



HAL
open science

Apport des techniques d'apprentissage (profond) à la microscopie holographique pour applications médicales.

Lyes Bouchama

► To cite this version:

Lyes Bouchama. Apport des techniques d'apprentissage (profond) à la microscopie holographique pour applications médicales.. Automatique / Robotique. Institut Polytechnique de Paris, 2023. Français. NNT : 2023IPPAS022 . tel-04480260

HAL Id: tel-04480260

<https://theses.hal.science/tel-04480260>

Submitted on 27 Feb 2024

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

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



INSTITUT
POLYTECHNIQUE
DE PARIS

NNT : 2023IPPAS022

Thèse de doctorat



Apport des techniques d'apprentissage (profond) à la microscopie holographique pour applications médicales

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom SudParis

École doctorale n°626 Ecole doctorale de l'Institut Polytechnique de
Paris (ED IP Paris)
Spécialité de doctorat : Informatique

Thèse présentée et soutenue à Palaiseau, le 21/12/2023, par

LYES BOUCHAMA

Composition du Jury :

M. François DELHOMMEAU Professeur, Sorbonne Université, Centre de Recherche Saint-Antoine	Président
Mme. Cécile MALLET Professeur, Université de Versailles Saint-Quentin-en-Yvelines	Rapporteur
M. Maxime JACQUOT Professeur, Université Bourgogne Franche-Comté	Rapporteur
M. Thomas OLIVIER Maître de conférence, Université Jean Monnet, Laboratoire Hubert Curien, UMR CNRS 5516	Examineur
M. Yaneck GOTTESMAN Professeur, Télécom SudParis (SAMOVAR)	Directeur de thèse
Mme. Bernadette DORIZZI Professeur émérite, Télécom SudParis (SAMOVAR)	Co-directrice de thèse
M. Jacques KLOSSA Docteur, TRIBVN/T-life	Invité

Apport des techniques d'apprentissage (profond)
à la microscopie holographique pour applications
médicales

Remerciements

Je tiens, avant tout, à exprimer ma profonde gratitude envers mes directeurs de thèse, Yaneck, Bernadette, et Jacques.

L'enthousiasme de Yaneck et de Bernadette pour inculquer une méthode scientifique de travail rigoureuse a profondément marqué mon évolution. Leur soutien inébranlable, leur dévouement constant et leur capacité à guider chaque phase de ce parcours avec patience et perspicacité se sont révélés d'une valeur incommensurable.

Jacques a joué un rôle crucial en me familiarisant avec le monde industriel, un domaine qui m'a permis d'élargir considérablement mes horizons. Grâce à son influence, j'ai non seulement acquis une expérience précieuse, mais j'ai aussi profité de sa vision avant-gardiste et de sa bienveillance.

Je souhaite exprimer également ma profonde reconnaissance à Vincent Leymarie, Marc Thellier et François Delhommeau, les médecins avec qui j'ai eu l'opportunité de travailler en collaboration durant cette thèse. Leur expertise en biologie a conféré à mon travail de recherche une dimension unique.

Je souhaite aussi exprimer mes remerciements à Houda, doctorante, et Arthur, post-doctorant, de notre équipe de recherche pour leur collaboration, qui a mis en lumière la richesse des échanges et l'importance fondamentale de la collaboration dans notre parcours scientifique. Je souhaite également remercier mes collègues et amis des équipes RST et EPH, en particulier les post-doctorants Aina et Nikson.

Je tiens à exprimer ma gratitude envers tous les collaborateurs avec lesquels j'ai eu le privilège de travailler chez TRIBVN, et à souligner spécialement le soutien remarquable de Guillaume.

Mes pensées reconnaissantes se tournent également vers mes enseignants du Master TRIED auprès desquels j'ai beaucoup appris : Cécile, Aymeric et Laurent.

Enfin, j'adresse une reconnaissance toute particulière à ma famille et à mes amis. Votre amour, votre soutien sans faille et votre confiance indéfectible en moi ont été mon phare dans cette aventure qu'a été ma thèse.

Ce parcours, riche en émotions et en défis, n'aurait certainement pas eu la même saveur sans votre présence à mes côtés. Du plus profond de mon cœur, recevez mes remerciements les plus sincères.

Resumé

Mon travail s'inscrit dans le cadre du partenariat stratégique Télécom Sud-Paris (TSP) et TRIBVN/T-life, dédié au développement de nouvelles approches en microscopie optique, couplées à l'intelligence artificielle, en vue d'identifier, de prédire et de monitorer les pathologies hématologiques et parasitologiques. C'est dans cette perspective que nous avons développé, dans le laboratoire, un prototype de microscope reposant sur un principe d'imagerie non conventionnelle à synthèse d'ouverture, basée sur l'approche FPM (Fourier Ptychographic Microscopy). Cette approche permet de dépasser les limites en résolution de l'optique conventionnelle ou, de façon équivalente, accéder à des champs de vue très larges (de 4 à 25 fois plus importants) à résolution fixée. Elle permet, en sus, de diversifier la nature des données acquises (avec l'enregistrement de la phase en complément des données d'intensité).

Toutefois, en raison de certaines contraintes, notamment le temps d'acquisition et de reconstruction des images pour obtenir une qualité optimale, cette technologie rencontre encore des difficultés pour trouver ses applications et être commercialisée par les différents acteurs du domaine de la microscopie.

Le travail réalisé dans cette thèse a permis des avancées significatives sur certains des aspects limitants de cette technologie grâce à la mise en œuvre de modèles à base de réseaux de neurones.

Nous avons proposé une refocalisation automatique des images bimodales efficace sur de grands champs de vue, grâce à un post-traitement basé sur un U-Net. Nous avons aussi proposé une approche originale, alliant apprentissage statistique et optimisation guidée par la physique pour réduire les temps d'acquisition des images et les temps de reconstruction.

Ces modèles ont démontré leur efficacité, par des diagnostics plus précis et discriminants, dans des applications de parasitologie et d'hématologie.

Les contributions apportées ont un potentiel d'application qui dépasse le domaine de la FPM, ouvrant des perspectives dans divers autres champs de l'imagerie calculatoire.

Abstract

This research is part of the Télécom SudParis (TSP) and TRIBVN/T-life strategic partnership, dedicated to the development of new approaches in optical microscopy, coupled with artificial intelligence, to identify, predict and monitor hematological and parasitological pathologies. In this regard, we developed a prototype microscope based on a computational imaging principle with a synthetic aperture, based on the FPM (Fourier Ptychographic Microscopy) approach. This approach makes it possible to overcome conventional optics' resolution limits, or equivalently access very large fields of view (from 4 to 25 times larger) at fixed resolution. It also enables us to diversify the nature of the data acquired (with phase recording in addition to intensity data).

However, despite its promise, the technology faces challenges in widespread adoption and commercialization within the microscopy field, primarily due to constraints such as the time-intensive process required for image acquisition and reconstruction to achieve optimal quality.

The research conducted in this thesis has led to substantial advancements in overcoming certain limitations of this technology, leveraging models based on neural networks.

We have proposed an efficient automatic refocusing of bimodal images over large fields of view, thanks to post-processing based on a U-Net. We have also proposed an original approach, combining statistical learning and physics-driven optimization to reduce image acquisition and reconstruction times.

These frameworks have validated their efficacy by yielding more precise and discriminating diagnoses in the fields of parasitology and haematology.

The potential applications of these contributions go far beyond the field of FPM, opening up perspectives in various other fields of computational imaging.

Table des matières

1	Introduction	21
1.1	Contexte général	21
1.2	Objectif de la thèse	23
2	Contexte pluridisciplinaire	27
2.1	Imagerie computationnelle	27
2.2	Microscope optique	29
2.2.1	La source de lumière	30
	a) Eclairage Köhler	31
	b) LED et laser (éclairage cohérent)	32
2.2.2	L'objectif du microscope	33
2.2.3	La caméra	34
2.2.4	Propriétés relatives à la résolution en microscopie	35
	a) Calcul de la résolution	36
	b) Échantillonnage au niveau de la caméra	37
2.3	Microscopie ptychographique de Fourier	38
2.3.1	Introduction	38
2.3.2	Principe de la Ptychographie	40
2.3.3	Principe de la FPM	41
2.3.4	Algorithme de reconstruction	44
	a) ePIE	45
	b) Algorithme initial de la reconstruction FPM	47
	c) Correction d'aberration FPM	49
	d) Algorithme EPRY	50
2.4	Les réseaux de neurones	51
2.4.1	Principe de fonctionnement	52
2.4.2	Rétropropagation et descente de gradient	54
2.4.3	Réseaux de neurones convolutifs	55
2.4.4	U-Net	56
2.4.5	YOLO	59
2.5	Contexte Biologique	63

2.5.1	Parasitologie	63
2.5.2	Leucocytes	65
2.6	Difficulté de la FPM	66
2.6.1	Correction numérique des aberrations globales du système optique	67
2.6.2	Correction numérique du focus	68
2.6.3	Temps d'acquisition	69
2.6.4	Temps de reconstruction	69
2.6.5	Bruit	70
3	Refocalisation des images FPM par réseaux de neurones	71
3.1	Introduction	71
3.2	Etat de l'art	72
3.3	Apport de la refocalisation en parasitologie	74
3.3.1	Configuration du microscope	74
3.3.2	Profondeur de champ en imagerie FPM	75
3.3.3	Caractéristiques recherchées dans les images refocalisées . . .	77
3.3.4	Apprentissage du U-Net	79
	a) Modèle neuronal	79
	b) Schéma d'apprentissage	80
3.3.5	Base de données de refocalisation	83
3.3.6	Résultats et évaluation de l'apport de la refocalisation . . .	83
	a) Évaluation visuelle de la refocalisation intensité phase	83
	b) Comparaison visuelle de la refocalisation intensité phase vs intensité seule	85
	c) Refocalisation intensité phase sur un grand champ de vue	86
3.3.7	Impact quantitatif sur la détection des parasites	86
	a) Base de donnée de détection	88
3.3.8	Résultats de détection	89
3.4	Apport de la refocalisation en hématologie	94
3.4.1	Configuration du microscope	94
3.4.2	Apprentissage et base de données	95
3.4.3	Resultats de la refocalisation	97
3.5	Conclusion	99
4	Reconstruction par réseaux de neurones cDIP	101
4.1	Introduction	101
4.2	Etat de l'art	101
4.3	Modèle cDIP informé par la physique	103
4.4	Configuration du dispositif en microscopie ptychographique de Fourier	105

4.5	Resultat cDIP $\Gamma = 60\%$	107
4.6	Resultat cDIP $\Gamma = 60\%$ avec correction d'aberrations	108
4.7	Resultat cDIP pour $\Gamma = 10\%$	110
4.8	Conclusion	110
5	Stratégie de reconstruction FPM dans les conditions de faible recouvrement	113
5.1	Etat de l'art	113
5.2	Configuration FPM	115
5.3	Schéma de la reconstruction par cDIP-DNN	116
5.3.1	Schéma d'apprentissage U-Net pour l'initialisation de cDIP (Étape 1)	116
5.3.2	Reconstruction cDIP sous conditions de faible recouvrement (Étape 2)	118
5.4	Résultats de simulation et discussion	118
5.4.1	Évaluation visuelle	120
5.4.2	Évaluation quantitative	121
5.5	Conclusion	123
6	Approche cDIP pour expérimentation avec faible taux de recouvrement et amélioration de focus	125
6.1	Introduction	125
6.2	Configuration du microscope	127
6.3	Modèle de reconstruction et refocalisation par apprentissage	130
6.4	Évaluation visuelle du modèle	132
6.5	Conclusion	135
7	Conclusion et Perspectives	137
7.1	Conclusion	137
7.2	Perspectives	139
7.2.1	Validation quantitative du modèle cDIP-L pour la réduction du nombre de LED et l'amélioration de la qualité d'image	139
a)	Évaluation complete cDIP-L	139
b)	Exploration de la configuration de LED optimale	140
7.2.2	Estimation du Focus	140
7.2.3	Reconstruction tomographique par cDIP-LO	140
7.2.4	Gestion du bruit	141
7.2.5	Extension à d'autres domaines	141

Table des figures

2.1	Processus de mesure et de reconstruction pour l'imagerie non conventionnelle	28
2.2	Structures du vivant présentées à diverses échelles [3].	30
2.3	Dispositif de microscopie optique intégrant une matrice de LED, un objectif, des oculaires et une caméra.	31
2.4	Exemples de dispositions des LED pour deux géométries de matrices : les LED sélectionnées sont représentées en vert, tandis que les LED non utilisées sont indiquées en noir.	33
2.5	Exemples des caractéristiques d'un objectif de microscopie	34
2.6	Schéma de la microscopie holographique numérique hors axe [60].	38
2.7	Montage pour l'imagerie diffractive cohérente. L'échantillon S est éclairé par un faisceau lumière cohérente et parallèle. L'intensité diffusée est enregistrée dans le champ optique lointain [75]	39
2.8	Représentation schématique du principe de fonctionnement de la ptychographie. On est ici dans le cadre d'un algorithme permettant de reconstruire à la fois l'objet et l'illumination.[9]	41
2.9	Dispositif FPM avec matrice de LED, Frottis, Objectif (lentille), Camera.	42
2.10	Illustration du modèle direct de la FPM	44
2.11	Résultats comparatifs [57] entre l'algorithme PIE : a,1) Intensité, b,1) phase et par l'algorithme ePIE a,2) intensité, b,2) phase	46
2.12	a), b) et c) illustrent les images de la pile FPM acquises avec différents angles d'éclairage. d) et e) représentent respectivement l'intensité et la phase reconstruite. d) module du spectre de l'objet complexe reconstruit.	49
2.13	À gauche : principe de fonctionnement d'un neurone. À droite : fonctions d'activation.	53
2.14	Principe de la convolution et du max pooling dans un CNN	57
2.15	Architecture utilisée pour la segmentation des images Biomédicales [77]	58

2.16	Architecture YOLOV3 [89] L'algorithme YOLO est basé sur une architecture neuronale convolutive avec une approche de régression. Il permet de réaliser une classification d'objets et une localisation simultanées en une seule passe.	61
2.17	Grand champ FPM sur un frottis de paludisme, avec un objectif 20xNA=0.75 avec 3 longueurs d'onde $\lambda = 625nm, 525nm, 465nm$. . .	65
2.18	Polynucleaire neutrophile, acquis avec un objectif 20xNA=0.75 avec 3 longueurs d'onde $\lambda = 625nm, 525nm, 465nm$	66
2.19	Illustration de la sensibilité de la microscopie FPM en fonction du réglage de la bague de correction du couvre-objet, en utilisant un objectif 40x à ouverture numérique (NA) de 0.95 et un facteur de super-résolution de 2x	68
3.1	a) Schéma d'un microscope équipé de sa matrice de LEDs. b) mire de résolution USAF1951 pour différentes positions de l'objectif 10xNA0.45, b ₁) led centrale brute, b ₂) images FPM d'intensité reconstruites. c) mire de résolution USAF1951 pour différentes positions de l'objectif 40xNA0.95 avec LED centrale. La comparaison de b) et c) montre une résolution comparable. d) Images d'intensité et de phase FPM obtenues sur un globule rouge parasité. La flèche pointe vers un compartiment qui n'est pas visible dans l'image d'intensité.	76
3.2	Images d'intensité (I) et de phase (Φ) FPM de globules rouges colorés acquises pour différents réglages de mise au point (allant de $z = -3\mu m$ to $z = +3\mu m$) pour $\lambda = 525nm$, notez l'évolution du contraste relatif d'un composant de la cellule parasite (hémozoïne) indiqué par la flèche. Les images encadrées dans la boîte de droite correspondent aux images optimales obtenues avec une compensation de mise au point numérique e-PIE ($\Delta z_{num} = 0.7\mu m$ for I, $\Delta z_{num} = -1.7\mu m$ for Φ).	78
3.3	Architecture du U-Net utilisée pour la tâche de refocalisation	81
3.4	Construction de la base de données d'entraînement U-Net. Les échantillons sont numérisés (1-3) avec FPM sous différentes conditions de mise au point (de $z = -8\mu m$ à $z = 8\mu m$). Pour toutes les valeurs de z , les images FPM sont reconstruites sans compensation de mise au point numérique (étapes 4a et 5a) pour produire les exemples d'entrée U-Net. Pour chacune de ces entrées, les images de phase de référence I et Φ sont obtenues après les étapes 4b et 5b (l'image de sortie souhaitée) et sont utilisées pour apprendre la correspondance et la transformation entre les entrées et les sorties.	82

3.5	(En haut au centre) Images FPM de frottis sanguin reconstruites pour différentes distances entre l'échantillon et l'objectif du microscope (de $z = -8\mu m$ à $z = 8\mu m$). Les images de référence (boîte en haut à droite) sont obtenues en suivant la branche de droite de la figure 3.4 en utilisant $\Delta z_{num} = -0.6\mu m$ $\Delta z_{num} = 0.8\mu m$ pour l'intensité et la phase respectivement; (en bas) les images correspondantes corrigées par U-Net.	84
3.6	Des exemples illustratifs d'images FPM de frottis sanguin enregistrées pour différentes distances entre l'échantillon et l'objectif du microscope (de $z = -8\mu m$ à $z = 8\mu m$). Les images de référence (boîte à droite) sont obtenues en suivant la branche de droite de la figure 3.4 en utilisant $\Delta z_{num} = -0.6\mu m$ $\Delta z_{num} = 0.8\mu m$ pour l'intensité et la phase respectivement; (en haut) les images correspondantes corrigées par U-Net bi-modal utilisant l'intensité et la phase; (en bas) les images d'intensité corrigées par U-Net qui a appris sur l'intensité seule.	85
3.7	Champ de vision complet d'un échantillon de frottis sanguin acquis en FPM. La superposition de la carte de couleurs indique la variation de la distance de séparation entre l'échantillon et l'objectif du microscope, causée par une inclinaison du porte-lames. Les deux encarts relatifs aux zones A et B présentent des images FPM avant et après correction par U-Net. La région A est naturellement bien mise au point ($z \sim 0$). La région B est initialement hors de mise au point ($z \sim 2.5\mu m$). Les flèches indiquent des compartiments sous-cellulaires de parasites dont la visibilité est renforcée grâce à U-Net.	87
3.8	Évolution de la sensibilité de détection de Plasmodium falciparum en fonction du paramètre de seuil sur la base de données de test. Les courbes en pointillés correspondent aux résultats obtenus sur des images brutes reconstruites par FPM sans le module de compensation U-Net (pour $z = 0\mu m$ et $z = 2\mu m$). Les courbes pleines correspondent aux résultats après la compensation des images par U-Net (bi-modal). b) Courbe ROC révélant les meilleurs points d'exploitation (cercles) définis par le meilleur compromis entre la sensibilité et la spécificité	91
3.9	Position des erreurs de détection pour $z = 2\mu m$ évaluée par l'architecture Yolo. a) Sans correction d'image U-Net (à droite). b) Avec correction U-Net bi-modal. La région délimitée par le cadre correspond à une région qui est correctement mise au point de manière native (pour $z = 2\mu m$).	92

3.10	Évaluation de la sensibilité et de la spécificité du modèle Yolo dans la plage de mise au point allant de $z = -4\mu m$ à $z = 4\mu m$, avec un intervalle de $2\mu m$. Nous avons comparé le modèle Yolo avec correction U-Net bimodal, le modèle YOLO sans correction, et le modèle YOLOMZ sans correction.	93
3.11	Nous avons évalué l'amélioration de la qualité d'image obtenue en utilisant le modèle U-Net pour une mise au point $z = 2.25\mu m$ et pour les 6 canaux RGB en intensité et en phase. Les images en haut sont les images optimales sélectionnées après empilement virtuel en profondeur (z-stacking) pour $z=0$. Au milieu, nous présentons les images de reconstruction obtenues pour une mise au point $z = 2.25\mu m$ sans amélioration. En bas, nous montrons les images qui ont été améliorées à l'aide du modèle U-Net.	97
3.12	Nous illustrons ici l'amélioration de la qualité d'image obtenue à l'aide du modèle U-Net pour différents niveaux de mise au point, allant de $z = -2.25\mu m$ à $z = 2.25\mu m$. La figure présente des images d'intensité combinées en une seule image RGB. En haut, nous montrons les résultats obtenus pour des polynucléaires neutrophiles, en bas pour des lymphocytes	98
4.1	Modèle cDIP : LR désigne la pile d'images FPM de faible résolution mesurées au niveau du plan de la caméra en utilisant les LEDs \mathcal{M} . Θ représente les paramètres du U-Net qui sont optimisés après chaque itération de calculs de rétropropagation. I' and Φ' représentent les images HR reconstruites avec U-Net à chaque itération. Les images reconstruites I, Φ sont obtenues lorsque le modèle a convergé. LR' est la pile d'images de faible résolution calculée par le modèle direct. La fonction de cout \mathcal{L}	104
4.2	L'architecture du U-Net utilisée pour résoudre le problème inverse selon les différents schémas.	106
4.3	Présentations schématiques du microscope équipé d'une matrice de LEDs $\mathcal{M} = \mathcal{M}_A \cup \mathcal{M}_B$ (a), région spectrale couverte dans le domaine de Fourier avec \mathcal{M}_A (b) et avec \mathcal{M} (c).	107
4.4	Comparaison visuelle des images reconstruites par ePIE et cDIP pour différentes conditions expérimentales. a) pour $z = 0\mu m$ et $\Delta z_{\text{num}} = 0\mu m$, b) pour $z = -70\mu m$ et $\Delta z_{\text{num}} = 0\mu m$, c) pour $z = -70\mu m$ et $\Delta z_{\text{num}} = -70\mu m$ et d) le couple intensité-phase de référence	108
4.5	Schema cDIP adapté pour inclure une correction d'aberration	109

- 4.6 Resultat cDIP $\Gamma = 60\%$ avec correction d'aberrations.a) Reconstruction cDIP de l'intensité et de la phase sans la reconstruction de la fonction d'aberration. b) Reconstruction cDIP de l'intensité, de la phase et de la fonction d'aberration. c) Intensité, phase et fonction d'aberration de référence. 111
- 4.7 Reconstruction FPM dans différentes expérimentales : Intensité et phase obtenues avec un recouvrement de $\Gamma = 60\%$ (130 LEDs).b) Intensité et phase obtenues avec un recouvrement de $\Gamma = 20\%$ (37 LEDs). c) Intensité et phase de référence. 112
- 5.1 Modèle cDIP-L (modèle d'apprentissage U-Net). LDS : ensemble de données d'apprentissage composé de nombreux couples d'images d'intensité I et de phase Φ . Le modèle FPM direct est utilisé pour calculer des images à basse résolution formées au niveau du plan de la caméra. \mathcal{M}_A et \mathcal{M}_B représentent la division de la matrice de LED en deux groupes. LR : pile d'images à basse résolution formées au niveau du plan de la caméra pour les différentes illuminations utilisées. Θ représente les paramètres du U-Net. LR' : images à basse résolution calculées à partir des images prédites par le U-Net I' et Φ' (en gris foncé pour \mathcal{M}_A , gris clair pour \mathcal{M}_B). \mathcal{L} : fonction de perte. 117
- 5.2 Comparaison des résultats de reconstruction obtenus avec des éclairages et un chevauchement d'environ $\sim 10\%$. Les images (**a-e**) correspondent aux images d'intensité, tandis que les images (**f-i**) correspondent aux images de phase. (**a**) Image brute acquise avec la LED centrale, (**b**) image de référence, (**c**) reconstruction ePIE (en utilisant les LED \mathcal{M}_A), (**d**) reconstruction cDIPL après entraînement (en utilisant les LED \mathcal{M}_A), (**e**) reconstruction cDIP-LO après entraînement et optimisation (en utilisant les LED \mathcal{M}_A). (**f**) correspond à la phase de référence, et (**g-i**) à la phase reconstruite avec e-PIE, cDIP-L et cDIP-LO, respectivement. Les flèches indiquent différents types d'artefacts (tels que le crosstalk ou la dégradation de la résolution). Les images de référence (**b,f**) sont obtenues à partir de la reconstruction ePIE en utilisant toutes les piles d'images (LEDs \mathcal{M}). 120

5.3	comparaison des reconstructions de la mire de résolution d'intensité USAF 1951 obtenues avec l'éclairage (recouvrement $\sim 10\%$). Pour les images (a–f), la mire de résolution est placée sur le plan focal du microscope. (a) Référence, (b) image brute acquise avec l'éclairage de la LED centrale, (c) ePIE, (d) cDIP-L, (e) cDIP-O, et (f) reconstructions cDIP-LO. Pour les images (g,h), la mire de résolution est placée à $8\mu m$ au-dessus du plan focal du microscope. (g) Référence (identique à l'image (a)), (h) LED centrale, (i) ePIE, (j) cDIP-L, (k) cDIP-O, et (l) reconstructions cDIP-LO. La référence est calculée en utilisant la reconstruction ePIE avec l'éclairage des LEDs \mathcal{M}	122
6.1	Régions spectrales couvertes dans l'espace de Fourier ;(a) la matrice de LEDs \mathcal{M} , présentant un recouvrement spectral de $\Gamma \approx 66\%$, et (b) la matrice de LEDs \mathcal{M}_A , avec un recouvrement spectral de $\Gamma \approx 20\%$	127
6.2	Illustration des images reconstruites d'un globule blanc et de globules rouges parasités à $z = 2.25\mu m$ avec et sans ajustements de la mise au point via la correction du front d'onde du Δz_{num} . a) Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 66\%$ en utilisant l'algorithme ePIE sont comparées à celles obtenues avec un $\Gamma = 20\%$, également en utilisant les algorithmes b) ePIE et c) cDIP-LO.	128
6.3	Illustration des images d'un globule blanc et de globules rouges parasités reconstruites à $z = 2.25\mu m$ avec et sans amélioration d'images par modèle U-Net FPM du Chapitre 3.a) Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 66\%$ en utilisant l'algorithme ePIE. Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 20\%$, b) en utilisant l'algorithme ePIE et c) en utilisant le modèle cDIP-LO.	129
6.4	Schéma d'apprentissage du modèle cDIP-L	131
6.5	Analyse comparative de la reconstruction d'un globule blanc polynucléaire neutrophile à $z = 0\mu m$. a) Résultats ePIE dans un cas de faible recouvrement $\Gamma = 20\%$ avec et sans correction de Δz_{num} . b) Résultats du modèle cDIP-L avec $\Gamma = 20\%$. c) Résultats ePIE dans un cas de fort recouvrement $\Gamma = 66\%$ avec Δz_{num} optimal.	133
6.6	Analyse comparative de la reconstruction d'un globule blanc polynucléaire neutrophile et des globules rouges parasités à $z = 2.25\mu m$. a) et c) Résultats ePIE dans un cas de faible recouvrement $\Gamma = 20\%$. b) et d) Résultats du modèle cDIP-L avec $\Gamma = 20\%$. c) Résultats ePIE dans un pour $\Gamma = 66\%$ et Δz_{num} optimal.	134

6.7	Analyse comparative des images à différentes mises au point pour la diode centrale (basse résolution) et des images d'intensité reconstruites à partir de 4 LEDs avec $\Gamma = 20\%$ en utilisant les algorithmes ePIE et cDIP-L	135
-----	---	-----

Liste des tableaux

3.1	Le nombre de globules rouges sains et infectés dans les ensembles d'apprentissage, de validation et de test.	89
3.2	Les performances de détection de <i>Plasmodium falciparum</i> sur l'ensemble de test, obtenues sur des images brutes reconstruites par FPM avec et sans le module de compensation U-Net bi-modal et U-Net intensité (pour $z = 0\mu m$ et $z = 2\mu m$).	90
5.1	Évaluation des métriques \mathcal{L} , SSIM et PSNR sur l'ensemble de données TDS. Les moyennes et écart-types des métriques sont présentés. Les reconstructions sont obtenues à partir de la pile d'images LR générées avec les LED de \mathcal{M}_A	122

Chapitre 1

Introduction

1.1 Contexte général

Mon travail s'inscrit dans le cadre du partenariat stratégique Télécom Sud-Paris "TSP" et TRIBVN/T-life, dédié au développement de nouvelles approches en microscopie optique, couplées à l'intelligence artificielle, en vue d'identifier, de prédire et de monitorer les pathologies hématologiques et parasitologiques.

La numérisation de spécimens biologiques appelée « Digital Pathology » (pathologie numérique) à l'aide de scanners de lames à une résolution microscopique sur de grandes dimensions connaît une popularité croissante indéniable depuis plusieurs années. Cet engouement s'explique par la capacité des scanners actuels à numériser des images de lames entières, à stocker, archiver et échanger numériquement des échantillons entre spécialistes médicaux. Il s'explique également par l'émergence de diagnostics automatisés utilisant, entre autres, des algorithmes de Machine-Learning (ML).

Le diagnostic biologique des hémopathies repose essentiellement sur l'héogramme, une énumération automatisée des éléments du sang; ce diagnostic est complété en première instance par un examen cytologique du frottis sanguin au microscope. Différents automates ont été développés pour faciliter l'examen du frottis au microscope. Ils visent à apporter aux médecins une aide au diagnostic en caractérisant et quantifiant morphologiquement les cellules du sang : leucocytes (globules blancs), hématies (globules rouges) et plaquettes. La « Digital Pathology » n'a, pour le moment, pas su introduire une captation de données suffisamment discriminantes pour qu'elle puisse être facilement exploitées de façon entièrement automatisée. Ce qui oblige les cliniciens à reprendre l'analyse manuellement, mais aussi les oblige à d'autres méthodes, tel que : examen microscopique de la moelle osseuse, cytométrie en flux, cytogénétique et génomique. Pour progresser, il devient primordial, d'une part, de promouvoir des techniques d'acquisition d'imagerie de

frottis beaucoup plus riches et donc plus massives et, d'autre part, de mettre en œuvre des méthodes et des outils qui permettent d'en extraire des informations pertinentes de façon automatisée.

Un vaste éventail de techniques d'imagerie non conventionnelles, telles que l'holographie numérique et la microscopie calculatoire, ont vu le jour. Ces approches facilitent l'accès à des modalités d'imagerie plus riches et quantitatives, comme la phase, promettant une meilleure caractérisation des objets biologiques avec des dispositifs simples.

Parmi celles-ci, la microscopie ptychographique de Fourier (FPM) [95] est assez singulière. Rapportée par Zheng et al en 2013, il s'agit d'une méthode de reconstruction séquentielle, qui combine des images d'intensité en basse résolution acquises selon différents angles d'éclairage, afin de pouvoir reconstruire des représentations en phase et intensité avec un facteur de super-résolution pouvant être élevé (jusqu'à 6) rendant accessible l'utilisation d'objectifs de faible grandissement (magnification en anglais), et donc de champ de vue étendu.

La reconstruction d'images de phase et d'intensité est réalisée en résolvant un problème inverse associé à une série d'images FPM, chacune acquise via une LED différente. Les algorithmes employés s'appuient sur des techniques d'optimisation, basées sur la descente de gradient, pour minimiser une fonction d'erreur de manière séquentielle pour chaque LED.

Toutefois, pour que cette technologie puisse être pleinement adoptée dans les processus diagnostiques et intégrée au sein de la Digital Pathology, il est vital de résoudre les défis spécifiques associés à l'utilisation efficace de la FPM. Cette démarche nécessite d'examiner et de traiter diverses limitations, telles que :

La Correction numérique du focus : Grâce aux techniques de super-résolution, la FPM s'impose comme une méthode permettant l'accès à de larges champs de vue en utilisant des objectifs à faible grandissement. Toutefois, l'obtention d'une netteté uniforme sur un large champ de vue se révèle être une tâche compliquée. Cette difficulté est accentuée par les aberrations chromatiques qui induisent des variations de mise au point spécifiques à chaque longueur d'onde. En conséquence, une correction numérique automatisée de la mise au point s'avère indispensable pour mettre en évidence les compartiments cellulaires étudiés.

Le temps d'acquisition en microscopie FPM est directement proportionnel au nombre de LED utilisées et au temps d'exposition alloué à chaque LED. Ce défi temporel devient d'autant plus critique lorsque l'on cherche à avoir des facteurs de super-résolution qui nécessitent l'acquisition d'images en champ sombre. Dans ces conditions, les temps d'exposition par LED tendent à être significativement prolongés. Pour surmonter cette contrainte, des études ont été menées, y compris celles réalisées par Laura Waller et al[84]. Ceci dit, cette question mérite encore d'être considérée pour des outils automatiques viables.

Temps de reconstruction : les algorithmes traditionnels de la ptychographie de Fourier sont itératifs et effectuent une mise à jour séquentielle de l'objet reconstruit, LED par LED. Cette approche nécessite une correction qui suppose l'existence d'un chevauchement spectral entre les LED successives. Cette condition rend difficile toute tentative de parallélisation informatique et conduit à des durées de reconstruction potentiellement longues.

La Correction numérique des aberrations globales du système optique peut s'avérer cruciale pour atteindre la résolution cible en microscopie FPM. Cette correction a pour but de compenser les imperfections de l'objectif et du système de mesure, une préoccupation commune à toutes les méthodes d'imagerie optique. Bien que des approches aient été proposées dans la littérature, cette correction demeure complexe, exigeant des ajustements spécifiques pour chaque objet sans pour autant garantir la résolution de tous les problèmes associés.

Le Bruit émane principalement des conditions d'éclairage et des caractéristiques de la caméra. Ce type de bruit affecte significativement les modèles de reconstruction, surtout lorsque des facteurs élevés de super-résolution sont visés.

Les applications visées dans la thèse concernent deux domaines dans lesquels l'imagerie microscopique ne permet pas encore un diagnostic efficace. Il s'agit de l'analyse des frottis sanguins pour la détection des hémopathies telles que les leucémies à partir des leucocytes (globules blancs) ou du paludisme à partir des parasites infectant les globules rouges.

Nos recherches sur le paludisme sont menées en partenariat avec le centre référence du paludisme, hôpital Pitié-Salpêtrière, tandis que les études portant sur les leucocytes sont réalisées conjointement avec l'hôpital Avicenne, dans le cadre du projet TAMIS[2]¹ et sont traitées dans le cadre du laboratoire commun OT4D avec l'hôpital Saint-Antoine [1]².

1.2 Objectif de la thèse

Le travail réalisé durant cette thèse a été motivé par deux objectifs principaux : d'une part, nous cherchons à pallier certaines des limitations des technologies précédemment mentionnées, d'autre part, nous explorons les avantages de l'information bimodale obtenue par la microscopie ptychographique de Fourier (FPM) pour la détection et la classification des cellules sanguines, dans le but de diagnostiquer

1. Projet financé par BPI France et HDH dans le cadre de travaux collaboratif de TSP, TRIBVN, hôpital Avicenne

2. Laboratoire commun financé par l'ANR entre TRIBVN et Sorbonne université/hôpital Saint-Antoine

certaines pathologies. Les limitations clés examinées comprennent la correction du focus, le temps d'acquisition et le temps de reconstruction, qui sont des aspects cruciaux pour la mise en pratique de cette technologie dans le domaine de la pathologie numérique.

Pour atteindre ces objectifs, nous avons développé des stratégies qui s'appuient sur des architectures de réseaux neuronaux. Ces stratégies exploitent à la fois l'apprentissage statistique et les équations physiques. En effet, nos modèles d'optimisation prennent en compte les équations physiques du modèle direct associées à la FPM et de ce fait permettent d'obtenir une solution de reconstruction adéquate. Nous avons, de plus, exploré l'apport de l'apprentissage statistique en complément de l'optimisation seule. L'apprentissage statistique nous permet de capturer des relations complexes qui sont difficilement modélisables par des formules mathématiques conventionnelles.

Ce travail est structuré en sept chapitres distincts, chacun consacré à un aspect spécifique des défis et des approches que j'ai explorées.

Dans le chapitre deux de cette thèse, je définis le cadre pluridisciplinaire dans lequel mon travail prend place. J'aborde un éventail de techniques, qui s'étendent de la microscopie conventionnelle à la microscopie à ptychographique de Fourier (FPM). Je présente aussi les modèles neuronaux qui sont mis en œuvre au cours de cette thèse. Le cadre biologique des applications visées est également mis en avant, centré spécifiquement sur l'étude des leucocytes et de la parasitologie.

Dans le chapitre 3, j'aborde le défi d'améliorer la qualité des images obtenues par microscopie ptychographique de Fourier (FPM), tant en phase qu'en intensité. Ceci permet une meilleure exploitation et analyse des larges champs de vue. J'ai mis en œuvre un réseau de neurones U-Net qui permet d'optimiser la focalisation et la qualité d'image, facilitant ainsi une observation plus précise des compartiments cellulaires. Ces travaux sont validés dans le contexte de parasitologie par une évaluation visuelle et quantitative sur la détection des parasites. J'applique également notre méthode au cas des leucocytes en utilisant un éclairage comprenant des longueurs d'ondes en rouge, vert et bleu (R, V, B). Cette approche permet non seulement d'améliorer la qualité des images pour chacune de ces longueurs d'ondes, mais elle prouve également son efficacité dans la correction des décalages de focus liés aux aberrations chromatiques.

Dans le quatrième chapitre, j'examine le changement de paradigme résultant de l'adoption de réseaux neuronaux de type cDIP régis par les équations physiques du modèle FPM, pour la reconstruction des images FPM. J'évalue cette nouvelle approche sous diverses conditions expérimentales, en prenant en compte plusieurs propriétés, telles que la refocalisation par correction de front d'onde et la correction des aberrations.

Le cinquième chapitre aborde une nouvelle stratégie qui repose sur un modèle

neuronal conçu pour permettre une reconstruction de haute qualité avec un nombre réduit de LED, dans le but d'accélérer le processus de production des images. Cette approche s'appuie sur une méthode d'apprentissage statistique destinée à la reconstruction d'images en utilisant un modèle cDIP. Cet apprentissage est ensuite affiné par une étape d'optimisation spécialement adaptée pour chaque cas étudié. Dans cette étude, je présente des résultats obtenus par cette méthode sur des données de simulations, caractérisées par une faible ouverture numérique initiale et un facteur de super-résolution élevé.

Le chapitre 6 s'intéresse à une configuration de microscope particulièrement adaptée aux problématiques d'analyse sanguine en hématologie.

Il s'agit d'une configuration visant une très haute résolution grâce à une Ouverture Numérique synthétique obtenue par la FPM (typiquement $NA_{\text{synthétique}} \geq 1.5$). Elle repose sur l'utilisation d'objectifs de microscope ayant une forte Ouverture Numérique nominale ($NA \geq 0.75$).

Le facteur de super-résolution est alors inférieur à ~ 2 conformément à la théorie. Pour cette configuration, j'étudie le comportement par simulation et expérimentalement des approches développées en chapitre 5 et chapitre 3. Je montre que certaines limitations sont rencontrées sur la qualité finale des images obtenues en raison de la présence de trous importants dans le spectre couvert par la matrice de LED utilisée (pour un facteur de recouvrement de l'ordre de 20%). Je détaille alors une dernière approche de reconstruction et refocalisation des images adaptée à cette situation expérimentale particulière. En particulier, je montre expérimentalement l'apport d'un apprentissage de reconstruction et refocalisation au sein d'un unique réseau CNN. Un tel apprentissage est rendu possible en raison du plus faible facteur de super-résolution visé.

Le chapitre 7 sert de conclusion à la thèse, résumant les principales contributions aux domaines de la ptychographie de Fourier et du diagnostic biologique par microscopie de nouvelle génération, révélant l'apport des techniques neuronales dans ce cadre. Ce chapitre énonce également les perspectives de recherche futures qui pourraient enrichir davantage ces domaines.

Chapitre 2

Contexte pluridisciplinaire

Dans ce chapitre, je présente les concepts et techniques autour desquels j'ai orienté ma thèse. En premier lieu, je définis l'imagerie computationnelle et la compare à l'imagerie classique. Ensuite, j'introduis quelques notions relatives à la microscopie optique. Dans une troisième partie, je présente la FPM et les algorithmes de reconstruction qui lui sont associés. Puis, je traite des réseaux de neurones et détaille les architectures que j'ai utilisées au cours de mes recherches. Enfin, j'expose les cas biologiques sur lesquels nous avons travaillé.

2.1 Imagerie computationnelle

Dans le monde physique et tangible qui nous entoure, l'homme est en constante quête d'observation de phénomènes physiques. Qu'il s'agisse d'observer des organismes microscopiques à travers la lentille précise d'un microscope, ou d'explorer l'immensité stellaire grâce à un télescope, l'homme a su inventer une multitude d'instruments qui lui permettent de capter directement l'image de phénomènes observables. Ces outils nous offrent un accès direct à diverses dimensions de notre monde, sans nécessiter aucun processus de calcul, que ce soit en l'observant à l'œil nu ou en visualisant le signal mesuré sur un écran d'ordinateur.

Afin d'accéder à des informations de plus en plus riches et de minimiser la complexité des instruments d'observation traditionnels, de grandes avancées ont pu être obtenues grâce à des approches numériques. Ces dernières reposent sur des méthodes d'observation indirectes obtenues par différentes manipulations de données. L'imagerie computationnelle en est un exemple.

Cette discipline permet notamment la reconstruction d'images quantitatives à partir de signaux mesurés en se basant sur des modèles physiques et mathématiques, utilisant des techniques informatiques avancées d'optimisation ou bien grâce à l'apprentissage automatique pour un meilleur traitement des vastes volumes de

données disponibles.

Les images de nouvelle génération, issues de l'imagerie computationnelle, sont calculées à partir des mesures effectuées sur des systèmes d'imagerie conçus conjointement sur le plan matériel et logiciel Figure 2.1. Ces images ont permis de surpasser les limites de l'imagerie conventionnelle, facilitant ainsi l'accès à une plus grande quantité d'informations : meilleure résolution spatiale, caractérisation physique, chimique et moléculaire des objets d'intérêt.

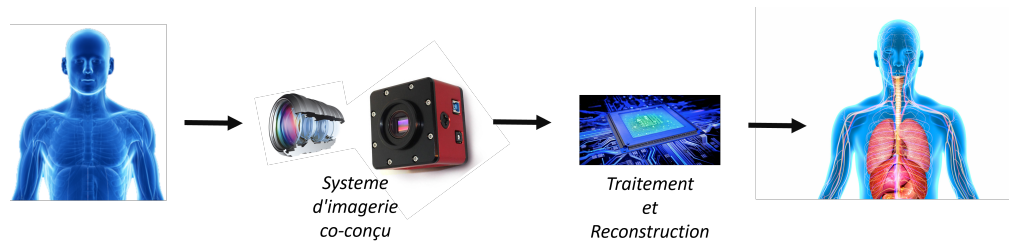


FIGURE 2.1 – Processus de mesure et de reconstruction pour l'imagerie non conventionnelle

Les types d'imagerie computationnelle couvrent un large spectre de domaines, notamment la microscopie numérique, tel que la Ptychographie[76], la microscopie par ptychographie dans l'espace de Fourier, ou FPM pour Fourier Ptychographic microscopy [95], l'imagerie tomographique mais aussi diverses autres applications : l'IRM [47], l'imagerie par ultrasons [22], le radar à ouverture synthétique (SAR)[24, 31], l'imagerie sismique [17], etc.

Les techniques déployés pour ces types d'imagerie sont généralement complexes et nécessitent une compréhension approfondie du système de mesure. Cela s'étend de la phase d'excitation de l'échantillon, utilisant l'onde d'éclairage, jusqu'à la capture du signal, en modélisant toutes les interactions au sein du système. Cette modélisation détaillée permet de développer des algorithmes capables de reconstruire, à partir du signal mesuré, une information plus riche.

Toutefois, plusieurs facteurs tels que la complexité des modèles physiques, la non-linéarité des processus physiques, les contraintes géométriques de l'imagerie, le bruit de mesure, les effets de diffusion, et bien d'autres, rendent la modélisation complète du modèle physique difficile. Cela rend souvent impossible la résolution de problème par des méthodes analytiques classiques et nécessite l'utilisation de techniques d'optimisation numérique. Ces dernières cherchent à minimiser une fonction d'erreur entre les données estimées par le modèle de reconstruction et les données mesurées par le système. En imagerie computationnelle, l'objectif des algorithmes d'optimisation est de résoudre le problème inverse, qui consiste à reconstituer l'image d'origine à partir des données recueillies par les capteurs. Par définition, un problème inverse est un problème mathématique qui vise à remonter

aux causes d'un effet observé. Il consiste à identifier les caractéristiques de l'image de l'objet recherché en utilisant le signal mesuré et en exploitant des informations sur le système. La résolution de problèmes inverses est souvent une tâche complexe. En effet, une seule mesure peut donner lieu à de multiples solutions possibles. De surcroît, les mesures sont fréquemment sujettes à des erreurs, des incertitudes et à divers bruits, complexifiant encore davantage la résolution du problème. Afin de résoudre ces problèmes inverses, on fait souvent appel à des techniques de régularisation. Celles-ci intègrent des fonctions de pondération pour contraindre le système et, par conséquent, réduire le nombre de solutions possibles. Ainsi, les solutions obtenues coïncident avec des informations partielles mais observables du système étudié (échantillon-microscope-éclairage).

Les problèmes inverses sont donc un domaine de recherche important pour de nombreux scientifiques, ingénieurs et chercheurs dans divers domaines. Pour résoudre un problème inverse, le point de départ consiste en la modélisation du problème direct, qui décrit comment les paramètres du modèle se manifestent sous forme d'observations expérimentales. La solution du problème inverse est traditionnellement obtenue soit par une approche d'optimisation numérique, soit par une résolution analytique. Par ailleurs, le problème à résoudre est souvent mal posé (optimisation non convexe), en grande partie en raison de la présence de modèles physiques incomplets et de la nécessité de conditions initiales précises.

Malgré les défis inhérents à la résolution de problèmes inverses dans le domaine de l'imagerie computationnelle, diverses stratégies et techniques mathématiques et informatiques ont été développées et mises en œuvre pour surmonter ces obstacles. De plus, les outils d'IA jouent un rôle de plus en plus important dans la pleine exploitation du potentiel de ces technologies. Ils permettent une analyse statistique plus avancée et plus discriminante des données, ainsi que le développement de nouvelles approches de reconstruction visant à dépasser les techniques d'optimisation classiques.

2.2 Microscope optique

L'émergence de microscopes de nouvelle génération marque une avancée significative dans le domaine de la science des matériaux et de la biologie. Ces outils innovants cherchent à fusionner les méthodes de calcul avancées et les techniques non conventionnelles, tout en s'appuyant sur les principes fondamentaux de la microscopie traditionnelle, y compris les microscopes optiques et électroniques. La compréhension du fonctionnement de la FPM, une de ces technologies émergentes, requiert une familiarité avec les principes fondamentaux de la microscopie optique et une connaissance approfondie du rôle de chaque composant du système optique. La microscopie est une technique d'observation qui a permis de grandes avancées

dans notre compréhension des petites structures, et ce, dans différents domaines. Elle a rendu possible l'exploration des organismes biologiques à diverses échelles Figure 2.2, permettant de voir des cellules, des tissus, des organes, des microbes, des virus, des protéines, des acides nucléiques et d'autres structures biologiques.

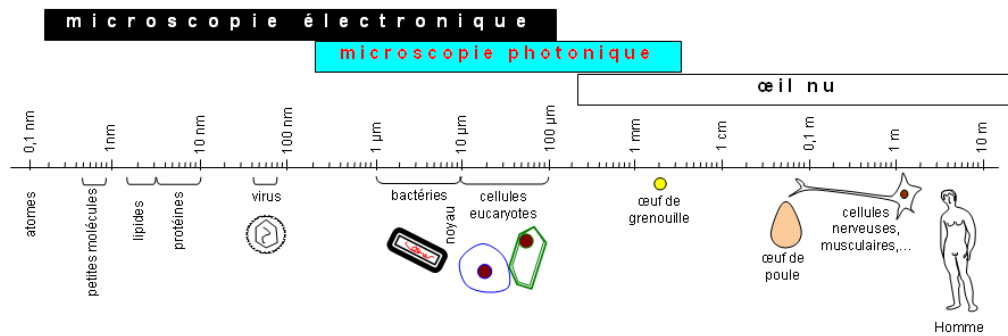


FIGURE 2.2 – Structures du vivant présentées à diverses échelles [3].

Le microscope photonique ou optique est le plus ancien des microscopes ; inventé à la fin du XVII siècle ; il permet de voir les échantillons biologiques qui peuvent être observés en temps réel, sans être endommagés par le rayonnement, voir des objets jusqu'à une résolution d'environ 200 nm. La théorie de la formation de l'image microscopique a été définie par Ernst Abbe (1873)[6].

Selon ce modèle théorique, la lumière traversant un objet subit une diffraction, se manifestant par une structure conique. Cette lumière, une fois déviée, devient un vecteur riche en informations, révélant les détails et les caractéristiques intrinsèques de l'objet. À l'inverse, la lumière qui ne subit pas cette déviation ou dispersion maintient sa trajectoire et structure originelles, ne reflétant ainsi aucune caractéristique de l'objet analysé. Cette interaction entre lumière et matière est cruciale pour la formation de l'image au microscope, où les nuances de diffraction sont essentielles pour l'interprétation et la compréhension des structures observées.

Ainsi dans la formation des images de microscopie, plusieurs éléments entrent en jeu comme l'illustre le dispositif 2.3. On peut les découper en trois niveaux : source de lumière, objectif et d'une caméra ou/et un oculaire.

