Automatic analysis of blood smears images : contribution of phase modality in Fourier Ptychographic Microscopy

Houda Hassini

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2024IPPAS014

Thèse de doctorat

Automatic Analysis of Blood Smears Images: Contribution of Phase Modality in Fourier Ptychographic Microscopy

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

École doctorale n°626 Dénomination (EDIPP)
Spécialité de doctorat: Informatique

Thèse présentée et soutenue à Palaiseau, le 25 Juin 2024, par

HOUDA HASSINI

Composition du Jury :

Elsa Angelini Professeure, Télécom Paris	Présidente
Laurent Barthes Maitre de conférences (HDR), Université de Versailles Saint Quentin en Yvelines	Rapporteur
Jihad Zallat Professeur, Université de Strasbourg	Rapporteur
Pauline Trouvé-Pelvoux Ingénieure de recherche, ONERA - The French Aerospace Lab	Examinatrice
Yaneck Gottesman Professeur, Télécom SudParis	Directreur de thèse
Bernadette Dorizzi Professeure émerite, Télécom SudParis	Co-encadrante de thèse
Jacques Klossa Directeur d'entreprise, TRIBVN/T-life	Invité

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Yaneck Gottesman, for his guidance. I thank him for teaching me the intricacies of research. His passionate vision has greatly influenced my own approach to research.
I would also like to extend my sincere thanks to my co-supervisor, Bernadette Dorizzi. She gives generously without counting and distinguishes herself not only through her scientific excellence but also by her exceptional human qualities. Her unwavering support and mentorship have helped me grow not only as a researcher but also as a person.
My sincere thanks go to Jacques Klossa, whose support was essential to this project's success, and to Vincent Leymarie and Mellody Sextius from Hospital Avicennes for their valuable contributions to the hematology aspect of my work.
I deeply appreciate my colleague Lyes Bouchama for sharing both the challenges and rewards of our research journey, and Arthur Baroni for his vital assistance in the field of optics.
I want to express my deepest gratitude to my partner Romain Haton. his encouragement, understanding, and belief in me have been invaluable, giving me the strength to persevere during the most challenging times.
Finally, I wish to extend my heartfelt thanks to my family, who have been my foundation throughout this entire journey. Their sacrifices, unconditional love, and unwavering belief in me have made this achievement possible. I am profoundly grateful for everything they have done to help me reach this milestone.

Abstract

Digital pathology presents today a fundamental tool for medical diagnosis, exploiting technological advances in digitalization to transform biological samples into digital data, thus facilitating their visualization and analysis. However, these methods, often based on conventional microscopy, encounter limitations that sometimes hinder their effectiveness.
From this perspective, unconventional imaging methods such as Fourier ptychographic microscopy offer promising prospects for overcoming these limitations. Indeed, FPM offers access to the phase in complement of the intensity and allows examining a large Field of View at a high resolution at a reasonable design cost.
This thesis explores Fourier ptychographic microscopy (FPM) 's potential in thin blood smear analysis. Several results have been obtained thanks to a multidisciplinary approach integrating deep learning and microscopy. We have first focused our attention on the problem of limited complexity of parasite detection for malaria diagnosis. The joint exploitation of intensity and phase is shown to improve the performance of a deep network detector. To this end, a complex-valued CNN has been introduced in Faster-RCNN architecture for efficient feature extraction.
Secondly, we have considered a more complex application, namely the classification of white blood cells, where the benefits of joint exploitation of intensity and phase were also confirmed. Furthermore, to reduce the imbalance of classes encountered in this task, we propose a novel physics-informed GAN model dedicated to generating intensity and phase images. This model avoids the mode collapse problem faced with usual GAN implementation.
Finally, we have considered optimizing the FPM microscope design. To this end, we explore strategies combining simulations, neural networks, and image formation modeling. We demonstrate that FPM can use low resolutions without significantly compromising performance.

This thesis underscores the interest in tailoring machine learning in connection to microscopy principles and highlights the potential of Fourier ptychographic microscopy for future automated diagnosis systems.

Résumé

La pathologie numérique constitue aujourd'hui un outil fondamental pour le diagnostic médical, exploitant les avancées technologiques en matière de numérisation pour transformer les échantillons biologiques en données numériques, facilitant ainsi leur visualisation et leur analyse. Cependant, ces méthodes, souvent basées sur la microscopie conventionnelle, rencontrent des limitations qui entravent parfois leur efficacité.
Dans ce contexte, des méthodes d'imagerie non conventionnelles telles que la microscopie ptychographique de Fourier (FPM) offrent des perspectives prometteuses pour surmonter ces limitations. En effet, la FPM offre un accès à la phase en complément de l'intensité et permet d'examiner un large champ de vision à haute résolution à un coût de conception raisonnable.
Cette thèse explore le potentiel de la microscopie ptychographique de Fourier dans l'analyse des frottis sanguins minces. Plusieurs résultats ont été obtenus grâce à une approche multidisciplinaire intégrant l'apprentissage en profondeur et la microscopie.
Nous nous concentrons d'abord sur le problème limité de la détection des parasites pour le diagnostic du paludisme. L'exploitation conjointe de l'intensité et de la phase permet d'améliorer les performances d'un détecteur de réseau neuronal profond. À cette fin, un CNN à valeurs complexes a été introduit dans l'architecture Faster-RCNN pour une extraction efficace des caractéristiques.
Ensuite, nous examinons une application plus complexe, à savoir la classification des globules blancs, où les avantages de l'exploitation conjointe de l'intensité et de la phase ont également été confirmés. Nous nous intéressons également au problème du déséquilibre des classes rencontré dans cette tâche, nous proposons un nouveau modèle GAN informé par la physique dédié à la génération d'images d'intensité et de phase. Ce modèle évite le problème de mode collapse rencontré avec l'implémentation habituelle des GAN.
Enfin, nous considérons l'optimisation de la conception du microscope FPM. À cette fin, nous explorons des stratégies combinant simulations, réseaux neuronaux et modélisation de la formation d'images. Nous démontrons que la

FPM peut utiliser des résolutions faibles sans compromettre significativement les performances.
Cette thèse souligne l'intérêt d'adapter l'apprentissage automatique en lien avec les principes de la microscopie et met en évidence le potentiel de la microscopie ptychographique de Fourier pour les futurs systèmes de diagnostic automatisés.

Synopsis en français

This synopsis is provided in compliance with the 1994 law on the use of the French language. It outlines the structure of the thesis and summarizes its chapters and contributions.
Ce synopsis est fourni en conformité avec la loi de 1994 relative à l'emploi de la langue française. Il reprend la structure de la thèse et résume les chapitres et les contributions de la thèse.

Introduction

La pathologie numérique se développe rapidement dans le domaine médical. Elles se base sur des technologies de numérisation pour convertir les spécimens biologiques, des tissus aux échantillons de cellules, en ensembles de données numériques complets. L'objectif premier de la pathologie numérique est de faciliter l'amélioration des capacités d'analyse. Avec la croissance exponentielle de l'information médicale, la pathologie numérique apparaît comme une solution permettant aux professionnels de la santé de gérer et d'interpréter efficacement de larges volumes de données numériques. Cette transition vers les plateformes numériques rationalise le processus de diagnostic et libère le potentiel des techniques informatiques avancées, principalement grâce aux réseaux neuronaux profonds. Ce processus analytique va au-delà des approches diagnostiques conventionnelles, ouvrant la voie à des techniques innovantes qui s'appuient sur la fusion d'information provenant de sources disparates. En intégrant divers ensembles de données et en exploitant des algorithmes avancés, la pathologie numérique permet aux cliniciens d'adopter une nouvelle perspective, facilitant ainsi des diagnostics plus nuancés et plus complets. L'analyse numérique des frottis sanguins est considérée comme un outil efficace pour diagnostiquer plusieurs pathologies, qu'il s'agisse de cancers, d'infections ou de troubles hématologiques. Plusieurs systèmes automatisés ont été développés à

cette fin. Ils utilisent des images de microscopie conventionnelle et des analyses d'apprentissage automatique pour compter et caractériser morphologiquement les différents types de cellules sanguines, telles que les globules blancs, les globules rouges et les plaquettes. Parmi ces systèmes :

- Cellavision commercialise un système d'hématologie complet capable de classer de nombreuses catégories de globules blancs. Malgré des progrès significatifs, les biologistes rencontrent des difficultés en raison de performances limitées, nécessitant des retours fréquents à la microscopie manuelle.

- Easyscango est un système commercial pour le diagnostic du paludisme combinant l'analyse de frottis sanguins épais et fins pour garantir une haute sensibilité, mais le traitement est complexe et chronophage.

- Noul propose un système mobile compatible avec l'hématologie et la parasitologie, acquérant des images à grand champ de vision, mais ses performances sont encore trop limitées pour une large diffusion.

Ces automates de pathologie numérique présentent des inconvénients qui limitent leur adoption par les biologistes :

- En microscopie conventionnelle, il est difficile de contrôler la qualité de l'image sur l'ensemble du champ de vision, ce qui peut entraîner une mauvaise interprétation des caractéristiques cellulaires.

- En raison d'une zone de mesure trop restreinte, le diagnostic fiable est compromis, notamment lors de la recherche d'événements rares.

- Les informations fournies sont uniquement qualitatives.

