005. (Cited on pages 35, 36, 37 and 43.)

[Dufour 2011] Alexandre Dufour, Roman Thibeaux, Elisabeth Labruyere, Nancy Guillen and Jean-Christophe Olivo-Marin. *3-D active meshes: fast discrete deformable models for cell tracking in 3-D time-lapse microscopy*. IEEE Transactions on Image Processing, vol. 20, no. 7, pages 1925–1937, 2011. (Cited on pages 35, 37 and 43.)

[Dzyubachyk 2007] Oleh Dzyubachyk, Wiro Niessen and Erik Meijering. *A variational model for level-set based cell tracking in time-lapse fluorescence microscopy images*. In 2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pages 97–100. IEEE, 2007. (Cited on pages 36 and 43.)

[Dzyubachyk 2008] Oleh Dzyubachyk, Wiro Niessen and Erik Meijering. *Advanced level-set based multiple-cell segmentation and tracking in time-lapse fluorescence microscopy images*. In 2008 5th IEEE International Sympo-

sium on Biomedical Imaging: From Nano to Macro, pages 185–188. IEEE, 2008. (Cited on pages 36, 37 and 43.)

[Dzyubachyk 2009] Oleh Dzyubachyk, Rob Jelier, Ben Lehner, Wiro Niessen and Erik Meijering. *Model-based approach for tracking embryogenesis in Caenorhabditis elegans fluorescence microscopy data*. In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 5356–5359. IEEE, 2009. (Cited on pages 37 and 43.)

[Dzyubachyk 2010] Oleh Dzyubachyk, Wiggert A Van Cappellen, Jeroen Essers, Wiro J Niessen and Erik Meijering. *Advanced level-set-based cell tracking in time-lapse fluorescence microscopy*. IEEE transactions on medical imaging, vol. 29, no. 3, pages 852–867, 2010. (Cited on pages 36, 37 and 43.)

[El Zehiry 2007] Noha El Zehiry, Steve Xu, Prasanna Sahoo and Adel Elmaghraby. *Graph cut optimization for the Mumford-Shah model*. In The Seventh IASTED International Conference on Visualization, Imaging and Image Processing, pages 182–187. ACTA Press, 2007. (Cited on page 38.)

[Elad 2006] M. Elad and M. Aharon. *Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries*. IEEE Transactions on Image Processing, vol. 15, no. 12, pages 3736–3745, Dec 2006. (Cited on pages 53, 54 and 87.)

[Gonzalez 2002] Rafael C Gonzalez, Richard E Woods*et al.* *Digital image processing*, 2002. (Cited on page 69.)

[Goodfellow 2014] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville and Yoshua Bengio. *Generative adversarial nets*. In Advances in neural information processing systems, pages 2672–2680, 2014. (Cited on page 32.)

[Gul-Mohammed 2014a] Jaza Gul-Mohammed, Ignacio Arganda-Carreras, Philippe Andrey, Vincent Galy and Thomas Boudier. *A generic classification-based method for segmentation of nuclei in 3D images of*

early embryos. BMC Bioinformatics, vol. 15, no. 1, page 9, 2014. (Cited on pages xxiv, xxix, 18, 25, 27, 28, 34, 80, 93, 98 and 105.)

[Gul-Mohammed 2014b] Jaza Gul-Mohammed and Thomas Boudier. *Classified region growing for 3D segmentation of packed nuclei*. In Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on, pages 842–845. IEEE, 2014. (Cited on pages 25, 27, 28, 30 and 34.)

[Haider 2016] SA Haider, A Cameron, P Siva, D Lui, MJ Shafiee, A Boroomand, N Haider and A Wong. *Fluorescence microscopy image noise reduction using a stochastically-connected random field model*. Scientific reports, vol. 6, page 20640, 2016. (Cited on pages 23 and 27.)

[Hartigan 1979] John A Hartigan and Manchek A Wong. *Algorithm AS 136: A k-means clustering algorithm*. Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 28, no. 1, pages 100–108, 1979. (Cited on page 57.)

[He 2017] Tao He, Hua Mao, Jixiang Guo and Zhang Yi. *Cell tracking using deep neural networks with multi-task learning*. Image and Vision Computing, vol. 60, pages 142–153, 2017. (Cited on pages 41 and 44.)