2.2.1 La source de lumière

En imagerie FPM comme en microscopie photonique traditionnelle, la source lumineuse assume une importance primordiale. Elle influe sur divers paramètres clés tels que la résolution de l'image et le contraste, facilitant ainsi la différenciation entre les divers éléments de l'échantillon observé. De plus, pour les spécimens

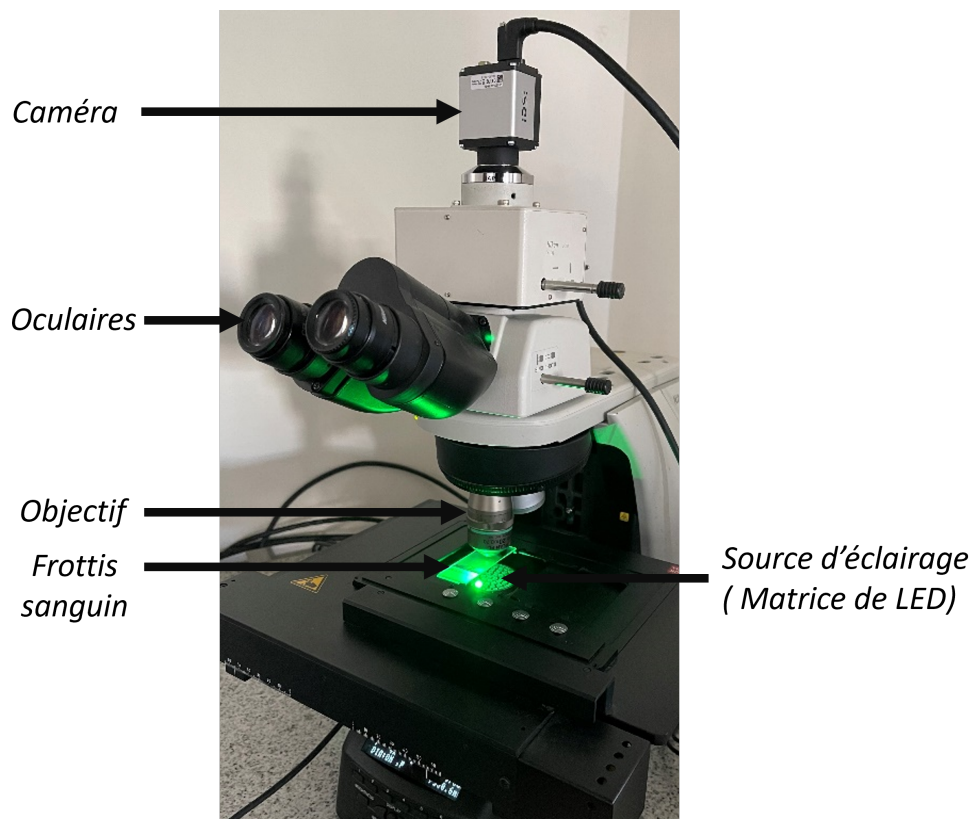


FIGURE 2.3 – Dispositif de microscopie optique intégrant une matrice de LED, un objectif, des oculaires et une caméra.

biologiques marqués¹, le choix des longueurs d'onde de la source lumineuse est crucial. En effet, cela permet une meilleure distinction des compartiments cellulaires. On identifie généralement deux catégories principales d'éclairage :

a) Eclairage Köhler

Les premiers microscopes se basaient sur l'utilisation de lentilles spécialement conçues pour capter et réfracter la lumière solaire en tant que source d'éclairage. Cette méthode présentait cependant des inconvénients, notamment en termes de variabilité et de complexité de contrôle de l'éclairage. Grâce à l'évolution technologique, une procédure plus standardisée et efficace, nommée "méthode Köhler", a été introduite par August Köhler en 1893 [46].

Cette méthode, largement utilisée en microscopie jusqu'à aujourd'hui, vise à

1. Le marquage désigne l'utilisation de marqueurs chimiques ou biologiques dans le but d'identifier et de différencier des composants spécifiques au sein de l'échantillon

optimiser l'éclairage des échantillons. Elle repose sur l'utilisation d'une lampe ou de LED, suivie d'une optique spécifique et d'un condenseur. En conséquence, chaque point de la source éclaire l'intégralité de l'objet, garantissant ainsi une illumination homogène sur tout le champ de vision. Cela implique que cette source de lumière ne possède pas de cohérence spatiale.

De plus, l'éclairage Köhler permet de régler de manière indépendante la taille du champ éclairé sur la préparation ainsi que l'angle du cône de lumière (l'ouverture numérique de l'éclairage) frappant chaque point de la préparation grâce au diaphragme d'ouverture. Lorsqu'un diaphragme d'ouverture est très fermé, au point d'être presque ponctuel, l'échantillon est illuminé par une onde pratiquement plane provenant d'une seule direction verticale, ce qui le rend partiellement cohérent.

b) LED et laser (éclairage cohérent)

Les autres types d'éclairage que l'on peut rencontrer consistent en des sources lumineuses basées sur des technologies telles que les lasers de LED (diodes électroluminescentes). Ces formes d'éclairage sont particulièrement prisées dans des domaines de microscopie non conventionnelle, en raison de leurs propriétés distinctes comparées à la méthode d'éclairage de Köhler.

Ces sources permettent de générer une lumière cohérente avec des ondes planes, ainsi qu'un rayonnement spectral plus pur (monochromatique). Ces types d'éclairage sont également employés en imagerie à contraste de phase ou en imagerie interférentielle telles que l'holographie ou la ptychographie.

Dans le cadre des travaux menés sur la ptychographie de Fourier au sein de cette thèse, les sources lumineuses sélectionnées reposent sur l'utilisation de LED, choisies pour leur cohérence spatiale suffisante ainsi que pour leur facilité de manipulation, offrant ainsi une combinaison optimale de performance et de praticité dans la conduite des expériences.

Les LEDs sont disposées dans des matrices qui peuvent être conçues selon plusieurs schémas, sous une forme circulaire, rectangulaire, en dôme, etc. Le choix de la géométrie des LED sélectionnées dépendra du facteur de la super-résolution que l'on souhaite obtenir, ainsi que des contraintes géométriques propres au dispositif.

Par exemple, le schéma (a) de la Figure 2.4 illustre une matrice de 225 LEDs, souvent utilisée avec des objectifs de microscopie à faible ouverture numérique, afin d'obtenir des images avec une variabilité d'angles et de distances importante, permettant ainsi d'obtenir des facteurs de super-résolution allant jusqu'à 6. Le schéma (b) de la Figure 2.4, quant à lui, montre une matrice de 59 LEDs avec 13 LEDs sélectionnées : une diode centrale et un anneau de 12 LEDs. Ce type de matrice est plus adapté aux applications pour lesquelles le facteur de super-résolution est plus restreint (typiquement un facteur 2).

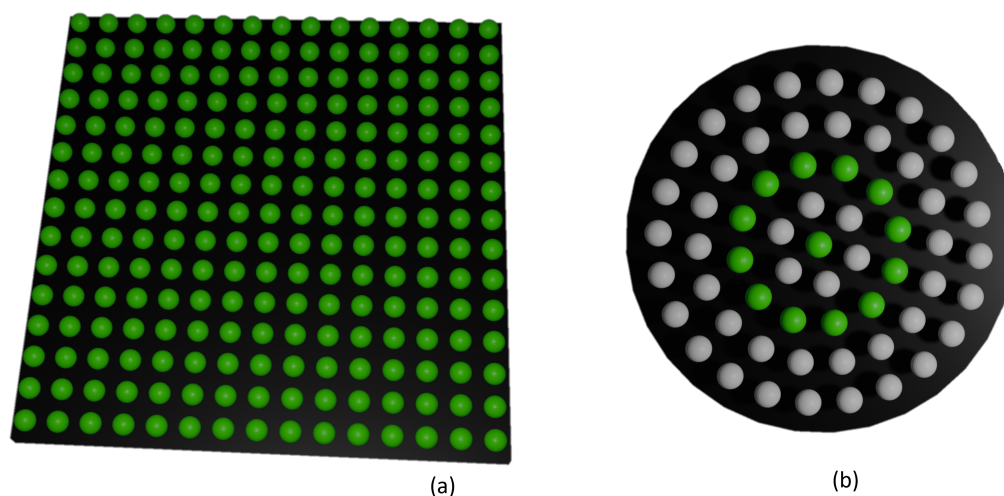


FIGURE 2.4 – Exemples de dispositions des LED pour deux géométries de matrices : les LED sélectionnées sont représentées en vert, tandis que les LED non utilisées sont indiquées en noir.

2.2.2 L'objectif du microscope

Dans le domaine de la microscopie optique ou celle de la FPM, l'objectif se distingue en tant qu'un de ses composants les plus complexes et essentiels. En effet, c'est l'objectif qui est en grande partie responsable de la formation de l'image. L'objectif est constitué de plusieurs lentilles en verre qui sont conçues pour grossir l'échantillon et projeter une image agrandie sur le plan focal du microscope. Les critères pour la sélection des objectifs varient selon plusieurs facteurs, notamment le grandissement et l'ouverture numérique, qui sont déterminants pour la résolution que l'on souhaite atteindre. D'autres paramètres spécifiques peuvent également être considérés, en fonction de l'application ciblée et des conditions expérimentales requises.

Dans la sélection des objectifs à utiliser, on va être sensible à divers paramètres. Parmi ceux-ci, l'ouverture numérique (NA) est un indicateur clé du pouvoir de résolution de l'objectif, c'est-à-dire sa capacité à distinguer les détails fins d'une image. Effectivement, l'objectif agit comme un filtre passe-bas dans le domaine fréquentiel, limitant de ce fait la quantité de détails fins qui peuvent être résolus. Le grandissement est aussi une caractéristique clé de l'objectif, il indique le facteur d'agrandissement de l'échantillon d'observation par l'objectif. D'autres facteurs interviennent également dans le choix d'un objectif en microscopie, comme la nécessité d'un milieu d'immersion (huile, eau, glycérine). Il est crucial de prendre en compte les corrections géométriques qui garantissent la planéité de l'image, les corrections chromatiques qui contrecarrent les aberrations dues aux variations de

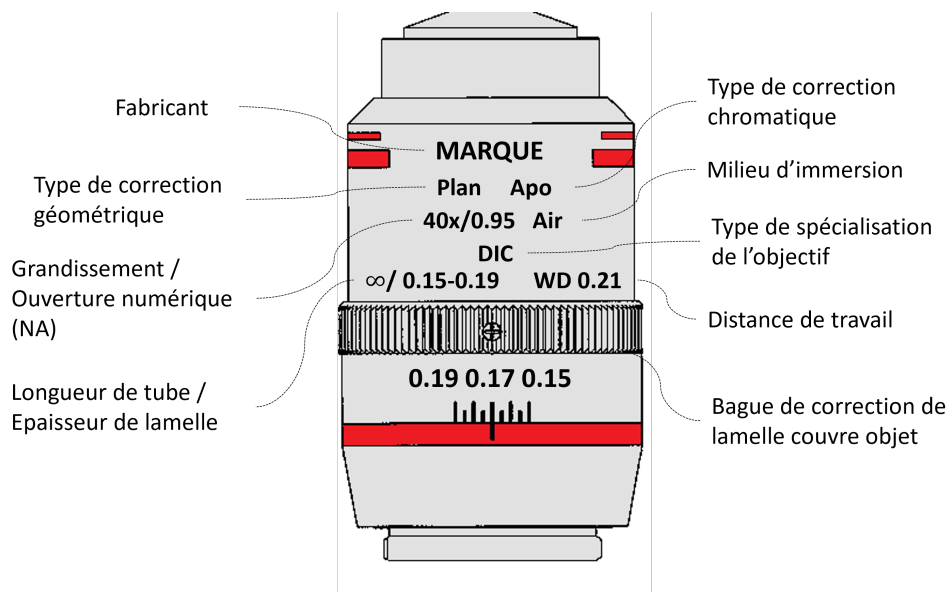


FIGURE 2.5 – Exemples des caractéristiques d'un objectif de microscopie

longueur d'onde lumineuse, et les bagues d'épaisseur de couvre-objet, indispensables pour les objectifs à grande ouverture numérique. Enfin, il est essentiel de noter que certains objectifs sont spécialement conçus pour diverses techniques de microscopie, telles que la fluorescence, le contraste de phase, la DIC, la polarisation, entre autres.

2.2.3 La caméra

En microscopie, la caméra représente le dernier élément de la chaîne de numérisation. Elle est souvent connectée directement à un ordinateur pour, d'une part, échantillonner le module du champ électromagnétique au plan focale objet, mais aussi pour une visualisation en direct. Il s'agit du principal moyen de collecte de données pour des méthodes comme l'imagerie FPM.

Les caméras sont équipées de capteurs d'images sophistiqués, tels que les CCD (Charge-Coupled Device) ou les CMOS (Complementary Metal-Oxide-Semiconductor). Chaque pixel du capteur enregistre la quantité de photons reçus pendant une durée d'exposition définie, traduisant ainsi l'intensité lumineuse. Cette information est ensuite transformée en un signal numérique, pouvant être quantifié selon différentes échelles de dynamique (8 bits/12 bits/16 bits).

Il convient de noter que les caméras modernes offrent souvent la possibilité de choisir la dynamique de quantification, ce qui permet aux utilisateurs de sélectionner la plus adaptée à leurs besoins.

L'une des caractéristiques essentielles à prendre en compte est le "pitch", défini comme la distance entre deux pixels adjacents sur un capteur. Cette distance joue un rôle crucial dans l'échantillonnage, c'est-à-dire la manière dont une image est numériquement capturée et reproduite. La taille des pixels a aussi un impact sur la qualité du signal capté par la caméra. D'autres facteurs, tels que la précision de la quantification numérique et les systèmes de refroidissement intégrés au capteur d'image, influencent également le signal mesuré.

L'être humain est habitué à percevoir des images en couleur. Ainsi, sur les pixels des capteurs, des filtres de couleur sont disposés pour séparer la lumière en ses composantes de base : rouge, vert et bleu. En mesurant l'intensité de chaque couleur, les informations colorimétriques sont ensuite fusionnées pour produire une image en couleur. Il existe différentes matrices de filtrage ; parmi les plus notables, on retrouve la matrice de Bayer, développée par Bryce E. Bayer en 1976, qui est la plus couramment utilisée. Cette matrice est composée de 50 % de filtres verts, 25 % de filtres rouges et 25 % de filtres bleus.

Pour certaines techniques d'imagerie où l'on utilise une seule longueur d'onde à la fois pour éclairer l'échantillon, comme c'est le cas avec la FPM (Fourier Ptychographic Microscopy), il est nécessaire d'éclairer l'objet plusieurs fois avec différentes longueurs d'ondes : bleu à 465 nm, vert à 525 nm et rouge à 635 nm. Ensuite, on combine numériquement ces images pour obtenir une image RGB. Cependant, il faut noter que le rendu des couleurs d'une telle image ne sera pas nécessairement le même que celui d'une image couleur obtenue avec un éclairage blanc ayant un large spectre.

On peut retrouver des caméras plus sophistiquées telles que des caméras de phase [13] qui permettent d'enregistrer une mesure de phase du spécimen qui représente l'épaisseur optique (vitesse avec laquelle la lumière traverse les structures selon leur indice de réfraction). Ces caméras sont souvent plus coûteuses, car elles sont électroniquement plus complexes, mais aussi, du fait que leur principe de fonctionnement induit une perte de résolution spatiale.

2.2.4 Propriétés relatives à la résolution en microscopie

En microscopie optique, qu'elle soit conventionnelle ou calculatoire, l'objectif principal est de mettre en évidence des détails de plus en plus fins, facilitant ainsi une caractérisation des spécimens grâce à des résolutions inédites. Il est donc important de présenter certaines méthodes de calcul pour déterminer la résolution, tant en microscopie optique que pour la microscopie FPM.

a) Calcul de la résolution

La capacité de distinguer des détails fins lors d'une observation microscopique est intrinsèquement liée à la performance de divers composants clés du système optique. Ces éléments, notamment la source lumineuse, l'objectif et la caméra, agissent en synergie pour assurer une caractérisation précise de l'échantillon.

La source lumineuse joue un rôle primordial en fournissant un éclairage ciblé sur l'échantillon. L'objectif, par sa conception, capte la lumière – qu'elle soit émise ou réfléchi – provenant de l'échantillon, puis la focalise pour former une image réelle qui est, par la suite, capturée par la caméra. Chacun de ces éléments, grâce à sa spécificité, contribue de manière déterminante à l'élaboration d'une image à la fois nette et dotée d'une résolution optimale.

En microscopie, la résolution désigne la distance minimale à laquelle deux objets ou points peuvent être distinctement séparés et identifiés comme deux entités séparées, plutôt que comme une seule.

Étant donné la variabilité des paramètres en microscopie optique, il n'existe pas une définition universelle de la résolution [92]. Bien que de nombreuses méthodes, telles que celles d'Abbe [6, 5], de Rayleigh [69] et de Sparrow [81], soient citées dans la littérature pour quantifier la résolution, elles se basent sur les caractéristiques optiques et physiques spécifiques du microscope et de l'échantillon.

La résolution d'Abbe, également connue sous le nom de critère d'Abbe, est une mesure de la plus petite distance entre deux points d'un échantillon qui peuvent être résolus en tant qu'objets séparés se basant sur le contraste. La résolution d'Abbe est donnée par la formule :

$$R_{Abbe} = \lambda / 2NA_{obj} \quad (2.1)$$

Avec λ qui est égal à la longueur d'onde de l'éclairage et NA_{obj} l'ouverture numérique de l'objectif.

La résolution de Rayleigh et la résolution d'Abbe diffèrent dans leur approche et leur interprétation. La résolution de Rayleigh (critère de Rayleigh) détermine la distance minimale à laquelle deux points d'un échantillon peuvent être perçus comme distincts plutôt que comme une tache floue unique, en se basant sur le profil d'intensité. Cette distance minimale est donnée par la formule :

$$R_{Rayleigh} = 1.22\lambda / (NA_{obj} + NA_{cond}) \quad (2.2)$$

Dans cette formule relative à la résolution, le NA_{cond} est intégré, symbolisant l'ouverture numérique du condenseur en présence d'une lumière de Kohler. Lorsque l'éclairage est cohérent et se produit sur l'axe optique, on considère que le NA_{cond} est équivalent à 0.

Le critère de Sparrow formule la capacité de distinguer deux longueurs d'onde adjacentes, particulièrement quand elles sont proches l'une de l'autre et émettent

avec une intensité similaire, en utilisant la courbe d'intensité conjointe comme référence. Si on compare cette approche à celle de Rayleigh où on considère la résolution comme la séparation minimale nécessaire pour percevoir deux points comme distincts, le critère de Sparrow, lui, la définit comme l'écart entre deux sources lumineuses ponctuelles pour lesquelles la zone intermédiaire ne montre aucun déclin en luminosité, mais affiche plutôt une intensité lumineuse stable entre les deux points lumineux principaux.

$$R_{Sparrow} = 0.47\lambda/NA_{obj} \quad (2.3)$$

b) Échantillonnage au niveau de la caméra

Afin de garantir une image de haute qualité et de réaliser la résolution prévue en fonction du type d'objectif et de la source lumineuse (en termes de longueur d'onde), il est primordial de respecter les principes d'échantillonnage pour les caméras.

Le théorème de Nyquist-Shannon nous indique que, pour reproduire fidèlement une caractéristique spatiale ayant une fréquence f , le taux auquel l'image est échantillonnée doit excéder $2f$. En contexte de microscopie, ce taux est défini par la disposition des pixels individuels du capteur, et est généralement proportionnel à la dimension du pixel. L'échantillonnage, tel que défini par Shannon, prend une importance capitale en microscopie. Les images de microscopie peuvent en effet présenter des détails à une variété de fréquences, certains étant particulièrement fins. Si le taux d'échantillonnage de la caméra ne convient pas, l'image risque une perte des détails ou de présenter des artéfacts dus au repliement spectral. D'où l'importance de choisir une caméra dont la dimension des pixels est en adéquation avec les détails que l'on souhaite capturer.

$$M_{obj} \cdot M_{cam} \cdot \frac{1}{\nu_{obj}} > 2 \cdot P_{capteur} \quad (2.4)$$

Avec M_{obj} pour le grandissement de l'objectif et M_{cam} pour celle de la caméra, ν_{obj} désignant la fréquence maximale de l'objectif, et $P_{capteur}$ indiquant le pitch de la caméra.

Pour une lumière de Köhler (non cohérente), $\nu_{obj} = \frac{2 \cdot NA_{obj}}{\lambda}$. Lorsque le diaphragme d'ouverture est très resserré (source presque ponctuelle), l'objet est illuminé par une onde pratiquement plane, ce qui rend l'éclairage spatialement cohérent. Dans ce cas, la fréquence de l'objectif devient : $\nu_{obj} = \frac{NA_{obj}}{\lambda}$.

Ainsi pour des lumières cohérentes, comme celles utilisées en FPM, lorsque la fréquence de l'objectif est réduite de moitié, la contrainte d'échantillonnage de la caméra devient alors deux fois moins stricte.

2.3 Microscopie Ptychographique de Fourier

2.3.1 Introduction

Avant d'aborder en détail la microscopie ptychographique de Fourier, permettez-moi de vous introduire succinctement à certaines technologies précurseuses qui ont ouvert la voie à la FPM. Ces technologies partagent, toutes, le but d'accéder à l'information de phase et d'intensité des échantillons analysés.

L'Holographie, découverte par Denis Gabor en 1947, a connu sa première application en 1967 avec les travaux de J.W. Goodman et R.W. Laurence [37] dont le but fut de reconstruire le champ optique à travers des images de phase (une épaisseur optique) et d'amplitude (absorption) de l'objet observé.

Cette méthode repose sur l'interférence entre un faisceau de référence qui ne traverse pas l'objet et un faisceau diffracté par l'objet en éclairage cohérent. Si ces faisceaux sont parallèles, on parle d'holographie en ligne (avec le problème des images jumelles). Si, en revanche, il existe un angle entre les deux faisceaux, objet et référence, on parle alors d'holographie hors axe. Cette interférence est ensuite mesurée à l'aide d'un capteur, tel qu'un CCD par exemple.

Puis par le calcul, on arrive à accéder à l'information sur le front d'onde et le champ électromagnétique issu de l'objet, ce qui permet d'avoir des représentations images 2D,3D.

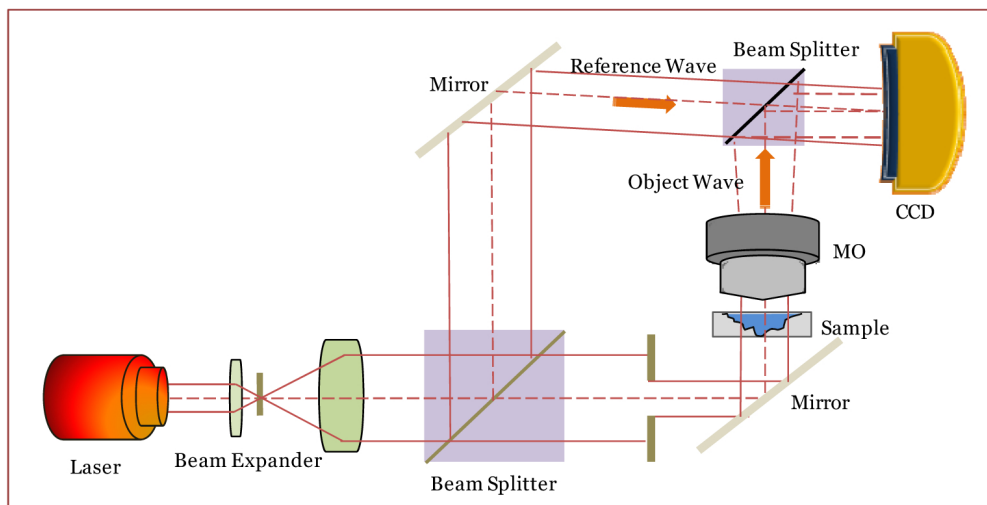


FIGURE 2.6 – Schéma de la microscopie holographique numérique hors axe [60].

Cette technique est utilisée dans de nombreuses applications [90], notamment en médecine, pour l'imagerie des spécimens biologiques, en ingénierie, pour l'inspection des matériaux. La résolution spatiale d'un dispositif holographique est

limitée par la longueur d'onde de la lumière utilisée ; de plus ce système est coûteux et assez contraignant, car il requiert une configuration bien calibrée, mais aussi une grande stabilité mécanique du dispositif pendant l'enregistrement des images.

Par ailleurs, il y a aussi l'Imagerie par Diffraction Cohérente (CDI) qui est une technique d'imagerie sans lentille, utilisée pour reconstruire les propriétés optiques des objets permettant d'obtenir des images 2D ou 3D, ainsi que des images à haute résolution spatiale sans utiliser de lentilles. La technique de CDI consiste à illuminer l'échantillon avec un faisceau de particules cohérentes (par exemple, laser, des rayons X cohérents...), puis à mesurer la diffraction de la lumière qui en résulte.

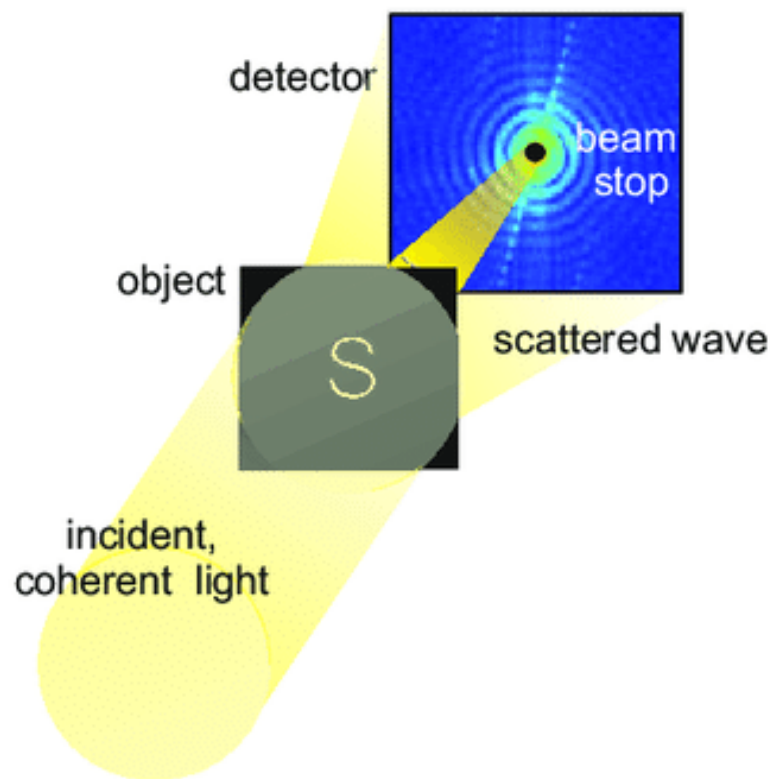


FIGURE 2.7 – Montage pour l'imagerie diffractive cohérente. L'échantillon S est éclairé par un faisceau lumière cohérente et parallèle. L'intensité diffusée est enregistrée dans le champ optique lointain [75]

Les motifs de diffraction obtenus sont ensuite traités algorithmiquement pour reconstituer l'image de l'échantillon. Contrairement à l'holographie, la reconstruc-

tion en Imagerie par Diffraction Cohérente (CDI) présente une complexité accrue en raison de la nécessité de résoudre un problème inverse "mal posé". Cette particularité rend possibles plusieurs solutions de phase et d'intensité pour une seule série de données de diffraction. De ce fait, l'incorporation de contraintes s'avère essentielle dans le processus de reconstruction.

Parmi les limitations de cette technique, le champ de sortie est contraint par la taille de l'illumination, ce qui rend la reconstruction de grands objets impossible.

2.3.2 Principe de la Ptychographie

Parmi les avancées les plus notables, la ptychographie, une technique d'imagerie calculatoire proposée par Walter Hoppe à la fin des années 60 pour la microscopie électronique. Toutefois, il a fallu attendre les années 1990 pour qu'elle se concrétise expérimentalement avec les travaux des équipes de Rodenburg [4].

La Ptychographie fait partie des techniques d'imagerie calculatoires qui partagent les mêmes principes que la Diffraction de Coherence Inverse (CDI), visant à dépasser les limites des microscopes conventionnels en se basant sur la reconstruction algorithmique de l'objet à partir de cliché de diffraction.

Dans cette technique de microscopie (Ptychographie), l'échantillon est éclairé par un faisceau lumineux cohérent et plusieurs motifs de diffraction sont enregistrés en déplaçant successivement l'échantillon. Cette méthode peut être employée avec ou sans l'utilisation de lentilles optiques. Les motifs de diffraction obtenus à partir de différentes zones de l'échantillon sont ensuite combinés afin de générer une image d'intensité et de phase plus étendue 2.8. En résolvant mathématiquement le problème inverse associé, il est possible de reconstruire l'image de phase associée à l'image d'intensité reconstruite.

Pour effectuer la reconstruction, il est nécessaire de bien connaître la longueur d'onde d'illumination, les informations sur le capteur, et ainsi connaître précisément le décalage des zones illuminées. Afin d'assurer une bonne convergence de l'algorithme d'optimisation, permettant d'obtenir une reconstruction de qualité optimale, exempte d'artefacts, tout en atteignant la résolution souhaitée du système, il est important de respecter une certaine superposition dans l'espace direct entre les zones illuminées avec un minimum de superposition spatiale de 60 % pour avoir un échantillonnage suffisant [18].

Outre une reconstruction de l'intensité et la phase, les algorithmes de reconstruction peuvent inclure la correction de la fonction d'illumination par optimisation. Le but est de corriger les aberrations d'illumination pour permettre l'obtention de la reconstruction de meilleure qualité. La ptychographie a également été déclinée dans une version vectorielle [29] où l'on cherche à accéder aux propriétés de polarisation d'une onde électromagnétique pour mesurer la polarisation des objets, tels que les matériaux biominéraux.

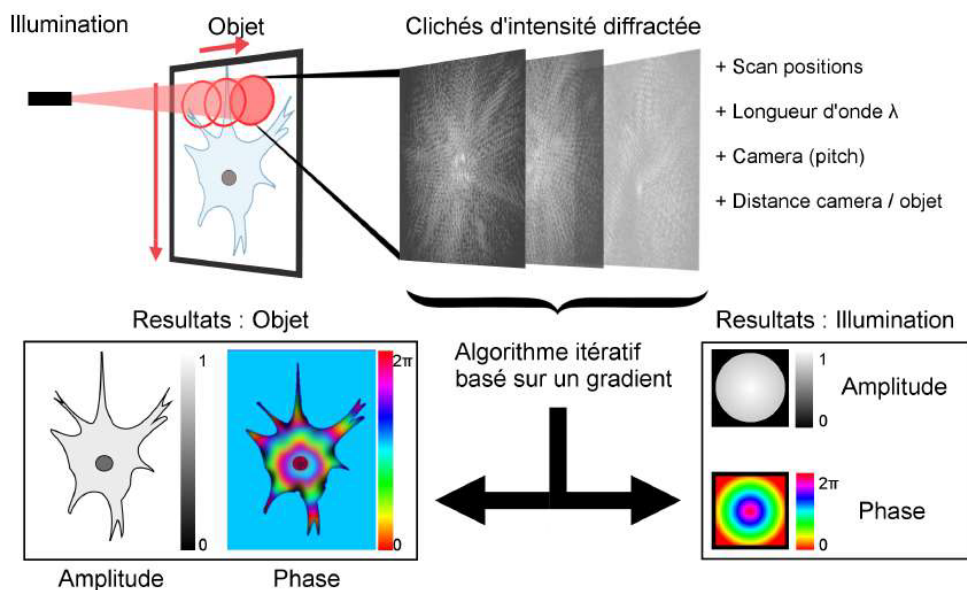


FIGURE 2.8 – Représentation schématique du principe de fonctionnement de la ptychographie. On est ici dans le cadre d'un algorithme permettant de reconstruire à la fois l'objet et l'illumination.[9]

Dans la ptychographie vectorielle, un montage polariseur / analyseur croisé est intégré au dispositif expérimental. ce montage agit comme un filtre optique qui n'autorise que la propagation des ondes lumineuses oscillant dans un plan spécifique, tout en bloquant celles oscillant dans les autres plans. Ce polariseur est positionné en aval de l'échantillon et génère un faisceau de lumière polarisée. Suite à cela, on utilise un second polariseur (analyseur) qui est en aval de l'échantillon. L'analyseur est orienté selon un angle particulier par rapport au premier polariseur et n'autorise le passage que des ondes lumineuses oscillant dans le plan correspondant à son propre axe de polarisation. Ce système à deux polariseurs offre une méthode plus nuancée pour étudier les propriétés de l'échantillon. Combiné à la ptychographie, il permet d'accéder aux caractéristiques optiques de l'échantillon, telles que sa biréfringence, son retard optique, ainsi que l'orientation de son axe optique, le tout avec une haute résolution.

2.3.3 Principe de la FPM

La ptychographie microscopique de Fourier (FPM) est une technique d'image avec lentille et une technique de microscopie visant à dépasser les limites de résolution la microscopie optique conventionnelle. Contrairement à la ptychographie,

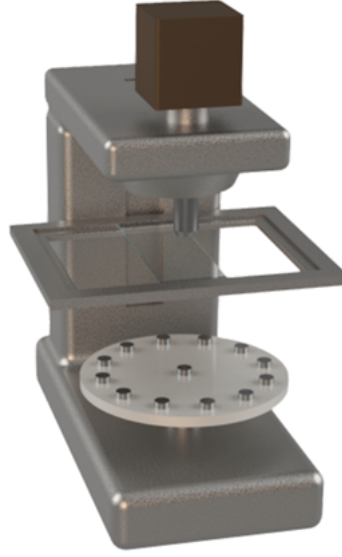


FIGURE 2.9 – Dispositif FPM avec matrice de LED, Frottis, Objectif (lentille), Camera.

l'échantillon est ici éclairé sous différents angles d'éclairage. De plus, en FPM on capte directement l'intensité de l'image de l'échantillon au niveau du détecteur pour chaque angle d'éclairage. Cette approche contraste avec la ptychographie classique, où ce sont les motifs de diffraction qui sont mesurés.

Puis grâce à la reconstruction par méthode de résolution du problème inverse associé à la formation des images de l'échantillon sous une diversité angulaire d'éclairage de l'échantillon, on peut accéder à la phase de l'échantillon en plus de son intensité. En plus du dispositif expérimental qui est moins complexe que la ptychographie, les images FPM sont reconstruites avec une super-résolution, le facteur de super résolution dépend de la diversité angulaire accessible à la source d'éclairage utilisée. Les matrices des LEDs utilisées sont composées de n sources lumineuses individuelles ponctuelles remplaçant l'éclairage Köhler traditionnel [6, 46]. Un ensemble de n images $I^{(i)}$ (i variant de 1 à n) est acquis, une image par incidence angulaire choisie. Dans la limite de l'assimilation de l'échantillon à un objet fin (thin sample approximation), l'échantillon est modélisé par une fonction de masque de l'objet complexe O et le champ E_o sortant.

$$E_o(x, y) = E_{in}(x, y) \cdot O(x, y) \quad (2.5)$$

Où E_{in} représente le champ électromagnétique immédiatement sous l'échantillon. Pour l' i^{eme} LED et en utilisant l'approximation de l'onde plane locale,

$$E_{in}^{(i)} = Ae^{j(k_x^{(i)} \cdot x + k_y^{(i)} \cdot y)} \quad (2.6)$$

Avec, $k_x^{(i)} = \frac{2\pi\nu_{x_0}^{(i)}}{\lambda}$, $k_y^{(i)} = \frac{2\pi\nu_{y_0}^{(i)}}{\lambda}$ la projection du vecteur \mathbf{k} le long des axes x et y associée à l'illumination i^{eme} LED. L'image de i^{eme} enregistrée par la caméra est :

$$I^{(i)} = |E_{in}^{(i)} \cdot O(r) * P(r)|^2 \quad (2.7)$$

où P représente la réponse impulsionnelle spatiale (PSF). Exprimée de manière équivalente dans le domaine de Fourier, l'équation 2.7 devient,

$$I^{(i)} = |A \cdot \mathcal{F}^{-1}[\widehat{O}(\nu - \nu_0^{(i)}) \cdot \widehat{P}(\nu)]|^2 \quad (2.8)$$

Dans ce contexte, A symbolise la transmittance de l'objet, tandis que $\nu_0^{(i)} = (\nu_{x_0}^{(i)}, \nu_{y_0}^{(i)})$ désigne le vecteur de fréquences spatiales.

C'est ici que le principe de la ptychographie trouve son application dans l'espace de Fourier. Selon l'équation 2.8, il est observé que chaque image acquise lors d'une expérience ptychographique correspond à une portion spécifique du spectre de l'objet O , centrée dans le domaine de Fourier sur $\nu_0^{(i)}$.

Les différents éclairages angulaires utilisés permettent ainsi de paver l'espace de Fourier et d'accéder à O sur un domaine spectral étendu par rapport à celui accessible par la seule lentille de l'objectif du microscope. Le mécanisme d'ouverture synthétique est obtenu par assemblage dans l'espace de Fourier des différentes images acquises. Cet assemblage est obtenu par l'algorithme de reconstruction. Du fait même des principes utilisés pour formuler le problème direct, cet algorithme d'inversion permet alors de reconstruire l'intensité et la phase de l'image sur toute la région couverte dans l'espace de Fourier. Il permet donc de remonter à l'image de l'échantillon avec une super-résolution 2.10, ainsi qu'à sa phase.

L'accès à cette information de phase exploite le fait que le support de la fonction de transfert cohérente \widehat{P} associée à l'objectif du microscope est borné.

De manière analogue à la ptychographie dans l'espace des objets, les algorithmes de reconstruction en microscopie FPM nécessitent également une superposition minimale entre les zones spectrales stimulées par différentes LED. Il est généralement admis qu'une superposition d'au moins 60 % est essentielle pour obtenir une reconstruction de bonne qualité. Cette exigence de redondance d'information spectrale, imposée par le caractère mal posé du problème inverse associé à la FPM, agit comme une contrainte qui favorise une convergence efficace des algorithmes de reconstruction.

Dans la ptychographie de Fourier, contrairement à la ptychographie classique, l'erreur est calculée dans le domaine spatial (direct) plutôt que dans le domaine spectral. Les algorithmes de reconstruction ont été adaptés en conséquence afin de corriger à la fois la fonction de pupille et l'objet lors de la reconstruction. De plus,

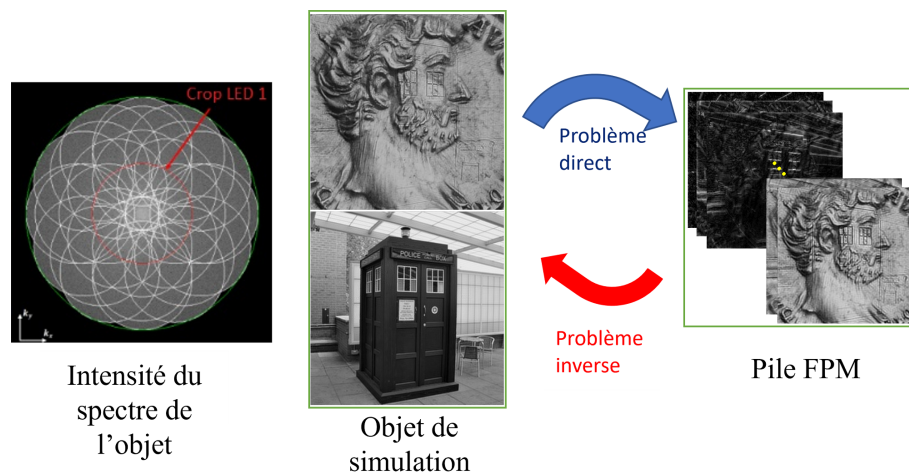


FIGURE 2.10 – Illustration du modèle direct de la FPM

dans la FPM il est possible d'introduire des fonctions de front d'onde pour réaliser une défocalisation d'une quantité Δz lors de l'étape de reconstruction.

2.3.4 Algorithme de reconstruction

Comme mentionné précédemment, les capteurs d'image conventionnels tels que les CMOS et CCD sont majoritairement adaptés pour la mesure d'intensités lumineuses et ne permettent pas de capturer l'information de phase. L'accès à cette information de phase à partir des images d'intensité constitue un défi majeur pour plusieurs techniques non conventionnelles. Cette information est reconstruite par la résolution d'un problème inverse, en minimisant une fonction d'erreur, associée à la somme des écarts entre les imagerie synthétiques (clichés simulés) produites par le masque de l'objet recherché O sous différents éclairages (problème direct), et les imagerie (clichés mesurés) obtenues expérimentalement.

De nombreuses études ont abordé la problématique des reconstructions de la phase à partir des images d'intensité (Schiske (1968), Hoppe (1970), Erickson et Klug (1970)). Leurs méthodes de reconstruction se sont appuyées sur des approximations linéaires, ce qui limite leur validité aux cas où il n'y a que de petites variations de phase et/ou d'amplitude du champ magnétique sondé.

Ainsi, le premier algorithme de reconstruction efficace pour retrouver l'image de phase a été celui de Gerchberg-Saxton, développé en 1972 [32]. Ne se limitant pas aux faibles écarts de phase, il nécessite deux mesures d'intensité. Le processus de reconstruction débute dans le plan de l'objet, où un objet complexe est généré en associant l'amplitude mesurée à une phase aléatoire. En appliquant une transformée de Fourier à cet objet, on obtient sa distribution complexe dans le plan

de diffraction. Dans ce plan, la distribution d'amplitude est remplacée par celle mesurée, ce qui donne une nouvelle distribution d'onde complexe mise à jour. Une transformée de Fourier inverse permet ensuite de retourner dans le plan objet. Là, l'amplitude est corrigée en utilisant l'amplitude mesurée, conduisant à une nouvelle version mise à jour de l'objet. Ces étapes sont répétées pour un nombre défini d'itérations.

Par la suite, en 1978, Fienup [30] a développé un algorithme assez général et plus robuste, face au bruit, que l'algorithme Gerchberg-Saxton (G-S.) L'algorithme assouplit les contraintes du G-S en se basant sur deux éléments : l'hypothèse que le module de l'objet est non négatif, ainsi que la connaissance du module de sa transformée de Fourier. Ces deux éléments sont utilisés pour reconstruire la phase de l'intensité de l'objet.

De 1989 à 2009, Rodenburg et son équipe ont concentré leurs efforts sur le défi de la phase manquante en créant différentes techniques de reconstruction. Dans les contributions notables de ce groupe, on trouve une technique itérative appelée "PIE" (Phase Retrieval by Iterative Extension) [76], et le modèle ePIE [57] de 2009, qui est reconnu comme une méthode de référence en ptychographie.

a) ePIE

L'algorithme ePIE (extended-Phase Retrieval and Integrated-Error) est une méthode itérative conçue pour traiter le problème de la phase en ptychographie. Cette approche est une extension de l'algorithme PIE et permet de retrouver la distribution de phase d'un échantillon en se basant sur des mesures de diffraction issues d'une série d'illuminations, le tout à travers un processus itératif.

L'algorithme ePIE permet non seulement de reconstruire l'intensité et la phase de l'objet mais aussi de corriger la fonction d'illumination pour une meilleure qualité de reconstruction et une convergence plus rapide. Il améliore progressivement l'estimation de la phase et de la fonction d'illumination en réduisant l'erreur entre les mesures expérimentales et les reconstructions à chaque itération.

Comme illustré dans la Figure 2.11, d'après les travaux de Maiden et al. [57], les images reconstruites en intensité (a,2) et en phase (b,2) grâce à l'algorithme ePIE offrent une résolution remarquable et des informations de phase et d'intensité nettement plus distinctes que les images (a,1) et (b,1) produites par l'algorithme PIE. On observe que les erreurs non corrigées par la fonction d'illumination dans l'algorithme PIE entraînent une perte d'informations et des images de moindre qualité plus facilement remarquable sur l'image d'intensité ; la qualité de l'image de phase est également améliorée. Cependant, étant donné que l'amplitude de phase est significative et que la phase est périodique, on observe une transition de certains pixels du blanc au noir. Cela illustre que l'emploi de l'algorithme ePIE permet d'exploiter pleinement les informations contenues dans les mesures, aboutissant à

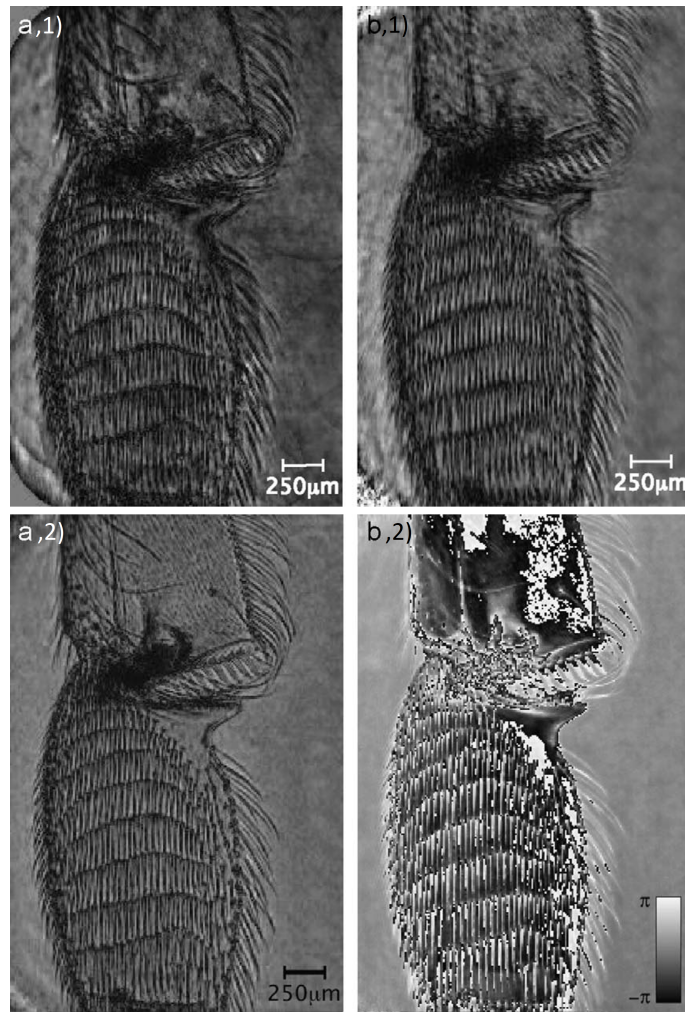


FIGURE 2.11 – Résultats comparatifs [57] entre l’algorithme PIE : a,1) Intensité, b,1) phase et par l’algorithme ePIE a,2) intensité, b,2) phase

des images de qualité supérieure.

Le processus de reconstruction des images par l'algorithme ePIE débute par l'initialisation de l'objet complexe $O_i(r)$ que l'on cherche à reconstruire, ainsi que de $P_i(r)$, représentant la forme d'onde de la sonde illuminant l'objet. En règle générale, l'objet initial est représenté par une fonction libre à trouver, tandis que la forme d'onde initiale est une fonction de support grossièrement dimensionnée selon la région d'illumination du front d'onde de la sonde. Par la suite, l'objet complexe est multiplié par la fonction d'illumination, en appliquant un décalage spécifique à chaque zone sondée $R(i)$

$$\psi_i(r) = O_i(r)P_i(r - R(i)) \quad (2.9)$$

Ensuite, sur le résultat $\psi_i(r)$, on applique une transformée de Fourier, où le module du spectre est remplacé par la racine carrée de l'intensité $I_i(u)$ qui a été mesurée.

$$\psi_i(\nu) = \sqrt{I_i(\nu)} \frac{F[\psi_i(r)]}{|F[\psi_i(r)]|} \quad (2.10)$$

Puis, on calcule un champ de sortie corrigé :

$$\psi_{t_i}(r) = F^{-1}[\psi_i(\nu)] \quad (2.11)$$

Enfin pour iteration i , on corrige O_{i+1} et P_{i+1} :

$$\begin{aligned} O_{i+1}(r) &= O_i(r) + \alpha \frac{P_i^*(r - R(i))}{|P_i(r - R(i))|_{max}^2} (\psi_{t_i}(r) - \psi_i(r)) \\ P_{i+1}(r) &= P_i(r) + \alpha \frac{O_i^*(r + R(i))}{|O_i(r + R(i))|_{max}^2} (\psi_{t_i}(r) - \psi_i(r)) \end{aligned} \quad (2.12)$$

Comme le montre l'équation 2.12, en Ptychographie, la correction de la fonction d'illumination et de l'objet s'effectue dans l'espace direct, où α et β représentent les pas de correction utilisés pour pondérer respectivement la correction de l'objet et de la fonction d'illumination. Ce processus est répété itérativement pour toutes les illuminations et j fois jusqu'à convergence, pour obtenir la phase et l'objet reconstruits de manière précise. En résumé, l'algorithme ePIE est une extension améliorée de l'algorithme PIE qui permet une reconstruction précise et rapide des images de phase en imagerie de Ptychographie grâce à la reconstruction simultanée de la fonction d'illumination et de l'objet.

b) Algorithme initial de la reconstruction FPM

La reconstruction de la FPM consiste à appliquer des contraintes, aussi bien dans l'espace direct que dans l'espace de Fourier, pour arriver à une estimation

de O , le champ complexe de l'objet examiné. Dans l'espace direct, ces contraintes sont basées sur la connaissance du module du champ $I(i)$ pour chaque LED. Dans l'espace de Fourier, les contraintes sont basées sur la connaissance du support du filtre réalisé par l'objectif.

Le premier algorithme utilisé dans la reconstruction FPM a été mis œuvre par Zheng, en 2013. Cet algorithme, reposant sur une méthode itérative de descente de gradient, s'inspire grandement de l'algorithme PIE. Toutefois, il a subi des adaptations pour tenir compte des domaines de mesure (spatial) et de correction (spectral) propres à la microscopie FPM.