La microscopie non conventionnelle représente une avancée significative dans le domaine de l'imagerie, capable de répondre à certaines limitations de la microscopie conventionnelle. Elle se réfère à des techniques et dispositifs optiques innovants qui s'écartent des optiques géométriques et physiques traditionnelles, incluant des méthodes avancées de traitement d'images, l'optique adaptative, l'imagerie computationnelle et l'holographie. Parmi les méthodes récentes, on peut citer l'imagerie de phase quantitative (QPI) basée sur des dispositifs interférométriques, présentant plusieurs avantages :

- Une résolution améliorée grâce à un mécanisme d'ouverture numérique synthétique.
- Une imagerie quantitative en intensité et en phase fournissant des informations supplémentaires sur les propriétés optiques des échantillons biologiques.
- Une correction des aberrations via la manipulation numérique des fronts d'onde.
- Une profondeur de champ étendue grâce à la manipulation des fronts d'onde post-acquisition.

Cependant, la mise en œuvre de la QPI dans des environnements non contrôlés est difficile en raison de la sensibilité de ces méthodes aux vibrations. Une alternative prometteuse est la microscopie ptychographique de Fourier (FPM), une technique d'imagerie computationnelle simple à mettre en œuvre, sans nécessiter de configuration interférométrique et offrant des avantages théoriques de la QPI avec des spécificités supplémentaires, comme une stabilité accrue et un coût réduit.
L'objectif de cette thèse est d'explorer le potentiel de la FPM dans l'analyse des frottis sanguins, à travers des expériences visant à détecter le parasite responsable du paludisme et à classifier différents types de globules blancs, en utilisant des méthodes d'apprentissage profond adaptées aux spécificités de la FPM. La thèse est structurée autour de plusieurs axes, notamment l'évaluation de l'impact de la FPM pour la classification automatique et l'optimisation de la configuration du microscope et des réseaux de neurones profonds pour ces tâches.
Les contributions de cette thèse incluent :

1. Le développement d'une méthodologie basée sur des réseaux neuronaux profonds pour traiter les images d'intensité et de phase.
2. La quantification de l'apport de la FPM pour améliorer la détection des globules rouges infectés et la classification des globules blancs.
3. La proposition d'un modèle GAN informé par la physique pour synthétiser de nouvelles images bimodales de FPM.
4. L'optimisation de la configuration du microscope pour l'amélioration de la tolérance aux résolutions faibles grâce à l'ajout d'images de phase.

Ces différentes contributions sont présentées de façon synthétique dans les sections suivantes.

Contribution de la phase pour l'analyse des objets colorés

Dans la littérature, l'introduction de la microscopie bimodale (intensité et phase) est liée principalement à la capacité de l'imagerie de phase à révéler des structures invisibles sur des échantillons, sans nécessité de marquage. Ceci est particulièrement intéressant pour les échantillons transparents. Dans le premier chapitre de ce travail, contrairement aux cas d'application couverts dans les travaux publiés précédemment, nous nous interrogeons sur la pertinence de la microscopie bimodale, plus précisement du couplage intensité et phase, dans le contexte des échantillons colorés où souvent l'information d'intensité est décrite comme suffisante.
Pour répondre à cette question, on considère une application biologique : la détection automatique de malaria qui consiste en une localisation et identification des globules rouges parasités sur des images de frottis colorés numérisées par un microscope. Ce sujet est choisi parce qu'il n'est pas encore résolu malgré sa complexité relativement limitée (deux classes d'objets à détecter, globule rouge sain / globule rouge parasité). Malgré plusieurs études , les sensibilités atteintes sont jugées encore insuffisantes pour des raisons qui ne sont pas complètement établies. Parmi les principales hypothèses, est évoquée la difficulté posée par la microscopie à produire des images de qualité parfaitement contrôlée sur un champ de vision complet avec une information suffisament discriminante.
Dans ce contexte, les images de phase sont donc intéressantes car elles peuvent révéler de fines variations dans l'épaisseur optique de l'échantillon mesuré, une information supplémentaire à celle fournie par les images d'intensité qui codent l'information d'absorption. Ce couplage pourrait faciliter la distinction entre un globule rouge sain et un globule rouge parasité.
Une étude de performance en termes de sensibilité et de spécificité est réalisée. Les performances obtenues avec le couple intensité et phase sont comparées à celles reposant sur les seules images d'intensité. Comme mentionné dans la partie introductive, tout au long de cette thèse, nous nous concentrons sur la technique la FPM comme méthode d'imagerie. Elle sera employé pour l'acquisition des images d'intensité et de phase car elle est facile à mettre en œuvre et très

tolérante aux vibrations.
Le protocole expérimental repose d'abord sur une acquisition et une reconstruction d'une base d'image spécifique à la tâche. Des frottis sanguins minces de 9 patients infectés par la malaria sont scannés avec un microscope FPM. Chaque champ de vision (FOV) est acquis avec 13 images brutes, où chaque image est obtenue en éclairant l'objet avec une LED avec un angle d'illumination spécifique. Les images bimodales (intensité et phase) sont reconstruites à l'aide de l'algorithme e-PIE. Les champs de vue reconstruits sont ensuite découpés en plus petites images contenant environ 200 à 300 globules rouges en moyenne. Un expert étiquette chaque image en indiquant la position de chaque globule rouge à l'aide d'une boite et en fournissant un label indiquant la présence ou l'absence de parasite.
Pour la tâche de détection, une architecture Faster-RCNN est employée. Plus précisément deux modèles Faster R-CNN sont développés et entraînés pour détecter et classer les globules rouges, il diffèrent dans le type d'entrée exploitée :

- Un modèle utilisant uniquement les images d'intensité ($\mathcal{I}$-RV).
- Un modèle utilisant les images d'intensité et de phase ($\mathcal{I}/\phi$-RV).

Les hyper-paramètres des modèles sont optimisés pour maximiser les performances de détection. Les modèles sont entraînés et évalués en utilisant une validation croisée à cinq plis. Les performances des modèles sont évalués en termes de taux de vrais positifs (TPR) et de taux de vrais négatifs (TNR).
Les résultats montrent que le modèle $\mathcal{I}/\phi$-RV, qui utilise les images d'intensité et de phase simultanement, améliore significativement la détection des globules rouges parasités par rapport au modèle $\mathcal{I}$-RV utilisant uniquement les images d'intensité. Le modèle $\mathcal{I}/\phi$-RV affiche un taux de détection des parasites de 95.22% contre 93.33% pour le modèle $\mathcal{I}$-RV. Les résultats illustrent donc la contribution de l'information de phase à l'amélioration de la qualité du diagnostic des événements rares. Dans cette étude de cas, nous observons que l'image de phase apporte une information complémentaire à l'intensité, même dans le cas d'objets non transparents, où la phase est habituellement considérée comme ayant un effet mineur. Ce premier chapitre montre l'impact positif des images bimodales intensité-phase et le valide dans une étude de cas relativement simple mais non résolue sur le diagnostic du paludisme. Un cas d'étude plus complexe sera étudié dans le chapitre suivant dans le but de chercher d'une généricité.

Classification des globules blancs

Dans le chapitre précédent, une preuve de concept a montré la contribution de l'utilisation conjointe de la phase et de l'intensité obtenues en FPM pour les échantillons colorés en étudiant le problème simple de la détection des parasites du paludisme dans les globules rouges. Ce chapitre étend ces résultats prometteurs à un problème plus complexe : la classification des globules blancs à partir de frottis sanguins fins colorés. Ce problème est considéré comme plus complexe que la détection des parasites car il implique 5 classes qui se ressemblent fortement, avec des déséquilibres biologiques naturels dans la représentation de ces classes dans le sang.
L'objectif principal est de classifier de manière automatique les types de leucocytes (globules blancs) que l'on trouve généralement dans le sang des individus sains, notamment les neutrophiles, les éosinophiles, les basophiles, les lymphocytes et les monocytes.
La première partie du travail vise à étendre la preuve de concept de l'efficacité de l'information bimodale dans la classification des échantillons colorés à un problème plus complexe.
La seconde partie propose une méthode innovante de génération d'images synthétiques pour compléter les classes de globules blancs sous-représentées, visant à améliorer les performances de classification. L'originalité de cette méthode réside dans l'introduction préalable du modèle de formation d'images physiques dans un modèle GAN qui permet d'augmenter la diversité des images générées et de réduire le phénomène d'effondrement de mode. Ce modèle est couplé à un protocole d'entraînement spécifique qui exploite des similarités entre certaines classes pour proposer une approche fine-tuning. Ce protocole offre la possibilité d'apprendre des GAN pour synthétiser des images de classes avec seulement un nombre très limité d'échantillons.

Classification des globules blancs en FPM

L'objectif principal est de développer et d'évaluer un système robuste de classification automatique des globules blancs chez les patients ne présentant pas de troubles hématologiques.
La classification porte sur cinq types de globules blancs : les neutrophiles, les lymphocytes, les monocytes, les éosinophiles et les basophiles.

L'étude vise également à surmonter les limites des systèmes automatiques actuels, qui nécessitent souvent une vérification manuelle par des spécialistes.

Protocole expérimental

Le protocole expérimental employé comprends plusieurs étapes :

- Acquisition de données FPM : Un ensemble d'images de globules blancs a été créé en utilisant la microscopie FPM. Des frottis sanguins de 129 patients en unité de soins intensifs ont été scannés à trois longueurs d'onde pour produire 3 couples des images d'intensité et de phase. Une base de 27326 globules blancs a été recueillie et étiquetée par un spécialiste en hématologie.