[Hernandez 2018] David E Hernandez, Steven W Chen, Elizabeth E Hunter, Edward B Steager and Vijay Kumar. *Cell Tracking with Deep Learning and the Viterbi Algorithm*. In 2018 International Conference on Manipulation, Automation and Robotics at Small Scales (MARSS), pages 1–6. IEEE, 2018. (Cited on pages 32 and 33.)

[Hilsenbeck 2016] Oliver Hilsenbeck, Michael Schwarzfischer, Stavroula Skylaki, Bernhard Schauberger, Philipp S Hoppe, Dirk Loeffler, Konstantinos D Kokkaliaris, Simon Hastreiter, Eleni Skylaki, Adam Filipczyk*et al. Software tools for single-cell tracking and quantification of cellular and molecular properties*. Nature biotechnology, vol. 34, no. 7, page 703, 2016. (Cited on page 45.)

[Johnson 2006] Don H Johnson. *Signal-to-noise ratio*. Scholarpedia, vol. 1, no. 12, page 2088, 2006. (Cited on page 80.)

[Kamentsky 2011] Lee Kamentsky, Thouis R Jones, Adam Fraser, Mark-Anthony Bray, David J Logan, Katherine L Madden, Vebjorn Ljosa, Curtis Rueden, Kevin W Eliceiri and Anne E Carpenter. *Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software*. Bioinformatics, vol. 27, no. 8, pages 1179–1180, 2011. (Cited on page 46.)

[Kanade 2011] Takeo Kanade, Zhaozheng Yin, Ryoma Bise, Seungil Huh, Sungeun Eom, Michael F Sandbothe and Mei Chen. *Cell image analysis: Algorithms, system and applications*. In 2011 IEEE Workshop on Applications of Computer Vision (WACV), pages 374–381. IEEE, 2011. (Cited on page 35.)

[Kapuscinski 1995] Jan Kapuscinski. *DAPI: a DNA-specific fluorescent probe*. Biotechnic & Histochemistry, vol. 70, no. 5, pages 220–233, 1995. (Cited on page 6.)

[Kesteven 1946] GL Kesteven. *The coefficient of variation*. Nature, vol. 158, no. 4015, page 520, 1946. (Cited on page 82.)

[Kushwaha 2012] H. S. Kushwaha, S. Tanwar, K. S. Rathore and S. Srivastava. *Denoising Filters for TEM (Transmission Electron Microscopy) Image of Nanomaterials*. In 2012 Second International Conference on Advanced Computing Communication Technologies, pages 276–281, Jan 2012. (Cited on page 25.)

[Li 2008] Kang Li, Eric D Miller, Mei Chen, Takeo Kanade, Lee E Weiss and Phil G Campbell. *Cell population tracking and lineage construction with spatiotemporal context*. Medical image analysis, vol. 12, no. 5, pages 546–566, 2008. (Cited on pages 37 and 43.)

[Li 2010] Fuhai Li, Xiaobo Zhou, Jinwen Ma and Stephen TC Wong. *Multiple nuclei tracking using integer programming for quantitative cancer cell cycle*

analysis. IEEE transactions on medical imaging, vol. 29, no. 1, pages 96–105, 2010. (Cited on pages 38 and 44.)

[Liao 2016] Miao Liao, Yu qian Zhao, Xiang hua Li, Peishan Dai, Xiao wen Xu, Jun kai Zhang and Beiji Zou. *Automatic segmentation for cell images based on bottleneck detection and ellipse fitting*. Neurocomputing, vol. 173, pages 615–622, 2016. (Cited on page 28.)

[Lichtman 2005] Jeff W Lichtman and José-Angel Conchello. *Fluorescence microscopy*. Nature methods, vol. 2, no. 12, page 910, 2005. (Cited on page 6.)

[Lima 2006] Rui Lima, Shigeo Wada, Ken-ichi Tsubota and Takami Yamaguchi. *Confocal micro-PIV measurements of three-dimensional profiles of cell suspension flow in a square microchannel*. Measurement Science and Technology, vol. 17, no. 4, page 797, 2006. (Cited on pages xxiii and 10.)