En premier lieu, l'objet complexe $O(r)$ est initialisé dans le domaine spatial. Ici, I_0 représente l'intensité mesurée pour une LED (i), qui est ensuite interpolée de manière à correspondre à la taille finale de l'objet que l'on souhaite reconstruire. Pour sa part, ϕ_0 peut-être initialisé avec une phase nulle. Puis pour avoir le spectre de l'objet initial, on applique une transformée de Fourier).

$$\widehat{O}(\nu) = F[\sqrt{I_0}e^{i\phi_0}] \quad (2.13)$$

Par la suite, pour chaque LED, une sous-région spectrale est sélectionnée grâce au décalage spectral ν_i , associé à la projection du vecteur correspondant à l'illumination de la i^{eme} LED. Sur cette sous-région, un filtre passe-bas est appliqué via la fonction de transfert cohérente \widehat{P} . Cette dernière correspond au filtrage effectué par l'ouverture numérique de l'objectif, en tenant compte du vecteur de décalage. En procédant ensuite à une transformation de Fourier inverse, on obtient le résultat souhaité.

$$\psi_i(r) = \sqrt{I_i}e^{i\phi_i} = F^{-1}[\widehat{O}(\nu - \nu_i)\widehat{P}(\nu)] \quad (2.14)$$

Puis on remplace dans $\sqrt{I_i}e^{i\phi_i}$, I_i par I_{im} qui est l'intensité mesurée par l'éclairage de la i^{eme} LED .

$$\psi_i I(r) = \sqrt{I_{im}} \frac{\psi_i(r)}{|\psi_i(r)|} \quad (2.15)$$

Enfin on applique une transformée de Fourier sur $\psi_i I(r)$ pour calculer le spectre qui nous servira à corriger la sous-région présélection.

$$\widehat{O}(\nu - \nu_i) = F[\psi_i I(r)] \quad (2.16)$$

Cette correction est effectuée LED par LED et autant de fois que nécessaire jusqu'à obtenir une reconstruction satisfaisante.

La Figure 2.12 illustre les images obtenues par la reconstruction FPM avec un gain de résolution de x2 dans une configuration un objectif 40x et (NA) de 0.95 et une configuration de 13 LED avec $\lambda = 535nm$. Grâce à la reconstruction du spectre électromagnétique de l'objet, nous avons pu récupérer l'information

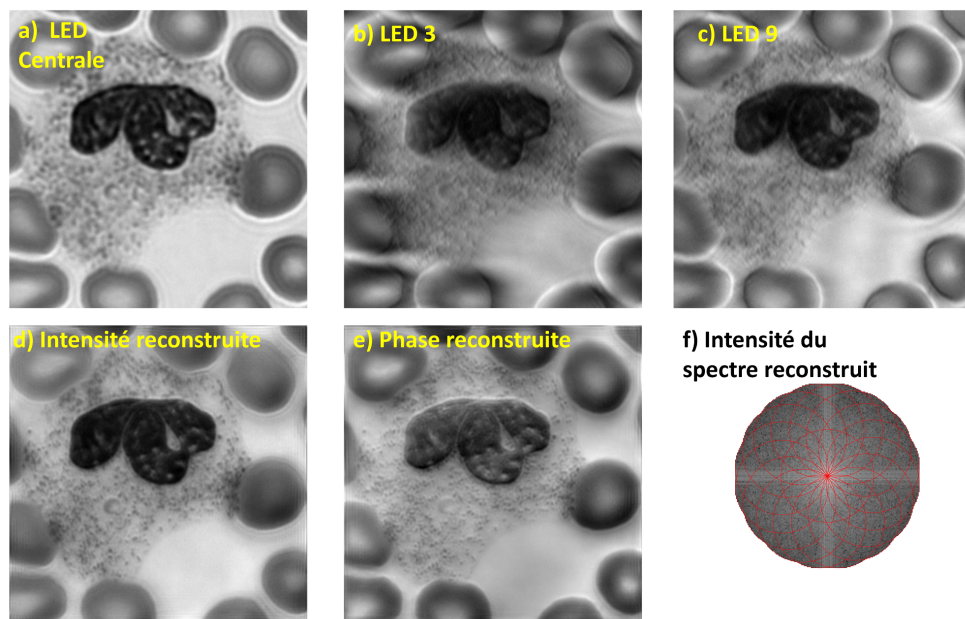


FIGURE 2.12 – a), b) et c) illustrent les images de la pile FPM acquises avec différents angles d'éclairage. d) et e) représentent respectivement l'intensité et la phase reconstruite. d) module du spectre de l'objet complexe reconstruit.

de phase, comme le montre la Figure 2.12 e). Cette information de phase a été obtenue avec une résolution doublée par rapport aux images d'intensité acquises au microscope 2.12 a,b,c).

c) Correction d'aberration FPM

Les objectifs et les systèmes optiques étant imparfaits, il est souvent crucial d'introduire des fonctions de correction pour compenser ces défauts et obtenir des images reconstruites de haute qualité. En ptychographie de Fourier, la correction des aberrations consiste à introduire une carte de phase dans la fonction de transfert cohérente du système noté \hat{P} lors de la reconstruction.

Cette correction se fait dans le domaine spectral. Lors de la reconstruction, la fonction de phase correspondant à l'aberration pré-estimée est introduite et multipliée par le résultat du filtrage effectué par la \hat{P} . Ainsi, le processus de reconstruction prend en compte cette aberration, ce qui permet de compenser l'objet reconstruit.

Parmi les corrections qui peuvent être introduites, on retrouve la correction de focus, qui est un élément crucial en imagerie optique. Cette correction permet d'augmenter (numériquement) la profondeur de champ du système de mesure, ce qui permet de reconstruire correctement des objets qui étaient initialement hors

du plan de focus lors de la capture des mesures.

Cette fonction de défocus est également une fonction qui se traduit par l'ajout d'un terme de décalage de phase par rapport au plan de la pupille. Elle s'écrit :

$$e^{i\phi(k_x, k_y)} = e^{i\sqrt{(2\pi/\lambda)^2 - k_x^2 - k_y^2} \cdot \Delta z}; k_x^2 + k_y^2 < NA \cdot (2\pi/\lambda) \quad (2.17)$$

Avec k_x, k_y qui représente le nombre d'onde dans le plan de la pupille, λ la longueur d'onde de l'éclairage utilisé, Δz la distance defocus désiré et NA l'ouverture numérique de l'objectif.

d) Algorithme EPRY

Tout comme l'algorithme PIE a été adapté pour l'optimisation conjointe de l'objet reconstruit et de la fonction d'illumination dans la ptychographie, dans le cas de la FPM, le premier algorithme développé par Zheng a été adapté pour permettre une reconstruction optimale de la phase et de l'intensité grâce à une correction conjointe de la fonction de pupille qui s'appelle EPRY [62] "Embedded pupil function recovery for Fourier ptychographic microscopy". Cela permet de se passer de l'estimation préalable des aberrations, qui est souvent complexe à réaliser et n'est pas souvent très précise.

Une telle imprécision dans l'estimation de la fonction de pupille peut être causée par les ordres limités d'aberration considérés dans le processus de pré-caractérisation ou par des changements mécaniques ou optiques dans le système de microscopie.

Comme expliqué dans la section 2.3.3 l'image enregistrée par le microscope en FPM correspond à l'équation suivante :

$$Im^{(i)}(r) = A \cdot \left| F^{-1} \left[\widehat{O}(\nu - \nu_i) \cdot \widehat{P}(\nu) \right] \right|^2 \quad (2.18)$$

Où $\widehat{O}(\nu - \nu_i)$ désigne la partie du spectre décalée par le vecteur de translation ν correspondant à l'éclairage par la LED (i), et $P(\nu)$ représente la fonction de transfert cohérente. Les fonctions \widehat{O} et $P(\nu)$ sont les deux fonctions cibles que nous chercherons à reconstruire à l'aide de l'algorithme EPRY

Tout comme l'algorithme de départ utilisé dans la reconstruction des images FPM, l'initialisation de $E_{out}^{(i)}$ se fait en effectuant la transformation de Fourier (FFT) d'une des images basse résolution interpolée avec une phase nulle. La fonction de transfert cohérente (\widehat{P}) est initialisée sous la forme d'un filtre passe-bas circulaire, avec des valeurs de un situées à l'intérieur d'un disque uniforme de rayon $NA \times 2\pi/\lambda$, où NA est l'ouverture numérique de l'objectif du microscope et λ la longueur d'onde de l'éclairage. L'extérieur du disque est, quant à lui, composé de zéros. La phase de \widehat{P} est initialisée avec des zéros uniformément répartis.

Pour chaque LEDs, on calcule le champ simulé au niveau de la caméra :

$$\psi_i(\nu) = \widehat{O}(\nu - \nu_i) \widehat{P}(\nu) \quad (2.19)$$

De façon à pouvoir appliquer la contrainte d'intensité sur $\psi_i(r)$ en remplaçant son module par le module de l'intensité $\sqrt{I_{im}}$ mesurée tel :

$$\psi_i(r) = \sqrt{I_{im}} \frac{\psi_i(r)}{|\psi_i(r)|} \quad (2.20)$$

Jusqu'à cette étape, l'algorithme EPRY est identique à celui de la reconstruction initiale. Ensuite, le champ de sortie corrigé est calculé par la transformée de Fourier, notée $\psi_i(n)^{(i)}(\nu) = F[\psi_i(n)^{(i)}(r)]$. Cette opération nous permet de corriger l'objet en lui ajoutant la différence entre l'onde corrigée et l'onde initiale.

De même, pour corriger la fonction de pupille, on inverse dans la fonction précédente les termes P et O, et le terme de correction est ici remplacé par β .

$$\begin{aligned} O_{i+1}(\nu) &= O_i(\nu) + \alpha \frac{P_i^*(\nu + \nu_i)}{|P_i(\nu + \nu_i)|_{max}^2} (\psi_i(\nu + \nu_i) - \psi_i(\nu + \nu_i)) \\ P_{i+1}(\nu) &= P_i(\nu) + \beta \frac{O_i^*(\nu - \nu_i)}{|O_i(\nu - \nu_i)|_{max}^2} (\psi_i(\nu) - \psi_i(\nu)) \end{aligned} \quad (2.21)$$

La correction effectuée par l'algorithme EPRY, conformément à l'équation 2.21, et celle réalisée via l'algorithme ePIE, décrite par l'équation 2.12, révèlent des similarités notables entre les deux méthodes. La différence majeure réside dans les domaines respectifs de correction et de mesure, qui changent du domaine spatial au domaine spectral, et inversement.

Ce processus de correction est poursuivi jusqu'à ce que nous ayons utilisé toutes les N images capturées lors de l'éclairage ν , en mettant constamment à jour les spectres de l'échantillon et de la pupille à chaque fois. Par la suite, nous répétons l'intégralité du processus itératif pour un nombre d'itérations afin d'atteindre la convergence et l'obtention des spectres finaux de l'échantillon et de la pupille. Finalement, nous appliquons une transformée de Fourier inverse sur le spectre de l'échantillon, ce qui nous permet de calculer le module et les phases de l'échantillon.

2.4 Les réseaux de neurones

Les réseaux de neurones représentent une forme d'intelligence artificielle inspirée par les modèles du cerveau humain. Les premiers travaux sur les réseaux de neurones ont été effectués par le neuroscientifique McCulloch et le mathématicien Pitts dans les années 1940. Le premier travail sur les réseaux de neurones par Rosenblatt en 1958 a permis de mettre en place le modèle neuronal « perceptron »

[78]. Rosenblatt a poursuivi ses travaux pour qu'en 1960 il aboutisse à la mise en œuvre du perceptron multicouche « MLP » une des bases de deep learning.

Par la suite, plusieurs travaux ont été effectués autour des réseaux multicouches, ceux élaborés par Rumelhart, Hinton et Williams (1986b) sur la backpropagation et son utilisation générale dans les réseaux de neurones [79]. Ces travaux ont permis par la suite à Yann LeCun en 1989 de concevoir le premier réseau de neurones à base de convolution CNN (convolution neural network)[48] dont le but fut la reconnaissance de l'image et la vision par ordinateur.

À la fin des années 2000, avec la disponibilité de GPU modernes plus puissants et l'explosion de la quantité de données numériques, l'engouement pour les réseaux de neurones et le deep learning a explosé. Cela a entraîné un développement et des avancées algorithmiques importants dans l'écosystème de l'apprentissage profond et de l'intelligence artificielle (IA). Des bibliothèques telles que TensorFlow, PyTorch et Keras, ainsi que des frameworks d'apprentissage automatique en libre accès, ont vu le jour. Cet engouement pour l'IA a permis le développement de modèles avancés, donnant une valeur importante aux données.

Les réseaux de neurone sont maintenant utilisés dans plusieurs tâches de reconnaissance et de classification des images, traitement de la parole et des textes, dans la prédiction financière, et même dans l'art. Selon les données disponibles et les objectifs recherchés, on est amenés à utiliser un type de modèle différent. Le choix du type d'apprentissage dépend également de la nature et des caractéristiques des données disponibles. Une tâche nécessitant une reconnaissance précise requiert une approche supervisée avec des données étiquetées. En revanche, une tâche de clustering d'une population sera réalisée par des méthodes non supervisées, qui n'exigent pas de références préalables, car la "vérité" n'est pas connue à l'avance.

En plus de l'évolution des modèles neuronaux pour améliorer leurs performances, nous constatons une intégration croissante des réseaux de neurones dans le domaine de la physique. Cette convergence a donné naissance à une nouvelle discipline axée sur la co-conception entre la physique et les modèles neuronaux. Cette alliance vise à concevoir de nouveaux dispositifs physiques en tirant parti du changement de paradigme offert par les réseaux de neurones et l'intelligence artificielle.

2.4.1 Principe de fonctionnement

Les réseaux de neurones multicouches reposent sur des modèles mathématiques et informatiques. Ils fonctionnent en traitant et en transmettant l'information à travers une série de neurones artificiels, chacun agissant comme un automate. Au fur et à mesure que l'information traverse ces couches, elle subit des transformations, des enrichissements et des abstractions. À un certain niveau de profondeur,

les réseaux de neurones sont si abstraits qu'ils deviennent difficilement interprétables. Cette difficulté à expliquer la relation entre les entrées et les sorties des réseaux de neurones conduit couramment à les qualifier de modèles « boîte noire ». Cette modélisation mathématique permet de découvrir des relations complexes qui sont difficiles à estimer et à automatiser avec des équations analytiques, surmontant ainsi différentes limitations.

L'apprentissage au moyen de réseaux de neurones s'appuie sur un jeu de données spécifique pour l'entraînement. Au sein de ce jeu, le réseau utilise une partie des données en tant qu'entrées (ou "inputs"). Les valeurs attendues associées à ces entrées servent de référence pendant le processus d'apprentissage statistique. Une fois l'entraînement terminé, le réseau est capable de donner des prédictions et d'appliquer ses connaissances à des données qu'il n'a jamais vues auparavant.

Un réseau de neurones multicouche peut être découpé en 3 niveaux : une couche d'entrée, une ou plusieurs couches cachées et une couche de sortie. Afin de calculer la sortie du modèle, les variables d'entrée sont propagées couche par couche.

Au niveau de la première couche, le réseau de neurones prend les données initiales en entrée. Ces données sont interconnectées par des poids, que l'on nomme également paramètres du modèle. On réalise une somme pondérée de ces entrées en utilisant des poids initialisés aléatoirement. Le résultat de cette somme pondérée est ensuite transmis à une fonction d'activation non linéaire. Cette fonction permet d'introduire une non-linéarité dans la sortie du neurone, qui est ensuite transmise à la couche suivante.

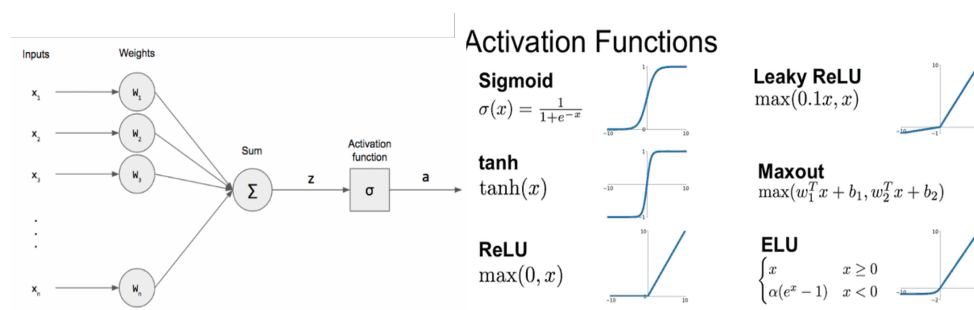


FIGURE 2.13 – À gauche : principe de fonctionnement d'un neurone. À droite : fonctions d'activation.

Nous avons $z = \sum_{i=1}^n W_i x_i + b_i$ et $a = \sigma(z)$ Avec :

- x_i l'entrée x pour la variable i .
- W_i , le poids associé à la variable x_i .
- b_i , le biais associé à la variable x_i .
- z , la combinaison linéaire des entrées.

- σ , fonction non linéaire.
- a , la sortie de l'activation d'un neurone.

Ce processus se répète couche par couche selon l'architecture du modèle défini jusqu'à ce que les données atteignent la couche de sortie ; cette dernière correspond à la prédiction du réseau de neurones.

Durant la phase d'apprentissage, la sortie générée par le réseau de neurones est comparée avec une valeur de référence. Cette étape vise à minimiser une fonction de coût pour ajuster de manière optimale les poids du réseau. Le processus comprend la propagation de l'information depuis l'entrée jusqu'à la sortie du réseau, ainsi que le calcul de la fonction de coût. Ces étapes permettent ensuite de calculer les dérivées partielles de la fonction de coût par rapport aux poids du réseau, ce qui sert à leur optimisation.

Le choix de la fonction d'erreur (coût) est dépendant de la nature du problème que l'on cherche à résoudre. Si les données à prédire sont de type continu, on est donc dans un cas de régression, on sélectionnera par les normes L1 ou L2... Selon le modèle de bruit qui peut être présent dans ses données, si les données sont de type catégorielle, on va séparer nos données en catégories de façon que le réseau prédise une probabilité d'appartenance à une telle catégorie grâce à des fonctions d'erreur, tel que la cross-entropy.

En minimisant cette fonction, les poids sont corrigés pour chaque exemple de la base de données et ainsi les poids corrigés à chaque étape permettent d'avoir des prédictions à la sortie du réseau de neurones de plus en plus proches de notre sortie désirée, cette mise à jour des poids dans un apprentissage supervisé est réalisée grâce au processus connu sous le nom de rétropropagation.

2.4.2 Rétropropagation et descente de gradient

La backpropagation, ou rétropropagation en français, est une technique essentielle pour l'apprentissage des réseaux neuronaux. En apprentissage supervisé, le but est de trouver les poids et les biais optimaux du réseau afin de minimiser la différence entre les prédictions du réseau et les résultats attendus.

Pour y parvenir, on commence par calculer l'erreur de chaque couche. Cette erreur est obtenue en utilisant la dérivée partielle de la fonction de coût par rapport aux poids et aux biais. Ensuite, cette erreur est rétropropagée à travers le réseau grâce à la méthode de la descente de gradient, couche par couche, depuis la sortie. Cette mise à jour permet d'améliorer les prédictions du réseau à chaque passage sur le jeu de données d'apprentissage.

Dans les modèles informatiques destinés à calculer ces dérivées partielles, la technique de la différenciation automatique est souvent employée. Cette méthode permet d'évaluer les dérivées de fonctions représentées par des algorithmes infor-

matiques. Il convient de noter que, quelle que soit la complexité de l'algorithme, il s'exécute en effectuant une séquence d'opérations élémentaires (telles que l'addition, la multiplication, sin, cos, log, exp, etc.).

Les outils disponibles permettent aux développeurs de calculer efficacement les gradients de fonctions arbitraires, sans avoir à écrire manuellement les gradients associés à chaque fonction utilisée. Les bibliothèques de deep learning sont particulièrement efficaces pour les calculs rapides, ce qui permet d'étendre leur utilisation en dehors des réseaux de neurones. Elles peuvent être utilisées pour résoudre des problèmes d'optimisation où les gradients sont difficiles à exprimer analytiquement. Ceci facilite le développement de nouveaux modèles et fonctions de coût pour la résolution de problèmes complexes. Par exemple, dans le cas de problèmes inverses où les dérivées du système sont difficiles à définir analytiquement, ces méthodes trouvent une application particulière.

Bien que la littérature propose une diversité d'architectures de réseaux de neurones, y compris les réseaux multicouches de perceptrons (MLP), je me limiterai, dans le cadre de cette thèse, à décrire les modèles de réseaux que j'ai utilisés. Plus précisément, j'aborderai les réseaux de neurones convolutifs (CNN), les architectures U-Net, et les modèles YOLO (You Only Look Once).

2.4.3 Réseaux de neurones convolutifs

Les réseaux de neurones convolutifs ont connu une croissance exponentielle au cours des dernières années. Ils ont été utilisés dans une grande variété d'applications telles que la reconnaissance d'images, la classification, la détection d'objets, la segmentation d'images, la traduction de langues, la synthèse d'images et bien plus encore.

Ces architectures tirent leur inspiration du fonctionnement biologique du cortex visuel animal, en simulant ce processus à travers un enchaînement hiérarchique de couches neuronales. En termes de performance, de nombreuses études ont montré que les réseaux de neurones convolutifs surpassent les autres méthodes traditionnelles dans de nombreux domaines de traitement d'images. Ils sont également connus pour leur capacité à apprendre des caractéristiques visuelles complexes et abstraites à partir de données d'entrée brutes ; ce qui les rend particulièrement adaptés aux problèmes de vision par ordinateur.

Un réseau de neurones convolutifs (CNN) utilise des opérations de convolution pour extraire des caractéristiques à partir de l'image d'entrée. La couche de convolution est l'une des principales composantes du CNN. Cette couche utilise des filtres (également appelés noyaux de convolution) pour extraire des caractéristiques de l'image d'entrée, permettant une détection efficace de motifs dans une image. Chaque filtre est une matrice de poids qui est appliquée à une petite zone de l'image d'entrée, appelée champ récepteur. Le filtre est glissé sur l'image d'en-

trée, calculant la somme pondérée des valeurs de pixels dans le champ récepteur à chaque position, ce qui produit une carte de caractéristiques. Chaque filtre peut être optimisé pour détecter une caractéristique spécifique de l'image, tel que des bords, des textures ou des motifs, ou bien pour des tâches de débruitages d'images pour améliorer la qualité.

Les CNN sont constitués de plusieurs couches de convolution qui permettent d'obtenir une hiérarchie de caractéristiques. Les caractéristiques extraites par chaque filtre sont ensuite combinées dans une carte de caractéristiques globale qui est utilisée comme entrée pour la couche suivante du réseau. Grâce à la propriété d'invariance à certaines transformations, le partage de poids et le pooling, ces réseaux limitent le nombre de paramètres, ce qui les rend efficaces pour les images.

En plus de la couche de convolution, le CNN utilise également des couches de pooling pour réduire la dimensionnalité de la carte de caractéristiques et des couches entièrement connectées pour effectuer la classification. Les couches de pooling réduisent la taille de la carte de caractéristiques en prenant la moyenne ou le maximum des valeurs dans des régions adjacentes. Dans certains modèles, on peut même avoir des couches entièrement connectées qui prennent en compte toutes les caractéristiques extraites de la carte de caractéristiques globale pour effectuer une classification.

L'apprentissage dans le CNN se fait de la même façon que les réseaux de neurones multicouches standards grâce à la rétropropagation, où l'erreur de classification est propagée en arrière à travers le réseau, et les poids des filtres sont ajustés pour minimiser l'erreur. Pendant l'entraînement, les filtres sont ajustés pour maximiser l'activation de certaines caractéristiques de l'image d'entrée, qui sont utiles pour la tâche optimisée.

2.4.4 U-Net

Il existe diverses architectures de réseaux neuronaux à convolution, adaptées à la tâche de réduction du bruit et d'amélioration de la qualité d'image. Dans la littérature scientifique, le modèle U-Net [77] se distingue comme l'une des solutions les plus populaires pour aborder ces défis.

Ce modèle a été initialement mis en place dans le but de traiter les images médicales dans une tâche de segmentation des structures neuronales[77]. L'architecture du U-Net 2.15 est, comme son nom l'indique, sous la forme d'un U. Cette forme d'auto-encoder permet de traiter l'information à différentes échelles ce qui lui donne une robustesse face au bruit ; à gauche on retrouve une partie compressive où les blocs de convolutions sont suivis par des couches de Max pooling 2x2 dont le rôle est de réduire la taille des cartes caractéristiques. Elles permettent de fournir une invariance aux translations locales dans l'image et une réduction du bruit. Puis, du côté droit du modèle, l'information est décompressée avec des

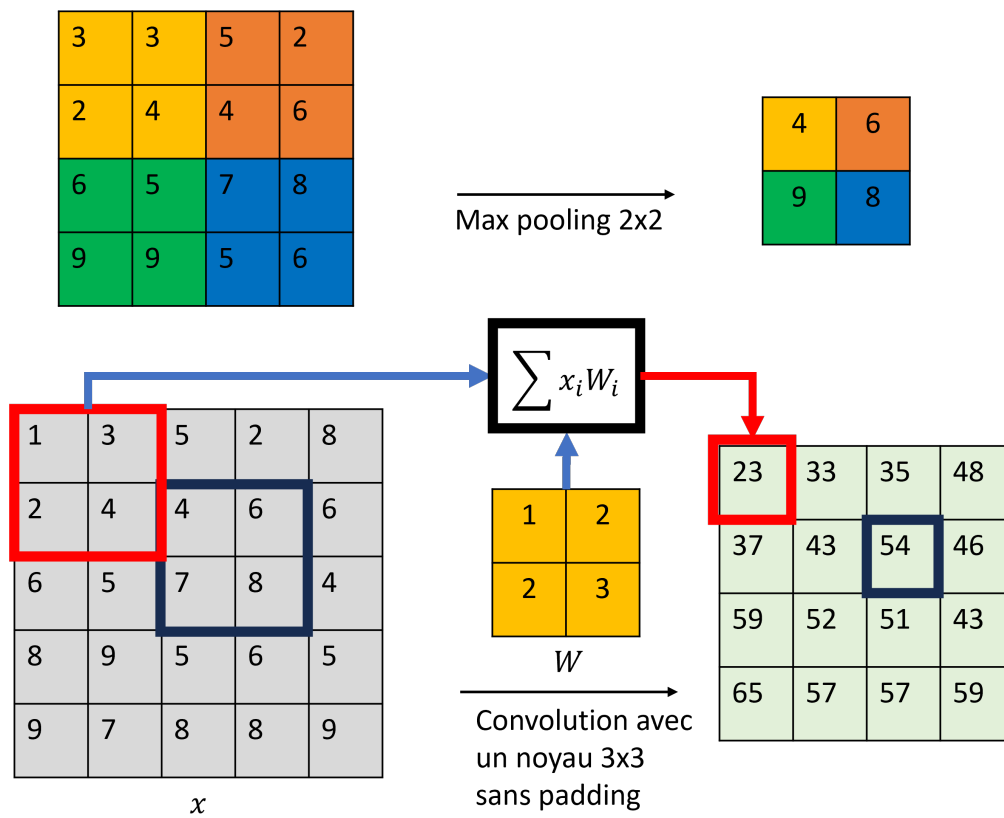


FIGURE 2.14 – Principe de la convolution et du max pooling dans un CNN

couches d'upsampling qui effectuent une interpolation des cartes caractéristiques pour augmenter leur résolution spatiale. Dans certains cas, les couches d'upsampling sont remplacées par des couches de déconvolutions (convolutions transposées) permettant d'effectuer une opération inverse à celle de la convolution, le but étant d'augmenter la résolution spatiale des cartes de caractéristiques par les poids appris contrairement aux couches d'upsampling qui sont de simples interpolations.

Les étages compressifs et décompressifs sont reliés à travers des "skips connections", qui sont des connexions permettant de récupérer les informations de haute résolution perdues lors de la phase de compression, mais aussi, ils permettent de combiner des informations de bas niveau et de haut niveau afin d'avoir des cartes caractéristiques plus riches.

Dans le papier original le U-Net est composé de 4 étages, chaque étage est composé de 2 blocs de convolutions avec n filtres de convolution (3×3) suivi d'une batch-normalisation et d'une activation Relu. Le nombre de filtres commence à 64 pour la couche la plus haute du modèle et augmente progressivement jusqu'à atteindre 1024 filtres au niveau de la couche la plus profonde.

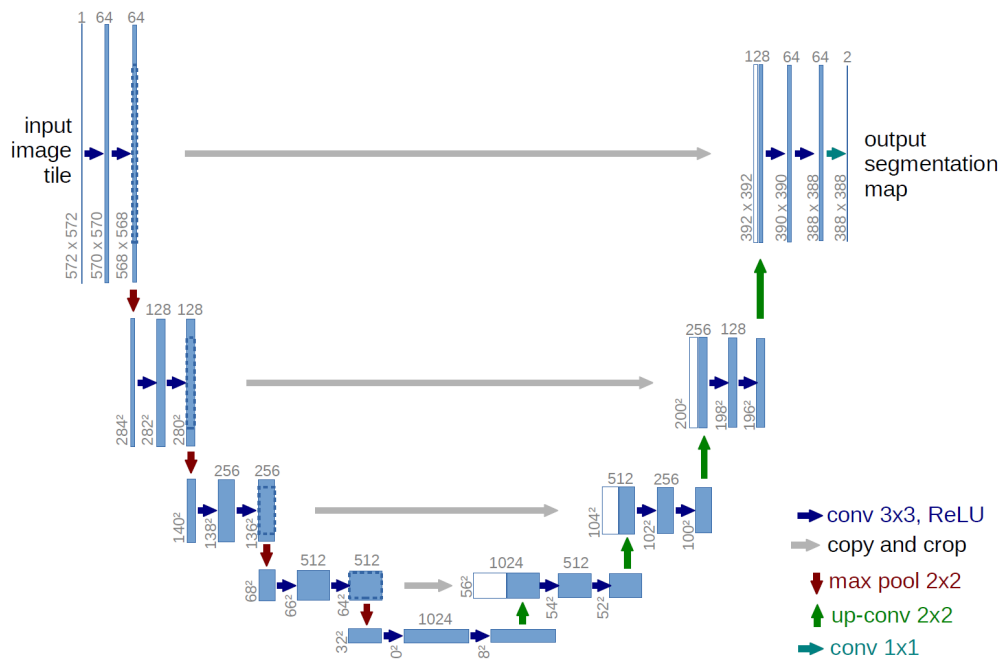


FIGURE 2.15 – Architecture utilisée pour la segmentation des images Biomédicales [77]

En résumé, les réseaux de neurones convolutifs, et plus spécifiquement le modèle U-Net, ont démontré leur efficacité dans les domaines de la segmentation et du débruitage, ainsi que dans le traitement d'images de manière plus générale. Ces

modèles sont utilisés dans un large éventail d'applications, y compris en médecine, pour la surveillance et la reconnaissance faciale, où ils contribuent de manière significative à l'amélioration des performances. En outre, ces modèles sont idéalement adaptés pour aborder des défis spécifiques tels que la reconstruction, la résolution de problèmes inverses, et la super-résolution.

Le choix de l'architecture U-Net pour cette thèse s'est basé sur une série de comparaisons avec des travaux existants. En confrontant le U-Net avec d'autres modèles neuronaux, en particulier ceux utilisés par Yair et al.[73] ainsi que le modèle SRRESNET [27], il est apparu que le U-Net offrait de meilleurs résultats.

Cette décision a aussi été renforcée par les résultats notables mentionnés dans la littérature scientifique en matière de reconstruction [33] et de débruitage d'images [39] dans le domaine de l'imagerie non conventionnelle, utilisant des architectures semblables à celle du U-Net.

2.4.5 YOLO

La détection et la classification des objets sont l'un des défis de la vision par ordinateur. Les modèles mis en place cherchent à détecter, localiser précisément et classer des objets spécifiques dans une image. Les réseaux de neurones convolutifs (CNN) sont souvent impliqués dans les modèles cherchant à résoudre ce type de question en raison de leur capacité à extraire des caractéristiques diverses et complexes à partir d'images.

Les modèles basés sur des réseaux neuronaux pour la détection et la classification d'objets se répartissent généralement en deux grandes catégories. La première comprend les modèles en deux étapes, qui réalisent initialement la détection d'objets avant de procéder à la classification des objets détectés. La seconde catégorie inclut les modèles qui accomplissent simultanément les deux tâches, à savoir la détection et la classification, en une seule étape.

Les modèles à deux étapes, tels que R-CNN, Fast R-CNN [34] et Faster R-CNN [72], sont généralement basés sur une approche de proposition de région (RPN). La première partie de ces modèles utilise des algorithmes de propositions de région pour identifier des régions potentielles contenant des objets dans l'image. Les régions proposées sont par la suite transmises au CNN afin d'effectuer la tâche de classification recherchée. La majorité de ces modèles utilise un apprentissage alternatif entre la partie RPN et la classification.

Les modèles à une étape, tels que « YOLO » (You Only Look Once) et SSD (Single Shot Detector), sont plus rapides que les modèles à deux étapes, car ils ne nécessitent pas de génération de propositions de région. Au lieu de cela, ils effectuent la détection d'objets et la classification en une seule étape en appliquant un CNN sur l'ensemble de l'image en utilisant une grille régulière. « YOLO » est un modèle de détection d'objets en temps réel, qui fonctionne en découpant une

image en une grille de cellules et en prédisant la présence et la classe d'objets dans chaque cellule. SSD est un modèle similaire à YOLO, mais qui utilise différentes tailles de boîtes d'ancrage pour améliorer la précision de la détection.

Le Yolo (You Only Look Once) a été présenté dans sa version originale dans [71]. Depuis lors, plusieurs versions améliorées ont été proposées, telles que Yolo V2, V3, V4 et une version légèrement plus petite, YOLO-tiny, qui est spécialement conçue pour atteindre une vitesse élevée de 220 images par seconde. Ce paragraphe est largement inspiré de [71].

Le choix de ce modèle a été guidé par sa rapidité, ainsi que par sa capacité à résoudre des problématiques de détection et de classification d'objets de tailles variées avec un seul modèle. Ceci est particulièrement utile pour détecter des éléments sanguins, qui sont au cœur de notre intérêt, tels que les leucocytes, mesurant environ entre 10 et 20 microns, les globules rouges d'environ 7 microns, et les plaquettes, plus petites, de l'ordre de 1 à 2 microns. Le choix s'est porté sur ce modèle, spécifiquement sur la version YOLO v3, qui est reconnue pour sa performance. Cette version se distingue par sa capacité à détecter et classifier de manière efficace des éléments de tailles variées, rendant cet outil particulièrement adapté à nos besoins.

Grâce à l'architecture multi-échelle du YOLO V3, illustrée dans la figure 2.16, ce modèle assure une détection efficace d'objets à différentes échelles. Exploitant divers blocs de convolution, le YOLOv3 utilise le Darknet-53 comme architecture de base ("backbone"), lequel est composé de blocs de convolution, de normalisation par lots (batch normalization) et d'activation Leaky ReLU. Il intègre également plusieurs couches de concaténation pour fusionner les informations à différents niveaux.

YOLO est conçu pour prédire à la fois la catégorie à laquelle appartient un objet et également le rectangle englobant cet objet. Chaque rectangle est caractérisé par quatre paramètres : les coordonnées du point central (t_x , t_y), sa largeur (t_w) et sa hauteur (t_h). En outre, l'algorithme attribue un score de probabilité, noté C , qui évalue la confiance que l'on peut avoir quant à la présence effective de l'objet dans le rectangle défini. YOLO génère aussi un coefficient, noté $p(c)$, qui mesure le degré de certitude de l'appartenance de l'objet à une classe spécifique.

L'image est divisée en cellules. Chaque cellule est alors responsable de la prédiction de K boîtes englobantes et avec un score de confiance pour chaque boîte. On considère qu'un objet se trouve dans une cellule spécifique uniquement si les coordonnées centrales de la boîte d'ancrage se trouvent dans cette cellule. Pendant la phase d'apprentissage, YOLO détermine la probabilité que la cellule contienne un objet d'une certaine classe. Le score de confiance reflète le degré de confiance du modèle dans le fait que la boîte contient un objet. Seules les boîtes avec un degré de précision supérieur à 0.5 sont retenues. Ce seuil est calculé selon la formule

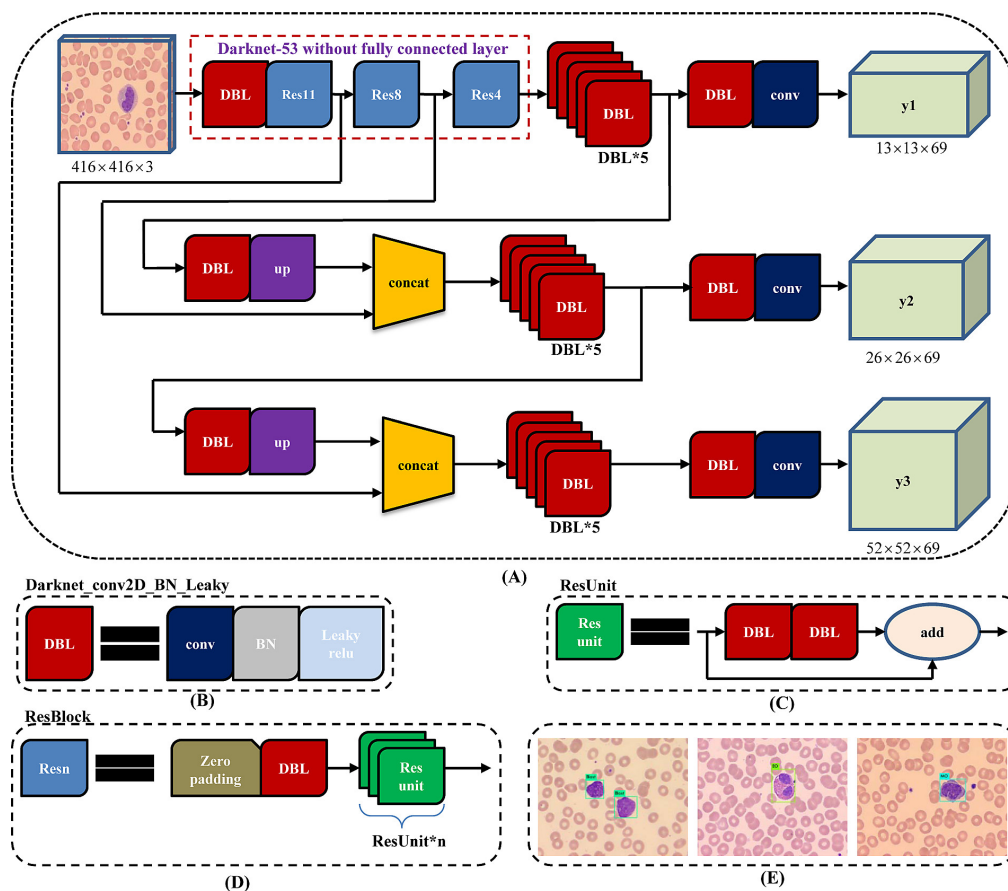


FIGURE 2.16 – Architecture YOLOV3 [89] L’algorithme YOLO est basé sur une architecture neuronale convolutive avec une approche de régression. Il permet de réaliser une classification d’objets et une localisation simultanées en une seule passe.

suivante 2.22 pour toutes les boîtes (j) de la cellule.

$$C_j(Object) = C(Object). (IOU(pred, true) > seuil_{limite}) \quad (2.22)$$

l’IOU mesure l’intersection sur l’Union entre la boîte prédite et la grille prédéfinie. Chaque cellule prédite a également des probabilités de classe conditionnelles pour chaque objet $p_i(Class = c/object)$

La probabilité Pr retenue est celle qui correspond à la valeur la plus élevée parmi toutes les classes dans la cellule. Suivant la règle de Bayes, la probabilité que l’objet soit dans class= i est donnée par :

$$p_i(c) = Pr(class = i/object).C(Object) \quad (2.23)$$

La classe de chaque boîte est attribuée en fonction d'un seuil de décision appliqué sur $p_i(c)$. Le processus reposant sur les probabilités de classe peut conduire à de nombreuses boîtes d'ancrage. Il est donc nécessaire de supprimer celles qui ne sont pas nécessaires. Pour résoudre ce problème, la suppression Non-max élimine les boîtes d'ancrage qui sont très proches en effectuant l'IoU (Intersection sur Union) avec celle qui a la plus grande probabilité de classe parmi elles ; l'algorithme rejette ensuite les boîtes dont la probabilité est supérieure à celle de la classe. Ce seuil est fixé empiriquement à 0,5 dans la littérature. Enfin, les sorties de l'algorithme sont le vecteur requis indiquant les paramètres de la boîte englobante associée à l'objet de l'image respective. Voici la formule de la fonction de coût du YOLOv3 détaillée :

$$\mathcal{L} = \lambda_{coord}\mathcal{L}_{coord} + \lambda_{obj}\mathcal{L}_{obj} + \lambda_{class}\mathcal{L}_{class} \quad (2.24)$$

où : λ_{coord} , λ_{obj} et λ_{class} sont des hyper paramètres qui pondèrent les différents termes de coût. \mathcal{L}_{coord} mesure l'erreur de prédiction de la position et de la taille des boîtes englobantes des objets. Elle est définie comme suit :

$$\mathcal{L}_{coord} = \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{K}_{ij}^{obj} \cdot \left[(\hat{t}_x - t_x)^2 + (\hat{t}_y - t_y)^2 + (\hat{t}_w - t_w)^2 + (\hat{t}_h - t_h)^2 \right] \quad (2.25)$$

où \mathbb{K}_{ij}^{obj} est un indicateur qui vaut 1 si la j-ème boîte englobante de la cellule i prédit un objet, et 0 sinon. Les termes \hat{t}_x , \hat{t}_y , \hat{t}_w et \hat{t}_h représentent les coordonnées et les dimensions de la boîte prédite, tandis que t_x , t_y , t_w et t_h représentent les coordonnées et les dimensions de la boîte réelle.

\mathcal{L}_{obj} mesure l'erreur de prédiction de la présence ou de l'absence d'objets dans les boîtes englobantes. Elle est définie comme suit :

$$\mathcal{L}_{obj} = \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{K}_{ij}^{obj} \cdot (\hat{C}_j - C_j)^2 \quad (2.26)$$

où \mathbb{K}_{ij}^{obj} est un indicateur qui vaut 1 si la j-ème boîte englobante de la cellule i prédit un objet, et 0 sinon. Les termes \hat{C}_j et C_j représentent la confiance prédite et la confiance réelle, respectivement.

\mathcal{L}_{class} mesure l'erreur de prédiction des classes des objets détectés. Elle est définie comme suit :

$$\mathcal{L}_{class} = \sum_{i=0}^{S^2} \mathbb{K}_{ij}^{obj} \sum_{c \in classes} (\hat{p}_i(c) - p_i(c))^2 \quad (2.27)$$

où $\hat{p}_i(c)$ et $p_i(c)$ représentent la probabilité prédite et la probabilité réelle de la classe c , respectivement. Le modèle YOLOv3 [70] 2.16 est une version améliorée de YOLO. Cette architecture utilise un réseau de neurones plus profond et plus complexe, où l'image d'entrée est analysée à plusieurs échelles, ce qui permet d'améliorer la détection d'objets de petite taille.

2.5 Contexte Biologique

L'analyse au microscope optique demeure un outil essentiel en cytohématologie, car elle offre la possibilité d'examiner les cellules sanguines de manière rapide et précise, permettant une caractérisation morphologique détaillée des cellules en présence.

L'objectif de ma recherche est de mettre en évidence l'apport de l'imagerie FPM grâce à la bi-modalité en intensité et phase, dans diverses applications biologiques.

Au cours de mes recherches, j'ai principalement porté mon attention sur deux domaines d'étude : la parasitologie, avec une emphase spécifique sur le paludisme, et la cytohématologie, en me focalisant particulièrement sur l'étude des images de leucocytes dans le cadre du projet TAMIS.

2.5.1 Parasitologie

Parmi les maladies qui nécessitent le recours au frottis sanguin et à l'analyse microscopique, nous retrouvons le paludisme. Le paludisme, également connu sous les noms de malaria ou de 'fièvre des marais', est une maladie infectieuse à potentiel mortel. Elle est causée par diverses espèces de parasites du genre Plasmodium. La propagation de cette maladie est principalement facilitée par les moustiques du genre Anophèle, qui agissent en tant que vecteurs clés, propageant l'infection aux humains par leurs piqûres. Le paludisme compte parmi les 20 principales causes de décès à l'échelle mondiale, avec environ 619 000 décès et une estimation de 247 millions de cas recensés par l'Organisation Mondiale de la Santé en 2021. Cette maladie est principalement présente en Afrique, où elle représente 95 % de l'ensemble des cas. Dans cette région, les enfants de moins de cinq ans constituent 80% des victimes. Un peu plus de la moitié de ces décès sont concentrés dans quatre pays : le Nigeria (31,3%), la République démocratique du Congo (12,6%), la Tanzanie (4,1%) et le Niger (3,9%). D'autres cas d'infection par le paludisme sont répertoriés en Asie et en Amérique Latine, régions où une réduction significative du nombre de cas a été observée. Cependant, l'éradication totale de cette maladie se révèle particulièrement ardue en raison du développement de la résistance du parasite face aux campagnes d'éradication. Dans les pays de l'Union Européenne, les cas de paludisme sont généralement associés aux voyages vers les

zones endémiques de la maladie, ainsi qu'à l'immigration en provenance de ces régions vers l'UE. Le paludisme est propagé par les femelles de moustiques anophèles lorsqu'elles se nourrissent de sang, injectant ainsi des organismes microscopiques mobiles appelés sporozoïtes. Ces derniers infiltrent les cellules du foie où ils se reproduisent pendant au moins une semaine. Lorsque cette première étape de reproduction atteint sa maturité, les nouveaux parasites, appelés mérozoïtes, sont libérés et envahissent les globules rouges, amorçant ainsi un cycle de reproduction non sexué dans le sang. Ce cycle varie en durée selon l'espèce du parasite : environ 48 heures pour *Plasmodium falciparum*, *P. vivax* et *P. ovale*, 72 heures pour *P. malariae* et seulement 24 heures pour *P. knowlesi*. Pendant cette phase, le parasite en croissance consomme l'hémoglobine des globules rouges. En l'absence de traitement, les globules rouges infectés par *Plasmodium falciparum*, notamment ceux qui hébergent le parasite depuis plus de 24 heures, subissent des modifications morphologiques qui les font adhérer à l'intérieur des parois des capillaires vasculaires.

Lorsque le nombre de globules rouges infectés atteint un niveau élevé, ce qu'on appelle "parasitémie élevée", ces cellules sanguines ont tendance à se séquestrer, ou s'accumuler, dans les capillaires. Ce phénomène peut entraîner des perturbations métaboliques sévères, contribuant ainsi à la gravité et au taux de mortalité associés à la maladie. Ces problèmes peuvent affecter divers organes, notamment le cerveau (entraînant un risque de coma), les reins, le foie, les poumons, entre autres.

L'établissement du diagnostic du paludisme s'effectue actuellement en laboratoire spécialisé. Face à une suspicion de paludisme, le laboratoire est tenu de livrer les résultats biologiques dans un délai de deux heures à partir de la réception de l'échantillon. Le gold standard reste toujours l'analyse par un biologiste du frottis au microscope optique pour vérifier la présence ou l'absence du paludisme et de quantifier la parasitémie, l'espèce et le stade du parasite. En complément, l'examen PCR est obligatoire pour obtenir une haute sensibilité.

Plusieurs méthodes d'intelligence artificielle basées sur des techniques d'imagerie conventionnelles ont été entreprises pour relever le défi du diagnostic du paludisme à partir de frottis sanguins [87, 80, 41, 67, 68, 42, 53]. Néanmoins, cette tâche reste un défi majeur dans le domaine médical.

L'utilisation d'un dispositif FPM permet ici d'analyser un grand nombre d'hématies (environ 3500), et ce avec une résolution équivalente à celle d'un objectif 100xNA1,5. Comme le montre la figure 2.17, nous avons acquis en microscopie FPM un large champ au grandissement 20x avec une ouverture numérique (NA) de 0,75. La configuration des LEDs, composée de 13 LEDs disposées en anneaux et d'une diode centrale, permet d'obtenir des images reconstruites avec une ouverture numérique synthétique NA_{synth} de 1,5.

En combinant l'intelligence artificielle et la bi-modalité introduite par la FPM,

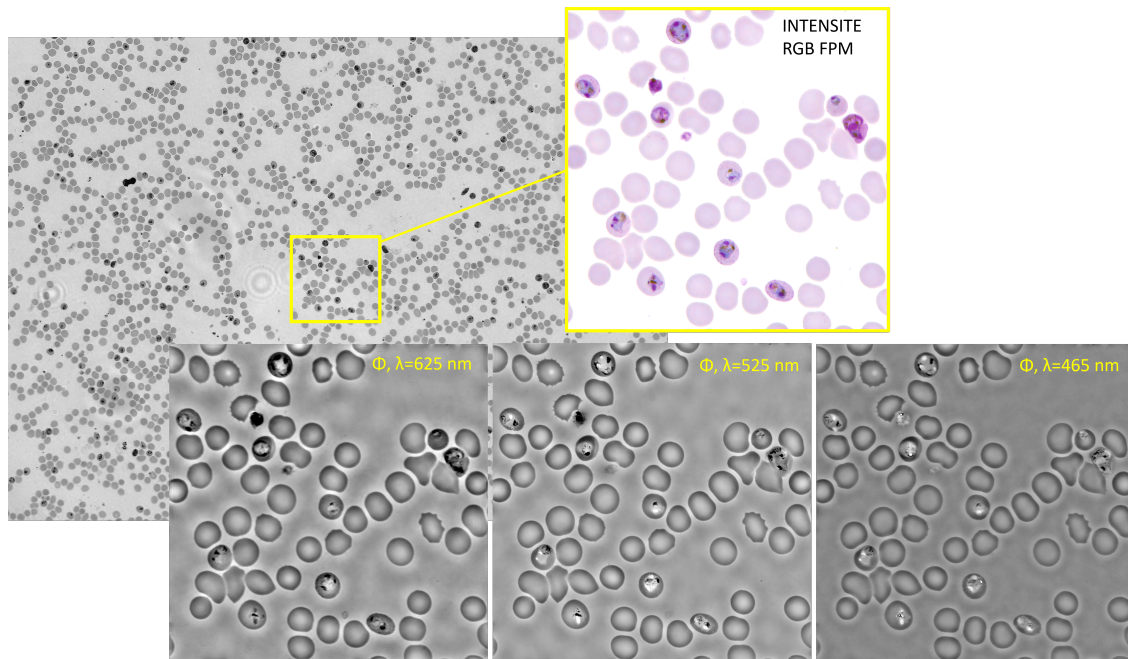


FIGURE 2.17 – Grand champ FPM sur un frottis de paludisme, avec un objectif $20\times\text{NA}=0.75$ avec 3 longueurs d'onde $\lambda = 625\text{nm}, 525\text{nm}, 465\text{nm}$

nous cherchons à faciliter l'analyse de larges champs de vue tout en générant des images de haute qualité. Cette approche étudie la possibilité de développer une méthode pragmatique pour le suivi de cette maladie, en utilisant un dispositif simple et performant.