- Acquisition d'une base d'images en microscopie classique : les mêmes frottis sanguins numérisés en FPM ont également été numérisés en utilisant un microscope classique (Di-60 de Cellavision) afin de permettre une comparaison avec les systèmes existants.

- Développement et entrainement des modèles: 4 modèles de réseaux neuronaux convolutionnels (CNN) ont été entrainés pour la tâche de classification, en utilisant une même architecture, inspirée du VGG-16 et simplifiée afin de l'adapter à la complexité du problème.

  1. I-CNN: Un modèle utilisant uniquement les images d'intensité.
  2. ϕ-CNN : Un modèle utilisant uniquement les images de phase.
  3. I-ϕ-CNN : Un modèle utilisant conjointement les images d'intensité et de phase.
  4. Di-60-CNN : Un modèle utilisant les images de microscopie classique.

  Les modèles sont comparés à travers trois indicateurs de performance : la précision totale, la précision moyenne par classe et le rappel moyen par classe.

Résultats et conclusion

Les résultats des expériences ont montré que le modèle I-ϕ-CNN, qui utilise conjointement les images d'intensité et de phase, surpasse les autres modèles enrainés sur une modalité des images FPM (I-CNN et ϕ-CNN). Le modèle I-ϕ-CNN est particulièrement efficace pour identifier les classes sous-représentées telles que les éosinophiles et les basophiles. L'analyse d'un t-SNE a confirmé que le modèle I-ϕ-CNN permet une meilleure séparation des classes de globules blancs dans l'espace des caractéristiques, par rapport aux autres modèles.
La comparaison du modèle I-ϕ-CNN avec le modèle Di-60-CNN met en évidence la supériorité notable du modèle I-ϕ-CNN par rapport au modèle Di-60-CNN, en particulier pour les classes des basophiles, des éosinophiles et des monocytes. Cette différence significative peut être attribuée à un problème de déséquilibre des données entre les classes, ce qui affecte davantage le modèle Di-60-CNN, entraînant des performances inférieures dans la reconnaissance et la séparation de ces classes, contrairement au modèle I-ϕ-CNN.
Les résultats montrent que l'exploitation conjointe des informations d'intensité et de phase améliore de manière significative les performances de la classification des globules blancs. L'amélioration de la précision et de la fiabilité suggère que le modèle I-ϕ-CNN a le potentiel de réduire la nécessité d'une vérification manuelle par des spécialistes, faisant ainsi progresser les systèmes de classification automatique.

Génération d'images bi-modales synthétiques en utilisant les GAN

L'étude de classification précédente a démontré une amélioration des performances en combinant les données d'intensité et de phase dans les réseaux neuronaux pour classer les globules blancs. Bien que le modèle soit performant dans la classification des classes les plus représentées, telles que les neutrophiles et les lymphocytes, sa précision doit encore être améliorée pour les classes moins représentées. En particulier, des performances insuffisantes peuvent être observées dans la classe des basophiles.
Pour améliorer les performances de cette classe, un protocole d'augmentation du nombre d'échantillons en utilisant les modèles génératifs de type GAN est employé. Compte tenu des difficultés d'implémentation et d'entraînement des

GAN classiques (dans le cas d'images bi-modales), tel que l'effondrement de mode (mode collapse) et l'instabilité, une approche innovante est explorée : un GAN informé par la physique, dans lequel les équations de formation de l'image (modèle direct de la FPM) sont insérées dans le formalisme neuronal. Combiné à une procédure de fine-tuning, ce modèle peut synthétiser des exemples supplémentaires qui enrichissent les classes sous-représentées.

Méthodologie et résultats

L'objectif principal de notre étude est de générer des images de basophiles à la fois en intensité et en phase. Étant donné le nombre limité d'échantillons disponibles pour cette classe spécifique, nous avons choisi de commencer par entraîner notre GAN sur une classe similaire plus abondante : les neutrophiles. Ce choix initial permet de former et d'optimiser un premier GAN avant de réaliser un fine-tuning par la suite sur les quelques exemples spécifiques de la classe des basophiles.
Initialement, cette approche du fine-tuning a été employée pour entrainer un GAN classique. Ce GAN classique est composé de deux modèles : un générateur et un discriminateur. Le générateur prend en entrée un bruit aléatoire et produit simultanément deux canaux. Le premier canal est associé à l'intensité de l'objet, tandis que le second correspond à l'image de phase. Le discriminateur prend en entrée les deux canaux représentant l'intensité et la phase générées par le générateur ou extraites de l'ensemble de données FPM. Ensuite, à l'aide d'une architecture neuronale, le discriminateur extrait les caractéristiques pertinentes des images d'entrée, qui sont ensuite utilisées pour classer les exemples comme réels ou synthétiques. La tâche de classification est ensuite utilisée pour calculer des fonctions de coût et effectuer la rétropropagation pour le générateur et le discriminateur.
Dans ce contexte, le principal défi consiste à concevoir un discriminateur qui interprète correctement les informations d'intensité et de phase des images. Le discriminateur doit apprendre à extraire les caractéristiques pertinentes de ces deux canaux afin d'encoder correctement les caractéristiques essentielles de l'intensité et de la phase et d'effectuer une classification qui tient compte de ces deux informations.
Il est important de noter que l'intensité et la phase sont deux types d'informations provenant de domaines de représentation différents (absorption et chemin optique). Malgré leur corrélation, elles ont des plages de valeurs et des distribu-

tions différentes, même après normalisation. Cette divergence dans les espaces de représentation peut rendre la tâche de génération d'images bimodales (intensité et phase) particulièrement complexe.
Les résultats ont été jugés insuffisants car ils ne présentaient qu'une variabilité limité. Malgré l'incorporation de la perte de Wasserstein et d'autres techniques de régularisation, les résultats obtenus restaient peu divers. Les modèles entrainés rencontraient un problème d'effondrement de mode (mode collapse), où le générateur produisait des images trop similaires et peu diversifiées, ne capturant pas la variabilité attendue dans les données réelles. De plus, ce type d'approche n'apporte aucune amélioration de performance pour les classes sous-représentées. Ce constat est en accord avec avec la littérature, qui indique une difficulté quant à l'apprentissage des données bi-modales à travers des modèles GANs.
Pour surmonter ce problème, une approche innovante est explorée : le GAN informé par la physique. Le générateur produit des canaux d'intensité et de phase similairement au GAN classique. Mais, le discriminateur ne traite pas directement la sortie du générateur. Dans notre cas, il est précédé d'un modèle direct de la FPM qui génère une pile d'images correspondante à ce que le microscope FPM capturera de l'objet observé en l'éclairant par les différentes LEDs. Le discriminateur prend cette pile d'images en entrée et les classe comme réelles ou synthétiques en utilisant l'architecture CNN.
Tout d'abord, en intégrant des connaissances physiques sur la relation entre l'intensité et la phase de l'image, le modèle est mieux équipé pour régulariser la génération d'images en assurant la cohérence entre ces deux aspects. En outre, le discriminateur du modèle informé par la physique est conçu pour prendre en compte plusieurs représentations des images, ce qui permet une évaluation plus précise de leur réalisme et de leur cohérence avec les images réelles. Cette approche facilite l'apprentissage de caractéristiques de données plus subtiles et plus complexes, améliorant ainsi la différenciation entre les images réelles et synthétiques.
L'approche informée par la physique a montré des résultats supérieurs dans la génération d'images bimodales . Les images produites étaient plus diversifiées et réalistes. Les scores FID pour le GAN informé par la physique étaient en moyenne 25% plus bas que ceux du GAN classique, ce qui indique une réduction significative de la différence entre les distributions des images générées et des images réelles.
De plus, les résultats montrent que l'augmentation des données sur les ba-

sophiles à l'aide d'un GAN a un impact significatif sur la précision de la classification. Le modèle avec augmentation grâce aux images issues du GAN montre une précision significativement plus élevée pour les basophiles par rapport au modèle sans augmentation. Plus précisément, le modèle sans augmentation atteint une précision de 94,00% avec un écart-type de 0,024, tandis que le modèle avec augmentation atteint une précision remarquablement améliorée de 99,37% avec un écart-type de 0,002. Cette augmentation substantielle de la précision, d'environ 5%, indique que l'ajout d'images synthétiques a considérablement amélioré la capacité du modèle à classer correctement les basophiles.
L'approche GAN informé par la physique se distingue par sa capacité à améliorer la variabilité du GAN tout en limitant le phénomène d'effondrement de mode. Cette approche innovante génère des images qui complètent les classes sous-représentées, telles que les basophiles, dans cette étude, améliorant ainsi de manière significative les performances de classification. Cette combinaison met en évidence l'importance des GANs informées par la physique et de l'approche de fine-tuning pour limiter les défis de la classification des cellules sanguines.