[Liu 2014] K. Liu, S. S. Lienkamp, A. Shindo, J. B. Wallingford, G. Walz and O. Ronneberger. *Optical flow guided cell segmentation and tracking in developing tissue*. In 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), pages 298–301, April 2014. (Cited on page 28.)

[Liu 2017] Zichuan Liu, Yifei Hu, Hang Xu, Lamees Nasser, Philippe Coquet, Thomas Boudier and Hao Yu. *NucleiNet: A convolutional encoder-decoder network for bio-image denoising*. In Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE, pages 1986–1989. IEEE, 2017. (Cited on pages xxx, 28, 31, 33, 105 and 106.)

[Long 2015] Jonathan Long, Evan Shelhamer and Trevor Darrell. *Fully convolutional networks for semantic segmentation*. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3431–3440, 2015. (Cited on page 31.)

[Lou 2014] Xinghua Lou, Martin Schiegg and Fred A Hamprecht. *Active structured learning for cell tracking: algorithm, framework, and usability.* IEEE transactions on medical imaging, vol. 33, no. 4, pages 849–860, 2014. (Cited on pages 40 and 44.)

[Luisier 2010a] Florian Luisier. *The SURE-LET approach to image denoising.* EPFL, Swiss Federal Institute of Technology Lausanne, 2010. (Cited on pages 10, 12 and 13.)

[Luisier 2010b] Florian Luisier, Cédric Vonesch, Thierry Blu and Michael Unser. *Fast interscale wavelet denoising of Poisson-corrupted images.* Signal Processing, vol. 90, no. 2, pages 415 – 427, 2010. (Cited on pages xxvii, 24, 27, 87, 88, 89 and 90.)

[Magnusson 2012] Klas EG Magnusson and Joakim Jaldén. *A batch algorithm using iterative application of the Viterbi algorithm to track cells and construct cell lineages.* In 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pages 382–385. IEEE, 2012. (Cited on pages 39 and 44.)

[Magnusson 2015a] K. E. G. Magnusson, J. Jaldén, P. M. Gilbert and H. M. Blau. *Global Linking of Cell Tracks Using the Viterbi Algorithm.* IEEE Transactions on Medical Imaging, vol. 34, no. 4, pages 911–929, April 2015. (Cited on pages 39 and 44.)

[Magnusson 2015b] Klas EG Magnusson and Joakim Jaldén. *Tracking of non-brownian particles using the viterbi algorithm.* In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pages 380–384. IEEE, 2015. (Cited on page 39.)

[Magnusson 2016] Klas EG Magnusson. *Segmentation and tracking of cells and particles in time-lapse microscopy.* PhD thesis, KTH Royal Institute of Technology, 2016. (Cited on pages 25, 27, 30, 34 and 71.)

[Maška 2014] Martin Maška, Vladimír Ulman, David Svoboda, Pavel Matula, Petr Matula, Cristina Ederra, Ainhoa Urbiola, Tomás España, Subramanian Venkatesan, Deepak M.W. Balak, Pavel Karas, Tereza Bolcková,

Markéta Štreitová, Craig Carthel, Stefano Coraluppi, Nathalie Harder, Karl Rohr, Klas E. G. Magnusson, Joakim Jaldén, Helen M. Blau, Oleh Dzyubachyk, Pavel Křížek, Guy M. Hagen, David Pastor-Escuredo, Daniel Jimenez-Carretero, Maria J. Ledesma-Carbayo, Arrate Muñoz-Barrutia, Erik Meijering, Michal Kozubek and Carlos Ortiz-de Solorzano. *A benchmark for comparison of cell tracking algorithms.* Bioinformatics, vol. 30, no. 11, pages 1609–1617, 2014. (Cited on pages xxix, xxx, xxxiv, 17, 35, 80, 91, 93, 96, 100, 102 and 104.)

[Mallat 1993a] S. Mallat and Z. Zhang. *Adaptive time-frequency transform.* In 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, volume 3, pages 241–244 vol.3, April 1993. (Cited on page 56.)

[Mallat 1993b] S. G. Mallat and Zhifeng Zhang. *Matching pursuits with time-frequency dictionaries.* IEEE Transactions on Signal Processing, vol. 41, no. 12, pages 3397–3415, Dec 1993. (Cited on page 56.)