2.5.2 Leucocytes

Les leucocytes, également appelés globules blancs, jouent un rôle primordial dans la défense immunitaire de l'organisme. Ils sont les acteurs principaux de la réponse immunitaire, protégeant le corps contre les infections et autres agents pathogènes. Il existe cinq types principaux de leucocytes : neutrophiles, lymphocytes, monocytes, éosinophiles et basophiles, chacun ayant des fonctions spécifiques dans la réponse immunitaire.

L'analyse morphologique de ces cellules est cruciale dans le diagnostic et le suivi de nombreuses maladies. Par l'observation de leur taille, de leur forme et de leur structure, on peut déterminer si ces cellules sont normales ou si elles présentent des anomalies. Ces anomalies peuvent inclure des changements dans la taille ou la forme des cellules, une proportion anormale de différents types de leucocytes, ou la présence de cellules immatures qui normalement ne devraient pas se retrouver dans le sang périphérique.

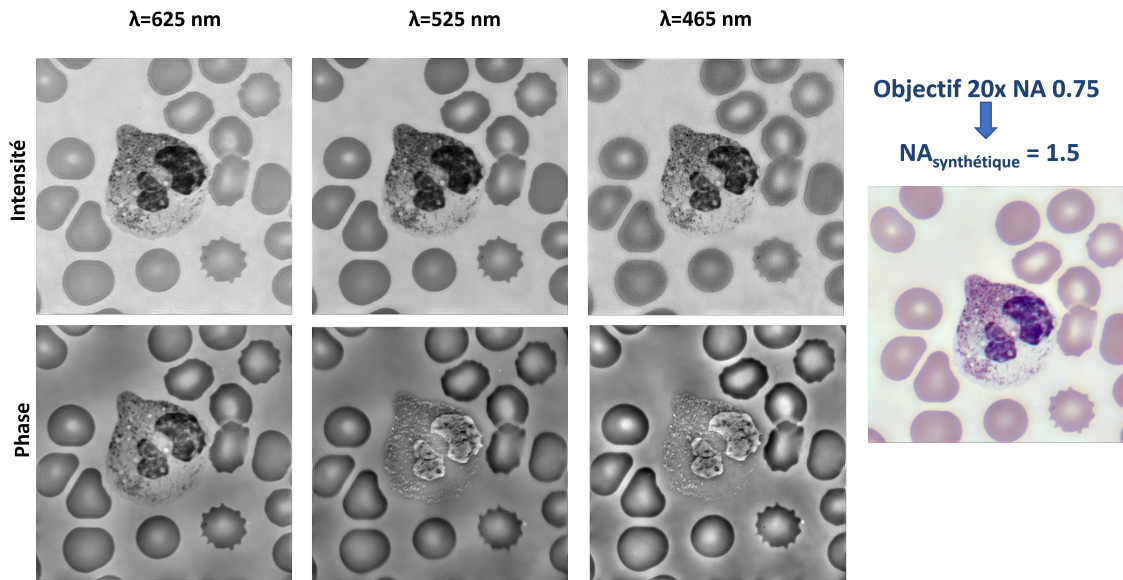


FIGURE 2.18 – Polynucleaire neutrophile, acquies avec un objectif 20xNA=0.75 avec 3 longueurs d’onde $\lambda = 625nm, 525nm, 465nm$

À l’heure actuelle, il existe plusieurs systèmes automatisés, tels que CellaVision, qui visent à analyser les leucocytes de manière automatique. Ces systèmes utilisent des images issues de microscopes optiques traditionnels et classifient les cellules grâce à des outils d’intelligence artificielle. Ce défi a fait l’objet de nombreuses études visant à améliorer la classification des leucocytes [8, 28, 58, 66, 11] et à établir un diagnostic complet basé sur la morphologie cellulaire.

La Figure 2.18 illustre la richesse de l’information apportée par la FPM, nous y présentons un exemple de leucocytes acquies à l’aide d’un microscope FPM doté d’une ouverture numérique synthétique NA_{sys} de 1,5. Nous disposons de six canaux reconstruits au point focal optimal pour une meilleure caractérisation des globules blancs. De plus, la combinaison des images d’intensité reconstruite en RGB pour les trois longueurs d’onde $\lambda = 625, nm, 525, nm, 465, nm$ est également disponible.

Dans le cadre de mon travail, l’objectif est de produire des images FPM de haute résolution et de qualité optimale, permettant une analyse morphologique précise des cellules.

2.6 Difficulté de la FPM

Malgré les nombreux avantages de la ptychographie de Fourier, tels que l’accès à des informations en très haute résolution et l’obtention de données de phase à partir d’un dispositif expérimental relativement simple, la FPM (Fourier Ptycho-

graphy Microscopy) peine à se populariser, en particulier dans le domaine de la pathologie numérique et du diagnostic automatisé. Dans ce secteur, où les images de microscope connaissent un renouveau grâce à l'émergence des réseaux de neurones, la FPM a encore du mal à s'imposer.

Les défis auxquels la ptychographie est confrontée partagent de nombreux points communs avec ceux rencontrés dans la microscopie optique en général, tels que le défocus et les aberrations optiques. Cependant, la ptychographie présente également des contraintes spécifiques, notamment en ce qui concerne le temps nécessaire à la reconstruction des images, le nombre d'acquisitions requises pour cette reconstruction, et leur sensibilité à la qualité des images brutes acquises par la caméra.

Dans cette section, je vais aborder ces différentes problématiques en les illustrant par des exemples concrets. Ensuite, je discuterai des limitations que j'ai traitées dans le cadre de ma thèse.

2.6.1 Correction numérique des aberrations globales du système optique

Afin d'obtenir une bonne reconstruction et d'atteindre la limite de résolution FPM synthétiques théoriques, la correction des aberrations des objectifs est essentielle. Plusieurs approches ont déjà été proposées, dont une intégrée à l'algorithme de reconstruction (embedded pupil function recovery -EPRY[62]) et reposant sur une optimisation conjointe de la fonction de phase et de la fonction de pupille. Ces approches, bien qu'efficaces, ne permettent pas de gérer les aberrations globales du système optique, en particulier celles introduites par la préparation de l'échantillon comme la colle, le couvre-objet, etc.

Lorsque l'on cherche à atteindre une résolution inédite, surpassant les limites de résolution des microscopes optiques traditionnels avec des ouvertures numériques ($NA > 1.5$), les exigences relatives aux conditions expérimentales de mesure s'intensifient considérablement.

Par exemple, pour des échantillons préparés avec un couvre-objet, il devient impératif d'ajuster avec précision la bague de compensation d'épaisseur du couvre-objet intégrée à l'objectif, et ce, pour chaque échantillon. Cette étape est cruciale pour obtenir des reconstructions de haute qualité. Cette sensibilité aux variations, de l'ordre de 0,01 mm, rend cette technique de microscopie particulièrement difficile à maîtriser pour un non-spécialiste.

La Figure 2.19 présente des images de l'échantillon, acquises en utilisant à la fois une diode centrale et une diode latérale, ainsi que les images reconstruites en intensité et en phase. Les images en haut 2.19 correspondent à une correction de l'anneau de couvre-objet de 0,19 mm en haut et 0,18 mm en bas.

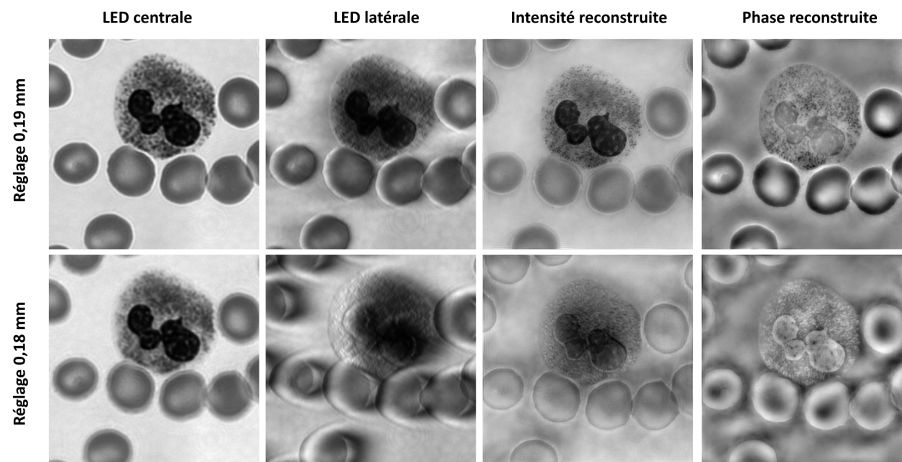


FIGURE 2.19 – Illustration de la sensibilité de la microscopie FPM en fonction du réglage de la bague de correction du couvre-objet, en utilisant un objectif 40x à ouverture numérique (NA) de 0.95 et un facteur de super-résolution de 2x

Il est évident que les images brutes de haute qualité sont obtenues lorsque le réglage de la bague de correction du couvre-objet est calibré à 0,19 mm.

Dans ces conditions, les images reconstruites, tant en termes de phase que d'intensité, surpassent nettement celles obtenues avec un réglage à 0,18 mm. Ce résultat souligne l'importance cruciale d'un ajustement mécanique précis de la correction du couvre-objet de l'objectif.

Nous ne savons actuellement pas compenser numériquement les aberrations introduites par ce réglage mécanique, de surcroît de façon automatisée. Il existe donc un enjeu applicatif réel de mieux étudier cette question, notamment en ré-analysant les fondations des modèles existants et/ou en proposant des approches calculatoires plus adaptées.

2.6.2 Correction numérique du focus

La microscopie ptychographique de Fourier offre l'avantage d'utiliser des champs de vision de caméra de grande taille avec une résolution élevée. Cependant, cette caractéristique engendre des défis particuliers. Notamment, il devient difficile de maintenir une netteté uniforme sur l'intégralité du champ de vision lorsque l'objectif possède une résolution native élevée et, par conséquent, une faible profondeur de champ. De plus, même une légère inclinaison de l'échantillon peut causer des variations de focalisation à travers le champ de vision (FOV). Ces limitations obligent souvent les techniques de microscopie conventionnelle à se concentrer sur un seul objet par champ de vision. Par conséquent, elles n'exploitent qu'une petite surface et nécessitent de multiples mesures et d'ajustements de mise au point pour

l'analyse de large surface.

Les variations de mise au point compliquent l'exploitation efficace de larges champs de vues (FOV) en microscopie. Bien que la FPM autorise des ajustements numériques du focus en post-acquisition, cette méthode nécessite une estimation a priori du décalage de focus δz souhaité. Alternativement, il est possible de tester plusieurs valeurs de δz afin d'identifier la correction optimale, cependant ces approches compliquent l'automatisation de la technique.

2.6.3 Temps d'acquisition

Actuellement, les temps d'acquisition sont très importants, de l'ordre de la minute par champ, pour atteindre des résolutions élevées bien que l'on se limite à un facteur de super-résolution de seulement $\times 2$.

De même, nous avons également des temps très longs, lorsque l'on cherche à reconstruire des images avec de grands facteurs de super-résolution par rapport à l'ouverture numérique de l'objectif, on est obligé d'éclairer l'échantillon avec un nombre élevé de LEDs (car il faut que la diversité angulaire d'illumination permette de couvrir des inclinaisons importantes, tout en respectant le facteur de recouvrement fixé à 60%). Pour les images en champ sombre où l'angle d'éclairage se situe à l'extérieur du cône optique de l'objectif, le flux de photons récolté par la caméra est faible, ce qui entraîne des temps d'exposition importants par LED.

Il est à noter que le nombre de LED à utiliser est imposé par l'algorithme de reconstruction (de l'ordre de 225 pour des facteurs de sur-résolution de 6). Ceci induit un dispositif peu adapté à un cadre clinique. Des travaux se sont intéressés à dépasser cette limitation dont ceux de Laura Waller et al [84] qui introduit une technique de multiplexage d'éclairage et une adaptation de l'algorithme FPM qui permet de réduire significativement le nombre de LED et le temps d'exposition par prise. Cependant, les résultats observés restent moins bons qualitativement qu'un éclairage LED par LED.

2.6.4 Temps de reconstruction

Les algorithmes utilisés en ptychographie sont des algorithmes itératifs qui effectuent la correction de l'objet reconstruit LED par LED ; cette correction est effectuée dans un ordre donné avec comme condition l'existence d'un recouvrement spectral entre les LED qui se suivent, ce qui rend difficile, toute parallélisation informatique, et conduit à des temps de reconstruction pouvant être longs. Bien que des stratégies antérieures aient envisagé le découpage de grands champs de vision (FOV) pour une reconstruction parallèle de zones plus petites, à assembler ultérieurement, l'adoption des unités de traitement graphique (GPU) et des réseaux de neurones offre de nouvelles perspectives pour optimiser ce processus.

2.6.5 Bruit

La reconstruction des images FPM repose sur un problème d'optimisation non convexe dont l'unicité ne peut être garantie. Toutefois, la régularité « physique » des solutions recherchées (relatives aux échantillons), bien que jamais explicitée ou formalisée, est indéniablement exploitée d'une façon ou d'une autre lors de la mise au point des algorithmes de reconstruction.

Or, un élément particulièrement délicat à modéliser concerne le modèle même de bruit à considérer. En effet, des variations du flux optique importantes conduisent à des modèles de bruit pouvant évoluer suivant l'angle d'éclairage. Par exemple, sous incidence normale (champ clair), le bruit dominant est vraisemblablement un bruit de grenaille, sous incidence oblique (champ sombre) un bruit thermique. Ainsi le SNR est indéniablement non constant et difficilement modélisable par une seule source de bruit. Il dépend également des conditions expérimentales et est particulièrement gênant lorsque l'on cherche à atteindre un facteur de super résolution élevée ($\gamma > 3$). Dès lors, la contribution du bruit sur le comportement des algorithmes de reconstruction me semble une question clé à l'origine des limitations d'ordre fondamental. Il s'agit d'une question encore peu abordée à notre connaissance dans la littérature qui gagnerait à être étudiée dans un cadre d'apprentissage statistique, propice à l'établissement expérimental des lois du bruit mises en œuvre. Ce même cadre pourrait également s'avérer propice à la réduction de leur impact.

Chapitre 3

Refocalisation des images FPM par réseaux de neurones

Ce chapitre s'inscrit dans la continuité de l'utilisation de la FPM, présentée dans le chapitre 2, pour la pathologie numérique. Plus précisément, il y est présenté une méthodologie originale que j'ai proposée pour la refocalisation des images FPM. Ce travail a fait l'objet d'une publication dans le journal *Biomedical Optics Express* [16] ainsi que d'un dépôt de brevet.

3.1 Introduction

La numérisation de spécimens biologiques avec des scanners de lames à résolution microscopique, est devenue indéniable ces dernières années. Ce dynamisme est alimenté en partie par le "Whole-Slide Imaging" (WSI), une technique qui révolutionne le stockage, l'archivage et le partage d'échantillons médicaux en format numérique. Cet élan est également dû à l'innovation dans les techniques de diagnostic automatisé, y compris l'utilisation de l'apprentissage profond (Deep Learning, DL). En conséquence, la pathologie numérique s'affirme comme un domaine en rapide évolution.

Toutefois, il faut noter que les technologies de numérisation actuelles, souvent basées sur des microscopes optiques traditionnels, rencontrent plusieurs limitations. Ces appareils, produisant des images selon l'intensité seule, sont limitées par plusieurs aspects tels que la résolution optique, la superficie du champ de vision (FOV) analysable, et la vitesse de capture et de traitement de l'information.

Dans le cadre biologique, l'observation à haute résolution est souvent une exigence, impliquant des spécifications d'objectifs particulières. Par exemple, en parasitologie, il est fréquent d'utiliser des objectifs à 40x de grandissement et une ouverture numérique (NA) de 0,9. En hématologie, les NA peuvent atteindre jusqu'à

1.5 grâce aux techniques d’immersion à l’huile. Ces spécificités ont pour conséquence de limiter considérablement la profondeur de champ, complexifiant ainsi la mise au point sur l’ensemble du FOV.

En revanche, comme expliqué dans le chapitre précédent, certaines techniques d’imagerie non conventionnelles, comme la Fourier Ptychographic Microscopy (FPM), offrent des avantages significatifs. La FPM permet, par exemple, d’élargir le FOV sans compromettre la résolution, grâce à une synthèse d’ouverture numérique optimisant le produit espace-bande passante. De plus, la profondeur de champ (DOF) dans la FPM est naturellement plus élevée et est inversement proportionnelle au carré de la NA de l’objectif, indépendamment du facteur de super-résolution utilisé.

La FPM est particulièrement intéressante pour les applications de diagnostic automatisé, notamment en hématologie et en parasitologie. Sa capacité à fournir des informations à la fois sur l’absorption et le chemin optique en fait un outil idéal pour améliorer la sensibilité et l’efficacité du diagnostic. Théoriquement, cette riche information bi-modale, permet de générer numériquement des images à différents plans de mise au point, conformément aux travaux de Goodman [38].

Toutefois, il est important de noter qu’en fonction de la nature du spécimen étudié, tous les objets présents dans le champ de vision (FOV) ne sont pas nécessairement au point optimal, même lorsque l’on utilise des techniques d’imagerie avancées comme la Fourier Ptychographic Microscopy (FPM).

3.2 Etat de l’art

Pour répondre à cette problématique de focus et qualité d’image, différentes techniques de traitement d’image post-acquisition et des méthodes d’adaptation du système d’acquisition ont été développées. Les techniques mises en œuvre permettent d’étendre la profondeur de champ du microscope (l’article de revue de Bian et al. [12]). Dans ces approches, les images présentant une profondeur de champ étendue sont obtenues en corrigeant numériquement les imprécisions de mise au point. Cette correction est effectuée au moyen de calculs post-acquisition ou de traitement d’image.

Une approche couramment utilisée en microscopie pour obtenir une profondeur de champ étendue est le z-stacking, où des images sont capturées à différents plans de mise au point (typiquement 10 plans ou plus). La profondeur de champ étendue est ensuite obtenue en combinant ces différentes images de manière appropriée à l’aide de divers algorithmes.

Une extension de l’approche avec un nombre de plans de mise au point plus faible a été rapportée dans l’article de Qian Li [51] où seulement deux images hors focus, prises autour du plan d’autofocus mécanique du microscope, sont fusionnées

en utilisant un réseau neuronal profond (DNN).

Ces limitations sont indéniablement à reconsidérer avec les récentes avancées en microscopie [64]. Parmi celles-ci, la microscopie ptychographique de Fourier [95](FPM) comme expliqué au Chapitre 2. En plus de l'intensité traditionnelle, cette microscopie permet d'obtenir l'information de phase des échantillons biologiques de manière simple [50] (c'est-à-dire sans utilisation de montage d'interférence).

Depuis la première démonstration du principe de fonctionnement du microscope en 2013 par Zheng [95] et al, de nombreux progrès ont été réalisés, comme en témoignent le nombre de résultats publiés et la variété des sujets abordés. Par exemple, on peut penser à la mise au point numérique des images lors de leur reconstruction [95], à la réduction du temps d'acquisition [84] permettant des taux de balayage élevés, à la correction intégrée des aberrations [62], à la tomographie reconstruite d'échantillons biologiques [85, 40, 65].

Néanmoins, l'exploitation complète des images FPM demeure un défi en raison des variations de mise au point sur un grand champ de vision (FOV). Dans l'article [52], une approche permettant d'obtenir des images avec une profondeur de champ étendue a été introduite. Leur méthode se base sur la reconstruction multi-Z de l'échantillon biologique, puis, pour chaque sous-région du champ de vision, on sélectionne la meilleure image en fonction d'un score de mise au point, que l'on réassemble à la fin pour produire un large FOV au bon focus. L'inconvénient de cette approche réside dans la nécessité de calculer plusieurs reconstructions pour différents plans de mise au point pour chaque objet d'intérêt, ce qui est assez lourd en termes de calcul.

Une approche assez similaire [23] reposant sur la sélection du meilleur plan de mise au point pour chaque globule rouge a été mise en place. Elle évite d'avoir à effectuer la reconstruction à différents plans : ici, on reconstruit une fois, la phase et l'intensité, puis on utilise cette information pour introduire une fonction de défocalisation globale du champ complexe de l'objet. Cependant, cette approche a été réfutée par les travaux de [96] qui indique qu'elle ne fonctionne pas pour les petites variations de mise au point et qu'il n'a jamais été prouvé expérimentalement ou théoriquement l'existence d'une fonction de phase globale en microscopie.

Dans un autre article [94], un effet de décalage latéral dans les images d'intensité brute est exploité pour estimer δz , la distance de defocus. Cette valeur est ensuite utilisée comme paramètre de reconstruction pour générer des images nettes. Un tel processus peut poser problème lorsque l'épaisseur de l'échantillon est comparable à la DOF du microscope. Dans ce cas, et comme nous le détaillerons, une variabilité importante dans le contraste et la netteté des composants sous-cellulaires peut être rencontrée dans les images de phase. Par conséquent, le paramètre δz estimé à partir des seules images d'intensité n'est pas nécessairement

adapté à la reconstruction des images de phase.

Approche proposée

Par la suite, nous avons mis en œuvre un réseau neuronal profond qui tire parti des images d'intensité et de phase (dans la même veine que Rivenson [74] et al.) ainsi qu'un z-stacking numérique. Ici un réseau de neurones bi-modal (exploitant conjointement image d'intensité et de phase) est entraîné pour renforcer le contraste et la netteté des compartiments sous-cellulaires ciblés.

À cette fin, le z-stacking numérique est effectué en calculant des images à différents plans de mise au point, qui sont minutieusement espacés. Cette opération est réalisée en ajustant le paramètre Δ_z correspondant à la position du plan imagé dans l'algorithme de reconstruction. Les images de la plus haute qualité sont ensuite isolées pour constituer une base de référence pour l'entraînement du DNN. Ici, le plan focal optimal est sélectionné en fonction de la modalité. Nous démontrons quantitativement l'impact du DNN du point de vue du diagnostic automatisé. Pour cela, la détection du *Plasmodium falciparum* dans les frottis sanguins est considéré comme exemple de mise en œuvre. Il est choisi, en notant, que le problème de la détection des parasites avec la sensibilité et la spécificité requises n'a jusqu'à présent pas été résolu de manière satisfaisante en utilisant uniquement des images de frottis sanguins. C'est aussi un problème typique où l'analyse d'échantillons sur une grande surface est nécessaire et donc où une correction locale est particulièrement appréciable pour une exploitation complète du FOV étendu de la FPM.

Puis, dans un second temps, j'exposerai quelques résultats préliminaires de cette même approche de refocalisation sur des leucocytes.

3.3 Apport de la refocalisation en parasitologie

3.3.1 Configuration du microscope

Comme mentionné, les échantillons considérés par la suite consistent en frottis sanguins parasités par *P. falciparum* Ils sont colorés avec la coloration May-Grünwald Giemsa. La résolution habituelle requise pour détecter des parasites d'environ $\sim 250nm$ [59, 21] peut être obtenue avec une ouverture numérique $NA=1$. En considérant les lentilles d'objectif de microscope disponibles sur le marché, un bon compromis entre la NA et le temps de mesure nous a conduit à utiliser une lentille d'objectif 10x avec une ouverture numérique de $NA=0.45$. La source lumineuse est constituée de 12 LEDs disposées en anneau [49, 63] et d'une LED centrale supplémentaire. Elle est réalisée avec un anneau NeoPixel RGB 5050 de rayon $r = 15mm$ avec 12 LEDs (de chez Adafruit) et une LED centrale NeoPixel

supplémentaire. Cette source lumineuse remplace l'éclairage Köhler traditionnel dans un microscope vertical Nikon. La distance des diodes à l'échantillon "d" est ajustée pour obtenir une ouverture numérique synthétique $NA_{synth} = 0.9$, avec $d = 30$ mm. Le chevauchement dans l'espace de Fourier est d'environ 67 % comme requis [26]. La caméra utilisée (modèle IDS UI-3200SE) intègre un capteur de grande surface de 4104×3006 pixels (avec un pas de pixel de $3.45 \mu m$). Cela correspond à un champ de vision (FOV) de $1.41 \times 1.03 \text{ mm}^2$. Le choix de cette configuration de microscope est principalement motivé par le fait qu'elle permet un bon compromis entre la résolution optique réalisable et la surface du FOV ainsi qu'un temps d'acquisition court de 0,5s par champ de vision [83].

3.3.2 Profondeur de champ en imagerie FPM

La profondeur de champ théorique d'un microscope conventionnel est dictée par

$$DOF = \frac{\lambda \cdot n}{NA^2} + \frac{n \cdot \Delta x}{M \cdot NA} \quad (3.1)$$

où n est l'indice de réfraction du milieu entre l'échantillon et l'objectif, Δx l'espacement entre chaque pixel du capteur de la caméra, M est le facteur de grandissement de l'objectif. Le DOF est évalué à $3.35 \mu m$ pour notre configuration de microscope et une longueur d'onde $\lambda = 525 \text{ nm}$.

Bien qu'à ma connaissance aucune formule de profondeur de champ n'ait été démontrée théoriquement en FPM, la formule 3.1 doit également fournir une bonne approximation de la profondeur de champ pour la raison suivante : considérons un échantillon placé à proximité de $z = 0$ le focus optimal, les images en diode centrale enregistrées pour n'importe quelle valeur de z entre $-DOF/2$ et $DOF/2$ ne peuvent montrer aucun changement notable. En conséquence, les images reconstruites ne peuvent pas présenter de changements significatifs non plus.

Pour illustrer ce point, nous étudions la DOF expérimentalement à l'aide d'une mire de résolution USAF 1951 pour deux configurations de microscope différentes. La première est liée à notre réglage FPM 10x et une ouverture numérique synthétique $NA_{sys} = 0.90$. La seconde consiste en un microscope conventionnel sous condition d'éclairage cohérent (dans la direction z) avec une lentille d'objectif 40x avec une résolution native correspondant à un $NA = 0.95$. Ces deux configurations sont utilisées pour comparer la DOF des microscopes conventionnels et FPM de résolution similaire.

La procédure utilisée est la suivante : Dans un premier temps, la mire de résolution est placée à $z=0$ et son détail le plus petit observable est identifié avec un éclairage de la diode centrale. Dans un second temps, on fait varier le focus

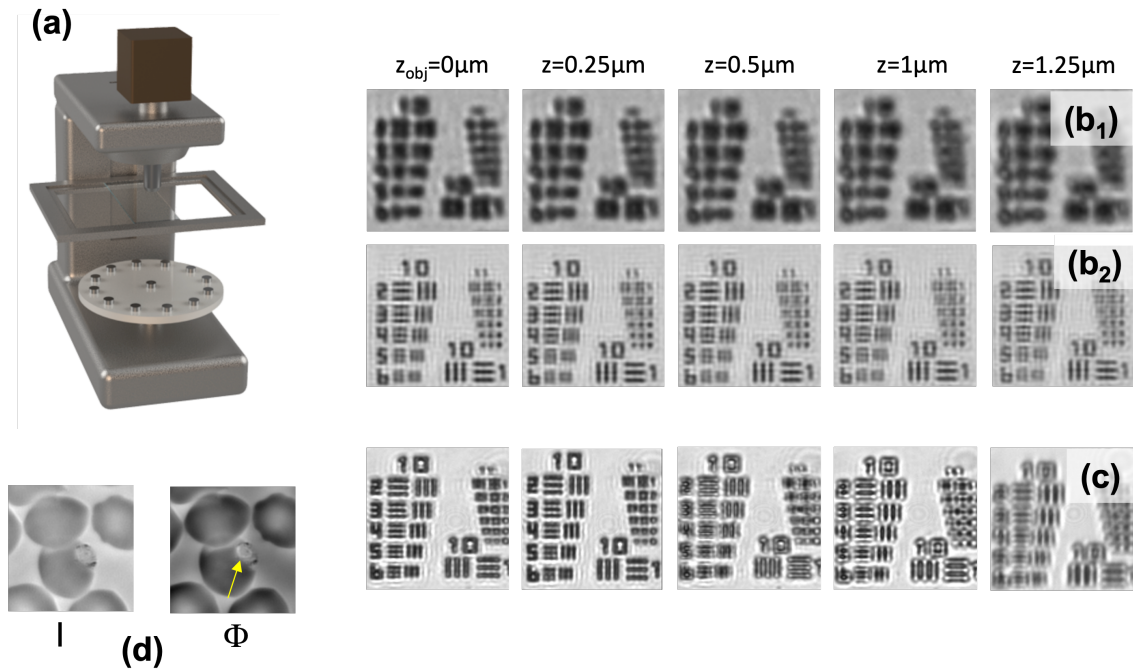


FIGURE 3.1 – a) Schéma d’un microscope équipé de sa matrice de LEDs. b) mire de résolution USAF1951 pour différentes positions de l’objectif 10xNA0.45, b₁) led centrale brute, b₂) images FPM d’intensité reconstruites. c) mire de résolution USAF1951 pour différentes positions de l’objectif 40xNA0.95 avec LED centrale. La comparaison de b) et c) montre une résolution comparable. d) Images d’intensité et de phase FPM obtenues sur un globule rouge parasité. La flèche pointe vers un compartiment qui n’est pas visible dans l’image d’intensité.

en déplaçant progressivement, la charte de résolution jusqu’à ce qu’elle devienne floue. La valeur z_{max} ainsi mesurée, fournit une valeur approximative de $DOF/2$.

Quelques résultats expérimentaux illustratifs sont indiqués dans la Figure 3.1 b) pour différentes valeurs de z . La Figure 3.1 b1) correspond à un éclairage LED central et la fig. 3.1b2) aux images reconstruites. Le détail le plus fin observable à $z = 0$ est lié à l’élément 5 du groupe 10 pour lequel la largeur de chaque barre est de 308 nm. À partir de l’observation des différentes images reconstruites, la valeur de $|z_{max}|$ est évaluée à environ $1.75 \mu m$ (l’image correspondante n’est pas présentée). La DOF est donc proche de $3.5 \mu m$, en accord avec l’équation 3.1. La même procédure est répétée pour la configuration du microscope conventionnel (sous éclairage cohérent). Des images expérimentales sont présentées à la Figure 3.1 c). Cette fois, la DOF est évaluée à environ $0.7 \mu m$ (également en accord avec l’équation 3.1).

Il est intéressant de noter que la DOF du FPM est essentiellement imposée par l'ouverture numérique native de l'objectif du microscope plutôt que par son ouverture numérique synthétique, cela suggère que la formule 3.1 est applicable pour évaluer le DOF en microscopie FPM. Ainsi, à résolution fixée, le FPM présente une DOF accrue par rapport à la microscopie conventionnelle de résolution équivalente.

Étant donné que l'épaisseur d'un frottis sanguin est d'environ 3-4 μm et que le DOF pour la configuration FPM choisie est de 3.35 μm , une correction de la mise au point de l'image est nécessaire pour obtenir des images nettes sur l'ensemble de son champ de vision. Des images brutes peuvent être utilisées pour calculer des images à un plan de mise au point Δz différent de celui utilisé lors de l'acquisition, à condition qu'un modèle de propagation électromagnétique d'une couche d'air d'épaisseur Δz_{num} est introduite dans le modèle direct de formation d'image de microscope utilisé pour la reconstruction. L'algorithme original e-PIE [56] qui a été adapté à la FPM inclut déjà cette possibilité en utilisant la correction de front d'onde numérique. Cette possibilité est utile pour compenser des conditions de mise au point approximatives. Elle permet également un ajustement fin des conditions de mise au point à posteriori lorsque nécessaire (par exemple pour l'exploration des images de phase) autour du plan focal optimal.

3.3.3 Caractéristiques recherchées dans les images refocalisées

Notre principale motivation est orientée vers l'amélioration des performances des classifieurs qui sont utilisés pour établir un diagnostic automatisé. En effet, comme cela peut être bénéfique, les modalités de phase et d'intensité sont complémentaires, car elles sont respectivement liées à la masse sèche de l'échantillon qui est mesuré et à son absorption. Un exemple illustratif est présenté à la Figure 3.1d) montrant une cellule de globule rouge parasitée. Comme on peut le constater, un compartiment de la cellule parasitaire presque invisible en image d'intensité est clairement observable en image de phase. Il est donc raisonnable d'imaginer que ces deux types d'images, si elles sont exploitées conjointement, devraient aider les performances des classifieurs, à condition que les compartiments d'intérêt du parasite soient bien visibles sur l'image d'intensité et de phase aussi.

La figure 3.2 présente des images FPM typiques obtenues sur un parasite pour différents réglages de mise au point, c'est-à-dire la position de l'objectif le long de l'axe vertical z (variant de $z = -3\mu m$ à $z = +3\mu m$, avec un pas incrémental de $1\mu m$). Ici, $z = 0$ est défini par la position optimale déduite de l'observation directe de l'échantillon à l'aide d'un éclairage à LED central. Il est à noter que ces images révèlent une variabilité de contraste frappante autour de la position z

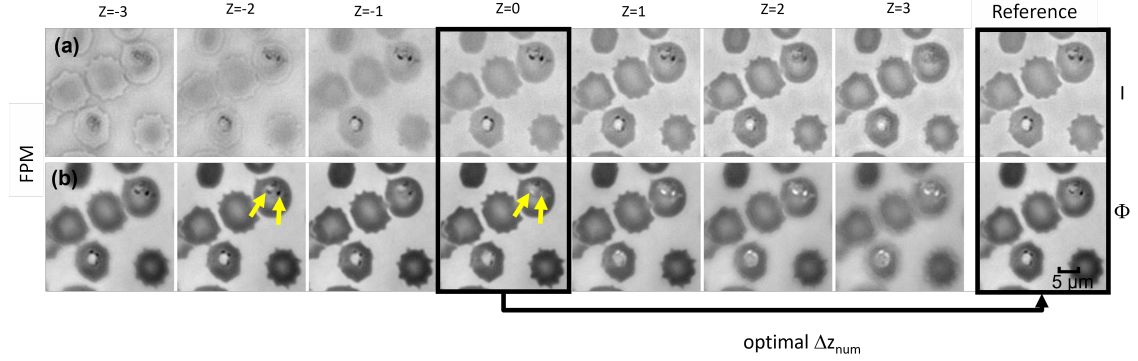


FIGURE 3.2 – Images d’intensité (I) et de phase (Φ) FPM de globules rouges colorés acquises pour différents réglages de mise au point (allant de $z = -3\mu m$ to $z = +3\mu m$) pour $\lambda = 525nm$, notez l’évolution du contraste relatif d’un composant de la cellule parasite (hémozoïne) indiqué par la flèche. Les images encadrées dans la boîte de droite correspondent aux images optimales obtenues avec une compensation de mise au point numérique e-PIE ($\Delta z_{num} = 0.7\mu m$ for I , $\Delta z_{num} = -1.7\mu m$ for Φ).

optimale de l’objectif du microscope. En particulier, pour les images de phase, un compartiment cellulaire du parasite (l’hémozoïne) passe du contraste noir au blanc avec z (voir flèche à la fig. 3.2b) sans modifications importantes de la netteté des globules rouges. Une telle variation de contraste est frappante, car une évolution faible de la fonction de phase est attendue pour de petites distances de défocalisation le long de l’axe z . Bien qu’une telle observation n’ait pas été rapportée à ce jour par d’autres auteurs (à notre connaissance), cet effet n’est pas si inhabituel. Nous l’avons également rencontré en utilisant d’autres objectifs de microscope (20xNA0.75 et 40xNA0.95). Nous pensons que cela est dû au fait que l’hypothèse d’échantillon mince au cœur du principe FPM n’est pas entièrement adaptée à la description des frottis sanguins. En conséquence, la variation de phase importante est très probablement à interpréter comme un artefact de reconstruction.

Pour définir davantage la meilleure image qui est exploitable par un parasitologue, les images obtenues pour $z = 0$ sont explorées en utilisant une compensation de mise au point numérique dans la reconstruction (empilement z virtuel) avec Δz_{num} allant de $-3\mu m$ à $+3\mu m$ avec un pas de $0.1\mu m$. Le choix des images calculées a été validé par un parasitologue expert pour la définition de la vérité terrain en tenant compte de la mise au point, du contraste et de la visibilité des différents compartiments cellulaires des parasites (cytoplasme, noyau, hémozoïne et vacuoles). Les images I et Φ choisies qui révèlent le mieux les parasites, sont indiquées à droite de la Figure 3.2 (référence). Elles ont été obtenues pour $\Delta z_{num} = 0.7\mu m$ et $\Delta z_{num} = -1.7\mu m$ pour les images d’intensité et de phase

respectivement. Il n'est pas surprenant que le réglage optimal de z pour l'éclairage à LED central et les images d'intensité reconstruites par FPM soient proches. La valeur du paramètre de reconstruction optimal Δz_{num} prise en compte pour l'image de phase FPM est cependant plus singulière et peut-être plus grande (ici $\Delta z_{num} = -1.7\mu m$). Nous attribuons probablement cela à une conséquence de la préparation de l'échantillon qui a été employée (fixation, colle, couvercle de verre de la lame) qui modifie les propriétés de phase des échantillons. De plus, cela peut être influencé par la position du compartiment cellulaire du parasite à l'intérieur du volume de la cellule de sang rouge.

En d'autres termes, et en ce qui concerne l'acquisition FPM, le positionnement en z de l'objectif du microscope peut être délicat pour plusieurs raisons :

1. La position optimale en z de l'objectif du microscope dépend de la modalité (intensité ou phase). Comme l'image de phase ne peut pas être observée directement, elle nécessite de nombreuses reconstructions d'images qui sont chronophages.
2. La position aléatoire des parasites à l'intérieur du volume des cellules ne permet pas nécessairement un ajustement systématique des réglages de mise au point.
3. La variabilité du contraste de phase, lorsqu'elle n'est pas adéquatement contrôlée, est susceptible de réduire les performances des classifieurs.
4. Lorsqu'on s'intéresse à produire une image RVB, l'objectif du microscope présente des aberrations chromatiques, on est alors obligé de compenser, pour chaque canal de couleur, l'intensité et la phase.

Dans notre perspective, l'utilisation de la compensation de mise au point numérique pour explorer des échantillons, bien qu'elle apporte des avantages, reste chronophage. Cette approche est donc peu adaptée pour des diagnostics automatisés. En effet, la reconstruction des images d'intensité et de phase implique des calculs avec une complexité C proportionnelle à N_f , le nombre de plans de mise au point à explorer avec $C = 4.n.log(n).N.K.N_f$ où n représente le nombre de pixels composant un petit patch d'image, N le nombre de LEDs utilisées et K le nombre de patches.

3.3.4 Apprentissage du U-Net

a) Modèle neuronal

J'ai introduit une stratégie basée sur un modèle de réseau neuronal profond de type U-Net combiné à des calculs de z -stacking numériquement pour la constitution de la base de données de référence.

Une fois appris, le modèle est chargé de corriger les images localement (c'est-à-dire à l'échelle des compartiments sous-cellulaires) plutôt que globalement, afin de produire des détails nets et contrastés des parasites.

L'aspect localisé de la correction de focus constitue une contribution importante de notre modèle. En effet, dans de nombreux cas, l'ajustement de la mise au point dépend spécifiquement de la position spatiale de la cellule au sein du champ de vision. La définition de ce qui est net et contrasté est ici apprise intrinsèquement par des exemples plutôt que défini par des propriétés formelles de l'image [36].

Ceci est particulièrement avantageux pour les images de phase. De plus, contrairement aux possibles corrections de focus par le calcul de la reconstruction à différentes mises au point, l'approche du réseau neuronal procède différemment. Elle ne cherche pas à identifier le plan focal optimal par modalité, mais produit plutôt une image nette pour chacune des deux modalités (intensité et phase) d'image obtenue après une unique reconstruction.

Le modèle de réseau neuronal apprend statistiquement la transformation qui doit être appliquée sur les images reconstruites grâce à une base de données d'images de référence. Cette base de données d'apprentissage est composée de paires d'images reconstruites associées à différentes mises au point et leurs références refocalisées correspondantes.

Une fois les paramètres du réseau, à savoir ses poids, appris, le réseau peut être utilisé sur des images FPM de n'importe quelle taille.

Parmi les différentes architectures de réseaux profonds de l'état de l'art, nous avons choisi une architecture U-Net [77]. Un avantage de ce modèle est qu'il est possible de considérer et de traiter conjointement les 2 images d'intensité-phase, qui pourront donc interagir pendant le processus de transformation. Le U-Net est un modèle encodeur-décodeur présenté à la Figure 3.3, qui prend en entrée des images non focalisées (I , Φ) et produit en sortie les images correspondantes, contrastées et mises au point sur les compartiments du parasite recherchés. Rappelons que cette correction est locale. Elle permet donc de traiter la netteté de chaque globule rouge indépendamment de celle de son voisin.

b) Schéma d'apprentissage

Toutes les étapes impliquées dans la construction de l'ensemble de données d'entraînement pour l'apprentissage des poids du modèle U-Net sont indiquées à la figure 3.4.

Dans un premier temps, le plan de focus optimal de l'objectif du microscope est trouvé avec la LED centrale du microscope FPM (étape 1). Cela définit la position de référence $z = 0$. Le frottis sanguin est ensuite acquis pour diverses conditions de mise au point (z allant de $z = -8\mu m$ à $8\mu m$ avec un pas de $2\mu m$) et en utilisant chaque fois 13 angles d'illumination successifs (étape 2°). Au cours de

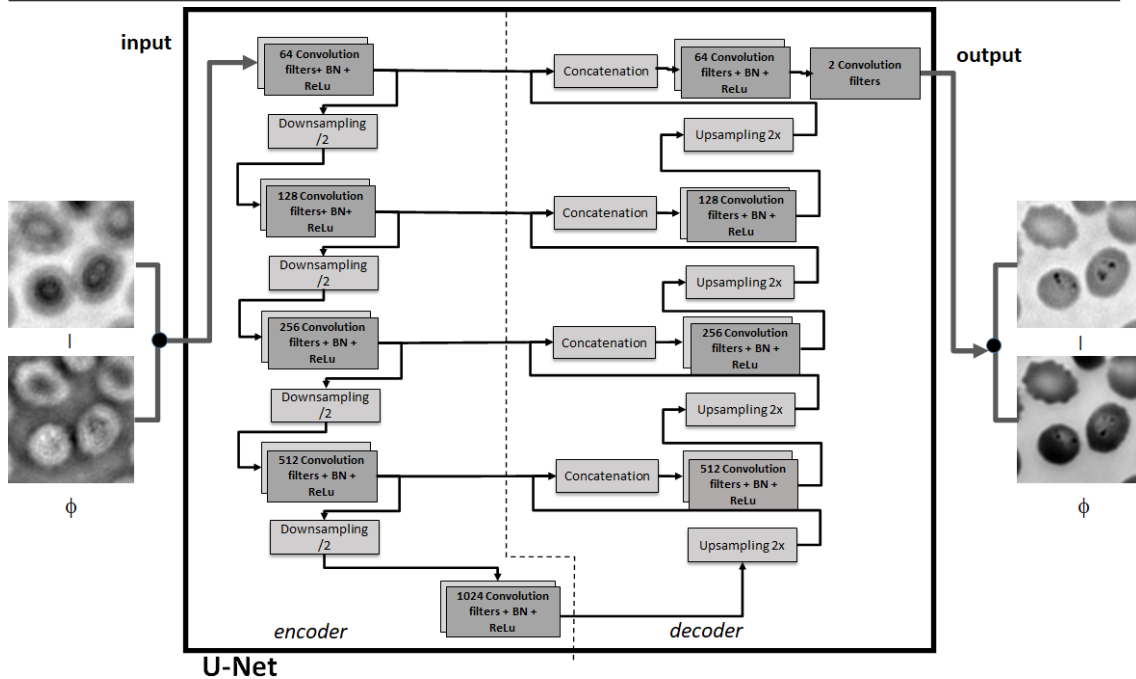


FIGURE 3.3 – Architecture du U-Net utilisée pour la tâche de refocalisation

la 3^{ème} étape, le champ de vision (FOV) restreint à ses 256x256 pixels centraux (après découpage des images brutes) est exploité. Ces dimensions ont été choisies pour assurer une large zone où les variations de netteté sont imperceptibles, garantissant ainsi une uniformité. Cela permet d’obtenir des images nettes des objets d’intérêt dans cette zone restreinte. Comme la loi d’aberration de défocalisation est spatialement invariante, l’entraînement réalisé à partir de cette région particulière est également applicable sur l’ensemble du FOV. C’est également la région où les autres aberrations sont généralement les plus petites, ce qui est bénéfique pour la construction d’images de référence de la plus haute qualité.

Chaque image recadrée contient un nombre suffisamment grand de globules rouges (typiquement ~ 100) pour apprendre les propriétés optimales des images de globules rouges (avec ou sans parasites). Après l’étape 3, le schéma des données est divisé en deux branches : la gauche est dédiée à la reconstruction des images à différentes valeurs de z (étape 4a-5a) avec $\Delta z_{num} = 0$; la droite est dédiée à la production des images de référence (étape 4b-5b). Les images reconstruites de la branche de référence sont ici explorées en variant finement le paramètre de reconstruction Δz_{num} . La sélection des images de référence pour les modalités d’intensité et de phase est finalement déterminée par un parasitologiste (étape 5b).

En phase de test, le réseau est appliqué sur des images complètes du champ de

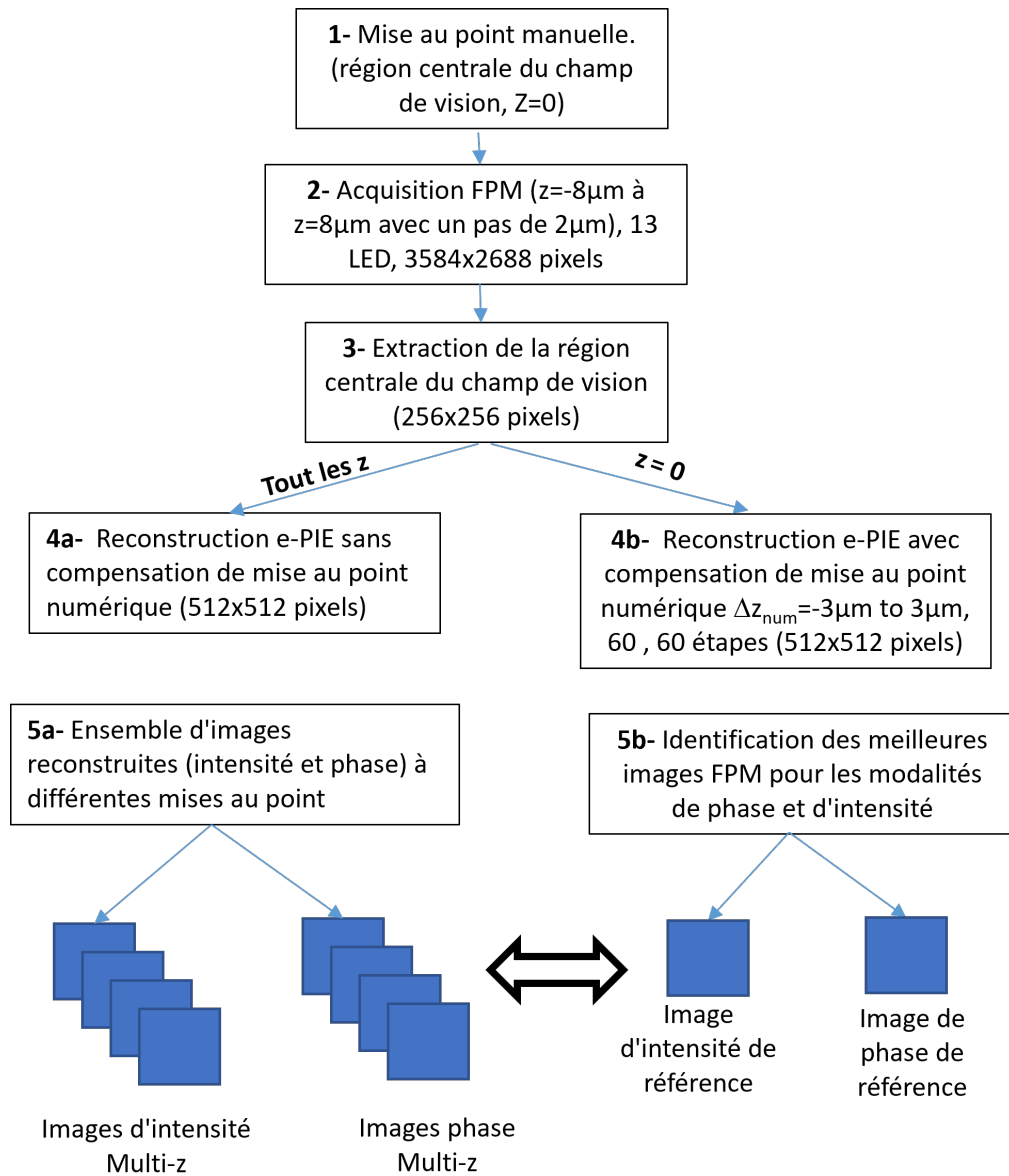


FIGURE 3.4 – Construction de la base de données d’entraînement U-Net. Les échantillons sont numérisés (1-3) avec FPM sous différentes conditions de mise au point (de $z = -8\mu m$ à $z = 8\mu m$). Pour toutes les valeurs de z , les images FPM sont reconstruites sans compensation de mise au point numérique (étapes 4a et 5a) pour produire les exemples d’entrée U-Net. Pour chacune de ces entrées, les images de phase de référence I et Φ sont obtenues après les étapes 4b et 5b (l’image de sortie souhaitée) et sont utilisées pour apprendre la correspondance et la transformation entre les entrées et les sorties.

vision (I, Φ) pour produire des images corrigées. Comme il sera montré dans la section suivante, le réseau neuronal est capable de corriger des images sur l'ensemble du champ de vision du microscope et pas seulement sur sa zone centrale.