Optimisation du système pour une application opérationnelle

Dans les chapitres précédents, la contribution de la microscopie FPM pour la classification des cellules sanguines a été explorée, démontrant son potentiel pour deux applications distinctes. En exploitant la bimodalité, la microscopie FPM a considérablement amélioré les performances du modèle, mettant en évidence une synergie entre l'intensité et la phase. Toutefois, pour passer à un système opérationnel, une optimisation globale du système est nécessaire.
La dernière partie du travail propose donc des études visant à optimiser le système selon deux axes principaux. Tout d'abord, l'accent est mis sur l'optimisation de l'exploitation des informations bimodales dans les réseaux de neurones. Bien que l'information bimodale ait révélé un gain de performance significatif, il reste à savoir comment exploiter pleinement son potentiel dans les modèles de traitement d'images. Deuxièmement, l'optimisation de la configuration optique du microscope, en particulier le choix de l'objectif, est abordé. L'objectif est de trouver le meilleur compromis entre la taille du champ de vision, la résolution et la performance de la classification. Cette approche combine des données expérimentales et un modèle de formation d'images optiques pour prédire

l'observation des échantillons et étudie les performances du modèle sur ces observations.
Cette approche d'optimisation du système par modélisation à partir de données expérimentales est illustrée par une application spécifique : la détection de la malaria. Cependant, les principes et les méthodes présentés s'appliquent à toute application de microscopie nécessitant une amélioration du champ de vision et des performances de détection.

Réseau de neurones à valeurs complexes

Le FPM permet d'accéder à des données bimodales en reconstruisant à la fois l'intensité et les informations de phase. L'approche la plus classique des réseaux neuronaux considère ces deux images comme des canaux d'information indépendants. En effet, elles représentent deux paramètres physiques distincts résultant de l'interaction lumière-matière. Typiquement, une image code les paramètres d'absorption (intensité) tandis que l'autre représente le chemin optique (phase). Dans ce contexte, il est courant d'appliquer des convolutions classiques à valeurs réelles pour extraire indépendamment les informations pertinentes de chaque image et les fusionner ensuite dans les différentes couches du réseau.
Toutefois, les images bimodales acquises peuvent également être considérées comme les composantes d'une matrice à valeurs complexes appelée matrice de transfert. Cette matrice est étroitement liée au champ électromagnétique complexe de l'échantillon observé. Dans cette optique, l'utilisation de convolutions à valeurs complexes devient une second option intéressante. Inspirée par le filtrage complexe en physique, qui s'est avéré révéler des détails (par exemple, dans l'imagerie en champ sombre), l'intégration de convolutions à valeurs complexes dans les réseaux neuronaux permet d'extraire simultanément des informations des deux modalités grâce à leur représentation complexe.
L'étude du formalisme le plus adapté pour l'exploration conjointe des images d'intensité et de phase dérivées de la reconstruction FPM implique une analyse utilisant des représentations à valeurs réelles et complexes dans le cadre du Faster-RCNN. Le Faster-RCNN à valeurs complexes est obtenu à partir de l'architecture à valeurs réelles en remplaçant les convolutions de la branche partagée par des convolutions à valeurs complexes, en remplaçant la fonction d'activation par CReLU et utilisant un type de pooling adapté. L'objectif ap-

plicatif est resté le même. Premièrement, localiser les globules rouges dans ces images et deuxièmement, les classer dans des catégories distinctes, à savoir les parasites ou les cellules saines. Deux modèles Faster-RCNN ont été mis en œuvre pour la détection des parasites *Plasmodium Falciparum*.

- Un Faster-RCNN classique à valeurs réelles utilisant l'intensité et la phase ($\mathcal{I}/\phi$-RV)
- Un Faster-RCNN à convolution à valeurs complexes utilisant l'intensité et la phase ($\mathcal{I}/\phi$-CV).

L'évaluation des modèles se fait en évaluant le taux de vrais négatifs (TNR) et le taux de vrais positifs (TPR). Les résultats mettent en évidence plusieurs avantages clés du modèle $\mathcal{I}/\phi$-convolutions à valeurs complexes par rapport au modèle $\mathcal{I}/\phi$-RV. I-ϕ-convolutions à valeurs complexes détecte 350 cellules infectées et 187 cellules saines de plus que $\mathcal{I}/\phi$-RV. ϕ-RV, ce qui démontre des capacités de classification supérieures, particulièrement utiles pour identifier les cellules infectées peu nombreuses parmi les cellules saines. $\mathcal{I}/\phi$-convolutions à valeurs complexes présente moins de cellules infectées mal classées, ce qui indique une meilleure différenciation entre les cellules infectées et les cellules saines. Les catégories saines et infectées présentent moins de cellules manquées avec $\mathcal{I}\phi$-convolutions à valeurs complexes, ce qui suggère une meilleure sensibilité de détection. Les deux modèles présentent un TNR robuste (>99%), reflétant une détection précise des globules rouges sains. Le réseau neuronal à valeurs complexes surpasse le réseau à valeurs réelles d'environ 2% en termes de TPR, ce qui démontre une sensibilité accrue dans la détection des cellules infectées par le paludisme. Les intervalles de confiance plus étroits avec les convolutions à valeurs complexes soulignent ses capacités de généralisation supérieures. L'étude démontre que l'intégration de réseaux neuronaux à valeurs complexes améliore la détection des globules rouges infectés par le paludisme à partir d'images bimodales FPM. Cette approche améliore la précision de la classification, la sensibilité et la vitesse de convergence, ce qui promet des avancées dans le domaine de l'imagerie biomédicale et des outils de diagnostic.

Optimisation du microscope ptychographique de Fourier pour la classification du paludisme : choix de l'objectif

Cette partie se concentre sur l'optimisation de la configuration du microscope en étudiant les performances du réseau neuronal pour différentes ouvertures numériques. Cette approche vise à déterminer l'ouverture numérique minimale viable pour l'application spécifique du diagnostic du paludisme. Notre étude se limite à la classification binaire des globules rouges en cellules infectées et saines.

Processus d'optimisation du choix de l'objectif

Différents réseaux neuronaux doivent être formés pour étudier les performances en fonction de la résolution de l'image dans le cadre du FPM. Pour entraîner ces classificateurs, il faut créer différents ensembles de données à diverses résolutions. Habituellement, cela nécessite l'acquisition de données FPM pour chaque objectif considéré et la reconstruction de leurs cartes intensité-phase. Pour éviter ce processus coûteux, un protocole spécifique pour dériver les bases de données nécessaires à partir d'une base de données initiale, reposant sur un modèle de formation d'image physique, est proposé.
Le processus commence par l'acquisition d'un ensemble de données de globules rouges recueillies auprès de 9 patients infectés. Le champ de vision de la zone de bonne diffusion a été capturé à l'aide du dispositif FPM décrit dans la section 3.3.1, ce qui a permis de reconstruire des images d'intensité et de phase pour l'ensemble du champ de vision avec une ouverture synthétique de 0.9 (NA = 0.45). Ensuite, les globules rouges ont été sélectionnés au hasard et extraits pour générer des images individuelles de taille 64 $\times$64. Cet ensemble de données comprend 12 000 cellules, réparties de manière égale entre 6 000 cellules saines et 6 000 cellules infectées par le parasite. Cet ensemble de données est ensuite combiné avec les modèles de formation et de reconstruction d'images pour déduire les ensembles de données à plus faible résolution.
Pour chaque valeur de NA (NA entre 0.1 et 0.45), les images des 6000 globules rouges sains et parasités sont dégradées en résolution en utilisant le modèle de formation d'image impliqué dans l'algorithme de reconstruction FPM. Cela correspond à un filtrage complexe de l'image dans l'espace de Fourier. La forme du filtre consiste en un disque centré sur (0,0) avec un rayon $r = 2\pi NA/\lambda$. Un ensemble de données contenant des images d'intensité et de phase à la résolu-

tion correspondante est généré pour chaque NA.
Les ensembles de données produits selon le protocole décrit ci-dessus sont utilisés pour former les deux modèles neuronaux ($\mathcal{I}$-RV-CNN et $\mathcal{I}/\phi$-CV-CNN) partageant la même architecture. Le premier modèle utilise exclusivement les données d'intensité, tandis que le second traite les informations d'intensité et de phase. Notamment, pour le second modèle, nous optons pour une architecture à valeurs complexes. Les résultats montrent que la performance de classification est systématiquement meilleure lorsque les images d'intensité et de phase sont utilisées dans un CNN à valeurs complexes. Deuxièmement, l'évolution des courbes est différente dans les deux cas. En particulier, les performances pour le modèle $\mathcal{I}/\phi$-convolutions à valeurs complexes est relativement stable sur une large gamme de NA (jusqu'à NA=0.2), et la pente de la tangente est proche de zéro. Ce n'est pas le cas du réseau I-RV, qui présente une détérioration avec la diminution de NA ; la valeur de la pente est égale à 7.34 et 3.82 pour la sensibilité et la spécificité, respectivement. Enfin, lorsque NA passe en dessous de 0.15, les performances chutent brutalement pour les deux modèles. Pour le diagnostic automatique du paludisme, les résultats suggèrent qu'un objectif avec une ouverture numérique aussi faible que 0,2 (accessible avec un objectif à grossissement 4x) pourrait être une option praticable. Avec un tel objectif, le champ de vision associé est environ 600 fois plus grand que celui d'un objectif 100x, qui est la norme pour l'examen visuel des frottis sanguins par les experts. Pour $NA = 0,2$, nous estimons la sensibilité et la spécificité à 98,1% et 99,3% respectivement. Ces chiffres pourraient être encore améliorés avec un ensemble de données d'apprentissage plus étendu.