[Maška 2010] Martin Maška, Pavel Matula, Ondřej Daněk and Michal Kozubek. *A fast level set-like algorithm for region-based active contours.* In International Symposium on Visual Computing, pages 387–396. Springer, 2010. (Cited on page 37.)

[Maška 2011] Martin Maška, Pavel Matula and Michal Kozubek. *Simultaneous tracking of multiple objects using fast level set-like algorithm.* In Sixth Doctoral Workshop on Mathematical and Engineering Methods in Computer Science (MEMICS'10)–Selected Papers. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2011. (Cited on page 38.)

[Maška 2012] Martin Maška, Arrate Munoz-Barrutia and Carlos Ortiz-de Solórzano. *Fast tracking of fluorescent cells based on the Chan-Vese model.* In 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), pages 1316–1319. IEEE, 2012. (Cited on pages 37 and 43.)

[Maška 2013] Martin Maška, Ondřej Daněk, Saray Garasa, Ana Rouzaut, Arrate Muñoz-Barrutia and Carlos Ortiz-de Solorzano. *Segmentation and shape*

tracking of whole fluorescent cells based on the Chan–Vese model. IEEE transactions on medical imaging, vol. 32, no. 6, pages 995–1006, 2013. (Cited on pages 38 and 43.)

[McQuin 2018] Claire McQuin, Allen Goodman, Vasiliy Chernyshev, Lee Kamentsky, Beth A Cimini, Kyle W Karhohs, Minh Doan, Liya Ding, Susanne M Rafelski, Derek Thirstrup*et al. CellProfiler 3.0: Next-generation image processing for biology*. PLoS biology, vol. 16, no. 7, page e2005970, 2018. (Cited on page 46.)

[Meijering 2008] Erik Meijering, Ihor Smal, Oleh Dzyubachyk and Jean-Christophe Olivo-Marin. *Time-lapse imaging*. Microscope Image Processing, pages 401–440, 2008. (Cited on pages 2, 6, 15 and 17.)

[Meijering 2012] Erik Meijering, Oleh Dzyubachyk and Ihor Smal. *Methods for cell and particle tracking*. In Methods in enzymology, volume 504, pages 183–200. Elsevier, 2012. (Cited on pages 35 and 45.)

[Meiniel 2018] William Meiniel, Jean-Christophe Olivo-Marin and Elsa D Angelini. *Denoising of Microscopy Images: A Review of the State-of-the-Art, and a New Sparsity-Based Method*. IEEE Transactions on Image Processing, vol. 27, no. 8, pages 3842–3856, 2018. (Cited on pages 10, 12, 13 and 22.)

[Mitchison 1971] John M Mitchison. The biology of the cell cycle. CUP Archive, 1971. (Cited on page 3.)

[Nasser 2017] Lamees Nasser, Paula Coronado, Etienne David and Thomas Boudier. *A dictionary-based approach to reduce noise in fluorescent microscopy images*. In Signal and Image Processing (ICSIP), 2017 IEEE 2nd International Conference on, pages 144–148. IEEE, 2017. (Cited on page 53.)

[Nhu 2017] Hoa Tran Thi Nhu, Rafael Arrojo E Drigo, Per-Olof Berggren and Thomas Boudier. *A novel toolbox to investigate tissue spatial organization*

applied to the study of the islets of Langerhans. Scientific reports, vol. 7, page 44261, 2017. (Cited on pages xxxi and 113.)

[Ogawa 2013] Takahiro Ogawa and Miki Haseyama. *Image inpainting based on sparse representations with a perceptual metric*. EURASIP Journal on Advances in Signal Processing, vol. 2013, no. 1, page 179, 2013. (Cited on page 53.)

[Ollion 2013] Jean Ollion, Julien Cochennec, François Loll, Christophe Escude and Thomas Boudier. *TANGO: a generic tool for high-throughput 3D image analysis for studying nuclear organization*. Bioinformatics, vol. 29, no. 14, pages 1840–1841, 2013. (Cited on page 77.)

[Otsu 1979] Nobuyuki Otsu. *A threshold selection method from gray-level histograms*. IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pages 62–66, 1979. (Cited on pages 54, 66 and 99.)