3.3.5 Base de données de refocalisation

Dans cette étude, des lames de frottis sanguins anonymisés provenant de deux patients infectés ont été utilisées pour l'entraînement du U-Net. Bien que limitée en termes de diversité de patients, la base de données qui a été réalisée est représentative des parasites rencontrés dans des situations réelles en termes de taille et d'âge.

Pour chaque patient, 14 régions centrales du champ de vision (de 512×512 pixels après reconstruction) ont été acquises, soit 28 régions de zone centrale. 20 régions ont été utilisées pour l'entraînement, 4 pour la validation et 4 pour le test. Chaque région est acquise avec 9 plans de mise au point, correspondant à un focus z allant de $z = -8\mu m$ à $z = 8\mu m$, avec un intervalle de $2\mu m$.

Chaque région reconstruite a ensuite été divisée en 25 petits patches de 128×128 pixels. Ce choix a été fait afin de contraindre le temps de traitement et d'assurer une quantité suffisante de données d'entraînement. Les données ont été davantage augmentées par des rotations et des rotations aléatoires pour augmenter la taille de l'ensemble de données d'entraînement.

La taille finale de la base de données d'entraînement (comprenant les transformations géométriques et l'augmentation des données) est de 18000 patches correspondant à 4 rotations \times 25 patches \times 9 plans de focus \times 20 régions. D'autres améliorations auraient pu être envisagées, notamment l'ajustement de la luminosité, qui se révèle être une méthode efficace pour modifier les teintes et les couleurs des frottis numérisés afin de simuler une variabilité de coloration MGG. Les ensembles de données de validation et de test sont chacun composés de 900 patches.

3.3.6 Résultats et évaluation de l'apport de la refocalisation

a) Évaluation visuelle de la refocalisation intensité phase

Quelques résultats illustratifs obtenus à la sortie du U-Net entraîné sont indiqués à la Figure 3.5. La rangée supérieure de la Figure affiche des images reconstruites pour un décalage de mise au point numérique $\Delta z_{\text{num}} = 0$ c), avec une échelle en z allant de $-8\mu m$ à $8\mu m$, par incréments de $2\mu m$.

Les images de référence, situées en haut à droite de la figure, sont extraites des reconstructions de z-stacking numérique à $z = 0\mu m$. Ces images de référence ont été obtenues pour $\Delta z_{\text{num}} = -0.6\mu m$ et $\Delta z_{\text{num}} = 0.8\mu m$ pour les images d'intensité et de phase, respectivement. Une approche de z-stacking numérique similaire est

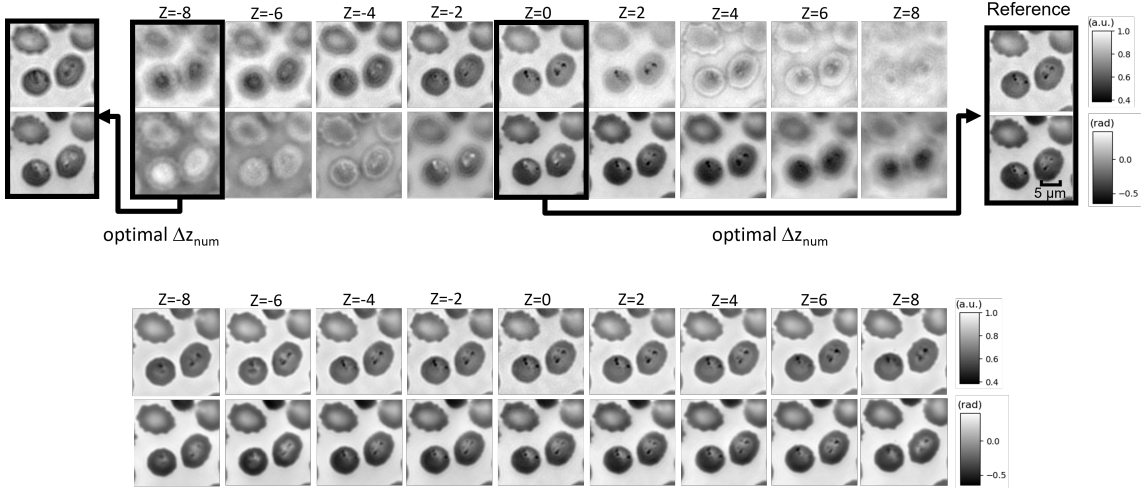


FIGURE 3.5 – (En haut au centre) Images FPM de frottis sanguin reconstruites pour différentes distances entre l'échantillon et l'objectif du microscope (de $z = -8\mu\text{m}$ à $z = 8\mu\text{m}$). Les images de référence (boîte en haut à droite) sont obtenues en suivant la branche de droite de la figure 3.4 en utilisant $\Delta z_{\text{num}} = -0.6\mu\text{m}$ $\Delta z_{\text{num}} = 0.8\mu\text{m}$ pour l'intensité et la phase respectivement ; (en bas) les images correspondantes corrigées par U-Net.

utilisée pour isoler l'image de référence qui serait obtenue à partir des images capturées à $z = -8\mu\text{m}$. Les images résultantes, affichées dans le coin supérieur gauche de la figure, ont été produites avec $\Delta z_{\text{num}} = 4.8\mu\text{m}$ et $\Delta z_{\text{num}} = 6.5\mu\text{m}$. À ces valeurs de Δz_{num} la correction de mise au point commence à être imparfaite.

Nous avons donc expérimentalement délimité les capacités de la mise au point par l'algorithme de reconstruction entre $z = -6\mu\text{m}$ et $z = 6\mu\text{m}$. À des fins comparatives, les sorties du U-Net appliquées aux images reconstruites (avec $\Delta z_{\text{num}} = 0\mu\text{m}$) pour diverses valeurs de z sont illustrées dans la partie inférieure de la Figure 3.5.

Comme on peut l'observer, les images produites révèlent une amélioration de la qualité en termes de mise au point et de contraste. Elles sont correctement et automatiquement corrigées avec le U-Net, indépendamment de z , un paramètre qui est inconnu pour le U-Net.

Il est intéressant de noter que toutes ces images sont très proches de la référence, même pour $z = \pm 6\mu\text{m}$. Au-delà de $z = 6\mu\text{m}$ ou en dessous de $z = -6\mu\text{m}$, certains détails sur les parasites commencent à disparaître (non présentés).

Ces premiers résultats permettent d'estimer la plage sur laquelle la compensation de mise au point du U-Net est efficace. En effet, cette plage est importante ($\sim 12\mu\text{m}$) et comparable à celle accessible à la reconstruction e-PIE avec compensation de la mise au point.

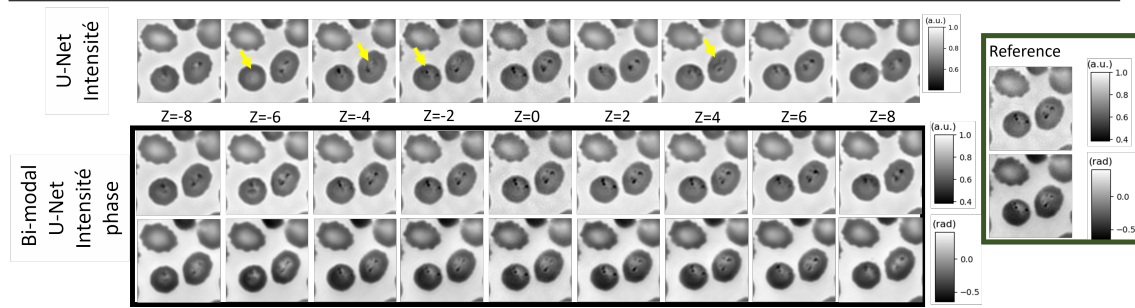


FIGURE 3.6 – Des exemples illustratifs d’images FPM de frottis sanguin enregistrées pour différentes distances entre l’échantillon et l’objectif du microscope (de $z = -8\mu m$ à $z = 8\mu m$). Les images de référence (boîte à droite) sont obtenues en suivant la branche de droite de la figure 3.4 en utilisant $\Delta z_{num} = -0.6\mu m$ $\Delta z_{num} = 0.8\mu m$ pour l’intensité et la phase respectivement ; (en haut) les images correspondantes corrigées par U-Net bi-modal utilisant l’intensité et la phase ; (en bas) les images d’intensité corrigées par U-Net qui a appris sur l’intensité seule.

b) Comparaison visuelle de la refocalisation intensité phase vs intensité seule

Pour évaluer l’avantage de l’utilisation d’un U-Net de refocalisation en intensité et phase par rapport à l’utilisation de l’intensité seule, j’ai adapté le modèle U-Net présenté dans la section 3.3.4 pour une entrée et une sortie uniques, représentant les images d’intensité. Ce modèle cherche à améliorer la netteté et contraste des images d’intensité en utilisant une information incomplète sur l’objet mesuré, sa phase. En théorie, l’absence de la phase ne permet pas de refocaliser les images.

En Figure 3.6, nous comparons des images d’intensité pour divers niveaux de mise au point, allant de $z = -8\mu m$ à $z = 8\mu m$. Deux modèles ont été considérés dans cette comparaison. Le premier modèle, que nous appelons U-Net intensité, prend en entrée des images d’intensité défocalisées (comme illustré dans la partie supérieure de la figure 3.5). Le second modèle, est U-Net bimodal, prenant en entrée à la fois des images d’intensité et de phase défocalisées.

En comparant les résultats obtenus par ces deux modèles, nous cherchons à comprendre l’impact de l’ajout des informations de phase dans le processus de refocalisation des images par réseaux de neurones.

Le modèle U-Net, qui a été appris uniquement sur des images d’intensité, réussit à produire des images avec une bonne mise au point. Cependant, des erreurs sont présentes et des informations manquent sur certains éléments, comme le montrent les flèches jaunes dans la figure 3.6.

D’autre part, nous observons que le modèle bimodal offre des images de meilleure qualité et plus fidèles à la référence par rapport au modèle formé uniquement sur

l'intensité. Ce résultat met en évidence l'importance de la prise en compte conjointe de l'intensité et de la phase pour obtenir des images plus précises et détaillées.

c) Refocalisation intensité phase sur un grand champ de vue

Nous considérons maintenant l'ensemble du champ de vision (FOV) du microscope. Un résultat représentatif obtenu avec et sans compensation du U-Net est visible sur la Figure 3.7. L'image reconstruite présente certaines variations de mise au point à travers le FOV particulièrement large (en raison de la taille du capteur).

Cela provient des imperfections horizontales du porte-lames du microscope qui varient de $\Delta z \sim -3\mu m$ (en bas à droite) à $\sim 3\mu m$ (en haut à gauche). La région encadrée **A** correspond à une région correctement mise au point, la région encadrée **B** à une région hors mise au point ($\Delta z = 2.5\mu m$). L'application du U-Net entraîné sur ces deux régions (voir les 2 encarts de la figure 3.7) confirme clairement le comportement attendu du réseau neuronal : sa capacité à corriger localement les images reconstruites. Plus précisément, la visibilité des parasites est clairement améliorée sur l'ensemble du FOV. Cela est particulièrement appréciable dans la région **B** ($\Delta z \sim 2.5\mu m$). De plus, dans la région **A** ($\Delta z \sim 0\mu m$), la netteté initiale des globules rouges, le contraste des compartiments cellulaires des parasites ciblés est amélioré dans les images d'intensité et de phase. Un cristal d'hémozoiné qui était à peine visible en image de phase devient maintenant clairement observable (pointé avec une flèche rouge).

Nous notons que le temps de traitement supplémentaire de U-Net est d'environ ~ 5 secondes en utilisant un GPU Nvidia RTX 3080. Cette valeur est à comparer au temps de reconstruction de ~ 7 secondes, également mis en œuvre sur le même GPU via le framework TensorFlow [43]; la reconstruction ePIE par CPU étant assez lente.

3.3.7 Impact quantitatif sur la détection des parasites

Après avoir évalué l'amélioration de la qualité des images obtenues grâce à l'application du réseau U-Net, notre objectif est d'examiner l'impact de cette amélioration sur l'efficacité de la détection automatique des globules rouges infectés par des parasites.

L'objectif principal de notre étude est de comparer les performances du modèle YOLO sur différentes modalités d'entrée, en particulier en évaluant l'impact du module U-Net sur notre système automatisé de détection des parasites. Afin de quantifier cet impact, nous avons mis en œuvre une série de comparaisons de performances. Nous avons choisi d'utiliser le modèle YOLO (You Only Look Once) pour cette tâche, comme décrit dans [70]. Ce modèle a été décrit en détail dans le Chapitre 2. C'est un modèle de détection d'objets en temps réel qui vise à prédire

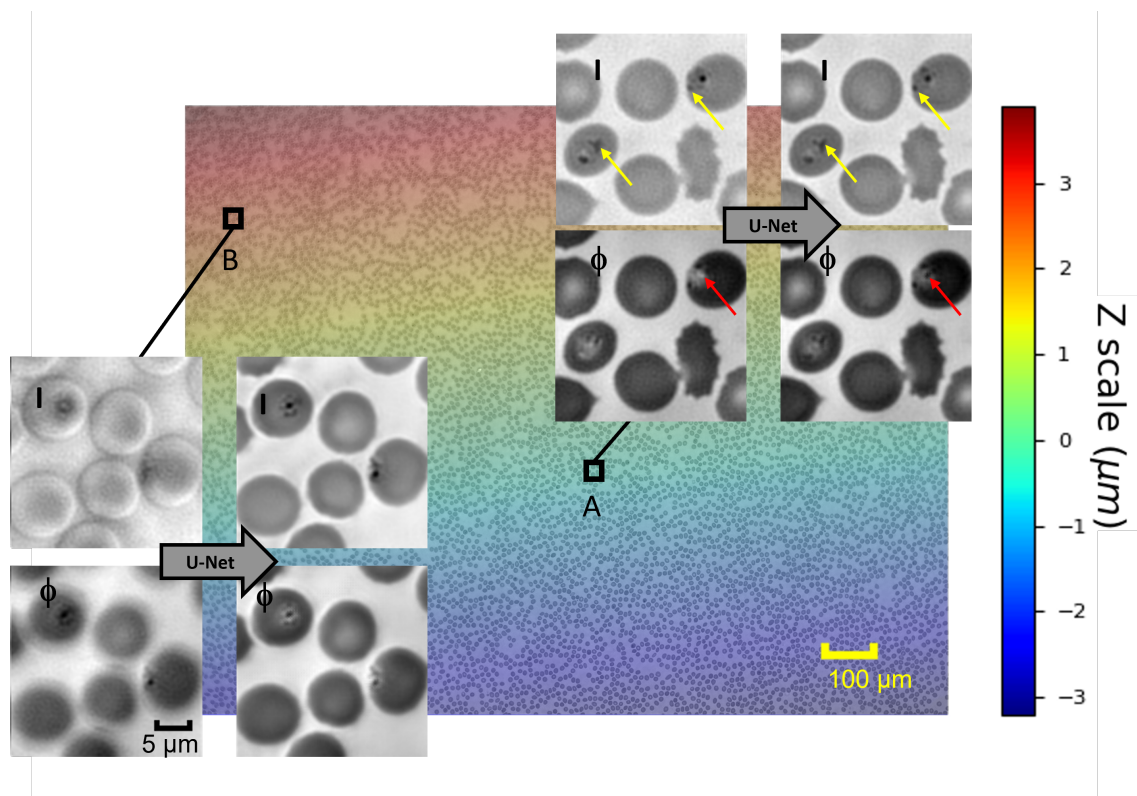


FIGURE 3.7 – Champ de vision complet d’un échantillon de frottis sanguin acquis en FPM. La superposition de la carte de couleurs indique la variation de la distance de séparation entre l’échantillon et l’objectif du microscope, causée par une inclinaison du porte-lames. Les deux encarts relatifs aux zones A et B présentent des images FPM avant et après correction par U-Net. La région **A** est naturellement bien mise au point ($z \sim 0$). La région **B** est initialement hors de mise au point ($z \sim 2.5\mu m$). Les flèches indiquent des compartiments sous-cellulaires de parasites dont la visibilité est renforcée grâce à U-Net.

simultanément la position, la probabilité de présence d'un objet d'intérêt, et la classe de cet objet dans une image donnée. Plus précisément, nous avons évalué la précision, la sensibilité et la spécificité du modèle avec des images originales et des images prétraitées avec le module U-Net.

Dans notre contexte spécifique, les objets d'intérêt sont les globules rouges, qui peuvent être parasités ou non. Ainsi, notre modèle YOLO est configuré pour détecter les globules rouges dans les images et classer chaque détection comme parasitée ou non parasitée.

Pour optimiser les hyperparamètres du modèle YOLO, nous avons expérimenté différentes stratégies d'apprentissage, telles que la variation de l'optimiseur, l'ajustement du taux d'apprentissage, le seuil IOU, et l'entraînement du modèle avec différentes stratégies. Pour la détection et la classification YOLO, les meilleurs résultats ont été obtenus pour une taille de lot de 48 images et un seuil de 0,5 (comme suggéré dans [70]), avec ou sans amélioration U-Net. Nous avons utilisé l'optimiseur Adam avec $\beta_1 = 0.9$ et $\beta_2 = 0.999$ qui représente le taux de décroissance du moment d'ordre 1 et d'ordre 2 de l'optimiseur Adam. Le processus d'entraînement a été effectué en deux étapes, de manière similaire à [71]. Notre processus d'entraînement a impliqué l'entraînement du modèle pendant 50 époques avec un taux d'apprentissage de 10^{-3} lors de la première étape, car les poids étaient initialisés au hasard. Nous avons ensuite poursuivi le processus d'apprentissage en utilisant un taux d'apprentissage plus faible de 10^{-5} , lors de la deuxième étape. Nous avons observé que la phase initiale aidait à obtenir une détection précise des globules rouges rapidement, tandis que la seconde phase améliorait les performances de classification du modèle.

a) Base de donnée de détection

Pour effectuer cette comparaison, nous avons constitué trois datasets, DS_1 , DS_2 et DS_3 ; 28 champs de vision ont été capturés sur des frottis sanguins appartenant aux mêmes 2 patients que ceux déjà utilisés dans la section 3.3.5 (14 FOV par patient). Le positionnement en altitude z de l'objectif du microscope est régulé par un système d'autofocus. Il est à noter que la région d'intérêt pour ce mécanisme d'autofocus est strictement limitée à la zone centrale du champ de vision total du dispositif microscopique. Après reconstruction, chaque FOV contient des images de 7168×5376 pixels. Chacune de ces grandes images a ensuite été divisée en 48 patches de taille 896×896 pixels. Ainsi, 1344 patches sont obtenus à partir de tous les champs de vision capturés.

Le jeu de données DS_1 est composé de tous ces patches. Le jeu de données DS_2 contient les mêmes images dont la qualité a été encore améliorée après leur traitement avec le U-Net bi-modal introduit dans la section 3.3.4; DS_3 contient les mêmes images améliorées après leur traitement avec le U-Net intensité.

	Apprentissage	Validation	Test
GR sains	212558	44333	41919
GR parasites	9597	2023	2048

TABLE 3.1 – Le nombre de globules rouges sains et infectés dans les ensembles d’apprentissage, de validation et de test.

L’augmentation des données a consisté en 3 rotations ($\pi/2$, π , $3\pi/2$) et a conduit à un total de 5376 patches pour DS_1 et 5376 patches pour DS_2 et le même nombre pour DS_3 . Tous les patches ont été étiquetés par un expert. L’étiquette consiste en la délimitation des boîtes englobant chaque globule rouge. Une classe indiquant la présence ou non d’un parasite est également définie pour chaque boîte. Pour chaque jeu de données, 3840 patches ont été utilisés pour l’entraînement, 768 pour le test et 768 pour la validation. La distribution des globules rouges sains et infectés contenus dans chaque division est détaillée dans le tableau 3.1.

Dans un premier temps, trois conditions expérimentales sont introduites pour explorer les limites de l’efficacité du module U-Net :

Dans la première, la position de l’objectif du microscope sur l’axe z est optimisée pour la région centrale du champ de vision (FOV). En raison d’inexactitudes dans l’horizontalité du porte-lames du microscope, z varie entre $-3\mu m$ et $3\mu m$ sur le champ de vision ($z = 0\mu m$ au centre).

Dans la seconde condition, l’objectif est déplacé de $2\mu m$ le long de l’axe z . Dans ce cas, z varie entre $-1\mu m$ et $5\mu m$ sur le champ de vision.

Dans la troisième condition expérimentale, nous avons constitué une base de données qui intègre les deux premières conditions, c’est-à-dire des images où la mise au point est réglée à $z = 0\mu m$ et $z = 2\mu m$. À cette base, nous avons ajouté une plus grande variabilité en incluant des images prises avec des réglages de mise au point supplémentaires : $z = -2\mu m$, $z = 4\mu m$ et $z = -4\mu m$. Ces nouvelles conditions permettent d’étendre la gamme des distances de mise au point et d’introduire des variations supplémentaires dans nos données.

3.3.8 Résultats de détection

Nous avons effectué l’apprentissage de trois modèles YOLO qu’on a comparé, en utilisant les datasets DS_1 (sans U-Net), DS_2 (U-Net bi-modal) et DS_3 (U-Net intensité) avec la première condition expérimentale $z = 0\mu m$.

Les résultats obtenus en utilisant le module U-Net sont comparés à ceux obtenus sans son utilisation à travers différents critères. De manière plus précise, la sensibilité (ou le taux de vrais positifs VPR) et la spécificité (ou le taux de vrais négatifs VNR) sont examinées.

TABLE 3.2 – Les performances de détection de *Plasmodium falciparum* sur l'ensemble de test, obtenues sur des images brutes reconstruites par FPM avec et sans le module de compensation U-Net bi-modal et U-Net intensité (pour $z = 0\mu m$ et $z = 2\mu m$).

YOLO	Sans U-Net		Avec U-Net bi-modal		Avec U-Net intensité	
	$z = 0\mu m$	$z = 2\mu m$	$z = 0\mu m$	$z = 2\mu m$	$z = 0\mu m$	$z = 2\mu m$
Sensibilité	97.998	93.311	98.34	96.826	94.484	91.895
Spécificité	99.288	98.324	99.424	99.329	98.734	97.033
AUC	0.987	0.977	0.992	0.987	0.976	0.965
Accuracy	98.885	98.091	99.373	99.212	98.5536	97.653
Précision	97.998	93.311	89.237	87.55	94.482	91.895
F1-score	89.101	81.964	90.018	88.96	85.714	78.433

Par définition, $VPR = \frac{VP}{VP+FN}$ et $VNR = \frac{VN}{VN+FP}$ où VP représente le Vrai Positif (parasites correctement détectés), FN Faux Négatif (parasites non détectés), FP Faux Positif (globule rouge sain détecté comme parasite), VN Vrai Négatif (globule rouge sain correctement détecté).

Le VPR et le VNR sont deux indicateurs de performance clés dans les tests de diagnostic. Le premier qualifie la capacité du système à déclencher de vraies alertes tandis que le second qualifie la capacité du système à ne pas déclencher de fausses alertes. Idéalement, le système a une sensibilité de 100 % et une spécificité de 100 %. De plus, nous introduisons 4 autres métriques afin de produire des comparaisons plus complètes : l'AUC (aire sous la courbe ROC : Courbe de Caractéristique de Fonctionnement du récepteur "Receiver Operating Characteristic"), l'accuracy globale (pourcentage de bonne classification globale (globules rouges infectés et sains)), la précision (performance du modèle à prédire la classe infectée) et le score F1 (moyenne harmonique de la précision et de la sensibilité du modèle) avec $F1\text{-score} = \frac{VP}{VP + \frac{1}{2}(FP+FN)}$

La sensibilité évaluée avec et sans U-Net pour $z = 0\mu m$ et $z = 2\mu m$ est tracée en figure 3.8 a) en fonction d'un seuil de confiance. Les courbes pleines ont été obtenues avec l'application du U-Net, les courbes en pointillés sans refocalisation. Comme on pouvait s'y attendre, les meilleures courbes de sensibilité sont obtenues lorsque le module de compensation U-Net est inclus dans le workflow ; les résultats obtenus avec $z = 2\mu m$ sont assez proches de ceux obtenus avec $z = 0\mu m$.

Comparativement, lorsque aucune correction d'image U-Net n'est utilisée, la sensibilité diminue même pour $z = 0\mu m$. Elle est très dégradée pour $z = 2\mu m$. La courbe ROC complémentaire (caractéristiques de fonctionnement du récepteur) qui représente le compromis entre sensibilité et spécificité en fonction du point de fonctionnement est tracée en figure 3.8 b). Les points de fonctionnement optimaux

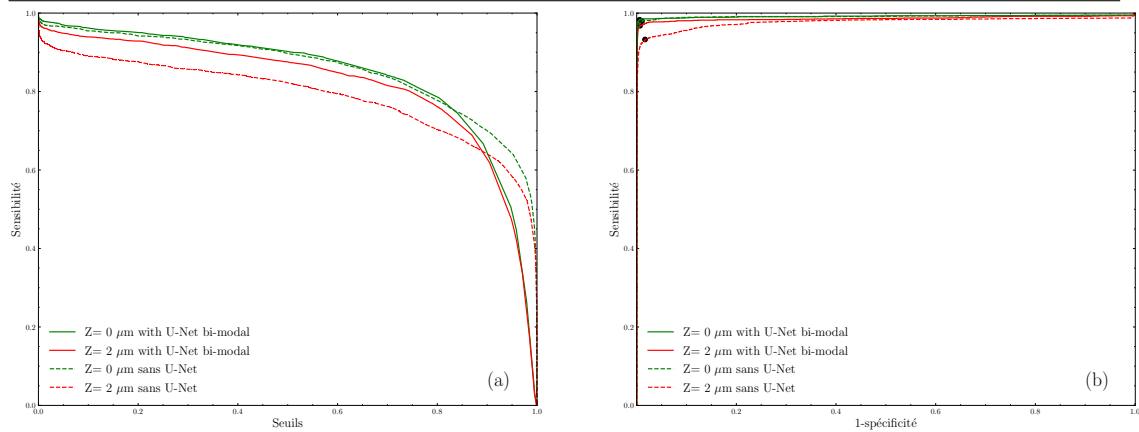


FIGURE 3.8 – Évolution de la sensibilité de détection de *Plasmodium falciparum* en fonction du paramètre de seuil sur la base de données de test. Les courbes en pointillés correspondent aux résultats obtenus sur des images brutes reconstruites par FPM sans le module de compensation U-Net (pour $z = 0\mu m$ et $z = 2\mu m$). Les courbes pleines correspondent aux résultats après la compensation des images par U-Net (bi-modal). b) Courbe ROC révélant les meilleurs points d'exploitation (cercles) définis par le meilleur compromis entre la sensibilité et la spécificité

y sont indiqués par des cercles.

Lors de l'analyse de ces résultats, nous devons tenir compte du fait que les 2 classes (sains, infectés) sont largement déséquilibrées (20 fois plus de globules rouges sains que de globules rouges infectés ; voir tableau 3.2). Les métriques qui se concentrent sur la performance de reconnaissance de la vraie classe (infectée), à savoir la sensibilité, la précision et le score F1, montrent une diminution entre $z = 0\mu m$ et $z = 2\mu m$, montrant que les parasites peuvent être difficiles à détecter en présence de défocalisation. Cela est particulièrement significatif sur la précision et le score F1.

Nos observations révèlent que le modèle YOLO, lorsqu'il est appliqué sur des images produites par le réseau U-Net bimodal, surpasse significativement le YOLO traitant des images issues de l'U-Net à intensité unique. Cette conclusion concorde avec les résultats visuels présentés dans la section 3.6, où nous avons remarqué que, à certains niveaux de mise au point, les informations concernant les parasites disparaissent totalement dans l'image produite par le U-Net exploitant l'intensité seule, soulignant ainsi l'importance cruciale de la bimodalité et de la phase dans le processus de refocalisation.

L'emploi d'un réseau de refocalisation U-Net bimodal permet d'augmenter significativement les performances du modèle YOLO à une distance focale de $z = 2\mu m$. Grâce à cette optimisation, les résultats à cette distance focale s'ap-

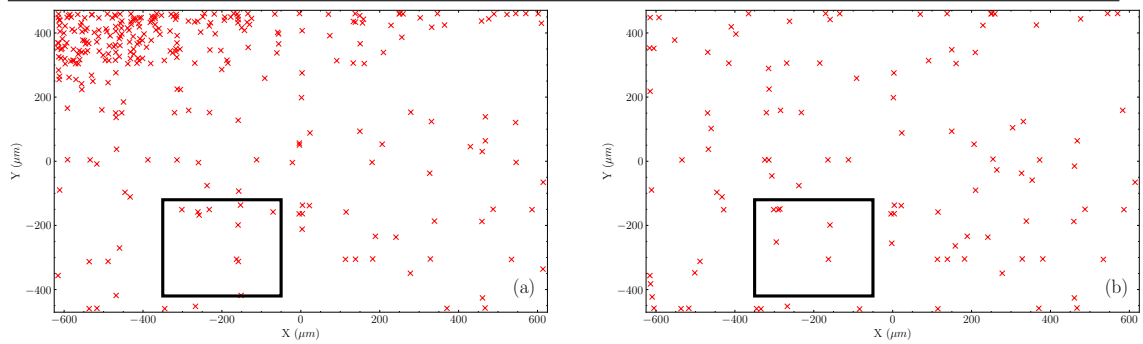


FIGURE 3.9 – Position des erreurs de détection pour $z = 2\mu m$ évaluée par l’architecture Yolo. a) Sans correction d’image U-Net (à droite). b) Avec correction U-Net bi-modal. La région délimitée par le cadre correspond à une région qui est correctement mise au point de manière native (pour $z = 2\mu m$).

prochent considérablement de ceux obtenus à $z = 0\mu m$. Plus spécifiquement, la compensation U-Net bi-modal conduit à une amélioration de la sensibilité sans dégradation de la spécificité. Considérant les métriques qui prennent en compte les deux classes sans les distinguer, comme la précision et l’AUC, elles ne révèlent pas de différences significatives entre $z = 0\mu m$ et $z = 2\mu m$, que ce soit refocalisé ou non. Compte tenu du grand nombre de cellules saines qui influent sur ces métriques, nous pouvons en déduire que les globules rouges sains peuvent être détectés même en présence d’une certaine défocalisation, contrairement à ce qui se passe avec les parasites.

Ces résultats démontrent l’impact positif du U-Net bi-modal. En conséquence, ils attestent que les tolérances concernant la précision du positionnement de z de l’objectif du microscope peuvent être considérablement relaxées. Rappelons que $z = 2\mu m$ au centre du FOV et le fait que le porte-lame est légèrement incliné (voir la carte de couleurs de la Figure 3.7) la tolérance est grossièrement évaluée à $\sim \pm 5\mu m$.

Pour analyser plus en détail ces résultats, les erreurs de détection des globules rouges sont indiquées en figure 3.9. Le jeu de données de test complet lié à $z = 2\mu m$ est ici considéré. Les croix identifient les positions des globules rouges non détectés dans les FOVs. Lorsqu’aucune correction U-Net n’est impliquée. (fig. 3.9 a), la plupart des erreurs se situent près du coin supérieur gauche du FOV, région où z est plus important. La situation change radicalement lorsque les images sont corrigées avec le U-Net (Figure 3.9 b). Ici, les erreurs sont très rares et presque uniformément réparties dans l’espace. De plus, la boîte noire en Figure 3.9 indique la région où la mise au point est nativement correcte. Dans cette région et sans compensation U-Net, la densité des erreurs est faible comme on pouvait s’y attendre. Elle diminue légèrement après compensation U-Net. Cela suggère que la qualité des images est

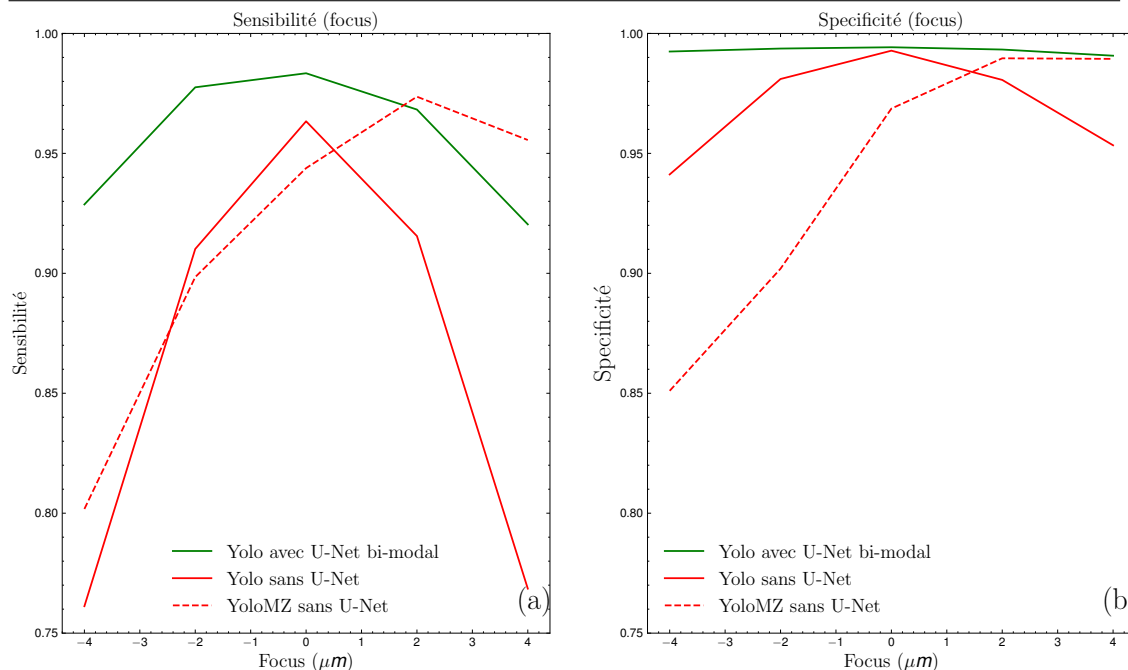


FIGURE 3.10 – Évaluation de la sensibilité et de la spécificité du modèle Yolo dans la plage de mise au point allant de $z = -4\mu m$ à $z = 4\mu m$, avec un intervalle de $2\mu m$. Nous avons comparé le modèle Yolo avec correction U-Net bimodal, le modèle YOLO sans correction, et le modèle YOLOMZ sans correction.

en effet améliorée même si la mise au point était déjà correcte.

Bien sûr, le réseau de neurones pourrait bénéficier d'optimisations supplémentaires (avec un ensemble de données d'entraînement plus large et une optimisation de certains de ses paramètres internes).

Nous avons utilisé un jeu de données d'entraînement comportant des conditions expérimentales plus variées afin d'analyser la robustesse du modèle de détection sur une plage de mise au point z plus étendue. Dans cette configuration, nous avons intégré un autre cas d'apprentissage YOLO. Contrairement à l'expérience précédente où nous avons entraîné le modèle YOLO sans compensation U-Net en utilisant le jeu de données DS_1 avec $z = 0\mu m$, ici pour le YOLOMZ, l'ensemble d'entraînement se compose du jeu de données DS_1 mais avec une variabilité de mise au point allant de $z = -4\mu m$ à $z = 4\mu m$ avec un intervalle de $2\mu m$. Cette augmentation de la variabilité de mise au point implique que le nombre de patches utilisés est multiplié par un facteur de 5 (4 plans de mise au point supplémentaires).

L'objectif de cette étude est de déterminer si le fait d'avoir un apprentissage multi- z permet au modèle de détection de réaliser une refocalisation implicite et d'extraire efficacement les informations nécessaires pour identifier si les cellules

sont parasitées ou non.

La figure 3.10 illustre les résultats en termes de sensibilité et de spécificité pour le modèle YOLO avec correction U-Net bimodal, le modèle YOLO sans correction, et le modèle YOLOMZ sans correction, en les évaluant sur la base de test pour la plage de mise au point allant de $z = -4\mu m$ à $z = 4\mu m$.

À $z = 0\mu m$, le modèle U-Net bimodal offre les meilleurs résultats en termes de sensibilité et de spécificité. Nous constatons que, malgré le fait d’avoir été entraîné sur plusieurs niveaux de mise au point, le modèle YOLOMZ sans correction U-Net affiche des résultats nettement moins bons que ceux du modèle YOLO sans correction U-Net. Cependant, pour le YOLOMZ, nous notons qu’il présente un comportement assez différent avec des résultats relativement médiocres pour les niveaux de mise au point $z = -4, -2$ et 0 , qui s’améliorent ensuite pour atteindre une sensibilité supérieure à celle du modèle avec correction U-net bimodal aux niveaux $z = 2$ et $4\mu m$.

Cette amélioration est probablement due à une meilleure discrimination des parasites du côté $z > 0$ et à des ondulations perturbant probablement leur détection du côté $z < 0$. Pour confirmer cette hypothèse, il serait nécessaire d’étudier ce cas sur un plus grand nombre d’exemples. Cependant, comme l’amélioration ne concerne qu’un nombre restreint de points de focus, son utilité dans une application courante semble limitée.

Les résultats présentés démontrent déjà quantitativement l’impact positif de la correction d’image apportée par le traitement U-Net bi-modal : il étend considérablement les tolérances du positionnement de l’objectif du microscope et permet également d’étendre la surface exploitable du FOV qui peut être bénéfique pour réduire l’acquisition d’échantillons avec un appareil FPM sur de grandes surfaces.

3.4 Apport de la refocalisation en hématologie

Dans cette section, je présente les résultats de nos travaux sur l’intégration de la refocalisation à l’aide de réseaux de neurones U-Net dans le contexte de l’hématologie. Les efforts déployés dans cette direction visaient à valider cette approche de refocalisation sur un échantillon et une base de données plus importante.

3.4.1 Configuration du microscope

Dans le cadre du projet TAMIS, spécialisé en hématologie, nous avons étendu notre approche de refocalisation aux données hématologiques capturées à plusieurs longueurs d’onde. Cette méthode repose sur un modèle U-Net multimodal, spécifiquement développé pour corriger non seulement le focus mais aussi les aberrations chromatiques associées à l’objectif. La configuration exploite trois longueurs

d’ondes distinctes : $\lambda = 625 \text{ nm}$, 535 nm et 465 nm , ce qui nous permet de combiner les images reconstruites avec des images en couleur RVB (Rouge-Vert-Bleu).

Les échantillons analysés dans cette étude se composent de frottis sanguins de patients sains où les zones d’intérêt sont principalement les leucocytes. Les frottis ont été colorés avec la coloration May-Grünwald Giemsa.

Les résolutions habituellement utilisées pour analyser les leucocytes sont de 40x ou 60x avec une ouverture numérique (NA) de 1.5. En tenant compte des objectifs de microscope disponibles, un bon compromis entre la NA et le temps de mesure nous a conduits à utiliser un objectif 20x avec une ouverture numérique de $NA=0.75$. La source lumineuse est composée de 12 LED disposées en anneau et d’une LED centrale supplémentaire, réalisée avec un anneau NeoPixel RGB 5050 de rayon $r = 15 \text{ mm}$ équipé de 12 LEDs (de chez Adafruit) et d’une LED centrale NeoPixel supplémentaire. La distance des diodes à l’échantillon, notée d , est ajustée pour obtenir une ouverture numérique synthétique $NA_{synth} = 1.5$, avec $\sim d = 13 \text{ mm}$. Le chevauchement dans l’espace de Fourier est d’environ 67 % comme requis dans les références [49, 63].

La caméra utilisée (modèle IDS UI-3200SE) est équipée d’un capteur de grande surface de 4104×3006 pixels (avec un pas de pixel de $3.45 \mu\text{m}$). Cela correspond à un champ de vision (FOV) de $0.705 \times 0.515 \text{ mm}^2$. Le choix de cette configuration de microscope est principalement motivé par le fait qu’elle offre un bon équilibre entre la résolution optique réalisable et la surface de l’échantillon. En effet, cette configuration nous permet de détecter entre 3 et 5 globules blancs par FOV, contrairement aux autres techniques classiques qui n’effectuent qu’une acquisition par globule blanc.

3.4.2 Apprentissage et base de données

La stratégie présentée ici est très similaire à celle introduite dans la section 3.3.4. Nous avons mis en œuvre un modèle U-Net qui, après apprentissage, vise à améliorer le contraste et la netteté des objets biologiques imagés.

Contrairement à l’approche précédente, où le modèle U-Net était utilisé pour traiter l’intensité et la phase pour une seule longueur d’onde, ici, les modèles traitent six modalités. Ils opèrent sur trois couples (I, Φ) , chacun correspondant à une longueur d’onde différente, afin d’améliorer les images et de les ramener au bon focus. L’architecture du modèle U-Net utilisée dans cette étude est presque identique à celle des versions précédentes, avec quelques différences notables. La couche d’entrée accepte 6 canaux au lieu de l’habituel 2, et la couche de sortie prédit également 6 canaux. Cette modification a été faite pour accommoder les différentes modalités de données que nous traitons pour les trois longueurs d’ondes.

Comme expliqué dans le schéma 3.4, le plan de focus optimal de l’objectif du microscope est trouvé avec la LED centrale du microscope FPM. Cela définit la

position de référence $z = 0$. Le frottis sanguin est ensuite acquis pour diverses conditions de mise au point (z allant de $z = -2.25\mu m$ à $2.25\mu m$ avec un pas de $0.75\mu m$) et en utilisant chaque fois 13 angles d'illumination successifs pour chaque longueur d'onde à savoir $\lambda = 625, 535$ et $465nm$.

La plage de focus choisie est liée à la lentille utilisée ici, à savoir un objectif de 20x avec une ouverture numérique (NA) de 0.75. Cette configuration est particulièrement sensible aux petites variations de focus car elle présente une profondeur de champ (DOF) qui est environ trois fois plus petite que celle de l'objectif 10xNA=0.45 utilisé dans la section précédente ($DOF_{20x,0.75} \approx 1.13\mu m$).

Les étapes 3, 4 et 5 demeurent identiques et sont réalisées pour trois longueurs d'ondes différentes.

Comme décrit dans le schéma 3.4, le plan de focus optimal de l'objectif du microscope est déterminé à l'aide de la LED centrale du microscope FPM. Cela définit la position de référence $z = 0$. Ensuite, le frottis sanguin est acquis sous diverses conditions de mise au point, z allant de $z = -2.25\mu m$ à $2.25\mu m$ avec un pas de $0.75\mu m$, en utilisant 13 angles d'illumination successifs pour chaque longueur d'onde, à savoir $\lambda = 625, 535$ et $465nm$.

Le choix de la plage de focus est lié à la lentille utilisée ici, 20xNA=0.75. Cette configuration est sensible aux petites variations de mise au point, car elle a une profondeur de champ (DOF) qui est trois fois plus petite que celle obtenue avec la lentille 10xNA=0.45 utilisée dans la section précédente ($DOF_{20x,0.75} \approx 1.13\mu m$).

Les étapes 3, 4 et 5 sont identiques à celles décrites précédemment, à l'exception qu'elles sont réalisées pour trois longueurs d'onde.

Dans cette étude, nous avons utilisé des lames de frottis sanguins anonymisées provenant de 20 patients sains. Ces échantillons ont servi à entraîner le modèle U-Net.

Pour chaque patient, nous avons acquis 200 régions centrales du champ de vision (mesurant 512x512 pixels après reconstruction) à sept niveaux de mise au point différents. Nous avons utilisé les échantillons de 16 patients pour l'entraînement, ceux de 2 patients pour la validation, et les échantillons des 2 patients restants pour le test. Chaque région reconstruite a ensuite été subdivisée en 25 petits patchs de 128x128 pixels. Cette démarche a été entreprise afin de réduire le temps de traitement et d'assurer une quantité suffisante de données pour l'entraînement.

La taille finale de la base de données d'entraînement est de 560 000 (200x16x7x25) patchs. Les transformations géométriques et l'augmentation des données ont été effectuées de manière dynamique, pour chaque batch, en effectuant aléatoirement soit des rotations, soit des symétries. Les ensembles de données de validation et de test sont chacun composés de 70 000 patchs.

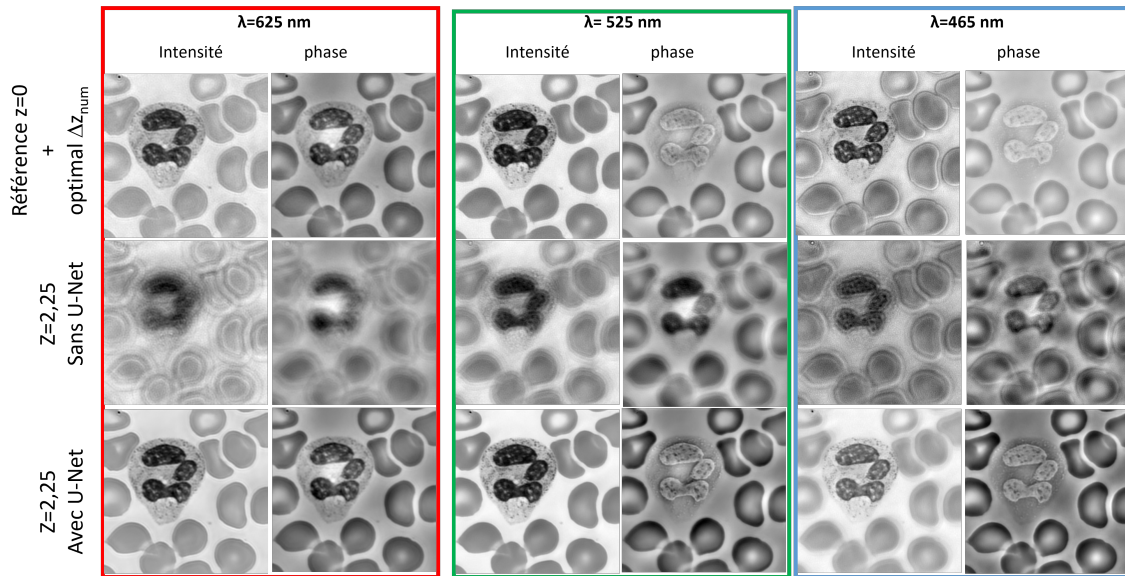


FIGURE 3.11 – Nous avons évalué l’amélioration de la qualité d’image obtenue en utilisant le modèle U-Net pour une mise au point $z = 2.25\mu m$ et pour les 6 canaux RGB en intensité et en phase. Les images en haut sont les images optimales sélectionnées après empilement virtuel en profondeur (z-stacking) pour $z=0$. Au milieu, nous présentons les images de reconstruction obtenues pour une mise au point $z = 2.25\mu m$ sans amélioration. En bas, nous montrons les images qui ont été améliorées à l’aide du modèle U-Net.

3.4.3 Resultats de la refocalisation

Nous avons évalué le modèle U-Net pour les différents niveaux de mise au point allant de $z = -2.25\mu m$ à $z = 2.25\mu m$. Les résultats visibles sur la Figure 3.11 obtenus montrent que le modèle d’amélioration de qualité d’image est capable de traiter les 6 canaux simultanément, en délivrant des images dont la qualité est améliorée, avec des détails aussi fins que ceux des images de référence. À mesure que le point de mise au point change, le contraste des globules blancs dans les images de phase varie également. Le modèle U-Net parvient néanmoins à produire des images dont le contraste correspond à celui qui a été sélectionné comme référence.

Afin d’évaluer la généralité de notre système, nous l’avons testé sur différentes catégories de globules blancs. Les résultats de ces tests sont illustrés dans la Figure 3.12.