Conclusion

Dans le cadre d'une recherche multi-disciplinaire, ce travail étudie le benécifice d'une méthode de microscopie non conventionnelle, la FPM, dans deux applications médicales confrontées aux limites de la microscopie classique. La FPM est une méthode d'imagerie de phase quantitative, stable et peu coûteuse, adaptée à une utilisation clinique de routine. Elle permet d'obtenir à la fois des images d'intensité et de phase en utilisant des modèles de reconstruction adéquats. L'objectif principal de cette thèse était de montrer comment l'information enrichie fournie par le cadre FPM peut aider à améliorer le diagnostic et le suivi des pathologies hématologiques et parasitologiques. Pour deux cas d'utilisation

spécifiques (détection de parasites pour le diagnostic de la malaria et analyse des globules blancs), nous avons exploré différents aspects liés à la conception de systèmes de détection basés sur le FPM dans les frottis sanguins. Dans la première partie, nous avons étudié l'intérêt de la phase pour l'analyse des objets colorés. Certaines hypothèses théoriques ont été avancées et validées sur le cas d'utilisation limité de la détection de parasites (classification en 2 classes). Les résultats démontrent que le couplage de l'intensité à la phase est plus efficace que l'utilisation de l'intensité seule. Cette question est étudiée sur une application plus complexe dans la seconde partie, à savoir la classification automatique des globules blancs. Les résultats montrent que le modèle combinant l'intensité et la phase est plus performant que ceux qui exploitent une seule modalité en termes de précision et de rappel où une amélioration allant jusqu'à 30% est observée, ce qui permet de mieux classer les différents types de globules blancs et de mieux séparer les classes. En outre, les images fournies par ces méthodes offrent de meilleures performances que celles des automates actuels, avec une amélioration des performances de 30% pour certaines classes. Les travaux ont également abordé d'autres problèmes limitants dans certaines applications biomédicales, comme le problème du déséquilibre des classes, fréquemment rencontré dans les applications biomédicales. Après des tentatives infructueuses de mise en œuvre d'un modèle GAN classique capable de générer des images FPM bimodales de bonne qualité, nous avons proposé une approche innovante, un GAN inspiré de la physique. En utilisant le modèle de formation d'images FPM couplé à des modèles génératifs adversaires, cette approche GAN informée par la physique a été capable de générer des images proches de la réalité avec une grande diversité, ce qui a permis de répondre au problème d'effondrement de mode. De plus, cette approche GAN informée par la physique, couplée à une procédure de fine-tuning a pu générer des images synthétiques pour une classe contenant initialement très peu de représentants. Nos résultats ont montré que l'intégration de données synthétiques améliorait significativement les performances de classification, en particulier pour les classes minoritaires, sans dégrader les performances des classes non augmentées. La dernière partie est consacrée à quelques tentatives d'optimisation du cadre FPM précédent vers un outil de diagnostic efficace. Tout d'abord, nous avons implémenté des couches de neurones convolutifs à valeurs complexes dans un RCNN plus rapide afin de détecter les globules rouges parasites. Les résultats montrent une amélioration des performances de détection des problèmes, portant la sensibilité à 97.15%. Il a également permis une convergence

plus rapide et une meilleure généralisation avec des intervalles de confiance complexes par rapport aux modèles à valeurs réelles. La deuxième partie a été consacrée à l'optimisation du choix de l'objectif du microscope en mettant l'accent sur l'équilibre entre la résolution et la vitesse d'acquisition des données, c'est-à-dire la taille du champ de vision. Les résultats ont démontré que les images FPM permettent de résister à une faible résolution et à une augmentation de la taille du champ de vision. Il est suggéré qu'une ouverture numérique aussi faible que 0.2 peut être suffisante pour obtenir des performances élevées tout en réduisant le temps d'acquisition.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contents

1.1 Motivation

Digital pathology rapidly expands within the medical field, leveraging scanning technology to convert biological specimens from tissue to cell samples into comprehensive digital datasets. The primary objective of digital pathology is to facilitate enhanced analytical capabilities. With the exponential growth of medical information, digital pathology emerges as a solution, enabling healthcare professionals to manage and interpret vast volumes of numerical data efficiently. This transition to digital platforms streamlines the diagnostic process and unlocks the potential for advanced computational techniques, mainly through deep neural networks. This analytical process extends beyond conventional diagnostic approaches, paving the way for innovative techniques that rely on cross-referencing information from disparate sources. By integrating diverse datasets and leveraging advanced algorithms, digital pathology enables clinicians to adopt a novel perspective, facilitating more nuanced and comprehensive diagnoses.
Digital blood smears analysis is foreseen as an efficient tool to diagnose several pathologies, ranging from cancers or infections to hematological disorders. Several automated systems have been developed to this end. They use conventional microscopy images and machine learning analysis to count and morphologically characterize the different types of blood cells, such as white blood cells, red blood cells, and platelets. Among those systems:

- Cellavision commercializes a complete hematology system that can classify many White Blood Cell classes. Despite significant progress over the last ten years, biologists report difficulties due to limited performance. As a result, frequent back and forth tomanuel microscopy is necessary.

- Easyscango is a commercial system for malaria diagnosis that combines thick and thin blood smear analysis to guarantee high sensitivity. The processing is, therefore, time-consuming and complex.

- Noul proposes a mobile system compatible with hematology and parasitology, acquiring large field-of-view images. However, its performance is still too limited for its widespread diffusion to biologists.

We can see from the above descriptions that these digital pathology automata suffer from some drawbacks that limit their appropriation by biologists:

- In conventional microscopy, controlling image quality across the entire field of view is difficult. Objects located at different depths within the specimen may appear blurry, which can lead to misinterpretation of cellular characteristics. This restriction makes distinguishing fine details and subcellular structures difficult, complecating the identification of certain blood elements.

- Due to a too ristricted measurement area, reliable diagnosis is fundamentally compromised, especially when searching for rare events.

- The provided information is qualitative.

Unconventional microscopy represents a significant advance in the imaging field that has the potential to answer some conventional microscopy limitations. It refers to optical techniques and devices that diverge from traditional geometrical and physical optics. This field involves the development of innovative ways to manipulate and generate images. These methods may include advanced image processing techniques and approaches that involve coherent manipulation of light. It includes adaptive optics, computational imaging, and holography. Among the most recent, we can cite Quantitative Phase imaging (QPI) methods based on interferometric devices, which present the following advantages:

- Improved resolution thanks to a synthetic numerical aperture mechanism (up to a factor 2).

- Quantitative intensity and phase imaging: An interferometry setup provides additional information on biological samples' optical properties.

- Aberration correction using numerical wavefront manipulation during the calculation step.

- Extended depth of field thanks to wavefront manipulation post-acquisition in the image formation model.

Although QPI has made significant advances, their implementation in uncontrolled environments is challenging due to their sensitivity to vibration.
A promising alternative for these QPI methods is Fourier ptychographic microscopy (FPM) [153, 133]. FPM is a computational imaging technique relying on inverse problem-solving. FPM is simple to implement as it does not require

an interferometric setup and offers the theoretical QPI advantages with additional specificities:

- **Extended space bandwidth product:** The synthetic numerical aperture can be very high and goes to factors up to 6.
- **Reduced cost:** FPM relies on a classical microscope where a LED matrix replaces the light source.
- **Stability:** The FPM is resistant to vibrations and can be used in various environmental conditions.

Fourier ptychographic microscopy (FPM) has gained significant interest over the last years (see Figure 1.1). FPM demonstrated its potential and relevance for observing objects of interest at high resolution. The demonstrated results rely on visual observation. However, phase imaging quality obtained through a reconstruction process has not yet been thoroughly investigated. That means that its interest in a digital pathology framework is not guaranteed. Therefore, in-depth experimental studies are needed. Such studies can also provide valuable insights for designing automatic FPM systems.

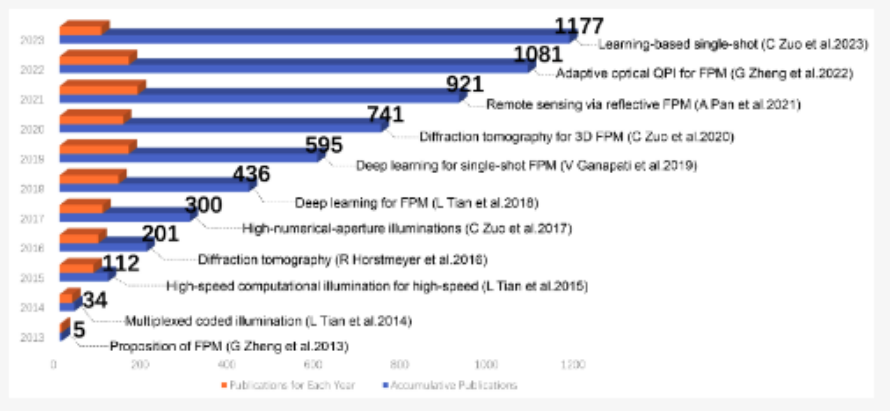


Figure 1.1: The development of FPM has experienced remarkable growth in the number of related publications since its introduction in 2013. (Images sourced from the scientific journal article [147])

1.2 Objective and contributions

The objective of this thesis is to explore the potential of FPM in blood smear analysis. Our approach consists of numerous experiments for advanced statistical analysis. More precisely, these experiments will be conducted on two specific tasks: detecting the parasite responsible for malaria and classifying different types of white blood cells. They were chosen based on the available biological specimens, and the results were analyzed and questioned by biologists' teams specialized in parasitology and hematology. The work aims to identify the possible contribution of the FPM, particularly the phase images for analyzing stained samples using deep learning methods for building classifiers. For this purpose, the thesis is structured around the following different axes :

- Evaluation of the impact of FPM for automatic classification in applications requiring increased sensitivity, such as malaria parasites detection and white blood cell classification. Particular focus is placed on proper exploitation of enriched information, namely bimodal (intensity and phase) images as compared to a classical intensity image.
- Choice and adaptation of Deep Learning models to the physics specificities of FPM framework.
- Investigate some design rules for an FPM-based digital pathology framework. This concerns the optimization of the microscope configuration (choice of objective lens) and the Deep Neural Networks (Information coding and processing).