[Padfield 2009] Dirk Padfield, Jens Rittscher, Nick Thomas and Badrinath Roysam. *Spatio-temporal cell cycle phase analysis using level sets and fast marching methods*. Medical image analysis, vol. 13, no. 1, pages 143–155, 2009. (Cited on pages 37 and 43.)

[Padfield 2011] Dirk Padfield, Jens Rittscher and Badrinath Roysam. *Coupled minimum-cost flow cell tracking for high-throughput quantitative analysis*. Medical image analysis, vol. 15, no. 4, pages 650–668, 2011. (Cited on pages 41 and 44.)

[Parvati 2008] K Parvati, Prakasa Rao and M Mariya Das. *Image segmentation using gray-scale morphology and marker-controlled watershed transformation*. Discrete Dynamics in Nature and Society, vol. 2008, 2008. (Cited on page 54.)

[Payne 2018] Samantha L Payne, Priya N Anandakumaran, Balazs V Varga, Cindi M Morshead, Andras Nagy and Molly S Shoichet. *In vitro maturation of human iPSC-derived neuroepithelial cells influences transplant survival*

in the stroke-injured rat brain. Tissue Engineering Part A, vol. 24, no. 3-4, pages 351–360, 2018. (Cited on page 6.)

[Powers 2011] D. M. W. Powers. *Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation.* Journal of Machine Learning Technologies, vol. 2, no. 1, pages 37–63, 2011. (Cited on page 74.)

[Qiu 2012] Zhen Qiu, Lei Yang and Weiping Lu. *A new feature-preserving nonlinear anisotropic diffusion for denoising images containing blobs and ridges.* Pattern Recognition Letters, vol. 33, no. 3, pages 319–330, 2012. (Cited on page 53.)

[Rapoport 2011] Daniel H Rapoport, Tim Becker, Amir Madany Mamlouk, Simone Schicktanz and Charli Kruse. *A novel validation algorithm allows for automated cell tracking and the extraction of biologically meaningful parameters.* PloS one, vol. 6, no. 11, page e27315, 2011. (Cited on page 6.)

[Roels 2018] Joris Roels, Jan Aelterman, HQ Luong, Saskia Lippens, A Pižurica, Yvan Saeys and Wilfried Philips. *An overview of state-of-the-art image restoration in electron microscopy.* Journal of microscopy, vol. 271, no. 3, pages 239–254, 2018. (Cited on page 22.)

[Rohlfing 2012] Torsten Rohlfing. *Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable.* IEEE transactions on medical imaging, vol. 31, no. 2, pages 153–163, 2012. (Cited on page 74.)

[Rohr 2010] Karl Rohr, William J Godinez, Nathalie Harder, Stefan Wörz, Julian Mattes, Wolfgang Tvaruskó and Roland Eils. *Tracking and quantitative analysis of dynamic movements of cells and particles.* Cold Spring Harbor Protocols, vol. 2010, no. 6, pages pdb–top80, 2010. (Cited on page 35.)

[Ronneberger 2015] Olaf Ronneberger, Philipp Fischer and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation.* In Nassir Navab, Joachim Hornegger, William M. Wells and Alejandro F. Frangi,

editors, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, pages 234–241, Cham, 2015. Springer International Publishing. (Cited on pages 28, 31, 32 and 33.)

[Rousson 2002] Mikaël Rousson and Rachid Deriche. *A variational framework for active and adaptative segmentation of vector valued images.* In Workshop on Motion and Video Computing, 2002. Proceedings., pages 56–61. IEEE, 2002. (Cited on page 36.)

[Rubinstein 2010] Ron Rubinstein, Alfred M Bruckstein and Michael Elad. *Dictionaries for sparse representation modeling.* Proceedings of the IEEE, vol. 98, no. 6, pages 1045–1057, 2010. (Cited on page 57.)

[Sadanandan 2017] Sajith Kecheril Sadanandan, Petter Ranefall, Sylvie Le Guyader and Carolina Wählby. *Automated Training of Deep Convolutional Neural Networks for Cell Segmentation.* Scientific reports, vol. 7, no. 1, page 7860, 2017. (Cited on pages 28 and 32.)

[Schiegg 2013] Martin Schiegg, Philipp Hanslovsky, Bernhard X Kausler, Lars Hufnagel and Fred A Hamprecht. *Conservation tracking.* In Proceedings of the IEEE International Conference on Computer Vision, pages 2928–2935, 2013. (Cited on pages 39 and 44.)