Il convient de noter que la qualité des couleurs est légèrement perturbée en raison des différences de rendu colorimétrique de l’algorithme que nous utilisons, en particulier lorsque les images sont hors focus. Néanmoins, en ce qui concerne la qualité des images, nous observons une amélioration assez marquée. Le modèle

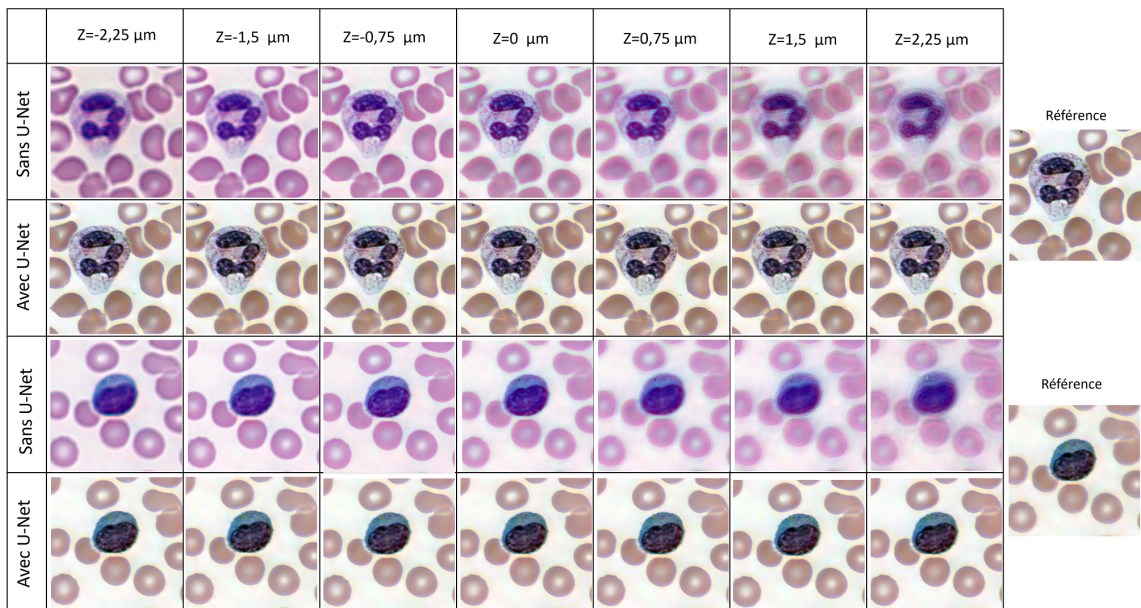


FIGURE 3.12 – Nous illustrons ici l’amélioration de la qualité d’image obtenue à l’aide du modèle U-Net pour différents niveaux de mise au point, allant de $z = -2.25\mu\text{m}$ à $z = 2.25\mu\text{m}$. La figure présente des images d’intensité combinées en une seule image RGB. En haut, nous montrons les résultats obtenus pour des polynucléaires neutrophiles, en bas pour des lymphocytes

U-Net parvient à refocaliser avec succès les images et à les rendre nettes et exploitables pour une future automatisation de la classification. Ceci est vrai pour les niveaux de mise au point allant de $z = -2.25\mu m$ à $z = 2.25\mu m$, ce qui suggère que notre système est robuste pour ce niveau de compensation.

3.5 Conclusion

Avec une résolution fixe, la profondeur de champ de FPM est nativement plus importante que celle obtenue avec un microscope conventionnel, ce qui rend ce type de microscopie attractif. Cependant, lorsque l'épaisseur de l'échantillon devient comparable ou supérieure à la profondeur de champ, une variabilité indésirable de la visibilité des compartiments cellulaires a été observée, en particulier pour les images de phase.

Pour surmonter cette limitation, nous avons introduit une approche d'apprentissage profond visant à améliorer la visibilité des compartiments cellulaires ciblés, dans une perspective de diagnostic automatisé. Le modèle U-Net est conçu pour renforcer simultanément la visibilité des compartiments cellulaires ciblés, mais aussi pour compenser toute imprécision de mise au point. La correction de l'image est effectuée à l'échelle cellulaire.

À cette fin, le réseau neuronal exploite conjointement les images d'intensité et de phase (U-Net bimodal), c'est-à-dire l'ensemble des informations d'échantillonnage accessibles à FPM. Pendant sa phase d'entraînement, la sélection des images de référence optimales est obtenue par exploration fine de la mise au point.

La reconstruction FPM autour du réglage optimal z ; à savoir le z-stacking numérique, est utilisée pour la phase d'apprentissage. Pendant la phase d'inférence, les images sont corrigées et le contraste du compartiment cellulaire ciblé est renforcé.

Pour juger de l'impact de cette approche et contrairement à ce qui se fait habituellement dans la littérature, le problème de parasitologie concret est testé quantitativement (courbes ROC) en complément de l'évaluation des métriques de qualité d'image visuelle. Les résultats expérimentaux révèlent que la sensibilité de détection est significativement améliorée sans dégradation de sa spécificité. De plus, la profondeur de champ de FPM est étendue.

Chapitre 4

Reconstruction par réseaux de neurones cDIP

4.1 Introduction

Dans le chapitre précédent, nous avons abordé et validé l'efficacité d'un réseau de neurones de type U-Net, pour la refocalisation et l'amélioration de la qualité d'image FPM. Cette avancée constitue une réponse à l'une des limitations inhérentes à l'imagerie par microscopie ptychographique de Fourier (FPM), surtout dans des situations où une forte ouverture numérique est impliquée. Toutefois, il est important de noter que cette stratégie, bien qu'efficace, demeure intrinsèquement liée à la technique de reconstruction ePIE traditionnelle. Cette dépendance introduit des contraintes et des limitations algorithmiques, qui se manifestent de manière plus importante lorsque l'on travaille avec des faibles ouvertures numériques. Ces défis algorithmiques limitent l'adoption de la FPM dans le domaine industriel, en dépit de ses avantages intrinsèques.

4.2 Etat de l'art

La FPM est une microscopie qui repose sur l'utilisation d'algorithmes d'optimisation basés sur des méthodes de descente de gradient pour accéder au champ électromagnétique de l'objet observé. L'algorithme classiquement utilisé pour cette technique exploite les données acquises sous différents angles d'illumination pour effectuer la reconstruction. Cependant, ce processus de reconstruction pose plusieurs défis. Il nécessite notamment une grande redondance d'informations, ce qui implique qu'un grand nombre de LED est nécessaire, avec un recouvrement spectral de 60%, rendant cette technique coûteuse tant en temps d'acquisition qu'en temps de reconstruction.

En réponse aux défis de la reconstruction FPM, l'utilisation des réseaux de neurones se présente comme une solution très prometteuse. De nombreux travaux ont proposé des techniques neuronales dans le but de surmonter certaines des contraintes inhérentes aux algorithmes classiques de reconstruction.

Les efforts entrepris dans ce domaine visent à obtenir une reconstruction plus rapide et plus robuste tout en préservant la qualité de l'image. On peut regrouper ces travaux en deux catégories, en fonction de leur approche : Dans le premier type d'approche [61, 93, 44, 14] le processus de reconstruction est obtenu à partir de l'entraînement d'un réseau neuronal profond (DNN). Il s'agit de faire une correspondance entre la pile d'images et l'image intensité et phase reconstruite. Cette transformation étant apprise à partir d'une base de données pour laquelle on connaît cette relation.

Les avantages supplémentaires exposés par la mise en œuvre de ces approches portent principalement sur des reconstructions plus rapides apportées par les GPU dans les calculs feed-forward (inférence), ou une meilleure tolérance au bruit.

Les performances de ces méthodes de reconstruction ont été jugées satisfaisantes dans des contextes expérimentaux optimaux, particulièrement lorsqu'un grand nombre d'images brutes est disponible. Cependant, il convient de souligner que les résultats antérieurement publiés, obtenus soit expérimentalement, soit par simulation, ne sont pas toujours satisfaisants lorsqu'un facteur de résolution élevé est impliqué ou lorsque le nombre d'images brutes est limité. Cette observation reste valide même en utilisant des méthodes complexes comme les GAN [35] pour augmenter la qualité d'image. Par exemple, une perte de résolution et une diminution du niveau de détail (perte des hautes fréquences) sont souvent observées dans de telles situations. De plus, des artefacts apparaissent fréquemment dans les hautes fréquences des images reconstruites, bien que celles-ci soient perceptuellement proches des images reconstruites en utilisant la méthode FPM originale dans les mêmes conditions.

L'autre type est la combinaison des réseaux de neurones aux équations physiques. Parmi ces approches, on retrouve celle de [43] où chaque couche du réseau utilisé représente une équation du modèle physique de la FPM, cette approche visant seulement à un gain temps de reconstruction grâce à l'utilisation des GPU et de tensorflow.

La situation est sur le point de changer avec l'arrivée des deep images prior (DIP) [86, 26]. Dans un tel schéma, un réseau neuronal est utilisé comme un à priori pour résoudre les problèmes inverses. L'une des spécificités remarquables du DIP est qu'il n'exploite aucune donnée d'entraînement préalable. Les caractéristiques de l'image sont capturées par la structure d'un générateur d'image convolutionnel plutôt que par des capacités précédemment apprises. En pratique, l'architecture neuronale introduit une certaine régularisation dans le schéma classique d'optimi-

sation et de l'inversion. Le DIP a prouvé son efficacité dans de nombreux domaines d'application de l'imagerie [91, 54, 25] tels que la réduction du bruit, la super-résolution et l'inpainting. Plus récemment, une reconstruction d'image avec une résolution axiale améliorée a été obtenue dans le domaine de la tomographie de diffraction optique [7, 97, 88] exploitant ses capacités de régularisation. D'autres approches ont également été mises en œuvre, telles que les travaux de Baroni [10] qui ont permis de résoudre le problème inverse en ptychographie de Fourier vectorielle en intégrant le modèle cDIP. J'ai moi-même participé à ce travail en intégrant le modèle physique sous-jacent dans la fonction de coût, ce qui était difficilement exprimable analytiquement.

Approche proposée

Ce chapitre est consacré à la présentation des travaux que j'ai réalisés en ptychographie de Fourier, en utilisant le Deep Image Prior (DIP) pour inverser le problème direct. Je décrirai en détail l'architecture du modèle basé sur le DIP, que j'ai conçu et mis en œuvre. Ensuite, je discuterai de mon approche pour la reconstruction conjointe de l'objet et de la fonction de la pupille, une méthode qui diffère de celles déjà publiées. J'exposerai également les résultats obtenus en utilisant uniquement le DIP, en me concentrant sur les conditions où le chevauchement des données est faible. Pour évaluer ces expériences, j'ai eu recours à des données de simulation, comportant des variations en phase ainsi qu'en intensité.

4.3 Modèle cDIP informé par la physique

Le DIP est un formalisme qui a été initialement introduit dans le contexte de l'inversion de problèmes en traitement d'images. Il repose sur un réseau neuronal non entraîné dont les poids sont optimisés pour chaque exemple à inverser. Cette approche est radicalement différente de l'apprentissage classique, car elle ne nécessite pas l'utilisation d'un ensemble d'entraînement. Elle ne s'intéresse qu'à résoudre le problème inverse associé à un seul exemple en incluant explicitement les équations du modèle direct dans un réseau neuronal profond. Le réseau neuronal (typiquement un CNN, une architecture d'encodeur-décodeur, ou un U-Net [77]) prend comme entrée initiale un bruit aléatoire δ . Sa structure fournit une régularisation implicite pendant la recherche d'une solution, ce qui permet de trouver une solution adéquate pour les problèmes mal posés. L'objectif est de trouver un ensemble de causes X d'un phénomène à partir des observations expérimentales de ses effets Y .

Nous dénoterons F le modèle direct reliant X à Y . Typiquement, X est obtenu à partir de Y en résolvant un processus d'optimisation dans tout l'espace de X , qui consiste à trouver l'élément qui correspond le mieux aux effets observés, une

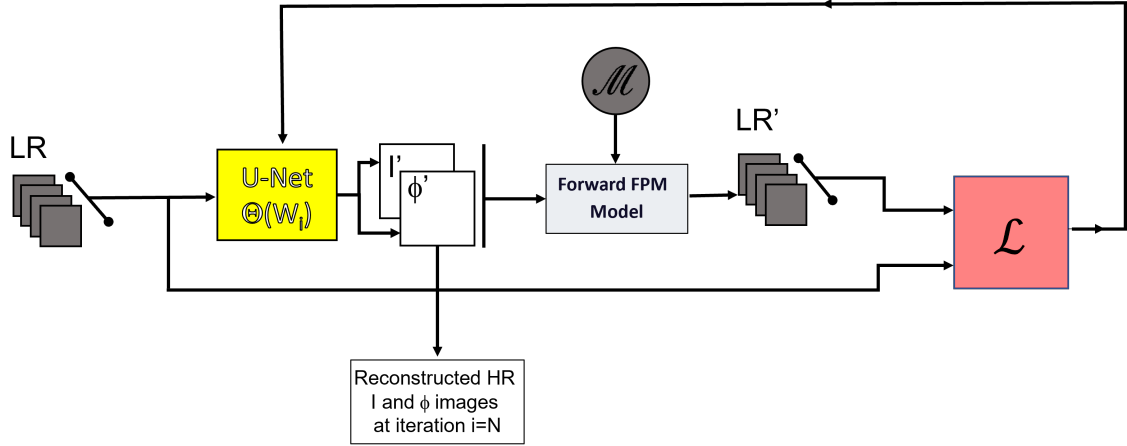


FIGURE 4.1 – Modèle cDIP : LR désigne la pile d’images FPM de faible résolution mesurées au niveau du plan de la caméra en utilisant les LEDs \mathcal{M} . Θ représente les paramètres du U-Net qui sont optimisés après chaque itération de calculs de rétropropagation. I' and Φ' représentent les images HR reconstruites avec U-Net à chaque itération. Les images reconstruites I, Φ sont obtenues lorsque le modèle a convergé. LR' est la pile d’images de faible résolution calculée par le modèle direct. La fonction de coût \mathcal{L} .

fois traité par le modèle direct F. Cette formulation est souvent insuffisante pour la résolution de problèmes mal posés et l’introduction de termes de régularisation supplémentaires est nécessaire pour contraindre les valeurs possibles de X . Cependant, ces termes de régularisation sont souvent ad hoc et peuvent être insuffisants pour résoudre les problèmes mal posés [86]. L’intérêt des réseaux DIP réside dans leur capacité à introduire un autre type de régularisation basé sur les poids du réseau lui-même.

Les approches DIP ont d’abord été employées pour résoudre des problèmes bien posés tels que les problèmes de débruitage ou d’inpainting. Elles ont ensuite été étendues pour résoudre des problèmes inverses mal posés où elles fournissent de meilleurs résultats que les méthodes classiques sur les données de simulation et expérimentales [86].

Nous avons utilisé dans cette thèse une approche DIP pour la reconstruction des images FPM, illustrée dans la (Figure 4.1). Par rapport au DIP, le modèle cDIP prend en entrée une pile d’images au lieu d’un bruit blanc. Le modèle est composé de deux blocs, à savoir un U-Net et un bloc FPM direct. Le U-Net prend en entrée l’ensemble complet des images de faible résolution (LR) expérimentales capturées avec les LEDs de \mathcal{M} . Sa sortie est un couple d’images d’intensité et de phase ou de manière équivalente la fonction recherchée $O(\nu) = F[I'e^{j\Phi'}]$

Le rôle du U-Net est d'extraire des informations des images présentées en entrée. Il fournit en sortie l'approximation optimale de la solution recherchée. De plus, la solution estimée $O(\nu)$, mise à jour à chaque époque, est transmise au modèle FPM direct. Ce module est chargé de calculer les images brutes associées à la sortie du U-Net, en utilisant les équations de formation d'image décrites dans la section 2.3.3. Pour déterminer la solution recherchée, il est ensuite nécessaire de minimiser l'Erreur Quadratique Moyenne (MSE) entre ces images et celles qui ont été obtenues expérimentalement. La perte résultante peut être formulée comme suit :

$$\mathcal{L} = \sum_{i=1}^{\mathcal{N}} \left| \sqrt{|F^{-1}[F[I'e^{j\Phi'}](k_x, k_y) \cdot \widehat{P}(k_x - k_x^{(i)}, k_y - k_y^{(i)})]|} - \sqrt{I^i(x, y)} \right|^2 \quad (4.1)$$

où F et F^{-1} sont respectivement la transformation de Fourier et sa transformation inverse, I' et Φ' représentent l'intensité et la phase en sortie du CNN, \widehat{P} est la transformation de Fourier de la fonction d'étalement du point, $k_x^{(i)}$ et $k_y^{(i)}$ représentent les projections du vecteur k le long des axes x et y , correspondant à l'éclairage LED i^{me} , et $I^i(x, y)$ désigne les images basse résolution.

L'intégration des équations de formation d'image du microscope au sein de la fonction de coût, comme indiqué en 4.1, nous permet de considérer le modèle comme faisant partie des modèles "Physics-informed" selon [45].

Le U-Net effectue différentes opérations de convolution et de pooling sur plusieurs couches et son architecture précise est détaillée en 4.2.

En pratique, les poids du réseau sont initialisés aléatoirement et optimisés grâce à un processus itératif visant à réduire la perte grâce à la descente de gradient. Le processus est arrêté lorsque la stabilité de la fonction de perte est obtenue.

L'architecture mise en œuvre ici est très similaire à celle du U-Net standard, tel que décrit dans le Chapitre 2 et présenté en Figure 4.2. Ce modèle prend en entrée une pile d'images à basse résolution (LR) et produit les images d'intensité et de phase reconstruites en sortie.

4.4 Configuration du dispositif en microscopie ptychographique de Fourier

Afin d'évaluer la reconstruction par la méthode cDIP, nous utilisons un ensemble de données simulées. L'objectif majeur de cette démarche est de générer une vérité terrain pour l'information d'intensité et de phase. De plus, il est essentiel d'obtenir une phase et une intensité qui sont complètement décorrélés. Cela nous permet d'apprécier si l'information est correctement séparée, et ainsi d'assurer une analyse indépendante.

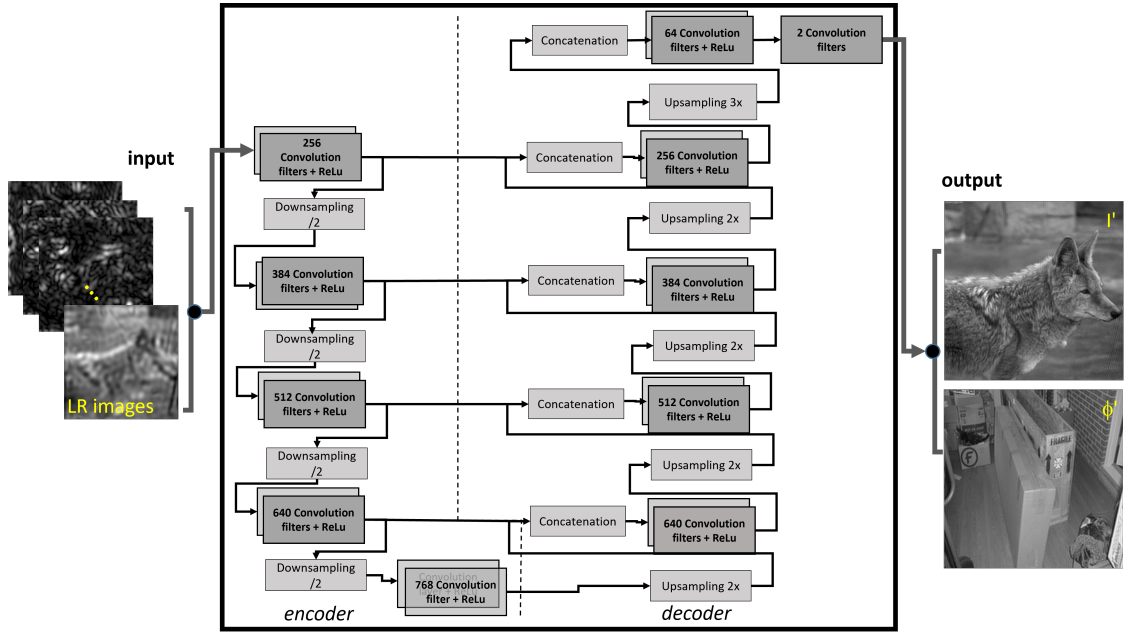


FIGURE 4.2 – L’architecture du U-Net utilisée pour résoudre le problème inverse selon les différents schémas.

Afin de réaliser cette tâche, nous employons le modèle direct de la microscopie ptychographique de Fourier pour convertir notre intensité I et phase Φ en une pile d’images en basse résolution. Ce principe est exploité dans la fonction de coût du cDIP, comme le montre de manière explicite le schéma référencé à la Figure 4.1.

Dans la suite, nous considérons une matrice circulaire de LEDs \mathcal{M} (voir Figure 4.3a). Cette matrice peut être divisée en deux groupes distincts de LEDs notés \mathcal{M}_A et \mathcal{M}_B avec $\mathcal{M} = \mathcal{M}_A \cup \mathcal{M}_B$. Leur disposition est indiquée dans le domaine de Fourier dans Figure 4.3b et Figure 4.3c, respectivement. La matrice de LEDs \mathcal{M}_A nous permet de paver une large région spectrale et donc d’atteindre un facteur de super-résolution important. \mathcal{M}_A est composée de 37 LEDs et conduit à un facteur de recouvrement spectral $\Gamma \sim 10\%$. La géométrie de \mathcal{M}_A a été conçue pour obtenir le facteur de recouvrement le plus faible possible sans aucune information manquante dans le domaine de Fourier. \mathcal{M}_B est composée de 93 LEDs complémentaires. Combinées avec les LEDs de \mathcal{M}_A , elles nous permettent d’atteindre $\Gamma = 60\%$ (voir la légende de la Figure 4.3 pour plus de détails sur la configuration du microscope).

Nous simulons la configuration d’un microscope équipé d’un objectif de grandissement $4\times$ et de $NA = 0.08$. La matrice de LEDs \mathcal{M} , composée de 130 LEDs à $\lambda = 525nm$, est introduite dans la Figure 4.3 et est placée à une distance $d = 55mm$

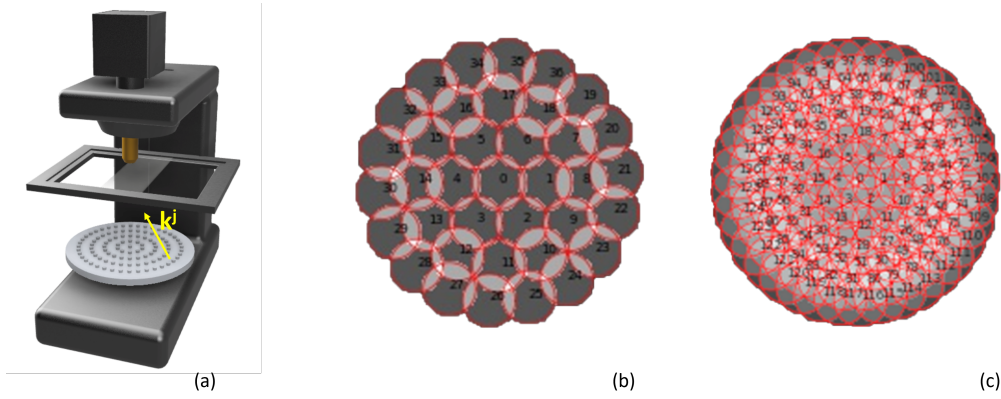


FIGURE 4.3 – Présentations schématiques du microscope équipé d’une matrice de LEDs $\mathcal{M} = \mathcal{M}_A \cup \mathcal{M}_B$ (a), région spectrale couverte dans le domaine de Fourier avec \mathcal{M}_A (b) et avec \mathcal{M} (c).

en dessous de l’échantillon. Une telle configuration nous permet d’atteindre un facteur de super-résolution théorique $\gamma = 5$.

Cette configuration nous permet d’obtenir une pile FPM de \mathcal{N} images, avec $\mathcal{N} = 130$. Chaque image de la pile FPM sera de taille 128×128 . Nous avons choisi cette taille d’image en basse résolution afin de produire des images reconstruites de taille 384×384 . Pour les études de la reconstruction cDIP qui suivent, nous utiliserons la configuration \mathcal{M} avec toutes les LEDs, avec un fort recouvrement $\Gamma = 60\%$ comme condition de référence. Puis, pour un faible recouvrement, nous n’utiliserons que 37 LEDs de la matrice \mathcal{M} , ce qui correspond à \mathcal{M}_A .

4.5 Resultat cDIP $\Gamma = 60\%$

Pour évaluer le modèle cDIP, nous nous sommes concentrés sur la comparaison entre le cDIP et le ePIE dans des conditions favorables (130 LEDs et $\Gamma = 60\%$).

Dans le cadre de cette étude comparative, notre objectif était d’évaluer la performance du modèle cDIP en termes de reconstruction d’image, en le comparant à l’approche ePIE traditionnelle sous différentes conditions expérimentales. Plus précisément, nous avons examiné le modèle cDIP en présence et en absence de fonctions de défocalisation afin de déterminer si les caractéristiques intrinsèques de l’ePIE tels que l’intégration d’un terme de phase pour la correction de la mise au point, étaient également pertinents et efficacement transposables au cadre méthodologique du cDIP.

La défocalisation sur les données simulées a été mise en œuvre par l’introduction d’un terme de phase dans la fonction de pupille $P(\nu)$ ce qui permet de générer des

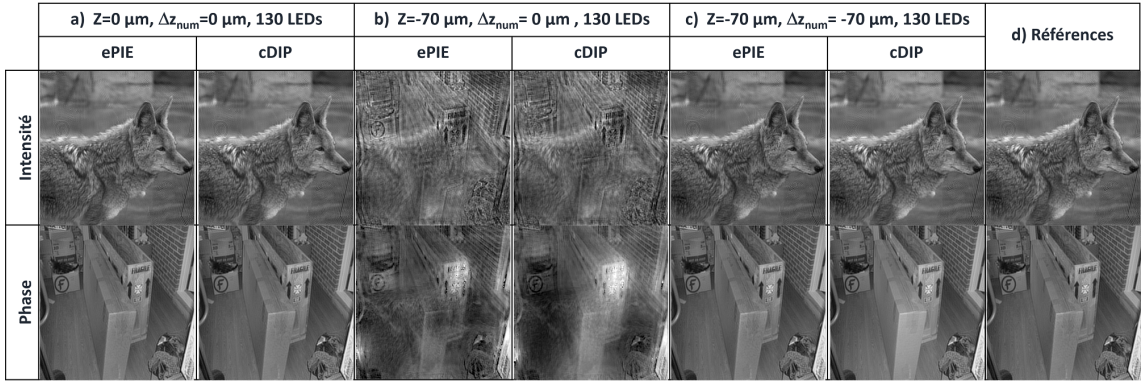


FIGURE 4.4 – Comparaison visuelle des images reconstruites par ePIE et cDIP pour différentes conditions expérimentales. a) pour $z = 0\mu\text{m}$ et $\Delta z_{\text{num}} = 0\mu\text{m}$, b) pour $z = -70\mu\text{m}$ et $\Delta z_{\text{num}} = 0\mu\text{m}$, c) pour $z = -70\mu\text{m}$ et $\Delta z_{\text{num}} = -70\mu\text{m}$ et d) le couple intensité-phase de référence

piles FPM (LR) avec le niveau de defocus désirées.

Les résultats obtenus par cDIP dans les conditions spécifiques de $z = 0\mu\text{m}$ et $z = -70\mu\text{m}$ démontrent une qualité similaire à ceux produits par ePIE.

Dans le cas des reconstructions à $z = -70$ et $\Delta z_{\text{num}} = 0$, l'analyse montre que les deux algorithmes partagent le même comportement pour les objets défocalisés. Ils peuvent facilement être remis au bon focus avec le bon paramètre de Δz_{num} . Cette observation confirme que l'opération de refocalisation numérique peut être effectivement intégrée dans le processus de reconstruction cDIP.

Cependant, un inconvénient majeur de cette méthode réside dans la vitesse de convergence du cDIP, qui se révèle être assez lente. Le principal facteur contribuant à cette lenteur est la manière dont les poids Θ_s sont initialisés, qui est faite aléatoirement. Pour ce jeu de données particulier, le cDIP converge au bout de 1500 itérations, chaque itération prenant 200 ms sur une Nvidia RTX 8000, pour un total de 5 minutes. En contraste, ePIE, en tant que méthode de référence, ne prend que 30 secondes pour converger (100 itérations) lorsqu'il est exécuté sur un CPU.

4.6 Resultat cDIP $\Gamma = 60\%$ avec correction d'aberrations

Au cours de notre étude sur les performances du modèle cDIP que nous avons mis en œuvre, une question essentielle s'est posée sur l'intégration d'une méthode de correction des aberrations optimisée lors du processus d'inversion, à l'instar de

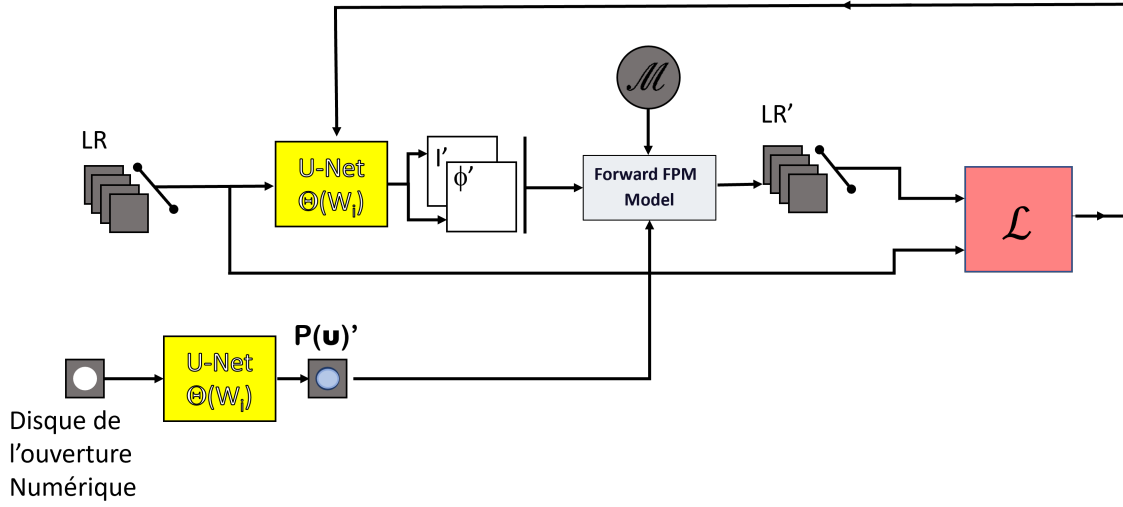


FIGURE 4.5 – Schema cDIP adapté pour inclure une correction d'aberration

l'algorithme EPRY. Cette correction a une importance capitale, notamment pour les systèmes optiques sujets à des aberrations considérables, lesquelles peuvent affecter significativement la qualité des images de phase et d'intensité. Il est d'ailleurs à noter que ces aberrations sont souvent propres au système optique lui-même et peuvent même nécessiter des ajustements en fonction de l'échantillon observé. Par conséquent, l'intégration d'un mécanisme de correction d'aberration en parallèle avec le processus de reconstruction d'image peut être nécessaire.

Pour aborder ce défi, nous avons conçu un modèle neuronal dédié qui fonctionne en parallèle avec notre modèle principal de reconstruction d'intensité et de phase. Ce modèle neuronal annexe a pour mission spécifique d'estimer les corrections d'aberrations nécessaires, comme illustré dans le schéma de la Figure 4.5.

Ces travaux ont été effectués au début de 2022. Récemment, en octobre 2022, une équipe a publié une approche similaire [20], où une reconstruction par cDIP intégrant une correction de pupille a été mise en avant. Leur approche repose sur un MLP (Multi-Layer Perceptron) qui estime les coefficients du polynôme de Zernike[55] afin de générer une approximation $\widehat{P}(\nu)'$, qu'ils intègrent, comme nous l'avons fait, dans la reconstruction avec correction d'aberration.

Dans notre modèle, l'adaptation du schéma cDIP s'est traduite par l'intégration d'un modèle U-Net [77] dans le processus de reconstruction. Ce modèle prend en entrée une image correspondant au disque de l'ouverture numérique et cherche en sortie à estimer la fonction de phase qui correspond à $\widehat{P}(\nu)'$ dans la fonction de coût 4.1.

Dans le cadre de l'évaluation de notre modèle, nous avons introduit une fonction

d’aberrations basée sur les cinq premiers ordres de polynômes intégrés à la fonction de phase durant la génération de notre séquence d’images pour la microscopie ptychographique de Fourier (FPM). L’objectif était de déterminer si notre modèle cDIP pouvait efficacement estimer ces corrections d’aberration tout en réussissant simultanément la reconstruction des images de phase et d’intensité associées.

Les résultats obtenus sont remarquablement satisfaisants. Comme nous le voyons dans 4.6a), si nous ne corrigeons pas l’aberration, les images de phase et d’intensité reconstruites sont de très mauvaise qualité. En revanche, l’introduction de la correction a permis d’obtenir une reconstruction 4.6b) très proche de l’image de référence, tout en estimant avec précision l’aberration du système.

Cette méthode confirme que la reconstruction conjointe de la pupille, de l’intensité et de la phase offre des résultats satisfaisants. Il est évident que cette intégration d’une correction de l’aberration améliore substantiellement la qualité de l’image, validant ainsi notre approche.

4.7 Resultat cDIP pour $\Gamma = 10\%$

L’un des plus grands défis en FPM est d’obtenir des résultats de haute qualité avec un nombre réduit de LEDs comme expliqué dans le Chapitre 2. Ainsi, nous avons expérimenté le modèle cDIP dans ce type de reconstruction, en utilisant une configuration réduite pour examiner la mise en œuvre à la place du e-PIE qui permet d’obtenir une meilleure qualité d’image.

Dans cette étude, nous avons utilisé la matrice \mathcal{M}_A avec 37 éléments, recouvrant le même spectre que les expériences précédentes avec 130 LEDs, mais avec un recouvrement qui permet une redondance ($\Gamma \sim 10$).

Les résultats de la Figure 4.7 montrent que les méthodes ePIE et cDIP, dans ces conditions de recouvrement, ne permettent pas une bonne reconstruction des images de phase et d’intensité. Les images reconstruites présentent de nombreux artefacts, et bien que la reconstruction cDIP permette d’obtenir une image de phase légèrement améliorée, la qualité reste nettement insatisfaisante par rapport à l’image de référence.

4.8 Conclusion

Le modèle cDIP que nous avons proposé pour la reconstruction est original ; c’est un DNN informé par la physique, en ce sens que le modèle de formation d’image de microscope (modèle direct) est explicitement introduit dans le modèle pour calculer la fonction de perte ; les images HR d’intensité et de phase reconstruites sont donc obtenues par ce modèle, mais leurs valeurs n’ont jamais

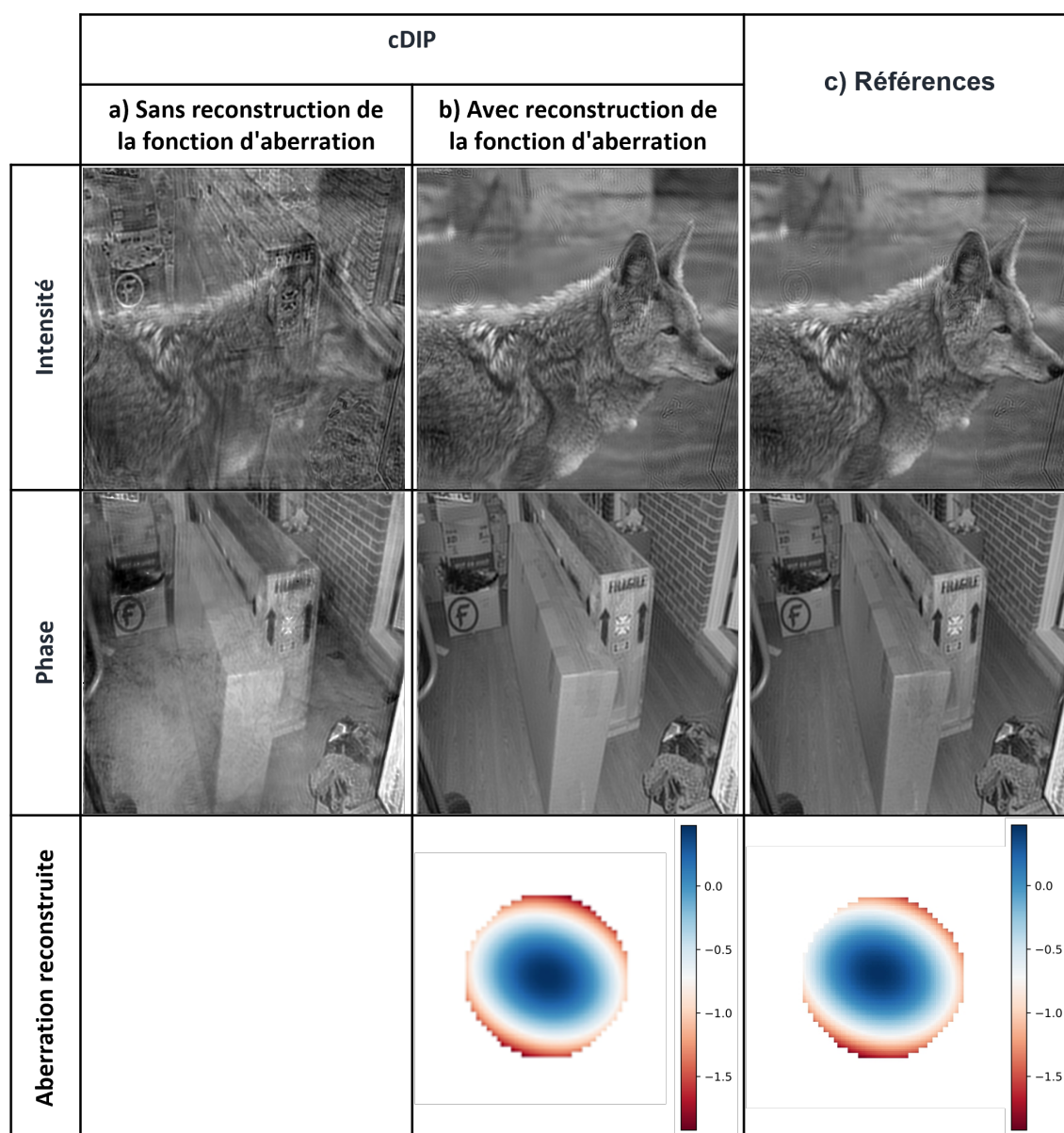


FIGURE 4.6 – Resultat cDIP $\Gamma = 60\%$ avec correction d'aberrations. a) Reconstruction cDIP de l'intensité et de la phase sans la reconstruction de la fonction d'aberration. b) Reconstruction cDIP de l'intensité, de la phase et de la fonction d'aberration. c) Intensité, phase et fonction d'aberration de référence.

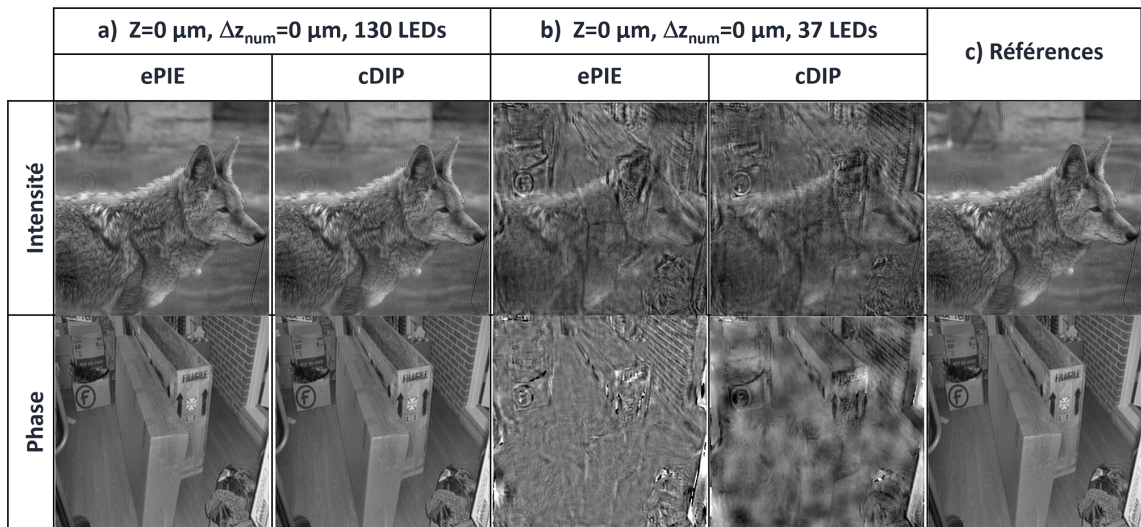


FIGURE 4.7 – Reconstruction FPM dans différentes expérimentales : Intensité et phase obtenues avec un recouvrement de $\Gamma = 60\%$ (130 LEDs). b) Intensité et phase obtenues avec un recouvrement de $\Gamma = 20\%$ (37 LEDs). c) Intensité et phase de référence.

été utilisées directement pour calculer l'erreur.

Nos études de simulation montrent que dans des conditions de fort recouvrement $\Gamma = 60\%$, le modèle cDIP offre une reconstruction de haute qualité en utilisant une stratégie de correction de gradient globale, laquelle s'appuie sur l'incorporation d'équations physiques dans la fonction de coût.

Ce modèle est également capable d'effectuer une correction de la fonction de la pupille, permettant ainsi de compenser efficacement les aberrations et de récupérer les informations en phase et en intensité de l'objet, même lorsque les images initiales sont fortement dégradées. Toutefois, il est important de noter que ces modèles nécessitent une optimisation spécifique des poids pour chaque objet étudié, requérant environ 2000 itérations. Ce temps de calcul est nettement supérieur à celui de l'approche ePIE, qui ne nécessite, environ qu'une vingtaine d'itérations.

Dans la condition de faible recouvrement Ce résultat révèle que le modèle cDIP que nous avons mis en place avec l'architecture U-Net ne permet pas une reconstruction FPM satisfaisante dans des conditions de faible recouvrement. Nous proposons dans le chapitre suivant (Chapitre 5), un protocole qui nous permettra de diminuer encore ce taux de recouvrement tout en gardant de bonnes propriétés de reconstruction.

Chapitre 5

Stratégie de reconstruction FPM dans les conditions de faible recouvrement

Dans le chapitre précédent, nous avons exposé les résultats obtenus pour un faible recouvrement, qui se sont avérés infructueux, que ce soit notre approche cDIP ou l'algorithme ePIE. Dans ce chapitre, nous allons présenter la stratégie mise en place pour dépasser cette limitation et ainsi obtenir une reconstruction de bonne qualité dans des conditions de faible recouvrement. Ce travail original a fait l'objet d'une publication dans le journal *Sensors* [15]

5.1 Etat de l'art

Comme expliqué dans le Chapitre 2, le temps d'acquisition en microscopie de Fourier ptychographique (FPM) est un facteur critique, car il est directement proportionnel au nombre d'images brutes nécessaires pour la reconstruction de l'image et ce nombre est nativement élevé. À cet égard, obtenir une haute qualité de reconstruction avec un nombre réduit d'images est un domaine de recherche de premier plan. La méthode initiale mise en avant par le groupe de Waller, telle que citée dans les travaux [85, 84], emploie un codage multiplex de l'éclairage. Dans cette méthode, l'échantillon sous observation est illuminé simultanément par plusieurs ondes planes angulaires. Chacune de ces ondes planes permet d'obtenir des informations de différentes régions du domaine de Fourier. Lorsque ces ondes interagissent avec l'échantillon et sont ensuite capturées par la caméra, elles se combinent en une seule image composite. L'algorithme de reconstruction utilise alors une technique de séparation de sources pour décomposer cette image composite en ses composantes individuelles dans le domaine de Fourier.

Ce processus de séparation est intégré dans l'algorithme de reconstruction, permettant ainsi une efficacité accrue en termes de temps d'acquisition et de qualité de l'image reconstruite. C'est une avancée significative qui contribue à rendre la FPM envisageable dans une variété d'applications en imagerie biomédicale et autres domaines nécessitant une haute résolution et un temps d'acquisition rapide.

Cette méthode présente deux avantages : le nombre d'images qui doivent être enregistrées pour une couverture complète de la région spectrale ciblée est considérablement réduit ; de plus, la puissance optique totale utilisée pour sonder l'échantillon est augmentée proportionnellement au nombre de LED utilisées simultanément, ce qui peut entraîner une possible réduction du temps d'exposition de la caméra. Cependant, les conséquences exactes de ce codage sur la qualité de l'image reconstruite ne sont pas encore claires puisque l'information acquise avec une dynamique de caméra donnée doit être distribuée sur de larges régions spectrales.

Des approches alternatives ont également été étudiées où le processus de reconstruction est obtenu à partir de l'entraînement d'un réseau neuronal profond (DNN) que nous avons cité dans le Chapitre précédent (Chapitre 4) tel que dans [44].

Dans ce travail, nous abordons ce problème de la reconstruction FPM avec peu d'images brutes dans un contexte d'apprentissage profond. La méthode que nous proposons ira au-delà des deux tendances précédentes ; plus précisément, nous introduisons ici une méthode générale basée sur le formalisme DIP. Dans une première étape, le réseau cDIP est entraîné afin de réaliser des reconstructions à partir d'un jeu de données simulées de Fourier Ptychographic Microscopy (FPM). Suite à cette phase d'entraînement et à l'apprentissage des poids du réseau, le modèle est en mesure de fournir une estimation moyenne pour les reconstructions FPM.

Dans une seconde étape, pour un objet donné, les poids sont davantage optimisés afin d'obtenir une estimation précise de la solution recherchée qui est spécifique à cet objet.

Le principe d'un tel schéma de reconstruction est démontré à l'aide de simulations. Une reconstruction avec un facteur de super-résolution d'environ $\gamma = 5$ sans artefact appréciable dans les hautes fréquences est obtenue avec seulement 37 images brutes. Un motif spécifique de matrice de LED d'éclairage de microscope est considéré dans les simulations. Son arrangement spatial a été choisi afin de paver l'espace de Fourier avec le moins de LED possible. L'approche est générale et applicable à d'autres motifs d'éclairage tels qu'une matrice de LED rectangulaire classique. Elle peut également ouvrir des perspectives intéressantes pour la reconstruction FPM sous un éclairage codé en multiplex afin de réduire encore le nombre nécessaire d'images capturées.

5.2 Configuration FPM

Comme expliqué dans la partie FPM du chapitre 2 2.3.3, il est important de noter que le problème inverse à résoudre en FPM est mal posé. Cela impose certaines conditions expérimentales sur la configuration : le bruit thermique doit être suffisamment faible, ce qui peut être réalisé en fixant une puissance de LEDs adéquate. De plus, le nombre de LEDs (ou d'images LR exploitées) doit être suffisant. Par ailleurs, et étant donné que la descente de gradient est réalisée de manière itérative, seules des données expérimentales partielles sont exploitées (image I^i) à chaque étape. Le facteur de recouvrement Γ entre deux régions successives mises à jour dans le domaine de Fourier aux itérations i et $i + 1$ doit être supérieur à 50–60%. Cette question a été étudiée par de nombreux auteurs et $\Gamma > 60\%$ a été déterminé à partir de simulations qui ont été confirmées par la suite expérimentalement [82],[19]. Ces régions de recouvrement sont indiquées par les régions grises de la Figure 4.3. Par conséquent, le domaine de Fourier doit être pavé avec une grande redondance, c'est-à-dire avec de nombreuses images différentes I^i . Cela limite le débit du FPM. Il est clair que la quantité théorique minimale de données qui devrait être acquise dépend quadratiquement du facteur de super-résolution γ . Ce facteur γ est défini dans le domaine de Fourier par le rapport entre la surface à reconstruire et la surface couverte par l'ouverture numérique de l'objectif du microscope. Par exemple, pour $\gamma = 5$, seuls 36 angles d'éclairage différents (ou images équivalentes) sont théoriquement nécessaires (pour $\Gamma \sim 0\%$). Cependant, pour assurer une convergence correcte de l'algorithme de reconstruction, Γ doit être supérieur à 60%. Les échantillons doivent donc être acquis avec une grande redondance, ce qui entraîne un débit lent pénalisant le microscope. Dans [95], la configuration expérimentale utilisée assure $\gamma = 6$ et 225 éclairages ont été nécessaires. Par conséquent, le temps minimal pour l'acquisition d'un FOV en couleurs RVB est d'environ 1'30 s en considérant seulement 20 ms de temps d'exposition pour chaque acquisition de pile (en conditions de champ clair et même sombre).

Comme indiqué précédemment, l'utilisation de \mathcal{M}_A seule ne peut pas nous permettre d'obtenir une reconstruction d'image HR FPM exploitable en utilisant l'ePIE, cDIP, ou des algorithmes de reconstruction connexes à cause de Γ . Cette propriété est probablement générale. De plus, nous avons observé (simulations non présentées ici) que la situation est inchangée même si un algorithme de descente de gradient global [43] est utilisé. Cela signifie que de nouvelles stratégies de reconstruction algorithmique sont nécessaires. Nous avons ainsi élaboré une méthodologie qui s'appuie sur le modèle cDIP, associé à une stratégie d'apprentissage, en vue de pallier les limitations liées aux conditions de faible recouvrement.

5.3 Schéma de la reconstruction par cDIP-DNN

Nous proposons d'incorporer une étape d'apprentissage supplémentaire dans le modèle avant l'optimisation DIP 4.1.

Dans une telle situation, les poids Θ_s du réseau sont optimisés pour un grand nombre d'exemples. Le but de cette première étape est d'approximer la fonction inverse elle-même (à partir de divers exemples) plutôt que de résoudre le problème inverse pour un exemple spécifique. Après la convergence de cette étape d'apprentissage, une seconde étape est entreprise. Là, la fonction inverse (par le biais des poids Θ_s) est résolue spécifiquement pour les données considérées qui doivent être inversées. La fonction inverse approximée obtenue à partir de l'étape d'apprentissage est utilisée pour initialiser l'étape d'optimisation 4.1. La même architecture cDIP (c'est-à-dire unifiée) est exploitée pendant ces deux étapes de calcul, ce qui assure une cohérence globale entre l'étape d'apprentissage et celle d'optimisation.

5.3.1 Schéma d'apprentissage U-Net pour l'initialisation de cDIP (Étape 1)

Pour constituer la base de données, le modèle électromagnétique direct F est appliqué à chaque image HR (I, Φ) d'un ensemble de données d'apprentissage (LDS), conduisant à de nombreuses piles d'images basses résolution LR formées au niveau du plan de la caméra pour les différentes illuminations utilisées (voir 4.3). La taille de chaque pile dépend du nombre de LEDs considérées. Plus précisément, à ce stade, nous appliquons F en considérant les $\mathcal{N} = 130$ LEDs composant \mathcal{M} . Nous divisons les images LR résultantes en deux sous-piles : celles liées aux LEDs de \mathcal{M}_A ($\mathcal{N}_A = 37$ images) et à \mathcal{M}_B ($\mathcal{N}_B = 93$ images), respectivement.