This thesis offers several contributions:

1. Developing a methodology based on deep neural networks for processing intensity and phase images to obtain performance comparison metrics. These DNN models have been adapted to the bimodal nature of the information by introducing complex convolutions for efficient feature extraction from images.
2. Quantifying the contribution of the coupling between intensity and phase images resulting from FPM to improve both the detection of red blood cells infected by malaria parasites and the classification of white blood cells, highlighting the importance of the exploitation of phase information in the case of stained samples.

3. Proposing a physics-informed GAN model (PI-GAN) to synthesize new FPM bi-modal images, improving the diversity of generated images and limiting the impact of the mode collapse problem encountered with classical GANs. This corresponds to an essential advance in the field thanks to an efficient regularization obtained by introducing the image formation model in the loss functions. This model is associated with an innovative strategy for fine-tuning GANs. This combination demonstrated its ability to complete minority classes even when only a few samples were available for learning the GAN.

4. Optimizing the microscope configuration by studying the best trade-off between the resolution and the size of the field of view, demonstrating that introducing the phase in addition to the intensity images allows better tolerance at a lower resolution.

1.3 Thesis outline

This manuscript is organized into 6 chapters describing the work carried out for various contributions.
Following this introductive chapter, the chapter 2 presents the multidisciplinary context of the thesis. It exposes the limits of classical microscopy and the contribution of unconventional microscopy to overcome them, the concepts of neural networks, and the two applications examined in this thesis.
Chapter 3 questions the interest of surplus information on a simple but yet unsatisfactory solved question (parasite detection in red blood cells for malaria detection). To this end, performance with intensity and phase are calculated and compared with intensity alone. This chapter sets up a detection network based on the Faster-RCNN architecture to compare these two models.
Chapter 4 is dedicated to a more complex application, namely white blood cell classification from thin-stained blood smears, since it can have a significant biological impact. In the first part of the work, we confirm the results obtained in chapter 3 and its general scope. A comparison with conventional imaging methods was also carried out. Moreover, we also investigated the problem of under-represented classes, often present in biology. To this end, we introduce a physics-informed GAN to produce synthetic bimodal images, as the classical GAN model gave unsatisfactory results in producing this bi-modal information.

Chapter 5 aims to optimize the FPM approaches regarding information processing and microscope configuration. It has been applied to malaria detection use-case but could be easily extended to other cases. In this chapter, we explore an implementation via complex-valued CNN in the Faster-RCNN, which seems well adapted to the bimodal specificity of the data. Next, the chapter looks at optimizing the choice of the microscope objective lens using a method that combines experimental data and imaging models to obtain varied databases at different resolutions. Performance shows increased stability when intensity and phase are exploited together.
Chapter 6 offers a retrospective of the research carried out in this thesis, highlighting the impact of FPM on advanced diagnostics while identifying relevant neural architectures, taking into account the bimodal nature of the information. It also explores avenues for future research that could expand on these advances.

A pre-study of the work presented in Chapter 4 was conducted and os presented in Annexe A. In particular, the methodology for implementing GANs for classes with a low number of representatives carried out on a classic image database is presented.

Chapter 2

Multidisciplinary context

Contents

This chapter aims to provide a general overview of the multidisciplinary context of the thesis. It begins by discussing the role of microscopy for medical diagnosis, before explaining classical microscopy concepts and limitations and those of unconventionnel microscopy. Subsequently, it outlines the fundamental principles of neural networks and briefly introduces the two applications explored in this thesis.

2.1 Advanced imaging Modalities in medical diagnosis

The significance of precise and reliable diagnostic procedures cannot be overstated within medical diagnosis. A precise diagnosis forms the cornerstone of patient care, guiding clinicians in tailoring treatment strategies and closely monitoring disease progression [25]. This crucial process commences with a thorough assessment of the patient's clinical presentation, incorporating initial hypotheses. However, the refinement and validation of these observations are pivotal, accomplished through meticulous laboratory investigations and advanced diagnostic imaging techniques [18]. These additional steps not only serve to corroborate suspected diagnoses but also provide invaluable insights essential for an accurate and comprehensive medical evaluation [104].
Various diagnostic techniques exist in contemporary medicine, each offering varying degrees of precision and yielding distinct observations. While some methods emphasize cellular morphology, revealing structural abnormalities, others explore intricate biochemical or molecular signatures, unraveling the underlying physiological processes. Certain diagnostic modalities offer a panoramic view of tissues, facilitating a holistic understanding, while others provide unprecedented detail into subcellular structures, aiding in nuanced disease characterization. These varying approaches collectively contribute to the holistic understanding required for accurate medical diagnoses.
Among these advanced diagnostic tools, several stand out as macroscopic tools. This includes Magnetic resonance imaging (MRI), an advanced medical technique for obtaining two or three-dimensional images. Based on the principles of nuclear magnetic resonance (NMR), it exploits the quantum properties of atomic nuclei to analyze the chemical composition of biological tissues. Since its first studies in 1973 [69], MRI has gradually evolved to become a common clinical technique in the 1980s. Today, it is widely used to study anatomical

structures and soft tissue pathologies, providing complementary information to other imaging modalities. Its ability to characterize and differentiate tissues based on their physical and biochemical properties makes it a crucial tool for medical diagnosis [106]. MRI is used in various fields of medicine; in neurology, it allows neuroanatomical structures to be explored with great precision, facilitating the diagnosis of brain tumors, strokes, and degenerative diseases such as Alzheimer's disease [105, 33, 102]. In oncology, MRI evaluates the extension of tumors, defines their vascularization and evaluates the response to treatments, playing an essential role in the monitoring and treatment of cancers [129], particularly in the areas of breast [92] and prostate oncology [94]. The advantages of MRI lie in its high spatial resolution for soft tissues, allowing the precise distinction of different anatomical structures without exposure to ionizing radiation. However, this technology remains expensive, requiring specialized facilities and potentially longer examination times than other imaging modalities. In addition, it has limitations in imaging bone tissue, restricting its usefulness for the fine detection of fractures or bone damage.
Another medical imaging technique called Positron Emission Tomography (PET) involves administering radionuclides to the patient, which emit positrons [141, 56]. This imaging technology is widely used in oncology to assess the presence and progression of tumors [23]. It also makes it possible to visualize the functioning of organs and tissues, thus helping diagnose and manage neurological and cardiac diseases, among others [96, 17, 46, 128]. The advantages of PET lie in its ability to provide precise functional and metabolic information, allowing early detection of abnormalities before they are detectable by other imaging methods. However, the high cost of the procedure, the need for radioactive isotopes, and the short lifespan of radioisotopes sometimes limit its use, in addition to the exposure to ionizing radiation associated with this technique.
Finally, ultrasound, or ultrasonography, is also a medical imaging technique that uses high-frequency sound waves to produce real-time images of the inside of the body [95, 75]. Ultrasound is used in various medical fields, including obstetrics to view the fetus during pregnancy, cardiology to examine the heart and blood vessels, gastroenterology to study the abdominal organs, and radiology for comprehensive diagnostics. It makes it possible to observe internal organs, detect abnormalities, guide medical procedures such as biopsies and punctures, and monitor the progression of diseases [112, 91, 10]. The advantages of ultrasound lie in its safety, non-invasiveness, and painlessness for the patient. It does not involve ionizing radiation, making it safe, especially for pregnant women.

However, its effectiveness may be limited in some cases due to the difficulty in obtaining clear images through bone or gas, and its interpretation may vary depending on the operator.
The advent of imaging techniques has undeniably transformed the field of medical diagnostics, equipping healthcare professionals with a range of indispensable tools to evaluate physiological conditions. Innovations such as MRI, PET, endoscopy, and ultrasound yield invaluable perspectives into the intricate landscapes of anatomical structures and pathological deviations. However, in some cases, macroscopic imaging approaches may not be sufficient to provide a complete diagnosis, for example, the observations of specific microscopic structures need to be observed microscopically. In such cases, microscopic observation is essential to establish the precise pathology.

2.2 Microscopy

Classical microscopy techniques have always played an essential role in medical diagnosis by providing healthcare professionals with a detailed and comprehensive view of cellular and subcellular structures. These techniques have proven highly effective in recognizing specific features associated with various illnesses. In this section, we will delve into the history and principles of the microscope, discuss its applications in medical diagnosis, and highlight the limitations of this technique.