[Schiegg 2014] Martin Schiegg, Philipp Hanslovsky, Carsten Haubold, Ullrich Koethe, Lars Hufnagel and Fred A Hamprecht. *Graphical model for joint segmentation and tracking of multiple dividing cells.* Bioinformatics, vol. 31, no. 6, pages 948–956, 2014. (Cited on pages 40 and 44.)

[Schindelin 2012] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid*et al. Fiji: an open-source platform for biological-image analysis.* Nature methods, vol. 9, no. 7, page 676, 2012. (Cited on page 46.)

[Schmid 2010] Benjamin Schmid, Johannes Schindelin, Albert Cardona, Mark Longair and Martin Heisenberg. *A high-level 3D visualization API for Java and ImageJ.* BMC bioinformatics, vol. 11, no. 1, page 274, 2010. (Cited on page 77.)

[Schneider 2012] Caroline A Schneider, Wayne S Rasband and Kevin W Eliceiri. *NIH Image to ImageJ: 25 years of image analysis.* Nature methods, vol. 9, no. 7, page 671, 2012. (Cited on page 45.)

[Sim 2016] K. S. Sim, V. Teh and M. E. Nia. *Adaptive noise Wiener filter for scanning electron microscope imaging system.* Scanning, vol. 38, no. 2, pages 148–163, 2016. (Cited on pages 24 and 27.)

[Soille 1990] Pierre Soille and Luc M Vincent. *Determining watersheds in digital pictures via flooding simulations.* In Visual Communications and Image Processing'90: Fifth in a Series, volume 1360, pages 240–251. International Society for Optics and Photonics, 1990. (Cited on page 69.)

[Sommer 2011] Christoph Sommer, Christoph Straehle, Ullrich Koethe and Fred A Hamprecht. *Ilastik: Interactive learning and segmentation toolkit.* In 2011 IEEE international symposium on biomedical imaging: From nano to macro, pages 230–233. IEEE, 2011. (Cited on page 47.)

[Stalling 2005] Detlev Stalling, Malte Westerhoff, Hans-Christian Hege*et al.* *Amira: A highly interactive system for visual data analysis.* The visualization handbook, vol. 38, pages 749–67, 2005. (Cited on page 48.)

[Stephens 2003] David J Stephens and Victoria J Allan. *Light microscopy techniques for live cell imaging.* science, vol. 300, no. 5616, pages 82–86, 2003. (Cited on pages 8 and 9.)

[Tarnawski 2013] W. Tarnawski, V. Kurtcuoglu, P. Lorek, M. Bodych, J. Rotter, M. Muszkieta, ?. Piwowar, D. Poulikakos, M. Majkowski and A. Ferrari. *A Robust Algorithm for Segmenting and Tracking Clustered Cells in Time-Lapse Fluorescent Microscopy.* IEEE Journal of Biomedical and Health

Informatics, vol. 17, no. 4, pages 862–869, July 2013. (Cited on pages 28, 30 and 34.)

[Theodorakopoulos 2014] Ilias Theodorakopoulos, Dimitris Kastaniotis, George Economou and Spiros Fotopoulos. *Hep-2 cells classification via sparse representation of textural features fused into dissimilarity space*. Pattern Recognition, vol. 47, no. 7, pages 2367–2378, 2014. (Cited on page 53.)

[Thorn 2016] Kurt Thorn. *A quick guide to light microscopy in cell biology*. Molecular biology of the cell, vol. 27, no. 2, pages 219–222, 2016. (Cited on pages xxxiii, 7, 8 and 9.)

[Tonti 2015] Simone Tonti, Santa Di Cataldo, Andrea Bottino and Elisa Ficarra. *An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test*. Computerized Medical Imaging and Graphics, vol. 40, no. Supplement C, pages 62 – 69, 2015. (Cited on pages 28, 30 and 34.)

[Türetken 2017] E. Türetken, X. Wang, C. J. Becker, C. Haubold and P. Fua. *Network Flow Integer Programming to Track Elliptical Cells in Time-Lapse Sequences*. IEEE Transactions on Medical Imaging, vol. 36, no. 4, pages 942–951, April 2017. (Cited on pages 28, 29, 34, 40 and 44.)