En effet, dans cette phase d'apprentissage, nous profitons de la disponibilité de ces deux sous-piles (à savoir la pile réduite et la pile complémentaire) d'images LR pour faciliter la convergence du réseau et donc obtenir une bonne qualité de reconstruction. Les \mathcal{N}_A images correspondant à la configuration à LEDs réduite \mathcal{M}_A sont l'entrée d'un U-Net, qui prédit un couple (I', Φ') d'images HR. Ses paramètres Θ_s (les poids) sont appris, en utilisant une optimisation DIP d'une fonction de perte, qui est l'erreur quadratique moyenne entre la pile initiale de \mathcal{N} images et la pile de \mathcal{N} images résultant du modèle direct appliqué sur (I', Φ') . À chaque itération i , un nouveau couple d'images de l'ensemble de données d'apprentissage est considéré ; de cette façon, les paramètres du U-Net, les Θ_s sont optimisés pour un grand nombre d'exemples, contrairement à ce qui se passe dans l'implémentation DIP classique où un seul exemple est considéré.

Il faut noter que, même si l'entrée du U-Net est limitée à la pile réduite d'images, la fonction de perte exploite la pile complète de \mathcal{N} images. De cette façon, l'image finale reconstruite bénéficie des informations riches obtenues à par-

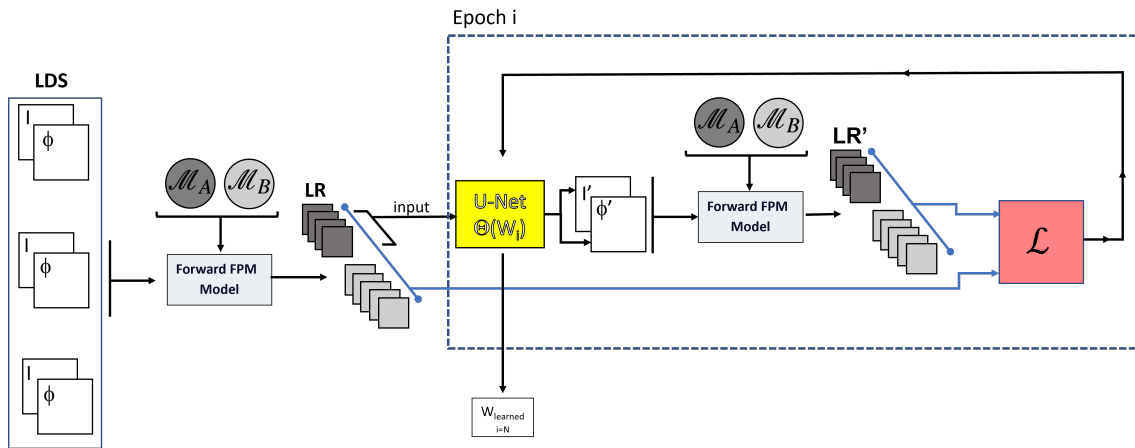


FIGURE 5.1 – Modèle cDIP-L (modèle d'apprentissage U-Net). LDS : ensemble de données d'apprentissage composé de nombreux couples d'images d'intensité I et de phase Φ . Le modèle FPM direct est utilisé pour calculer des images à basse résolution formées au niveau du plan de la caméra. \mathcal{M}_A et \mathcal{M}_B représentent la division de la matrice de LED en deux groupes. LR : pile d'images à basse résolution formées au niveau du plan de la caméra pour les différentes illuminations utilisées. Θ représente les paramètres du U-Net. LR' : images à basse résolution calculées à partir des images prédites par le U-Net I' et Φ' (en gris foncé pour \mathcal{M}_A , gris clair pour \mathcal{M}_B). \mathcal{L} : fonction de perte.

tir des images HR originales. De plus, la fonction de perte ne considère pas une MSE entre (I, Φ) et (I', Φ') comme c'est généralement le cas dans la littérature. En effet, dans l'approche actuelle, la fonction de perte résulte de la comparaison des images LR résultant du modèle direct avec le nombre complet de LEDs \mathcal{N} , appliqué aux deux couples d'images HR. Cela garantit que les équations physiques sont explicitement résolues pendant l'étape d'apprentissage. Nous dénotons cDIP-L comme le modèle de l'étape 1 (pour cDIP-Learned)

5.3.2 Reconstruction cDIP sous conditions de faible recouvrement (Étape 2)

Une fois que les paramètres du réseau sont appris (après un nombre approprié d'itérations), le modèle est prêt à reconstruire une pile inconnue d'images LR FPM acquises sous des conditions de chevauchement réduit Γ . La pile réduite de \mathcal{N}_A images (liées aux illuminations \mathcal{M}_A) est présentée à l'entrée du U-Net. Comme expliqué précédemment, nous avons une dernière étape (étape 2) afin d'affiner la solution. Cette étape est nécessaire de manière à améliorer encore la qualité de la reconstruction, comme cela sera visible dans la section 5.4. Cette étape est obtenue en résolvant le modèle direct qui est attaché à un seul ensemble d'images LR seulement, comme détaillé dans Chapitre 4.

Il est à noter que cette optimisation est effectuée sans aucun changement dans l'architecture et la même fonction de perte qui a été introduite à l'étape 1 est exploitée. La seule différence réside dans le fait que les poids du réseau $\Theta_{learned}$, obtenus après la phase d'apprentissage, sont utilisés pour l'initialisation du U-Net. De plus, ces poids sont maintenant optimisés en utilisant uniquement la pile réduite, car c'est la seule information disponible à ce stade. L'image HR reconstruite finale est obtenue à la sortie du réseau après la convergence de l'étape d'optimisation.

Nous notons que l'intérêt de notre approche est que les modèles utilisés dans la phase d'apprentissage et d'optimisation sont identiques. Les deux étapes sont donc unifiées du point de vue du modèle DNN. Par la suite, ce modèle est désigné cDIP-LO (pour cDIP appris et optimisé) lorsque le modèle est initialisé avec des poids appris et cDIP-O lorsque l'initialisation des poids DNN est seulement aléatoire.

5.4 Résultats de simulation et discussion

La configuration du microscope utilisée pour évaluer le modèle de reconstruction cDIP-LO est un microscope équipé d'un objectif de grandissement $4\times$ et de $NA = 0.08$. Le modèle est testé uniquement par des simulations pour évaluer ses performances ultimes. La matrice de LED \mathcal{M} composée de 130 LED à $\lambda = 525$

nm introduite dans la Figure 4.3 est placée à une distance $d = 55$ mm en dessous de l'échantillon. Les différentes régions du spectre de l'échantillon qui sont ainsi sondées correspondent aux nombreuses régions indiquées dans la Figure 4.3 b,c. Une telle configuration nous permet d'atteindre un facteur de super-résolution théorique $\gamma = 5$. Le pas de pixel de la caméra simulée est de $3.45\mu\text{m}$.

Les poids d'initialisation du modèle cDIP-LO sont obtenus en utilisant l'ensemble de données d'apprentissage LDS (cf. Figure 5.1) après l'entraînement du modèle cDIP-L. LDS est construit à partir du Défi de Reconnaissance Visuelle à Grande Échelle ImageNet 2019 (ILSVRC'19) et contient 10 000 masques complexes O_j (voir Équation (1)). Chacun d'eux représente un échantillon numérique utilisé pour simuler des images LR FPM. Plus précisément, O_j est obtenu à partir d'images prises dans la base de données ImageNet. Pour chaque j (j variant de 1 à 10 000), deux images de 384×384 pixels (respectivement I_j et Φ_j) sont extraites aléatoirement du catalogue ImageNet. O_j est alors formé avec $\sqrt{I_j}e^{j\Phi_j}$.

Pour chaque O_j , une pile de 130 images LR simulant les images qui seraient acquises par la caméra sur l'échantillon O_j est calculée en utilisant un modèle direct pour les différentes LED de la matrice d'éclairage $\mathcal{M} = \mathcal{M}_A \cup \mathcal{M}_B$. Le facteur de chevauchement Γ est d'environ 60%. Il est important de souligner ici que le LDS ainsi construit nous permet de simuler une grande variété d'échantillons qui seraient difficilement accessibles aux mesures FPM expérimentales. En particulier, ImageNet contient un nombre considérable de photos avec de larges variations dans leurs caractéristiques spatiales (contraste, fréquences spatiales, etc.).

De plus, les fonctions I et Φ sont non corrélées par construction. Ce n'est généralement pas le cas dans les vraies expériences puisque l'absorption des échantillons biologiques et l'épaisseur optique sont généralement étroitement liées. Un tel LDS pourrait donc être bénéfique pour une bonne généralisation du modèle. En raison du pas de la caméra utilisé dans les simulations (de $3.45\mu\text{m}$) et des caractéristiques de l'objectif ($4 \times \text{NA} = 0.08$), les images LR produites sont échantillonnées par la caméra avec un facteur de sur échantillonnage important par rapport aux critères de Shannon. Dans notre cas, chaque image LR est de 128×128 pixels.

En complément du LDS, un ensemble de données de test TDS est également constitué. Il se compose de 1000 masques complexes qui sont obtenus en suivant la même construction que celle utilisée pour le LDS. Les images tirées d'ImageNet sont cependant différentes de celles utilisées dans LDS. Pour chaque masque, deux piles d'images LR sont également calculées en utilisant le modèle prédictif FPM. La première pile est liée aux LED de \mathcal{M}_A , et la seconde aux LED de \mathcal{M} .

Le TDS est exploité pour comparer les différents modèles de reconstruction (ePIE, cDIP-L, et cDIP-LO). Un résultat typique est présenté à la Figure 5.2, où les images (a) et (g) correspondent à (I et Φ) HR-images qui sont reconstruites avec l'algorithme ePIE en utilisant la pile d'images LR relatives aux LED de \mathcal{M} .

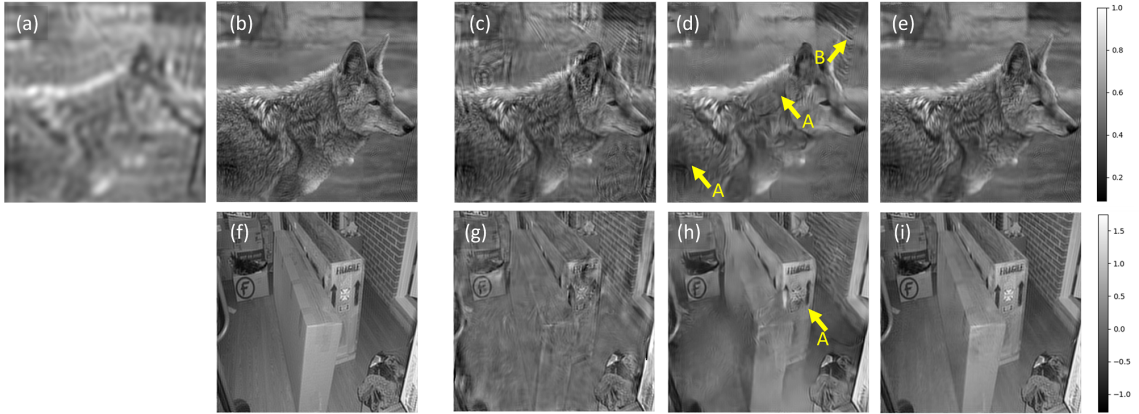


FIGURE 5.2 – Comparaison des résultats de reconstruction obtenus avec des éclairages et un chevauchement d'environ $\sim 10\%$. Les images (a–e) correspondent aux images d'intensité, tandis que les images (f–i) correspondent aux images de phase. (a) Image brute acquise avec la LED centrale, (b) image de référence, (c) reconstruction ePIE (en utilisant les LED \mathcal{M}_A), (d) reconstruction cDIP-L après entraînement (en utilisant les LED \mathcal{M}_A), (e) reconstruction cDIP-LO après entraînement et optimisation (en utilisant les LED \mathcal{M}_A). (f) correspond à la phase de référence, et (g–i) à la phase reconstruite avec e-PIE, cDIP-L et cDIP-LO, respectivement. Les flèches indiquent différents types d'artefacts (tels que le crosstalk ou la dégradation de la résolution). Les images de référence (b,f) sont obtenues à partir de la reconstruction ePIE en utilisant toutes les piles d'images (LEDs \mathcal{M}).

Ces deux images constituent les images HR d'intensité et de phase ciblées (ou de référence). L'image (b) correspond à l'image brute liée à la LED centrale de la matrice. Là, on peut apprécier la résolution native de l'objectif.

5.4.1 Évaluation visuelle

Les autres images de la Figure 5.2 sont les images qui sont reconstruites en utilisant la pile d'images LR relatives aux LED de \mathcal{M}_A uniquement. En particulier, les images (c) et (h) correspondent à la reconstruction ePIE, les images (d) et (i) à la reconstruction cDIP-L, et (e) et (j) à cDIP-LO (intensité et phase). Les images (c) et (h) illustrent les artefacts typiques qui peuvent être obtenus avec la reconstruction ePIE dans des conditions de faible recouvrement. Un mélange (crosstalk) important entre l'intensité et la phase est observé. De plus, le contraste de phase est fortement dégradé. Il est clair que la qualité de ces images est trop basse pour envisager leur exploitation pour des applications réelles (avec un faible

γ). En revanche, les images I et Φ obtenues avec cDIP-L sont de meilleure qualité perceptuelle à la fois pour l'intensité et la phase. Cependant, des artefacts sont encore perceptibles. Certains d'entre eux sont indiqués par des flèches dans la Figure 5.2. Par exemple, la flèche A correspond à des dégradations dans les hautes fréquences qui sont pour la plupart perdues. La flèche B correspond à une séparation incomplète entre l'image de phase et l'image d'intensité (crosstalk). Bien que cDIP-L nous permet de récupérer des informations de phase avec des avantages par rapport à ePIE, la résolution finale des images reconstruites est discutable : la résolution de l'image d'intensité est presque comparable à l'image d'intensité brute relative à l'éclairage LED central. Ceci est conforme aux résultats publiés de l'état de l'art [61], qui reposent également sur des systèmes basés sur un apprentissage statistique seul. Ceci semble être inhérent aux modèles qui sont seulement estimés statistiquement. Fait intéressant, la situation est très différente avec le modèle cDIP-LO. Les différences entre les images reconstruites et les images de référence sont trop faibles pour être visuellement appréciables.

5.4.2 Évaluation quantitative

Cette amélioration apportée par la reconstruction neuronale cDIP-LO est davantage confirmée en évaluant ses performances quantitativement à l'aide de la fonction d'erreur \mathcal{L} , calculée sur l'ensemble du jeu de données TDS. Les métriques traditionnelles SSIM et PSNR sont évaluées pour les images reconstruites par rapport aux images de référence pour l'intensité et la phase. Les résultats sont indiqués dans le Tableau 1 où la moyenne et l'écart-type de \mathcal{L} , SSIM et PSNR sont calculés sur l'ensemble du TDS. La colonne ePIE désigne la reconstruction ePIE en utilisant les LED de \mathcal{M}_A .

De manière non équivoque, l'algorithme cDIP-LO conduit à une meilleure reconstruction d'image. En particulier, la perte évaluée pour ePIE et cDIP-L est presque comparable, alors qu'elle est 10 fois plus faible pour cDIP-LO. La métrique de similarité SSIM est proche de 1 pour cDIP-LO (comme souhaité) et bien plus élevée que cDIP-L et ePIE SSIM. Le PSNR révèle également une importante augmentation pour cDIP-LO par rapport à cDIP-L et ePIE (au-dessus de 10 dB).

Il est donc clair que le modèle cDIP-LO améliore sensiblement les performances de reconstruction. Nous attribuons une telle amélioration au fait que l'initialisation des poids de cDIP-LO bénéficie d'informations relatives à toutes les LED de \mathcal{M} pendant l'étape d'apprentissage et au fait que les calculs de reconstruction cDIP-L sont ensuite résolus en utilisant les équations du modèle direct. On peut également observer dans le Tableau 5.1 que les écarts-types des métriques relatifs à cDIP-LO sont faibles, indiquant la capacité de cette approche DNN à généraliser. Nous rappelons que TDS a été constitué avec divers masques complexes artificiels sans corrélation entre l'intensité et la phase.

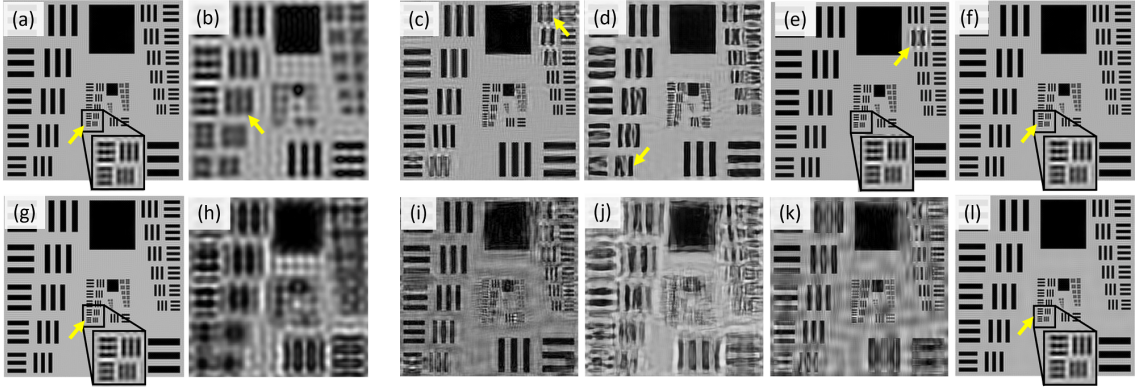


FIGURE 5.3 – comparaison des reconstructions de la mire de résolution d’intensité USAF 1951 obtenues avec l’éclairage (recouvrement $\sim 10\%$). Pour les images (a–f), la mire de résolution est placée sur le plan focal du microscope. (a) Référence, (b) image brute acquise avec l’éclairage de la LED centrale, (c) ePIE, (d) cDIP-L, (e) cDIP-O, et (f) reconstructions cDIP-LO. Pour les images (g,h), la mire de résolution est placée à $8\mu m$ au-dessus du plan focal du microscope. (g) Référence (identique à l’image (a)), (h) LED centrale, (i) ePIE, (j) cDIP-L, (k) cDIP-O, et (l) reconstructions cDIP-LO. La référence est calculée en utilisant la reconstruction ePIE avec l’éclairage des LEDs \mathcal{M} .

TABLE 5.1 – Évaluation des métriques \mathcal{L} , SSIM et PSNR sur l’ensemble de données TDS. Les moyennes et écart-types des métriques sont présentés. Les reconstructions sont obtenues à partir de la pile d’images LR générées avec les LED de \mathcal{M}_A .

	ePIE	cDIP-L	cDIP-LO
$\mathcal{L} \times 10^{-4}$	$\bar{\mu} = 4.43$ $\sigma = 2.85$	$\bar{\mu} = 1.47$ $\sigma = 0.724$	$\bar{\mu} = 0.244$ $\sigma = 0.0941$
$SSIM_{\text{Intensity}}$	$\bar{\mu} = 0.569$ $\sigma = 0.143$	$\bar{\mu} = 0.671$ $\sigma = 0.105$	$\bar{\mu} = 0.919$ $\sigma = 0.029$
$SSIM_{\text{Phase}}$	$\bar{\mu} = 0.19$ $\sigma = 0.109$	$\bar{\mu} = 0.491$ $\sigma = 0.167$	$\bar{\mu} = 0.915$ $\sigma = 0.06$
$PSNR_{\text{Intensity}}$ (dB)	$\bar{\mu} = 17.28$ $\sigma = 4.38$	$\bar{\mu} = 25.03$ $\sigma = 2.68$	$\bar{\mu} = 34.8$ $\sigma = 2.71$
$PSNR_{\text{Phase}}$ (dB)	$\bar{\mu} = 18.32$ $\sigma = 3.53$	$\bar{\mu} = 25.05$ $\sigma = 1.86$	$\bar{\mu} = 39.14$ $\sigma = 2.81$

Les métriques \mathcal{L} , SSIM et PSNR ne fournissent pas d’information concernant la résolution finale de l’image produite. Pour cette raison, nous avons entrepris des

simulations complémentaires de reconstruction de la mire de résolution d'intensité USAF 1951 en utilisant les LED \mathcal{M} . Des résultats sont présentés à la Figure 5.3. La première ligne d'images correspond à la situation où l'objet USAF1951 est placé sur le plan focal du microscope. Pour la deuxième ligne, l'objet est hors foyer et placé à $80\mu m$ au-dessus. Les figures 5a,b correspondent respectivement à l'image de référence et à l'image brute relative à la LED centrale. Le dernier élément résolu est indiqué par une flèche. Les reconstructions obtenues avec ePIE, cDIP-L et cDIP-O (c'est-à-dire sans initialisation du modèle cDIP) sont illustrées sur les figures 5.3c, d et e, respectivement. Comme on peut le voir, le dernier élément résolu ne change pas beaucoup. Cela signifie qu'aucune amélioration significative de la résolution (comparée à la résolution de l'objectif natif) n'est observable avec la reconstruction, bien que la qualité perceptuelle globale des images reconstruites soit améliorée. En revanche, la reconstruction cDIP-LO (voir figure 5f) révèle un changement significatif dans la position du dernier élément résolu. Elle nous permet d'évaluer que le facteur de super-résolution atteint est $\gamma = 5$, comme on peut s'y attendre théoriquement. Notez également qu'aucune différence notable par rapport à l'image de référence n'est observable. Bien que la qualité perceptuelle de l'objet USAF1951 soit légèrement améliorée dès que la reconstruction est effectuée (comparer les figures 5c-f à l'image 5b), la situation est très différente lorsque l'objet est hors foyer (seconde ligne d'images de la figure 5.3). Les images des figures 5.3i-l correspondent à des images reconstruites avec une correction de mise au point numérique (comme expliqué dans la section 2.3.3). On peut noter que la qualité des images obtenue avec ePIE, cDIP-L, et cDIP-LO est dégradée. En revanche, cDIP-LO démontre sa capacité à reconstruire correctement la mire de résolution sans aucune dégradation. Cet algorithme de reconstruction est donc pleinement fonctionnel et capable de gérer correctement l'intégration de la correction de mise au point dans le modèle direct, ce qui est très souhaitable pour des situations pratiques.

5.5 Conclusion

Dans ce travail, nous avons introduit un algorithme basé sur le DNN, à savoir le cDIP-LO DNN, pour reconstruire des images acquises avec la FPM sous un nombre réduit d'illuminations de l'échantillon (c'est-à-dire $\Gamma < 60\%$). Les propriétés de reconstruction de cette approche ont été comparées aux algorithmes de pointe ePIE et au modèle DNN appris statistiquement (cDIP-L). Les résultats simulés pour une configuration FPM avec un facteur Γ très faible ($\Gamma \sim 10\%$) démontrent les avantages conceptuels du cDIP-LO en termes de qualité d'image et de résolution.

Les poids du modèle cDIP-LO sont initialisés après une étape d'apprentissage statistique cDIP-L; dans cette étape, le modèle direct de la FPM a été introduit comme en remplacement de toute exploitation d'une solution de référence

reconstruite par une autre approche. Cela permet l'utilisation d'un modèle unique et cohérent (ou architecture DNN unifiée) dans l'étape d'apprentissage et dans l'étape d'optimisation.

Nous notons que l'étape d'apprentissage est obligatoire cDIP-L ; la reconstruction avec cDIP-LO sans initialisation du modèle conduit à des images inexploitable. Dans les travaux futurs, le test du cDIP-LO sur des données expérimentales sera entrepris. D'un point de vue pratique, nous pensons que les données simulées devraient suffire lors de l'étape d'apprentissage sans avoir besoin de données expérimentales. Ce point sera spécifiquement étudié par rapport à la tolérance au bruit expérimental ou à une éventuelle imprécision de position des LED. L'idée de coupler une étape d'apprentissage statistique et une optimisation DNN informée par la physique est générale.

Chapitre 6

Approche cDIP pour expérimentation avec faible taux de recouvrement et amélioration de focus

6.1 Introduction

Dans ce chapitre nous étudions la question de reconstruction des images FPM dans le cas d'un faible recouvrement du spectre pour une configuration de microscope adaptée à l'obtention d'images de très haute résolution. La résolution visée est celle utilisée pour l'observation des frottis sanguins en hématologie. En microscopie traditionnelle, elle est atteinte avec un objectif immersion à forte ouverture numérique et fort grossissement (magnification), typiquement un objectif 100xNA1.3

En microscopie FPM, nous utilisons un objectif 20xNA0.75 qui permet d'atteindre une ouverture synthétique théorique égale 1.75 . D'un point de vue expérimental, nous considérons une situation pour laquelle le facteur de super-résolution atteint est limité à un facteur 2, c-a-d $NA_{synth} = 1.5$. Cela permet d'éviter tout recours à des acquisitions d'images en champ sombre et donc de limiter le temps des mesures. La matrice de LEDs utilisée contient 13 LEDs. Le dispositif de diagnostic s'en trouve ainsi grandement simplifié. Les échantillons utilisés correspondent aux frottis sanguins recueillis dans le cadre du projet TAMIS.

Bien que la configuration retenue nécessite déjà un nombre relativement restreint de LEDs, l'objectif de ce chapitre est de réduire encore davantage ce nombre. En effet, même pour ce nombre restreint de LEDs, la cadence actuelle de numérisation d'un frottis sanguin complet est encore élevée. Pour notre système, il est

typiquement de 4 lames/heure. Cette cadence est encore trop faible pour être raisonnablement utilisée en condition clinique (typiquement, la cadence souhaitée est de l'ordre de 15 lames / heure). Pour atteindre cet objectif, nous nous inspirons de l'approche détaillée dans le chapitre 5, que nous mettons en œuvre sur des données expérimentales. De plus, la réduction du nombre de LED doit s'effectuer sans compromettre la qualité de l'imagerie. L'image obtenue doit aussi être améliorée et optimisée, tel qu'énoncé dans le chapitre 3.

je présente des résultats encore à un stade préliminaire, car récents. Pour ce faire je mets en œuvre les modèles neuronaux présentés au Chapitre 5 (le cDIP-LO pour la reconstruction, le réseau CNN pour la refocalisation des images) pour un nombre de LEDs correspondant à un facteur de recouvrement Γ de 20%. Les résultats présentés permettent d'identifier des limitations spécifiques à la résolution ciblée (présence de « trous » non négligeables dans le spectre couvert par l'éclairage utilisé). Pour y remédier j'introduis une alternative de reconstruction basée sur une reconstruction par apprentissage seul qui n'était pas fonctionnelle pour des facteurs de super-résolution plus élevés (à facteur de recouvrement similaire $\sim 20\%$). Plus précisément, l'approche mise en place consiste en un unique réseau CNN en charge de la reconstruction et refocalisation, et ce, de façon simultanée. Cette dernière est rendue viable pour les raisons suivantes : 1) elle profite du facteur de super-résolution limité visé. Il s'agit d'une situation pour laquelle le problème inverse est malgré tout plus facile à résoudre bien qu'il existe des trous plus importants dans le spectre imagé de l'échantillon. 2) La connaissance du modèle direct associé à la formation des images brutes de FPM dans la fonction de coût pour un réseau appris (cDIP-L introduit en chapitre 4-5) est mise à profit pour contraindre l'apprentissage du modèle statistique. 3) Une base de données spécifique est construite et exploitée en apprentissage. Cette dernière repose sur un mixte entre des images brutes FPM acquises pour une matrice de LEDs avec un facteur de recouvrement de 20% et celles acquises pour les mêmes objets avec un facteur de recouvrement de 60%. L'idée sous-jacente est d'apprendre au réseau à compléter l'information manquante dans le spectre. Par ailleurs, la base intègre différentes images acquises pour le même échantillon avec des conditions de focus variées pour apprendre ce qu'est une image nette. Bien que ces résultats soient encore préliminaires et nécessitent encore une meilleure compréhension du rôle des différents mécanismes indiqués en points 1 à 3, les résultats obtenus sont d'ores et déjà intéressants. Ils doivent permettre de répondre à la problématique applicative visée par une validation quantitative des performances de détection (travail qui reste à faire au plan expérimental).

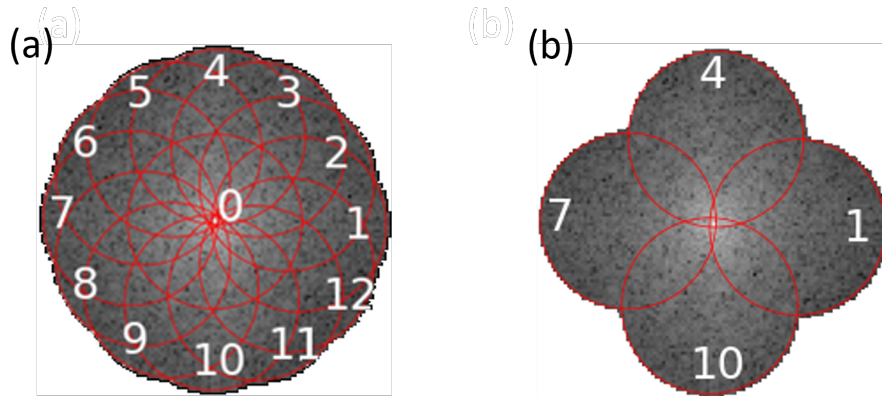


FIGURE 6.1 – Régions spectrales couvertes dans l’espace de Fourier ;(a) la matrice de LEDs \mathcal{M} , présentant un recouvrement spectral de $\Gamma \approx 66\%$, et (b) la matrice de LEDs \mathcal{M}_A , avec un recouvrement spectral de $\Gamma \approx 20\%$.

6.2 Configuration du microscope

La source lumineuse est composée d’un anneau de 12 LEDs et d’une LED centrale additionnelle. Cette configuration est réalisée en utilisant un anneau NeoPixel RGB 5050 d’un rayon de $r = 15mm$, équipé de 12 LEDs fournies par Adafruit, ainsi qu’une LED centrale NeoPixel additionnelle.

Nous avons ajusté la distance entre les diodes et l’échantillon afin d’obtenir une ouverture numérique synthétique $NA_{synth} = 1.5$, correspondant à une distance $d \sim 13mm$. La caméra employée, un modèle IDS UI-3200SE, est dotée d’un capteur haute résolution de 4104×3006 pixels, avec une taille de pixel de $3.45\mu m$.

La Figure 6.1 (a) illustre la disposition spectrale associée à la matrice de LEDs \mathcal{M} , composée d’un anneau de 12 LEDs et d’une LED centrale. Cet agencement satisfait les conditions de recouvrement, $\Gamma \approx 66\%$ facilitant ainsi la convergence des algorithmes de reconstruction classique de la FPM et permettant d’atteindre la résolution cible.

De ce fait, le choix de réduire le nombre de LED dans la matrice \mathcal{M} soulève la question du nombre optimal à conserver : 6, 4 ou 3, en éliminant soit 1 LED sur 2, 1 LED sur 3 ou 1 LED sur 4 des 12 LED circulaires initiales. Après analyse, nous avons déterminé que la conservation de 4 LEDs serait la plus adéquate. Cette configuration permet une réduction significative du temps de production, avec une diminution supérieure à 3, tout en minimisant la perte d’information par rapport à la configuration initiale.

Cette configuration permet de couvrir $\sim 87\%$ du spectre de l’échantillon accessible à la matrice complète (avec ses 13 LEDs). En choisissant de n’utiliser que 3 LED, nous nous serions trouvés dans une situation avec une perte d’information

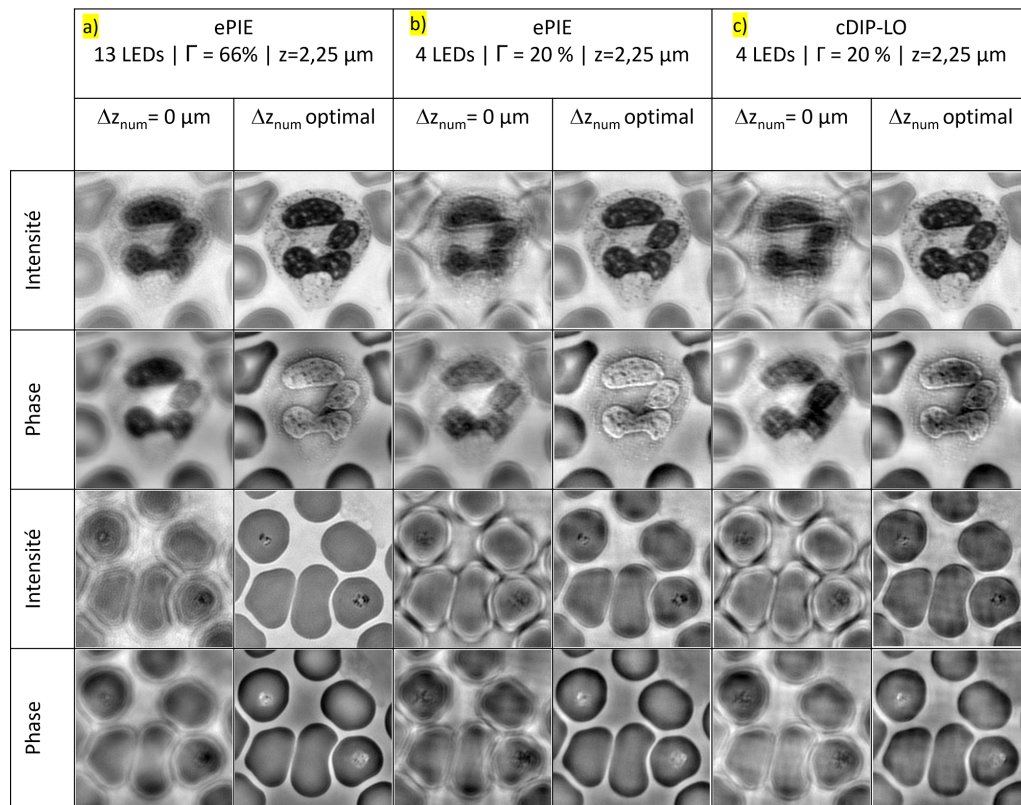


FIGURE 6.2 – Illustration des images reconstruites d’un globule blanc et de globules rouges parasités à $z = 2.25 \mu\text{m}$ avec et sans ajustements de la mise au point via la correction du front d’onde du Δz_{num} . a) Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 66\%$ en utilisant l’algorithme ePIE sont comparées à celles obtenues avec un $\Gamma = 20\%$, également en utilisant les algorithmes b) ePIE et c) cDIP-LO.

plus importante. Cette situation soulève des questions sur la capacité des modèles de reconstruction à compléter ces données, en particulier dans le contexte sensible du diagnostic médical.

La configuration retenue \mathcal{M}_A , est donc constituée de 4 LEDs. Il s’agit d’un sous ensemble de \mathcal{M} , avec un chevauchement spectral de $\Gamma \approx 20\%$ comme l’illustre la Figure 6.1 (b).

La figure 6.2 présente les résultats de reconstruction obtenus sur un globule blanc avec l’algorithme e-PIE et cDIP-LO pour la matrice de LED \mathcal{M}_A . Il est intéressant de remarquer que les algorithmes ePIE et cDIP-LO permettent une reconstruction nativement intéressante à condition de connaître Δz_{num} . Pour ces reconstructions, le paramètre Δz_{num} est fixé après un z -stacking numérique. Cela montre que la difficulté de reconstruction est nativement plus simple que dans le

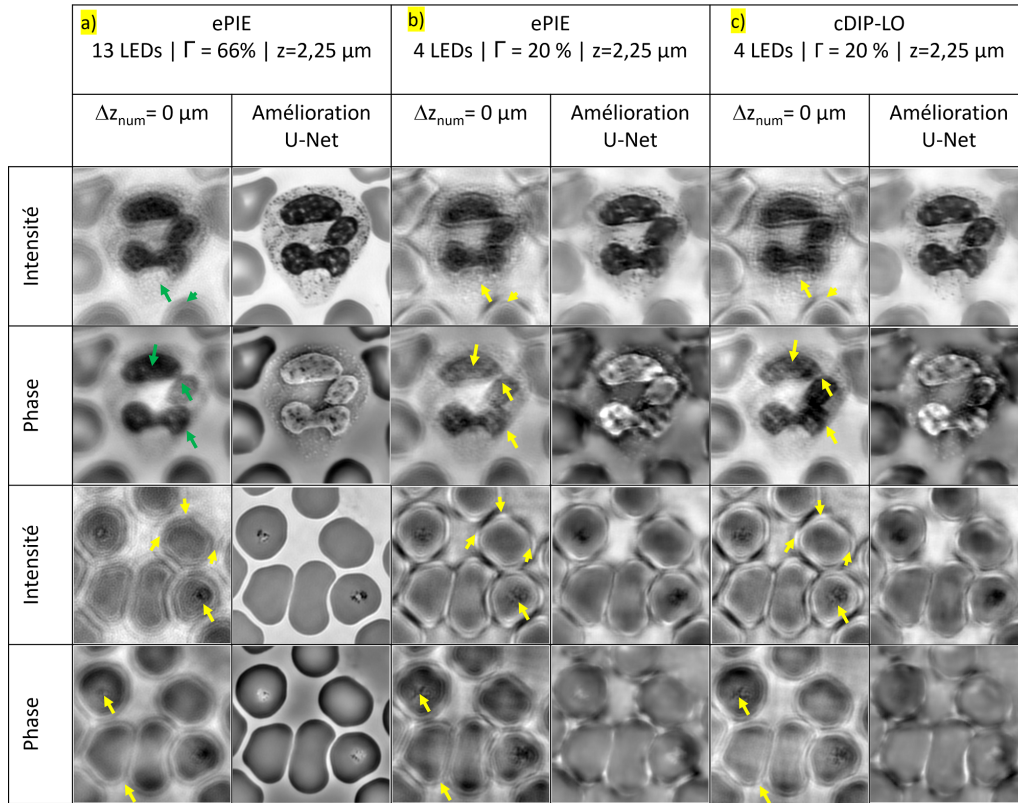


FIGURE 6.3 – Illustration des images d'un globule blanc et de globules rouges parasités reconstruites à $z = 2.25 \mu\text{m}$ avec et sans amélioration d'images par modèle U-Net FPM du Chapitre 3. a) Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 66\%$ en utilisant l'algorithme ePIE. Les images obtenues sous des conditions de recouvrement spectral de $\Gamma = 20\%$, b) en utilisant l'algorithme ePIE et c) en utilisant le modèle cDIP-LO.

cas traité dans le chapitre 5 puisque de tels résultats ne peuvent notamment pas être obtenus pour un facteur de super-résolution égale à 6 (cf. résultats présentés en figure 5.2 et 5.3).

Bien que les images reconstruites pour $\Delta z_{\text{num}} = \Delta z_{\text{optimal}}$ paraissent de bonne qualité avec cette configuration, on peut tout de même noter une légère différence de qualité comparées aux images obtenues avec 13 LED $\Gamma = 66\%$."

Ces observations montrent que les informations nécessaires pour une mise au point correcte sont bien présentes dans les images. Il s'agit de trouver une stratégie pour les exploiter de manière efficace.

Comme déjà évoqué, l'estimation du paramètre Δz_{num} est contraignante et cela nous avait conduit en chapitre 3 à développer un module de compensation de focus par U-Net. Nous le mettons en œuvre pour les mêmes conditions expérimentales

que celles utilisées en figure 6.2. Nous pouvons observer que les résultats issus de la compensation par le U-Net tels que présentés en figure 6.3, ne sont pas exploitables.

En effet, les images obtenues sont très détériorées. À la sortie du U-net, la résolution est dégradée. Les détails fins sont perdus. L'image présente de plus un tramage additionnel gênant. Par ailleurs, les contrastes de l'image de phase sont assez éloignés de la référence (pour l'algorithme ePIE et cDIP-LO). Cette situation est probablement attribuable à la déformation significative observée dans les images reconstruites hors focus. En effet, pour $\Delta z_{\text{num}} = 0$, les images reconstruites présentent des structures anormales par comparaison aux mêmes images reconstruites avec les 13 leds.

Les anomalies sont indiquées par les flèches jaunes. Il peut notamment être observé que les anomalies sont plus prononcées au niveau des bords des cellules GB et GR, là où les variations de contraste sont importants.

Pour contourner cette difficulté, et profitant du fait que la reconstruction peut être obtenue de façon adéquate par ePIE ou cDIP-LO à condition de connaître Δz_{num} , nous cherchons à nous affranchir de cette méconnaissance dans un réseau d'apprentissage. Pour cela, le réseau U-Net mis en place exploite un apprentissage statistique de type cDIP-L. Ce réseau doit également prendre en charge la compensation du focus par apprentissage en s'inspirant des démarches mise en place au Chapitre 3.

Pour relever ce défi, nous avons mis en œuvre une méthodologie originale basée sur l'apprentissage statistique en utilisant un modèle cDIP. Cette stratégie vise à optimiser pleinement l'exploitation de la pile d'images avec un taux de recouvrement $\Gamma = 22\%$ en permettant une reconstruction au bon focus des images de phase et d'intensité en une seule étape. Cette optimisation est réalisée grâce à des références contenant toutes les informations de la matrice \mathcal{M} , ainsi qu'à l'utilisation d'images de référence au bon focus.

6.3 Modèle de reconstruction et refocalisation par apprentissage

J'utilise l'approche cDIP telle que décrite dans le Chapitre 4. Ce modèle spécifique permet une reconstruction en employant une fonction de coût qui incorpore le modèle direct de la microscopie ptychographique de Fourier (FPM). De plus, cDIP présente la flexibilité d'intégrer un apprentissage statistique basé sur cette même fonction de coût, comme nous l'avons montré dans le chapitre 5.

Nous utilisons ici le modèle cDIP-L, illustré en Figure 6.4. La différence par rapport au modèle du chapitre 5 réside dans le fait que la base d'apprentissage est composée d'images à différents niveaux de focus.

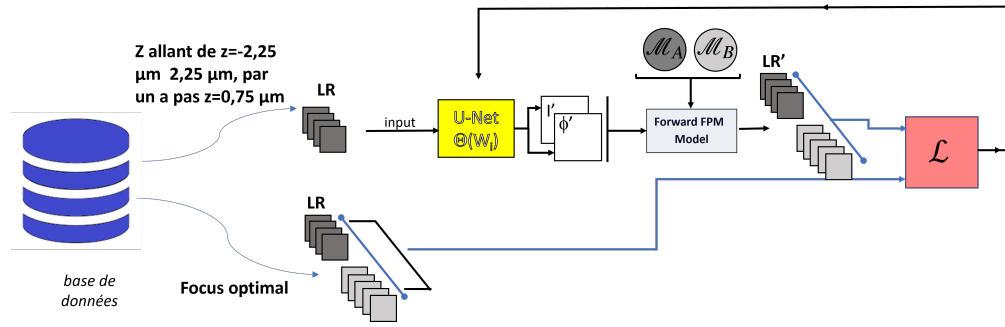


FIGURE 6.4 – Schéma d'apprentissage du modèle cDIP-L

Le modèle prend en entrée des images à basse résolution (LR), éclairées par quatre LED et sélectionnées à partir de la matrice \mathcal{M}_A . Ces images peuvent être prises à différentes distances focales z , allant de $z = -2.25 \mu m$ à $z = 2.25 \mu m$, avec un incrément de $0.75 \mu m$. En sortie, le modèle U-Net 4.2 génère une estimation de l'intensité I' et de la phase Φ' , tout en doublant le facteur de résolution.

Afin de calculer l'erreur durant la phase d'apprentissage, nous employons le modèle direct de la Microscopie par FPM. Ce modèle utilise les images I' et Φ' et s'appuie sur la configuration LED définie par $\mathcal{M} = \mathcal{M}_A + \mathcal{M}_B$, permettant ainsi de générer 13 images à basse résolution, notées LR'. Ces images LR' sont ensuite comparées à un ensemble de 13 images LR et centrées à $z = 0 \mu m$ et Δ_z optimale. Cet ensemble a été spécifiquement constitué pour servir de référence, permettant ainsi d'ajuster les poids du réseau neuronal.

L'objectif est qu'à l'issue de la phase d'apprentissage, le modèle soit capable de fournir des images de phase Φ' et d'intensité I' avec le focus approprié et à la résolution attendue.

Pour la réalisation de ces travaux, nous avons constitué une base de données conforme à la méthodologie décrite dans le Chapitre 3.

Pour construire notre pile d'images de référence, nous sélectionnons à la fois des images de phase (Φ) et des images d'intensité (I), sélectionnées de manière à ce qu'elles soient au focus optimal, $\Delta z_{\text{num}} = \Delta z_{\text{opt}}$. Sur ces paires d'images I et Φ , nous appliquons le modèle direct de la FPM en utilisant une configuration de LED, désignée par \mathcal{M} . Cette démarche produit une pile FPM composée de 13 images à basse résolution (LR), avec un taux de recouvrement spectral de $\Gamma = 66\%$, et avec une mise au point optimale.

Par la suite, pour chaque pile d'images de référence, nous générons sept piles d'images à basse résolution en utilisant la configuration de LED \mathcal{M}_A , avec un taux de recouvrement d'information de $\Gamma = 20\%$. Pour ces 7 piles, la distance focale (z) varie dans un intervalle allant de $z = -2,25 \mu m$ à $z = 2,25 \mu m$, avec des incréments de $z = 0,75 \mu m$.

Par conséquent, pour toutes les mises au point et pour chaque pile composée de 4 images, la sortie des modèles U-Net et du modèle direct permet d’obtenir une pile d’images de 13 LEDs LR’, que nous comparons aux piles de 13 images de référence à la mise au point idéale, avec un taux de rétention d’information de $\Gamma = 66\%$.

La base de données est composée d’un mélange de données expérimentales et de données de simulations, dans le but d’obtenir un apprentissage statistique diversifié et une meilleure généralisation sur différents types d’objets.

La base de données complète pour l’apprentissage se compose de 4 000 objets, chacun avec des couples d’intensité et de phase (I, Φ) . Chaque couple est représenté à sept niveaux de mise au point différents, ce qui permet d’avoir un nombre total de piles d’images LR de 28 000. Chaque image de la pile a une dimension de 256x256 pixels et est associée à son image de référence, correctement mise au point.

6.4 Évaluation visuelle du modèle

Nous avons évalué les performances du modèle cDIP après l’apprentissage en utilisant un ensemble de données composé de 1000 objets simulés et 1000 objets issus d’expérimentations, chacun étant évalué à différents niveaux de mise au point. Dans cette section, nous présentons les résultats obtenus sur les données expérimentales, qui constituent notre principal centre d’intérêt.

Les résultats obtenus pour le problème le plus simple de reconstruction lorsque $z = 0\mu m$ sont satisfaisants. Ils sont présentés en figure 6.5. Nous observons qu’ici, avec un facteur de super-résolution de x2, le modèle basé exclusivement sur l’apprentissage, cDIP-L, permet d’obtenir des images en haute résolution. Le piqué de l’image et les détails relatifs aux structures de l’objet semblent très proches de ceux correspondant à la référence (pour lequel le facteur de super-résolution a été mesuré par ailleurs expérimentalement sur une mire de résolution, et correspond à un facteur x2). Ces images sont comparables en résolution à celles de la référence ePIE avec une configuration à un taux de recouvrement spectral de $\Gamma = 66\%$. Il est également notable que le modèle cDIP-L est efficace pour mettre en évidence les compartiments cellulaires d’intérêt. Ceci est illustré par la différence dans l’image de phase à $\Delta z_{num} = 0\mu m$ entre ePIE et le modèle cDIP-L. Avec le cDIP-L, nous observons davantage de détails au niveau du noyau et des granules du globule blanc.

Nous avons également analysé les résultats du modèle cDIP-L pour $z = 2.25\mu m$, comme le montre la Figure 6.6. Cette analyse inclut des images de globules blancs ainsi que des images de globules rouges infectées du paludisme de type *P. falciparum*.

Nous constatons que le modèle cDIP-L permet une reconstruction du leucocyte

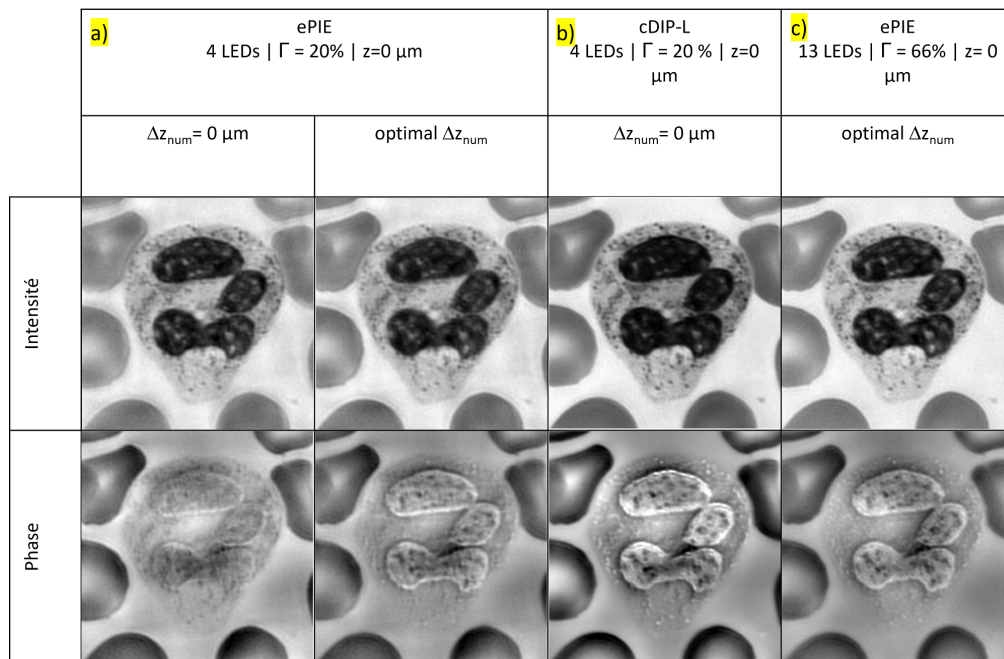


FIGURE 6.5 – Analyse comparative de la reconstruction d'un globule blanc polynucléaire neutrophile à $z = 0 \mu\text{m}$. a) Résultats ePIE dans un cas de faible recouvrement $\Gamma = 20\%$ avec et sans correction de Δz_{num} . b) Résultats du modèle cDIP-L avec $\Gamma = 20\%$. c) Résultats ePIE dans un cas de fort recouvrement $\Gamma = 66\%$ avec Δz_{num} optimal.