2.2.1 Brief history

The history of microscopy dates back to antiquity, but its significant development occurred in the 17th century, marking the beginnings of modern microscopy. In 1590, brothers Hans and Zacharias Janssen designed the compound microscope, a rudimentary device using lenses to increase the size of objects observed. However, with the publication of "Micrographia" by Robert Hooke in 1665, microscopy was brought to life [49]. Antonie van Leeuwenhoek, a Dutch merchant, made a significant contribution in the 1670s by making his own simple but remarkably powerful microscopes. Constant improvements in the objectives and lenses of optical microscopes have marked successive advances. The 19th century saw significant developments, with the formulation of the theory of image formation in microscopy in 1873 by Ernst Abbe [2]. The

real revolution occurred in the 20th century with the invention of the electron microscope by Ernst Ruska in the 1930s [61]. This device allowed exploration on a smaller scale than that is accessible by optical microscopy, thus opening new perspectives in cell biology and materials science. Since then, microscopy has expanded rapidly [11, 89, 154].

2.2.2 Basic Concepts of Microscopy

The device

The microscope device shown in Figure 2.1 can be divided into two fundamental components that serve distinct purposes.

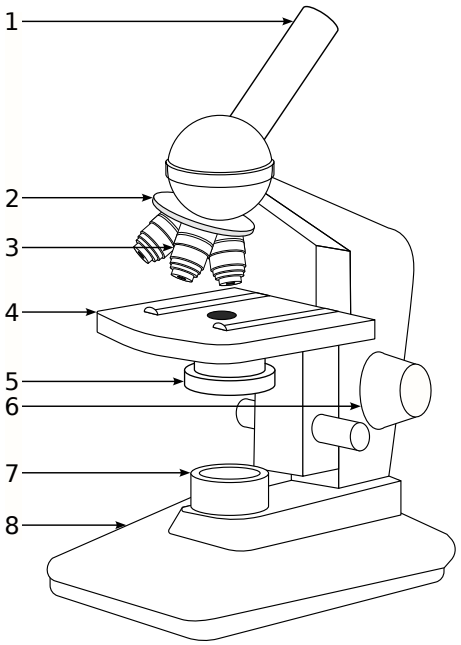


Figure 2.1: Basic microscope scheme. 1-Eyepiece tube; 2-Objectives support; 3-Objective; 4-Stage; 5-Condenser; 6-Adjustment knob; 7-Light source; 8-Base.

The first part, the optical component, is primarily responsible for enabling sample visualization. it includes an illumination source necessary to illuminate the sample. Köhler illumination [63] is the current standard for research and diagnostic microscopes, providing even illumination and optimal resolution in conventional microscopes. Following the illumination source, a condenser focuses the light onto the sample. The sample is then placed between the condenser and the objective, which is the most crucial component of the microscope. An objective is a converging lens system with a short focal length that projects a magnified, inverted image of the sample onto the lower focal plane of the eyepiece for further magnification and observation. The objective's characteristics are generally imprinted on the objective itself. The second part, the mechanical component, provides stability and comfort while using the microscope. It includes a base for stability, an objective holder, an eyepiece holder tube, a stage that moves in two dimensions, and an adjustment knob for fast and precise focusing.

Microscopy principle

When the light sources illuminate the sample, the light interacts with the components of the sample. Some elements absorb light, while others reflect or transmit it. The light modified by these interactions then enters the microscope objective. The objective lens is the first to capture the light coming from the sample. This lens is designed to magnify the sample image and focus the modified light to form a sharp, magnified image on the microscope's focal plane. This enlarged image, formed by the objective, is then observed by the eye through the microscope's eyepieces.

Classical microscopy resolution

The most critical aspect to consider when using a microscope is the resolution. Distinguishing and visualizing the smallest details in the observed samples relies heavily on it. However, light diffraction, caused by the wave nature of light, inherently limits this capability. Instead of a precise point representation, the image of a point viewed through an optical microscope appears as a light spot, known as an Airy spot or Point Spread Function (PSF). This light diffusion causes neighboring points' images to overlap, making it difficult to differentiate between them and ultimately decreasing the microscope's resolution capacity. There are many definitions of resolution in microscopy depending on the microscopic setup [2, 110]. The most used definition is based on Abbe's theory, which proposes a resolution limit equation, expressed as the smallest distance below which two adjacent points can no longer be discerned. This formula is determined by the illumination wavelength λ, the refractive index at the lens exit n, and the half angle of the maximum light cone accessible α. The formula for calculating the resolution is in the form:

$$d_{Abbe} = \frac{\lambda}{2 \cdot n \cdot \sin \alpha} = \frac{\lambda}{2 \cdot \mathrm{NA}} \tag{2.1}$$

Where λ is the wavelength used, n is the refractive index, and α is the aperture angle. The numerical aperture of the lens (NA) represents the product of the refractive index and the sine of the aperture angle, playing a crucial role in determining the resolution. Two main approaches can be considered to improve the resolution: increasing the refractive index and decreasing the wavelength. Using immersion lenses, where the front of the lens is immersed in a liquid with a high refractive index, can increase this index or the increase of the objective

numerical aperture.
Rayleigh's formulation suggests that resolution is proportional to half the wavelength of light used, multiplied by a factor that depends on the numerical aperture of the objective.

$$d_{Rayleigh} = \frac{1.22 \cdot \lambda}{NA_{obj} + NA_{cond}} \tag{2.2}$$

In this formulation, the condenser numerical aperture NA_{cond} also impacts resolution by controlling the amount of incident light.

Imaging requirement

The microscope is commonly used to capture images of samples being studied. When integrating a camera into the microscope, it is important to ensure that the images captured accurately reflect what is observed under the microscope. The Shannon-Nyquist theorem is a principle that states that in order to reconstruct a signal from samples accurately, the sampling frequency must be at least twice as high as the maximum frequency present in the original signal. In the case of capturing images via a camera connected to a microscope, the Shannon-Nyquist theorem dictates that the image sampling rate must be high enough to capture the fine details of the observed sample effectively. Important information could be lost if the sampling rate is too low, even if the other microscope components can observe it.

Sample preparation

Sample preparation is also critical for precise and in-depth observations in microscopy. Appropriate techniques preserve cellular and tissue structures and provide essential morphological and functional details. Specific probes or fluorescent dyes can selectively label biological constituents, allowing targeted visualization of particular structures or processes within the samples.
After the specimen collection, specimens may undergo different preparation pathways depending on the desired microscopic examination technique. The first common step is fixation. The fixation stage aims to stop cellular and biochemical processes, thereby preventing cell degradation and deterioration. It thus makes it possible to maintain the morphology of the cells, thus preserving their characteristics and spatial organization.
For blood smears, the fixation stage involves spreading blood on a glass slide,

which is then fixed to stabilize the cells and prevent their degradation. The most common fixation is carried out with methanolic or ethanolic alcohol, which allows the precipitation of cellular proteins. For other samples, such as biological tissues, commonly used solutions for fixation include formaldehyde and glutaraldehyde.
Secondly, it is essential to ensure that the observed parts are distinguishable. This often requires a marking process, which can be accomplished using various methods. Among these, we find particular staining techniques. Staining involves the Application of specific dyes to biological samples to improve the visibility and differentiation of cellular structures. It relies on specific chemical reactions between dyes and cellular components. Some dyes selectively bind to proteins, lipids, or nucleic acids, thus making it possible to highlight different cellular structures. Different dyes are used in microscopy to highlight specific structures. Common examples include the dye H&E (Hematoxylin and Eosin) [134, 22], known to stain cell nuclei blue and cytoplasm pink. Giemsa is used for staining chromosomes and detecting blood parasites, while MGG (May-Grunwald Giemsa) [42] is used to stain blood cells. Microscope examination may also rely on fluorescence, which is based on the ability of certain materials, called fluorochromes, to absorb light at a specific wavelength and reemit light at a different wavelength [5]. Fluorescence is based on the excitation of a fluorochrome with a specific light, followed by light emission at a higher wavelength. Excitation occurs at a specific wavelength, and emission occurs at a longer wavelength. This technique provides excellent sensitivity and specificity for visualizing specific structures in cells or tissues. Fluorescent labeling involves using specific antibodies or probes labeled with fluorochromes to target specific proteins or cellular structures. Samples are incubated with these labeling agents to allow the fluorochromes to bind to the targets of interest. Fluorochromes are selected according to their affinity for the structures to be visualized and their emission properties. Specific antibodies, when used, selectively bind to target antigens, allowing fluorochromes to label these structures specifically. Overall, preparing samples for microscopy is a multi-step and intricate task involving several meticulous steps that must be undertaken with precision.

2.2.3 Microscopy applications in biomedical fields

The microscope has played a vital role in the progression of biology and medicine. As a result, microscopes have now become an indispensable tool in modern medicine [30]. In oncology, it is employed to identify various cancer types, including lymphoma, basal cell carcinoma, and squamous cell carcinoma [79, 53]. Likewise, in hematology, microscopy plays a significant role in diagnosing conditions like sickle cell anemia, leukemia, and thrombocytopenia [27, 99, 131]. It is also helpful in identifying malaria by detecting Plasmodium parasites in the blood [145]. In microbiology, microscopy is utilized to diagnose parasitic infections such as giardiasis, amoebiasis, and toxoplasmosis. Additionally, it is importante for identifying the bacteria responsible for infections like tuberculosis, syphilis, and gonorrhea [58, 37]. In dermatopathology, microscopy examines skin cells and structures to diagnose a range of skin conditions [93].