[Türetken 2015] Engin Türetken, Xinchao Wang, Carlos J. Becker and Pascal Fua. *Detecting and Tracking Cells using Network Flow Programming*. CoRR, vol. abs/1501.05499, 2015. (Cited on pages 28 and 29.)

[Ulman 2017] V. Ulman, M. Maška, K. E. G. Magnusson, O. Ronneberger, C. Haubold, N. Harder, P. Matula, P. Matula, D. Svoboda, M. Radojevic, I. Smal, K. Rohr, J. Jaldén, H. M. Blau, O. Dzyubachyk, B. Lelieveldt, P. Xiao, Y. Li, S.-Y. Cho, A. Dufour, J. C. Olivo-Marin, C. C. Reyes-Aldasoro, J. A. Solis-Lemus, R. Bensch, T. Brox, J. Stegmaier, R. Mikut, S. Wolf, F. A. Hamprecht, T. Esteves, P. Quelhas, Ö. Demirel, L. Malmström, F. Jug, P. Tomancák, E. Meijering, A. Muñoz-Barrutia, M. Kozubek and C. Ortiz de Solor. *An objective comparison of cell-tracking algorithms*.

Nature Methods, 2017. (Cited on pages xxiv, xxix, xxx, xxxiv, 17, 18, 35, 80, 91, 93, 96, 100, 102 and 104.)

[Van De Ville 2009] Dimitri Van De Ville and Michel Kocher. *SURE-based non-local means*. IEEE Signal Processing Letters, vol. 16, no. 11, pages 973–976, 2009. (Cited on page 23.)

[Wählby 2004] Carolina Wählby, I-M Sintorn, Fredrik Erlandsson, Gunilla Borgefors and Ewert Bengtsson. *Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections*. Journal of microscopy, vol. 215, no. 1, pages 67–76, 2004. (Cited on pages 28, 29, 34 and 71.)

[Wang 2017] Yangxu Wang, Hua Mao and Zhang Yi. *Stem cell motion-tracking by using deep neural networks with multi-output*. Neural Computing and Applications, pages 1–13, 2017. (Cited on pages 41 and 44.)

[Weigert 2017a] Martin Weigert, Loic Royer, Florian Jug and Gene Myers. *Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks*. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 126–134. Springer, 2017. (Cited on pages 25 and 27.)

[Weigert 2017b] Martin Weigert, Uwe Schmidt, Tobias Boothe, M Andreas, Alexander Dibrov, Akanksha Jain, Benjamin Wilhelm, Deborah Schmidt, Coleman Broaddus, Siân Culley*et al.* *Content-aware image restoration: pushing the limits of fluorescence microscopy*. bioRxiv, page 236463, 2017. (Cited on pages 25 and 27.)

[Wiesmann 2015] V Wiesmann, D Franz, C Held, C Münzenmayer, R Palmisano and T Wittenberg. *Review of free software tools for image analysis of fluorescence cell micrographs*. Journal of microscopy, vol. 257, no. 1, pages 39–53, 2015. (Cited on page 45.)

[Wilson 1900] Edmund Beecher Wilson. The cell in development and inheritance. Number 4. The Macmillan Company, 1900. (Cited on page 2.)

[Yang 2010] Lei Yang, Richard Parton, Graeme Ball, Zhen Qiu, Alan H. Greenaway, Ilan Davis and Weiping Lu. *An adaptive non-local means filter for denoising live-cell images and improving particle detection.* Journal of Structural Biology, vol. 172, no. 3, pages 233 – 243, 2010. (Cited on pages 22 and 27.)

[Yang 2015] Sejung Yang and Byung-Uk Lee. *Poisson-Gaussian noise reduction using the hidden Markov model in contourlet domain for fluorescence microscopy images.* PloS one, vol. 10, no. 9, page e0136964, 2015. (Cited on pages 24 and 27.)

[Zhang 2018] Jing Zhang, Emilia Moradi, Michael G Somekh and Melissa L Mather. *Label-Free, High Resolution, Multi-Modal Light Microscopy for Discrimination of Live Stem Cell Differentiation Status.* Scientific reports, vol. 8, no. 1, page 697, 2018. (Cited on page 6.)