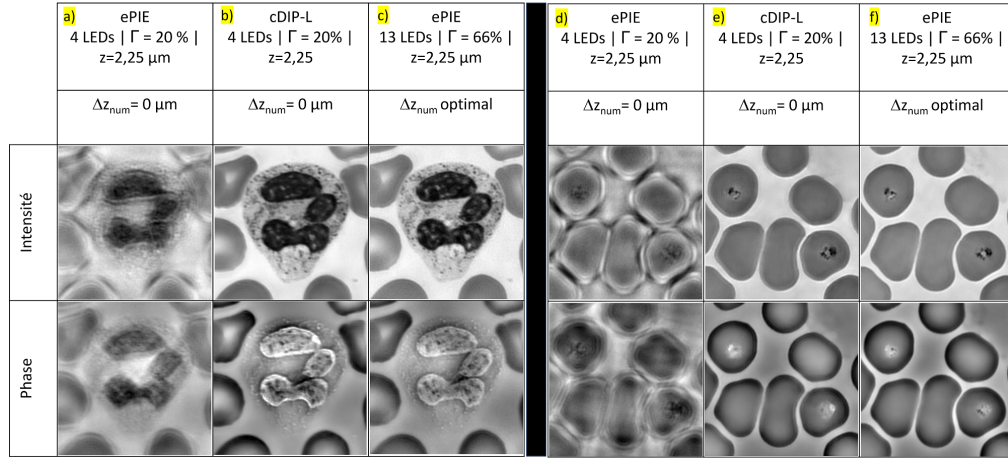


FIGURE 6.6 – Analyse comparative de la reconstruction d'un globule blanc polynucléaire neutrophile et des globules rouges parasités à $z = 2.25 \mu m$. a) et c) Résultats ePIE dans un cas de faible recouvrement $\Gamma = 20\%$. b) et d) Résultats du modèle cDIP-L avec $\Gamma = 20\%$. c) Résultats ePIE dans un pour $\Gamma = 66\%$ et Δz_{num} optimal.

à $z = 2.25 \mu m$ qui est équivalente, en termes de résolution et de qualité d'image, à celles obtenues via le modèle ePIE avec $\Gamma = 66\%$, tout en optimisant le réglage du focus. En revanche, les images générées par ePIE en utilisant la configuration \mathcal{M}_A et une compensation numérique suite à une sélection manuelle ne parviennent pas à atteindre cette qualité. On constate une perte de détails et la présence de quelques légers artefacts dans le plasma.

Les résultats observés sur les globules rouges parasités sont tout aussi satisfaisants que ceux obtenus sur les leucocytes. Notamment, le modèle cDIP-L fournit des informations discriminantes et complémentaires en phase et intensité à $z = 2.25 \mu m$, malgré un facteur de recouvrement de $\Gamma = 20\%$. Ceci est d'autant plus remarquable que le modèle n'a jamais été exposé à ce type d'objet biologique durant la phase d'apprentissage.

Nous avons également effectué des tests sur une mire de résolution (intensité) USAF1951. Les images correspondant à différentes mises au point sont présentées dans la Figure 6.7. À $z = 0 \mu m$, nous pouvons clairement distinguer le groupe 10,3 en utilisant la diode centrale. Un début de séparation du groupe 10,4 est également observable. En doublant la résolution, l'algorithme ePIE permet de séparer efficacement les informations du groupe 11,3, même lorsqu'il est limité à 4 LEDs mais cette information est perdue pour $z \neq 0 \mu m$

Avec le modèle cDIP-L, nous parvenons à reconstruire et refocaliser efficacement les images de la mire pour des valeurs de z allant de $\pm 1,5 \mu m$, atteignant ainsi la résolution cible de 11,3.

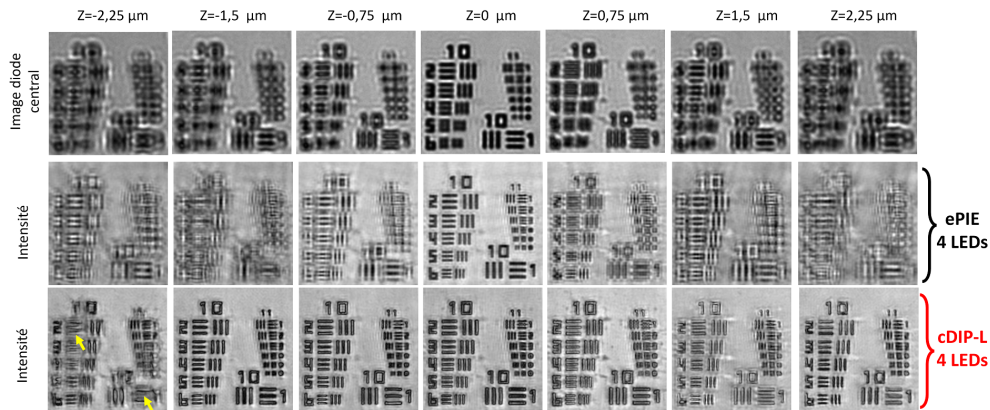


FIGURE 6.7 – Analyse comparative des images à différentes mises au point pour la diode centrale (basse résolution) et des images d’intensité reconstruites à partir de 4 LEDs avec $\Gamma = 20\%$ en utilisant les algorithmes ePIE et cDIP-L

Pour les valeurs de mise au point $z \neq 0\mu m$, le modèle cDIP-L permet d’accéder à des informations qui sont détériorées et perdues lors de la reconstruction effectuée par l’algorithme ePIE dans les conditions de $\Gamma = 20\%$.

En revanche, pour les valeurs $z = -2,25\mu m$ et $z = 2,25\mu m$, nous commençons à observer les limites de notre algorithme sur ce type d’objet. Des erreurs sont indiquées par des flèches jaunes dans les images.

Il est important de noter que la mire USAF1951 est un objet spécifique, essentiellement caractérisé par un motif d’intensité et sans phase associée. Cette particularité pourrait expliquer pourquoi les résultats présentent de légères différences par rapport à ceux illustrés dans la Figure 6.6, étant donné qu’elle ne coïncide pas pleinement avec les conditions d’apprentissage du modèle cDIP-L.

6.5 Conclusion

Dans ce chapitre, nous avons montré que la méthode de reconstruction cDIP-L permet une bonne reconstruction et une extension de la profondeur de champ directement à partir des images FPM, acquises en utilisant un éclairage à quatre LED, en une seule étape. Dans le cas d’un facteur de super-résolution $\times 2$ et dans un contexte expérimental, cette approche a significativement réduit le temps de production par un facteur de 3, en passant de 13 LED (12 annulaires et une centrale) à seulement 4 LED. Cela représente un avantage considérable, notamment pour les applications médicales.

Cette méthode, grâce à son utilisation du GPU, permet une reconstruction nettement plus rapide que les approches ePIE sur CPU. Avec le modèle cDIP-L, nous

pouvons reconstruire un champ complet en à peine 5 secondes, un temps comparable à celui de l'approche développée par Jiang et al [43]. En outre, l'avantage de l'approche cDIP-L est qu'elle élimine le besoin de tout traitement supplémentaire, qui pourrait autrement augmenter le temps de traitement.

L'approche mise en œuvre a permis d'étendre la profondeur de champ de $\sim \pm 0.5 \mu m$ à $\pm 2.25 \mu m$. Cette amélioration a été validée par des évaluations visuelles effectuées sur des objets biologiques. En utilisant une mire pour calibrage, la profondeur de champ obtenue a été évaluée comme étant supérieure à $\pm 1.5 \mu m$.

Une validation quantitative du modèle présenté dans ce chapitre est essentielle et pourrait faire l'objet de travaux ultérieurs.

Chapitre 7

Conclusion et Perspectives

7.1 Conclusion

Dans le cadre de cette thèse multidisciplinaire, nous avons étudié en profondeur les limitations associées à la Microscopie Ptychographique de Fourier (FPM) afin de les surmonter et de développer de nouvelles méthodes d'exploitation de cette technique d'imagerie. Notre objectif est d'améliorer le diagnostic et le suivi des pathologies hématologiques et parasitologiques. Plus précisément, nous avons examiné les défis liés à la profondeur de champ, au temps d'acquisition, au temps de reconstruction et à la qualité d'image. Pour aborder ces questions en microscopie FPM, nous avons associé des méthodes d'intelligence artificielle basées sur des approches d'optimisation neuronale (cadre de réseau informé par un modèle physique) à des approches neuronales d'apprentissage (cadre de réseaux de neurones avec apprentissage statistique). Cette combinaison de techniques offre des perspectives prometteuses pour optimiser à la fois l'efficacité et la précision du diagnostic.

La microscopie ptychographique de Fourier (FPM) offre une profondeur de champ supérieure à celle de la microscopie classique pour la même résolution. Grâce à l'intégration de méthodes neuronales, nous avons réussi à étendre davantage cette profondeur de champ comme montrée dans le Chapitre 3 en traitant les problématiques liées à la mise au point, et en corrigeant les décalages de focus causés par les aberrations chromatiques des objectifs. Ceci facilite l'exploitation d'un large champ de vision. De plus, la FPM permet une reconstruction de l'image qui offre des références fiables sur le comportement cellulaire ciblé par z-stacking numérique. Cette approche s'est appuyée sur un modèle U-Net bimodal, qui met en évidence les comportements cellulaires souhaités. Nous avons pu valider quantitativement cette méthode sur des problèmes de classification de cellules parasitées en observant une amélioration des résultats de classification, caractérisée notamment

par une augmentation de la sensibilité et de la spécificité.

Nos travaux sur les méthodes de reconstruction pour la Microscopie Ptychographique de Fourier (FPM) (Chapitre 4) ont mis en évidence l'efficacité du changement de paradigme apporté par l'utilisation de modèles neuronaux de type cDIP. Nous avons démontré que ces méthodes possèdent les mêmes propriétés que les algorithmes en termes de correction numérique de la mise au point et de correction des aberrations optiques rencontrés dans les objectifs de microscopie.

De plus, les approches basées sur cDIP-LO (Chapitre 5), qui combinent un apprentissage statistique à une phase d'optimisation, offrent la possibilité de relever des défis majeurs dans ce domaine d'imagerie. Sur des données simulées, elles permettent de réduire le nombre de LEDs nécessaires à des niveaux jamais atteints par les algorithmes classiques, avec des taux de recouvrement de l'ordre de $\Gamma \sim 10\%$ et seulement 37 LEDs pour atteindre des facteurs de super-résolution de 5. Toutefois, ces méthodes cDIP requièrent un nombre d'itérations plus élevé pour l'étape d'optimisation, étant donné le nombre élevé des paramètres du réseau de neurones.

Le chapitre 6 est principalement axé sur l'application de la FPM dans un contexte expérimental en hématologie ou parasitologie, nous avons spécifiquement ciblé des applications requérant une imagerie en haute résolution.

Dans un tel contexte, le modèle que nous avons retenu est le cDIP-L qui permet par un apprentissage efficace la reconstruction d'images (I et Φ) avec un nombre réduit de LEDs. Cette méthode a l'avantage non seulement d'améliorer la qualité de l'image, mais aussi d'étendre la profondeur de champ. Par conséquent, elle permet d'obtenir des images nettes sur un large champ de vision et offre la possibilité d'une imagerie bimodale discriminante. Plus spécifiquement, cette technique est capable de produire des images en un temps d'acquisition très court, nécessitant seulement 4 LEDs par longueur d'onde. De plus, comme méthode, nous n'avons pas le besoin d'une étape d'optimisation, ce qui implique que le temps de traitement global est considérablement réduit.

L'originalité de cette thèse se manifeste à travers plusieurs points clés. Premièrement, elle présente un modèle de refocalisation basé sur un réseau neuronal, tout en proposant une méthodologie pour la génération de bases de données, ne nécessitant pas des acquisitions coûteuses, car reposant sur des calculs numériques. Deuxièmement, le travail démontre les avantages spécifiques du modèle cDIP dans divers scénarios de reconstruction d'image. Ce modèle se distingue par l'intégration d'une stratégie d'apprentissage en amont cDIP-L, suivie d'une optimisation dans les cas nécessaires cDIP-LO qui, ensemble, permettent de réduire le nombre de LED nécessaires en imagerie FPM. Enfin, la thèse valide quantitativement les contributions de la FPM dans certains domaines de la biologie, en s'appuyant sur une base de données pour quantifier les apports de cette technologie dans un

contexte expérimental.

En conclusion, ce travail apporte une contribution importante au plan des principes, mais aussi au plan applicatif non seulement à la pathologie numérique et à la ptychographie de Fourier, mais il ouvre également la voie à des applications potentielles dans d'autres types d'imagerie. Il offre des outils robustes et automatisés pour l'acquisition et l'analyse d'échantillons biologiques grâce à la mise en œuvre stratégique de modèles neuronaux. Les méthodes que nous avons développées présentent un potentiel considérable pour améliorer non seulement la précision des diagnostics, mais également pour accélérer significativement le processus d'analyse. Cette efficacité est particulièrement bénéfique dans des contextes médicaux où le temps est un facteur critique.

7.2 Perspectives

7.2.1 Validation quantitative du modèle cDIP-L pour la réduction du nombre de LED et l'amélioration de la qualité d'image

Les travaux abordés au Chapitre 6 concernent la réduction du nombre de LEDs et l'amélioration de la qualité d'images, en particulier lors de l'utilisation d'objectifs à grande ouverture numérique. Ces travaux ont déjà fait l'objet d'une validation visuelle préliminaire. Le modèle cDIP-L s'est révélé particulièrement efficace pour des opérations impliquant un faible facteur de super-résolution.

Il serait intéressant de compléter les premiers résultats par une validation quantitative. Cette dernière sera réalisée en classifiant des leucocytes et en détectant des parasites à partir de larges bases de données de manière à évaluer le potentiel d'intégration de notre approche dans des systèmes automatisés d'imagerie médicale.

a) Évaluation complète cDIP-L

Un objectif supplémentaire, si les résultats sont confirmés, serait de poursuivre cette évaluation dans un objectif de commercialisation (TRIBVN/T-life) et ceci en menant une analyse comparative entre cette méthode et d'autres systèmes de diagnostic existants, et en se concentrant sur les mêmes types de frottis sanguins. Cette évaluation comparative permettra de discerner les atouts distinctifs que la microscopie ptychographique de Fourier (FPM) apporte au domaine, ainsi que les avantages spécifiques découlant de l'intégration de la bi-modalité dans ces techniques.

b) Exploration de la configuration de LED optimale

Un autre axe prometteur pour la poursuite des recherches sur le modèle cDIP-L consisterait à explorer la configuration optimale de la matrice de LEDs. Cette exploration pourrait s'étendre aux travaux sur l'éclairage et la réduction du nombre de LED permettant une réduction du nombre de LED. Une alternative à la configuration des LEDs présentée dans le chapitre 6 est actuellement en cours d'exploration et affiche des performances préliminaires comparables. Dans cette nouvelle configuration, un éclairage séquentiel est réalisé avec des groupes de 3 LEDs à chaque fois, ce qui représente un éclairage par quadrant dans l'anneau composé de 12 LEDs. Par exemple, on commence avec un éclairage de 3 LEDs situées au nord, suivi de 3 LEDs situées au sud, et ainsi de suite. Dans cette configuration, le rôle du cDIP-L devient principalement celui de la séparation des sources lumineuses. L'utilisation simultanée de trois LED multiplie la puissance lumineuse par trois, ce qui a pour conséquence de réduire encore davantage le temps nécessaire pour la capture d'image, en l'occurrence par un facteur de trois.

De plus, il serait intéressant d'examiner des approches qui combinent des reconstructions à partir de multiples longueurs d'ondes, en intégrant ces données directement dans un seul modèle cDIP-L.

7.2.2 Estimation du Focus

Un axe de recherche pertinent à étudier serait l'élaboration de méthodes basées sur l'apprentissage profond pour l'estimation précise du décalage de focalisation Δz_{num} . Cette estimation pourrait être complémentaire aux algorithmes d'autofocalisation proposés dans cette thèse dans le cadre de systèmes de diagnostic automatisés basés sur la microscopie ptychographique de Fourier (FPM). L'objectif de cette initiative est double. Premièrement, cette approche servira de référence pour le réglage du focus mécanique du microscope, permettant ainsi une optimisation du microscope. Deuxièmement, le Δz_{num} pourrait également être intégré en tant que paramètre dans les algorithmes de reconstruction comme cDIP-LO ou ePIE. Cela permettrait non seulement de générer des images directement focalisées, mais aussi d'étendre considérablement la profondeur de champ.

7.2.3 Reconstruction tomographique par cDIP-LO

L'une des perspectives visées serait d'étendre la stratégie actuelle du modèle cDIP-LO, qui repose sur des modèles physiques pour la reconstruction en microscopie ptychographique de Fourier 2D, vers une reconstruction 3D [97]. Dans ce contexte, nous envisageons d'explorer le potentiel des méthodes basées sur l'apprentissage statistique et l'optimisation. L'objectif est d'atteindre une qua-

lité d'image qui dépasse les limitations actuelles identifiées inhérents à l'imagerie 3D, telle que l'amélioration de la résolution suivant l'axe Z. Par ailleurs, cette approche pourrait également permettre de réduire le nombre de LEDs nécessaires qui est actuellement, dans la littérature existante, très important.

7.2.4 Gestion du bruit

Les réseaux de neurones ont démontré leur efficacité dans la gestion du bruit dans plusieurs domaines. L'étude des méthodes neuronales permettant une reconstruction FPM de bonne qualité avec des données bruitées présente un intérêt particulier pour les systèmes d'acquisition exploitant des facteurs élevés de super-résolution notamment en utilisant des configurations d'imagerie exploitant du champ sombre. Cette approche exploratoire viserait à alléger les contraintes sur la configuration expérimentale. Par exemple, elle pourrait réduire la nécessité d'équipements électroniques de haute qualité ou d'acquisitions à haute gamme dynamique (HDR) avec des temps d'acquisition prolongés. Ces méthodes pourraient également être envisagées pour atténuer le bruit introduit par le gain de la caméra, notamment lorsqu'on cherche à réduire les temps d'exposition.

7.2.5 Extension à d'autres domaines

Les principes qui soutiennent les modèles de reconstruction sont très fortement reliés à la résolution de problèmes inverses, mal posés.

De ce fait, il serait pertinent d'évaluer leur contribution dans d'autres domaines de la microscopie, tels que la ptychographie directe ou l'holographie, qui opèrent selon des principes similaires et offrent, tout comme la FPM, un accès à la phase en plus de l'intensité pour la caractérisation de l'objet étudié.

Bibliographie

- [1] OT4D-ANR project, (French national research agency), dedicated to the development of innovative optical microscopy and artificial intelligence algorithms for medical diagnosis (2021-2025).
- [2] TAMIS- BPI and HDH project,(Banque Publique d'Investissement), (Health Data Hub), Dedicated to the development of innovative optical microscopy techniques and artificial intelligence algorithms for the analysis of leukocytes, platelets, and schizocytes. (2021-2023).
- [3] http://svt.ac-dijon.fr/schemassvt/IMG/gif/bio_echell.gif.
- [4] The theory of super-resolution electron microscopy via wigner-distribution deconvolution. *Philosophical Transactions of the Royal Society of London. Series A : Physical and Engineering Sciences*, 339(1655) :521–553, June 1992. doi : 10.1098/rsta.1992.0050. URL <https://doi.org/10.1098/rsta.1992.0050>.
- [5] Abbe. The relation of aperture and power in the microscope (continued). *Journal of the Royal Microscopical Society*, 2(4) :460–473, 1882.
- [6] E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1) :413–468, December 1873. doi : 10.1007/bf02956173. URL <https://doi.org/10.1007/bf02956173>.
- [7] Daniel Otero Baguer, Johannes Leuschner, and Maximilian Schmidt. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9) :094004, sep 2020. doi : 10.1088/1361-6420/aba415. URL <https://dx.doi.org/10.1088/1361-6420/aba415>.
- [8] Dana Bani-Hani, Naseem Khan, Fatimah Alsultan, Shreya Karanjkar, and Nagen Nagarur. Classification of leucocytes using convolutional neural net-

- work optimized through genetic algorithm. In *Proc. of the 7th Annual World Conference of the Society for Industrial and Systems Engineering*, 2018.
- [9] Arthur Baroni. *Développement de la ptychographie vectorielle pour la microscopie quantitative de matériaux anisotropes*. PhD thesis, Aix-Marseille, 2020.
- [10] Arthur Baroni, Lyes Bouchama, Bernadette Dorizzi, and Yaneck Gottesman. Angularly resolved polarization microscopy for birefringent materials with fourier ptychography. *Opt. Express*, 30(21) :38984–38994, Oct 2022. doi : 10.1364/OE.469377. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-21-38984>.
- [11] Vitoantonio Bevilacqua, Antonio Brunetti, Gianpaolo Francesco Trotta, Domenico De Marco, Marco Giuseppe Quercia, Domenico Buongiorno, Alessia D’Introno, Francesco Girardi, and Attilio Guarini. A novel deep learning approach in haematology for classification of leucocytes. *Quantifying and Processing Biomedical and Behavioral Signals 27*, pages 265–274, 2019.
- [12] Zichao Bian, Chengfei Guo, Shaowei Jiang, Jiakai Zhu, Ruihai Wang, Pengming Song, Zibang Zhang, Kazunori Hoshino, and Guoan Zheng. Autofocusing technologies for whole slide imaging and automated microscopy. *Journal of Biophotonics*, 13(12), September 2020. doi : 10.1002/jbio.202000227. URL <https://doi.org/10.1002/jbio.202000227>.
- [13] Pierre Bon, Guillaume Maucort, Benoit Wattellier, and Serge Monneret. Quadrature lateral shearing interferometry for quantitative phase microscopy of living cells. *Optics express*, 17(15) :13080–13094, 2009.
- [14] Lokesh Boominathan, Mayug Maniparambil, Honey Gupta, Rahul Baburajan, and Kaushik Mitra. Phase retrieval for fourier ptychography under varying amount of measurements, 2018. URL <https://arxiv.org/abs/1805.03593>.
- [15] Lyes Bouchama, Bernadette Dorizzi, Jacques Klossa, and Yaneck Gottesman. A physics-inspired deep learning framework for an efficient fourier ptychographic microscopy reconstruction under low overlap conditions. *Sensors*, 23(15) :6829, July 2023. ISSN 1424-8220. doi : 10.3390/s23156829. URL <http://dx.doi.org/10.3390/s23156829>.
- [16] Lyes Bouchama, Bernadette Dorizzi, Marc Thellier, Jacques Klossa, and Yaneck Gottesman. Fourier ptychographic microscopy image enhancement with bi-modal deep learning. *Biomedical Optics Express*, 14(7) :3172, June 2023. doi : 10.1364/boe.489776. URL <https://doi.org/10.1364/boe.489776>.

- [17] Romain Brossier, Stéphane Operto, and Jean Virieux. Seismic imaging of complex onshore structures by 2d elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6) :WCC105–WCC118, 2009.
- [18] Oliver Bunk, Martin Dierolf, Søren Kynde, Ian Johnson, Othmar Marti, and Franz Pfeiffer. Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy*, 108(5) :481–487, April 2008. doi : 10.1016/j.ultramic.2007.08.003. URL <https://doi.org/10.1016/j.ultramic.2007.08.003>.
- [19] Oliver Bunk, Martin Dierolf, Søren Kynde, Ian Johnson, Othmar Marti, and Franz Pfeiffer. Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy*, 108 :481–487, 05 2008. doi : 10.1016/j.ultramic.2007.08.003.
- [20] Quan Chen, Duolan Huang, and Rui Chen. Fourier ptychographic microscopy with untrained deep neural network priors. *Opt. Express*, 30(22) :39597–39612, Oct 2022. doi : 10.1364/OE.472171. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-30-22-39597>.
- [21] Sangyeon Cho, Soomin Kim, Youngchan Kim, and YongKeun Park. Optical imaging techniques for the study of malaria. *Trends in biotechnology*, 30(2) : 71–79, 2012.
- [22] Kirsten Christensen-Jeffries, Olivier Couture, Paul A Dayton, Yonina C Eldar, Kullervo Hynynen, Fabian Kiessling, Meaghan O’Reilly, Gianmarco F Pinton, Georg Schmitz, Meng-Xing Tang, et al. Super-resolution ultrasound imaging. *Ultrasound in medicine & biology*, 46(4) :865–891, 2020.
- [23] Remy Claveau, Petru Manescu, Muna Elmi, Vijay Pawar, Michael Shaw, and Delmiro Fernandez-Reyes. Digital refocusing and extended depth of field reconstruction in fourier ptychographic microscopy. *Biomed. Opt. Express*, 11 (1) :215–226, Jan 2020. doi : 10.1364/BOE.11.000215. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-11-1-215>.
- [24] John C Curlander and Robert N McDonough. *Synthetic aperture radar*, volume 11. Wiley, New York, 1991.
- [25] Tao Dai, Yan Feng, Dongxian Wu, Bin Chen, Jian Lu, Yong Jiang, and Shu-Tao Xia. Dipdefend : Deep image prior driven defense against adversarial examples. In *Proceedings of the 28th ACM International Conference on Multimedia*, MM ’20. ACM, October 2020. doi : 10.1145/3394171.3413898. URL <http://dx.doi.org/10.1145/3394171.3413898>.

- [26] Sören Dittmer, Tobias Kluth, Peter Maass, and Daniel Otero Bager. Regularization by architecture : A deep prior approach for inverse problems. *Journal of Mathematical Imaging and Vision*, 62(3) :456–470, October 2019. doi : 10.1007/s10851-019-00923-x. URL <https://doi.org/10.1007/s10851-019-00923-x>.
- [27] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks, 2015. URL <https://arxiv.org/abs/1501.00092>.
- [28] Jan-Niklas Eckardt, Jan Moritz Middeke, Sebastian Riechert, Tim Schmittmann, Anas Shekh Sulaiman, Michael Kramer, Katja Sockel, Frank Kroshinsky, Ulrich Schuler, Johannes Schetelig, Christoph Röllig, Christian Thiede, Karsten Wendt, and Martin Bornhäuser. Deep learning detects acute myeloid leukemia and predicts NPM1 mutation status from bone marrow smears. *Leukemia*, 36(1) :111–118, September 2021. doi : 10.1038/s41375-021-01408-w. URL <https://doi.org/10.1038/s41375-021-01408-w>.
- [29] Patrick Ferrand, Arthur Baroni, Marc Allain, and Virginie Chamard. Quantitative imaging of anisotropic material properties with vectorial ptychography. *Opt. Lett.*, 43(4) :763–766, Feb 2018. doi : 10.1364/OL.43.000763. URL <https://opg.optica.org/ol/abstract.cfm?URI=ol-43-4-763>.
- [30] J. R. Fienup. Reconstruction of an object from the modulus of its fourier transform. *Opt. Lett.*, 3(1) :27–29, Jul 1978. doi : 10.1364/OL.3.000027. URL <https://opg.optica.org/ol/abstract.cfm?URI=ol-3-1-27>.
- [31] Nicolas Gebert, Gerhard Krieger, and Alberto Moreira. Digital beamforming on receive : Techniques and optimization strategies for high-resolution wide-swath sar imaging. *IEEE Transactions on Aerospace and Electronic Systems*, 45(2) :564–592, 2009. doi : 10.1109/TAES.2009.5089542.
- [32] Ralph W Gerchberg. A practical algorithm for the determination of plane from image and diffraction pictures. *Optik*, 35(2) :237–246, 1972.
- [33] Vahid Ghodrati, Jiaxin Shao, Mark Bydder, Ziwu Zhou, Wotao Yin, Kim-Lien Nguyen, Yingli Yang, and Peng Hu. Mr image reconstruction using deep learning : evaluation of network structure and loss functions. *Quantitative Imaging in Medicine and Surgery*, 9(9) :1516–1527, September 2019. ISSN 2223-4306. doi : 10.21037/qims.2019.08.10. URL <http://dx.doi.org/10.21037/qims.2019.08.10>.

- [34] Ross Girshick. Fast r-CNN. In *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, December 2015. doi : 10.1109/iccv.2015.169. URL <https://doi.org/10.1109/iccv.2015.169>.
- [35] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.
- [36] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.
- [37] J. W. Goodman and R. W. Lawrence. DIGITAL IMAGE FORMATION FROM ELECTRONICALLY DETECTED HOLOGRAMS. *Applied Physics Letters*, 11(3) :77–79, August 1967. doi : 10.1063/1.1755043. URL <https://doi.org/10.1063/1.1755043>.
- [38] Joseph W Goodman and P Sutton. Introduction to fourier optics. *Quantum and Semiclassical Optics-Journal of the European Optical Society Part B*, 8 (5) :1095, 1996.
- [39] Mattias P. Heinrich, Maik Stille, and Thorsten M. Buzug. Residual u-net convolutional neural network architecture for low-dose ct denoising. *Current Directions in Biomedical Engineering*, 4(1) :297–300, September 2018. ISSN 2364-5504. doi : 10.1515/cdbme-2018-0072. URL <http://dx.doi.org/10.1515/cdbme-2018-0072>.
- [40] Roarke Horstmeyer, Jaebum Chung, Xiaoze Ou, Guoan Zheng, and Changhuei Yang. Diffraction tomography with fourier ptychography. *Optica*, 3(8) :827–835, Aug 2016. doi : 10.1364/OPTICA.3.000827. URL <https://opg.optica.org/optica/abstract.cfm?URI=optica-3-8-827>.
- [41] Jane Hung and Anne Carpenter. Applying faster r-CNN for object detection on malaria images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE, July 2017. doi : 10.1109/cvprw.2017.112. URL <https://doi.org/10.1109/cvprw.2017.112>.
- [42] Tayyaba Jameela, Kavitha Athota, Ninni Singh, Vinit Kumar Gunjan, and Sayan Kahali. Deep learning and transfer learning for malaria detection. *Computational Intelligence and Neuroscience*, 2022 :1–14, June 2022. doi : 10.1155/2022/2221728. URL <https://doi.org/10.1155/2022/2221728>.
- [43] Shaowei Jiang, Kaikai Guo, Jun Liao, and Guoan Zheng. Solving fourier ptychographic imaging problems via neural network modeling and tensorflow.

- Biomed. Opt. Express*, 9(7) :3306–3319, Jul 2018. doi : 10.1364/BOE.9.003306. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-9-7-3306>.
- [44] Armin Kappeler, Sushobhan Ghosh, Jason Holloway, Oliver Cossairt, and Aggelos Katsaggelos. Ptychnet : Cnn based fourier ptychography. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 1712–1716, 2017. doi : 10.1109/ICIP.2017.8296574.
- [45] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6) :422–440, Jun 2021. ISSN 2522-5820. doi : 10.1038/s42254-021-00314-5. URL <https://doi.org/10.1038/s42254-021-00314-5>.
- [46] August Köhler. New method of illumination for photomicrographical purposes. *Journal of the Royal Microscopical Society*, 14 :261–262, 1894.
- [47] P. C. LAUTERBUR. Image formation by induced local interactions : Examples employing nuclear magnetic resonance. *Nature*, 242(5394) :190–191, Mar 1973. ISSN 1476-4687. doi : 10.1038/242190a0. URL <https://doi.org/10.1038/242190a0>.
- [48] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4) :541–551, 1989. doi : 10.1162/neco.1989.1.4.541.
- [49] Jiaji Li, Qian Chen, Jialin Zhang, Yan Zhang, Linpeng Lu, and Chao Zuo. Efficient quantitative phase microscopy using programmable annular led illumination. *Biomed. Opt. Express*, 8(10) :4687–4705, Oct 2017. doi : 10.1364/BOE.8.004687. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-8-10-4687>.
- [50] Jiaji Li, Qian Chen, Jialin Zhang, Yan Zhang, Linpeng Lu, and Chao Zuo. Efficient quantitative phase microscopy using programmable annular led illumination. *Biomed. Opt. Express*, 8(10) :4687–4705, Oct 2017. doi : 10.1364/BOE.8.004687. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-8-10-4687>.
- [51] Qiang Li, Xianming Liu, Junjun Jiang, Cheng Guo, Xiangyang Ji, and Xiaolin Wu. Rapid whole slide imaging via dual-shot deep autofocusing. *IEEE Transactions on Computational Imaging*, 7 :124–136, 2021. doi : 10.1109/TCI.2020.3046189.

- [52] Mingshu Liang, Cory Bernadt, Soon Boon Justin Wong, Changsoon Choi, Richard Cote, and Changhui Yang. All-in-focus fine needle aspiration biopsy imaging based on fourier ptychographic microscopy. *Journal of Pathology Informatics*, 13 :100119, 2022. ISSN 2153-3539. doi : <https://doi.org/10.1016/j.jpi.2022.100119>. URL <https://www.sciencedirect.com/science/article/pii/S2153353922007131>.
- [53] Zhaohui Liang, Andrew Powell, Ilker Ersoy, Mahdiah Poostchi, Kamolrat Silamut, Kannappan Palaniappan, Peng Guo, Md Amir Hossain, Antani Sameer, Richard James Maude, Jimmy Xiangji Huang, Stefan Jaeger, and George Thoma. Cnn-based image analysis for malaria diagnosis. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 493–496, 2016. doi : 10.1109/BIBM.2016.7822567.
- [54] Jiaming Liu, Yu Sun, Xiaojian Xu, and Ulugbek S. Kamilov. Image restoration using total variation regularized deep image prior. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7715–7719, 2019. doi : 10.1109/ICASSP.2019.8682856.
- [55] Virendra N. Mahajan. Zernike circle polynomials and optical aberrations of systems with circular pupils. *Appl. Opt.*, 33(34) :8121–8124, Dec 1994. doi : 10.1364/AO.33.008121. URL <https://opg.optica.org/ao/abstract.cfm?URI=ao-33-34-8121>.
- [56] Andrew Maiden, Daniel Johnson, and Peng Li. Further improvements to the ptychographical iterative engine. *Optica*, 4(7) :736–745, Jul 2017. doi : 10.1364/OPTICA.4.000736. URL <https://opg.optica.org/optica/abstract.cfm?URI=optica-4-7-736>.
- [57] Andrew M. Maiden and John M. Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109 (10) :1256–1262, September 2009. doi : 10.1016/j.ultramic.2009.05.012. URL <https://doi.org/10.1016/j.ultramic.2009.05.012>.
- [58] Christian Matek, Simone Schwarz, Karsten Spiekermann, and Carsten Marr. Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. *Nature Machine Intelligence*, 1(11) :538–544, 2019.
- [59] Courosh Mehanian, Mayoore Jaiswal, Charles Delahunt, Clay Thompson, Matt Horning, Liming Hu, Travis Ostbye, Shawn McGuire, Martha Mehanian, Cary Champlin, Ben Wilson, Earl Long, Stephane Proux, Dionicia Gamboa, Peter Chiodini, Jane Carter, Mehul Dhorda, David Isaboke, Bernhards Ogutu, Wellington Oyibo, Elizabeth Villasis, Kyaw Myo Tun, Christine Bachman,

- and David Bell. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops*, Oct 2017.
- [60] Inkyu Moon, Faliu Yi, Yeon H. Lee, Bahram Javidi, Daniel Boss, and Pierre Marquet. Automated quantitative analysis of 3d morphology and mean corpuscular hemoglobin in human red blood cells stored in different periods. *Optics Express*, 21(25) :30947, December 2013. doi : 10.1364/oe.21.030947. URL <https://doi.org/10.1364/oe.21.030947>.
- [61] Thanh Nguyen, Yujia Xue, Yunzhe Li, Lei Tian, and George Nehmetallah. Deep learning approach for fourier ptychography microscopy. *Opt. Express*, 26(20) :26470–26484, Oct 2018. doi : 10.1364/OE.26.026470. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-20-26470>.
- [62] Xiaoze Ou, Guoan Zheng, and Changhui Yang. Embedded pupil function recovery for fourier ptychographic microscopy. *Opt. Express*, 22(5) :4960–4972, Mar 2014. doi : 10.1364/OE.22.004960. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-22-5-4960>.
- [63] An Pan, Yan Zhang, Kai Wen, Meiling Zhou, Junwei Min, Ming Lei, and Baoli Yao. Subwavelength resolution fourier ptychography with hemispherical digital condensers. *Opt. Express*, 26(18) :23119–23131, Sep 2018. doi : 10.1364/OE.26.023119. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-18-23119>.
- [64] Jongchan Park, David J. Brady, Guoan Zheng, Lei Tian, and Liang Gao. Review of bio-optical imaging systems with a high space-bandwidth product. *Advanced Photonics*, 3(04), June 2021. doi : 10.1117/1.ap.3.4.044001. URL <https://doi.org/10.1117/1.ap.3.4.044001>.
- [65] William Pierré, Lionel Hervé, Chiara Paviolo, Ondrej Mandula, Vincent Remondiere, Sophie Morales, Sergei Grudinin, Pierre F. Ray, Magali Dhellemmes, Christophe Arnoult, and Cédric Allier. 3d time-lapse imaging of a mouse embryo using intensity diffractiontomography embedded inside a deep learning framework. *Appl. Opt.*, 61(12) :3337–3348, Apr 2022. doi : 10.1364/AO.453910. URL <https://opg.optica.org/ao/abstract.cfm?URI=ao-61-12-3337>.
- [66] Lorenzo Putzu, Giovanni Caocci, and Cecilia Di Ruberto. Leucocyte classification for leukaemia detection using image processing techniques. *Artificial intelligence in medicine*, 62(3) :179–191, 2014.

- [67] Hussin K. Ragb, Ian T. Dover, and Redha Ali. Deep convolutional neural network ensemble for improved malaria parasite detection. In *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–10, 2020. doi : 10.1109/AIPR50011.2020.9425273.
- [68] Sivaramakrishnan Rajaraman, Sameer K. Antani, Mahdiah Poostchi, Kamolrat Silamut, Md. A. Hossain, Richard J. Maude, Stefan Jaeger, and George R. Thoma. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*, 6 :e4568, April 2018. doi : 10.7717/peerj.4568. URL <https://doi.org/10.7717/peerj.4568>.
- [69] Rayleigh. Xxxi. investigations in optics, with special reference to the spectroscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 8(49) :261–274, 1879. doi : 10.1080/14786447908639684. URL <https://doi.org/10.1080/14786447908639684>.
- [70] Joseph Redmon and Ali Farhadi. Yolov3 : An incremental improvement, 2018. URL <https://arxiv.org/abs/1804.02767>.
- [71] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once : Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2016. doi : 10.1109/cvpr.2016.91. URL <https://doi.org/10.1109/cvpr.2016.91>.
- [72] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-CNN : Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6) :1137–1149, June 2017. doi : 10.1109/tpami.2016.2577031. URL <https://doi.org/10.1109/tpami.2016.2577031>.
- [73] Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light : Science amp ; Applications*, 7(2) :17141–17141, October 2017. ISSN 2047-7538. doi : 10.1038/lssa.2017.141. URL <http://dx.doi.org/10.1038/lssa.2017.141>.
- [74] Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light : Science & Applications*, 7(2) :17141–17141, Feb 2018. ISSN 2047-7538. doi : 10.1038/lssa.2017.141. URL <https://doi.org/10.1038/lssa.2017.141>.

- [75] Anna-Lena Robisch. *Phase retrieval for object and probe in the optical near-field*. Göttingen University Press, 2016. doi : 10.17875/gup2016-901. URL <https://doi.org/10.17875/gup2016-901>.
- [76] J. M. Rodenburg and H. M. L. Faulkner. A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20) :4795–4797, November 2004. doi : 10.1063/1.1823034. URL <https://doi.org/10.1063/1.1823034>.
- [77] 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. ISBN 978-3-319-24574-4.
- [78] F. Rosenblatt. The perceptron : A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6) :386–408, 1958. doi : 10.1037/h0042519. URL <https://doi.org/10.1037/h0042519>.
- [79] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088) :533–536, October 1986. doi : 10.1038/323533a0. URL <https://doi.org/10.1038/323533a0>.
- [80] Divyansh Shah, Khushbu Kawale, Masumi Shah, Santosh Randive, and Rahul Mapari. Malaria parasite detection using deep learning : (beneficial to humankind). In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 984–988, 2020. doi : 10.1109/ICICCS48265.2020.9121073.
- [81] Carroll Mason Sparrow. On spectroscopic resolving power. *Astrophysical Journal*, vol. 44, p. 76, 44 :76, 1916.
- [82] Jiasong Sun, Qian Chen, Yuzhen Zhang, and Chao Zuo. Sampling criteria for fourier ptychographic microscopy in object space and frequency space. *Opt. Express*, 24(14) :15765–15781, Jul 2016. doi : 10.1364/OE.24.015765. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-24-14-15765>.
- [83] Jiasong Sun, Chao Zuo, Jialin Zhang, Yao Fan, and Qian Chen. High-speed fourier ptychographic microscopy based on programmable annular illuminations. *Scientific Reports*, 8(1) :7669, May 2018. ISSN 2045-2322. doi : 10.1038/s41598-018-25797-8. URL <https://doi.org/10.1038/s41598-018-25797-8>.

- [84] Lei Tian, Xiao Li, Kannan Ramchandran, and Laura Waller. Multiplexed coded illumination for fourier ptychography with an led array microscope. *Biomed. Opt. Express*, 5(7) :2376–2389, Jul 2014. doi : 10.1364/BOE.5.002376. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-5-7-2376>.
- [85] Lei Tian, Ziji Liu, Li-Hao Yeh, Michael Chen, Jingshan Zhong, and Laura Waller. Computational illumination for high-speed in vitro fourier ptychographic microscopy. *Optica*, 2(10) :904–911, Oct 2015. doi : 10.1364/OPTICA.2.000904. URL <https://opg.optica.org/optica/abstract.cfm?URI=optica-2-10-904>.
- [86] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. *International Journal of Computer Vision*, 128(7) :1867–1888, March 2020. doi : 10.1007/s11263-020-01303-4. URL <https://doi.org/10.1007/s11263-020-01303-4>.
- [87] Muhammad Umer, Saima Sadiq, Muhammad Ahmad, Saleem Ullah, Gyu Sang Choi, and Arif Mehmood. A novel stacked cnn for malarial parasite detection in thin blood smear images. *IEEE Access*, 8 :93782–93792, 2020. doi : 10.1109/ACCESS.2020.2994810.
- [88] Fei Wang, Yaoming Bian, Haichao Wang, Meng Lyu, Giancarlo Pedrini, Wolfgang Osten, George Barbastathis, and Guohai Situ. Phase imaging with an untrained neural network. *Light : Science & Applications*, 9(1), May 2020. doi : 10.1038/s41377-020-0302-3. URL <https://doi.org/10.1038/s41377-020-0302-3>.
- [89] Qiwei Wang, Shusheng Bi, Minglei Sun, Yuliang Wang, Di Wang, and Shaobao Yang. Deep learning approach to peripheral leukocyte recognition. *PLOS ONE*, 14(6) :e0218808, June 2019. doi : 10.1371/journal.pone.0218808. URL <https://doi.org/10.1371/journal.pone.0218808>.
- [90] Yichen Wu and Aydogan Ozcan. Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring. *Methods*, 136 :4–16, 2018. ISSN 1046-2023. doi : <https://doi.org/10.1016/j.ymeth.2017.08.013>. URL <https://www.sciencedirect.com/science/article/pii/S1046202317301974>. *Methods in Quantitative Phase Imaging in Life Science*.
- [91] Jun Xu, Yuan Huang, Ming-Ming Cheng, Li Liu, Fan Zhu, Zhou Xu, and Ling Shao. Noisy-as-clean : Learning self-supervised denoising from corrupted image. *IEEE Transactions on Image Processing*, 29 :9316–9329, 2020.

- [92] Anthony Yen. Rayleigh or abbe? origin and naming of the resolution formula of microlithography. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 19(04), November 2020. doi : 10.1117/1.jmm.19.4.040501. URL <https://doi.org/10.1117/1.jmm.19.4.040501>.
- [93] Jizhou Zhang, Tingfa Xu, Ziyi Shen, Yifan Qiao, and Yizhou Zhang. Fourier ptychographic microscopy reconstruction with multiscale deep residual network. *Opt. Express*, 27(6) :8612–8625, Mar 2019. doi : 10.1364/OE.27.008612. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-27-6-8612>.
- [94] Shaohui Zhang, Guocheng Zhou, Chuanjian Zheng, Tong Li, Yao Hu, and Qun Hao. Fast digital refocusing and depth of field extended fourier ptychography microscopy. *Biomed. Opt. Express*, 12(9) :5544–5558, Sep 2021. doi : 10.1364/BOE.433033. URL <https://opg.optica.org/boe/abstract.cfm?URI=boe-12-9-5544>.
- [95] Guoan Zheng, Roarke Horstmeyer, and Changhui Yang. Wide-field, high-resolution fourier ptychographic microscopy. *Nature Photonics*, 7(9) :739–745, Sep 2013. ISSN 1749-4893. doi : 10.1038/nphoton.2013.187. URL <https://doi.org/10.1038/nphoton.2013.187>.
- [96] Haowen Zhou, Cheng Shen, Mingshu Liang, and Changhui Yang. Analysis of postreconstruction digital refocusing in Fourier ptychographic microscopy. *Optical Engineering*, 61(7) :073102, 2022. doi : 10.1117/1.OE.61.7.073102. URL <https://doi.org/10.1117/1.OE.61.7.073102>.
- [97] Kevin C. Zhou and Roarke Horstmeyer. Diffraction tomography with a deep image prior. *Opt. Express*, 28(9) :12872–12896, Apr 2020. doi : 10.1364/OE.379200. URL <https://opg.optica.org/oe/abstract.cfm?URI=oe-28-9-12872>.

Titre : Apport des techniques d'apprentissage (profond) à la microscopie holographique pour applications médicales

Mots clés : Apprentissage profond, Microscopie ptychographique de Fourier, Imagerie biomédicale

Résumé : Mon travail s'inscrit dans le cadre du partenariat stratégique Télécom SudParis (TSP) et TRIBVN/T-life, dédié au développement de nouvelles approches en microscopie optique, couplées à l'intelligence artificielle, en vue d'identifier, de prédire et de monitorer les pathologies hématologiques et parasitologiques.

C'est dans cette perspective que nous avons développé, dans le laboratoire, un prototype de microscope reposant sur un principe d'imagerie non conventionnelle à synthèse d'ouverture, basée sur l'approche FPM (Fourier Ptychographic Microscopy). Cette approche permet de dépasser les limites en résolution de l'optique conventionnelle ou de façon équivalente accéder à des champs de vue très larges (de 4 à 25 fois plus importants) à résolution fixée. Elle permet, en sus, de diversifier la nature des données acquises (avec l'enregistrement de la phase en complément des données d'intensité).

Toutefois, en raison de certaines contraintes, notamment le temps d'acquisition et de reconstruction des images pour obtenir une qualité optimale, cette tech-

nologie rencontre encore des difficultés pour trouver ses applications et être commercialisée par les différents acteurs du domaine de la microscopie.

Le travail réalisé dans cette thèse a permis des avancées significatives sur certains des aspects limitants de cette technologie grâce à la mise en œuvre de modèles à base de réseaux de neurones.

Nous avons proposé une refocalisation automatique des images bimodales efficace sur de grands champs de vue, grâce à un post-traitement basé sur un U-Net. Nous avons aussi proposé une approche originale, alliant apprentissage statistique et optimisation guidée par la physique pour réduire les temps d'acquisition des images et les temps de reconstruction.

Ces modèles ont démontré leur efficacité, par des diagnostics plus précis et discriminants, dans des applications de parasitologie et d'hématologie.

Les contributions apportées ont un potentiel d'application qui dépasse le domaine de la FPM, ouvrant des perspectives dans divers autres champs de l'imagerie calculatoire.

Title : Contribution of (deep) learning techniques to holographic microscopy for medical applications

Keywords : Deep Learning, Fourier Ptychographic Microscopy, Biomedical Imaging

Abstract : This research is part of the Télécom SudParis (TSP) and TRIBVN/T-life strategic partnership, dedicated to the development of new approaches in optical microscopy, coupled with artificial intelligence, to identify, predict and monitor hematological and parasitological pathologies. In this regard, we developed a prototype microscope based on a computational imaging principle with a synthetic aperture, based on the FPM (Fourier Ptychographic Microscopy) approach. This approach makes it possible to overcome conventional optics' resolution limits, or equivalently access very large fields of view (from 4 to 25 times larger) at fixed resolution. It also enables us to diversify the nature of the data acquired (with phase recording in addition to intensity data).

However, despite its promise, the technology faces challenges in widespread adoption and commercialization within the microscopy field, primarily due to constraints such as the time-intensive process re-

quired for image acquisition and reconstruction to achieve optimal quality.

The research conducted in this thesis has led to substantial advancements in overcoming certain limitations of this technology, leveraging models based on neural networks.

We have proposed an efficient automatic refocusing of bimodal images over large fields of view, thanks to post-processing based on a U-Net. We have also proposed an original approach, combining statistical learning and physics-driven optimization to reduce image acquisition and reconstruction times.

These frameworks have validated their efficacy by yielding more precise and discriminating diagnoses in the fields of parasitology and haematology.

The potential applications of these contributions go far beyond the field of FPM, opening up perspectives in various other fields of computational imaging.