2.2.4 Classical microscopy limits

Although microscopy has found numerous applications across various domains over the years, its progress has been hindered by several limitations that have restricted advancements in this technique. The diffraction limit constraint is one of the most significant limitations, as it sets a maximum resolution for the microscope and affects its ability to observe fine details. When the features are smaller than this limit, they cannot be resolved, reducing the precision of microscopic observations. Additionally, optimizing the resolution, field of view, and depth of field, known as the space bandwidth product, presents inherent challenges. Improving one of these parameters usually comes at the expense of the others, which can restrict the microscope's ability to provide a wide field of view and fine detail simultaneously. Another challenge lies in the qualitative nature of the information obtained by microscopes. Although they allow for precise visualizations, classical microscopes do not always provide precise quantitative data, limiting their applicability in studies that require precise measurements. Sample preparation processes, such as fixation, staining, and marking, can be time-consuming and complicated. These processes require technical expertise and time, which can be a constraint for studies that require quick or large sample volumes. Finally, only nontransparent objects can be studied.

2.3 Unconventional optics

The limitations of conventional microscopy have spurred the creation of specialized techniques designed to overcome these challenges. Consequently, unconventional optics has emerged as a field encompassing novel methods and approaches in microscopy. Unconventional optics are optical techniques and devices that diverge from traditional geometrical and physical optics methods. This field involves the development of innovative ways to manipulate, generate, and detect light by utilizing non-trivial optical concepts. These methods may include optimized structures, advanced image processing techniques, unusual optical properties of materials, and approaches that involve coherent manipulation of light. Unconventional optics are powerful alternatives to conventional methods, focusing on overcoming resolution limitations, Space Bandwidth Product (SBP), and sample preparation. This section briefly overviews some of these techniques, highlighting their potential by showcasing the remarkable boundaries they have already surpassed. These approaches include but are not limited to, holography, lensless microscopy, and structured illumination microscopy.

2.3.1 Holographic microscopy

Basic principles of holography

The origin of holography is rooted in the groundbreaking work of Dennis Gabor in 1948. He discovered that the diffraction pattern of the electron beam contained information regarding the amplitude and phase of the electron wave. Gabor created an optical synthesis of the object's field by recording the electron wave's diffraction. Consequently, image formation using visible light optics was more advanced than electron optics. This revolutionary imaging principle was appropriately named "holography" due to its ability to capture the complete optical field. Over time, there have been significant advancements in the field, such as the emergence of potent coherent light sources, commonly referred to as lasers, and the successful validation of Gabor's holographic storage and reconstruction principle. In the late 1960s, J.W. Goodman introduced digital holography. This approach entails electronically capturing holograms, which are then subjected to numerical processing to recreate the digital object. Digital holography, or DH, generates holographic interference through optical processes. Digital holographic microscopy stands out among other microscopy

techniques as it does not capture the projected image of the object. Instead, it records the light wavefront information that emanates from the object as a hologram, which is used in a numerical reconstruction algorithm to calculate the object image. Therefore, the computer algorithm replaces the image-forming lens in traditional microscopy. The information is captured on a digital image sensor or photodetector from which an image of the object is reconstructed. In order to produce a hologram, a coherent light source, such as a laser, is needed to create the necessary interference pattern. The laser light is split into an object beam and a reference beam. The object beam is expanded to illuminate the sample, creating the object wavefront. Once the object wavefront is collected by a lens or microscope objective, it is combined with the reference wavefront using a beam splitter to create the hologram. Using a digitally recorded hologram, a computer can use a numerical reconstruction algorithm to produce a viewable image of the object wavefront.

Digital holography in biomedical applications

Several studies have used digital holography in the field of biomedical applications. It allows detailed observation of cells and molecular processes at the subcellular level. It makes it possible to easily measure cellular properties that were previously very difficult to study in living cells, such as the thickness, volume and refractive index of cells, thanks to the quantitative phase obtained by reconstruction and without labeling [84, 108, 90]. In [107], fluctuations in the membrane of red blood cells depending on their cellular state are highlighted using a digital holographic microscopy approach thanks to its high resolution. In, [8, 136] digital holographic microscopy was used to identify red blood cells accurately (RBCs) infected by malaria. To go further, [137] proposes a flow cytometry approach that relies on digital holographic microscopy to differentiate leukocytes without the use of labels or reagents, thus offering automated diagnosis of hematological disorders based on on morphological criteria. In oncology, [83] have proposed a significant advance in digital holography by improving the quality of images obtained via quantitative phase contrast microscopy. Their approach made it possible to achieve a lateral resolution limited by diffraction at 0.5 μm. A notable feature of this method is its ability to significantly reduce coherent noise, thereby providing precise phase holographic images with a phase profile reaching an optical thickness of approximately 30 nm. These improvements enabled clear visualization and quantitative preci-

sion of intracellular and intranuclear organelles of SKOV-3 ovarian cancer cells. [3] explores phase-shift digital holographic microscopy to measure deformation without invasiveness to explore and characterize the biomechanical properties of cancer cells, particularly in the context of breast cancer, where differences between invasive and non-invasive methods were examined. These studies revealed significant variations in cellular deformation, suggesting a potential link between these mechanical properties and the metastatic behavior of cancer cells. [29] harnessed the power of digital holographic microscopy combined with quantitative phase imaging (QPI) to quantify the effects of three types of organic nanoparticles on cellular cytotoxicity. This approach, free of labeling, made it possible to monitor cell growth in real-time and to analyze, over 24 hours, the evolution of the dry mass of cells exposed to nanoparticles and to a control agent. [82] uses advanced imaging techniques to examine cervical smears. Researchers combined color and phase images to identify important information by analyzing cervical cells. They found that using these two images together improved the ability to group cells accurately.

2.3.2 Lensless microscopy

Basic principles of lensless microscopy

Lensless microscopy is an innovative imaging technique that captures image samples without conventional optical lenses, typically used in conventional microscopes. Unlike traditional microscopes, this technique is based on the principles of light diffraction. It is based on the direct recording of diffraction patterns produced by the sample. The operating process of lensless microscopy is based on several fundamental principles. It generally requires a coherent light source, such as a laser, to illuminate the sample. This light coherence creates complex diffraction patterns as the light interacts with the structure of the sample. When coherent light passes through or interacts with the sample, it is diffracted. A detector, often a digital camera, captures these diffraction patterns. The recorded diffraction patterns undergo digital processing to reconstruct an image of the sample. Image processing algorithms and diffraction-based reconstruction techniques convert these patterns into an interpretable image.

Lensless microscopy in biomedical applications

Lensless microscopy represents a significant advancement in the biomedical field, offering versatile solutions for various cell and tissue imaging applications. Huang et al. [52] proposed a lensless holographic microscope based on online holography. This approach offers a portable, affordable, and high-precision alternative for biological imaging, overcoming the limitations of expensive and bulky traditional microscopes. Using an iterative phase recovery algorithm, this technique improves the resolution of holographic images, including automatic identification and precise counting of cells. Similarly, [151] explored the integration of lensless microscopy with a microfluidic chip to monitor and sort sperm. Their system targeted the identification of the most motile sperm in samples with concentration or motility deficiencies, potentially improving the success rates of assisted reproductive techniques, particularly for cases of male infertility. [73] revealed a lightweight, portable, lensless holographic microscope capable of imaging samples in reflection and transmission modes. This versatility, combined with a spatial resolution of up to 2 µm, makes it ideal for resource-constrained environments, providing precise imaging solutions in global health and water quality monitoring contexts. Additionally, [41] have developed a lensless microscope enabling the imaging of entire human tissues with a field of view approximately 100 times larger than conventional high-resolution microscopes. This device provides sufficient spatial resolution and contrast for clinical diagnosis, representing a valuable tool for pathology laboratories in developed and developing countries. Finally, [26] presented a lensless microscope based on online holography, providing high-resolution label-free imaging with a wide field of view. This versatile microscope enables real-time observations of cellular dynamics without fluorescent labels or extensive sample preparation, paving the way for various biomedical applications from individual cell analysis to tissue imaging.

2.3.3 Structured illumination microscopy

Basic principles of structured illumination microscopy

Structured illumination microscopy (SIM) represents a super-resolution microscopy method that overcomes the resolution limitations of diffraction. By acquiring multiple images of the same sample under various lighting patterns, SIM computationally combines these images to obtain a single reconstruction with up

to 2x improvement over diffraction-limited resolution. The key concept behind SIM is based on the well-known moiré fringes. When two patterns are superimposed in a multiplying manner, moiré fringes appear in their product. A series of images acquired at various orientations of the lighting pattern are mathematically deciphered to extract this information. Indeed, high spatial frequencies, usually not visible, can be captured using a structured lighting model. To achieve near-isotropic resolution expansion, the illumination pattern must be rotated at least two other equally spaced angles, typically 60°, and additional data must be acquired for these pattern orientations. All these components, represented by circles in Fourier space, are assembled according to their original positions to form a final reconstructed image with extended resolution. Structured illumination microscopy takes advantage of the concept of reciprocal space, providing a more informative representation of physical reality, especially in terms of spatial resolution. The maximum resolution expansion limitation is 2x the conventional resolution limit due to the limitation imposed by diffraction in generating the illumination pattern.

Structured illumination microscopy in biomedical applications

In turn, structured illumination microscopy has offered significant advances over the limits of microscopy. [43] addressed the lateral resolution limit of optical microscopes using structured illumination microscopy. Their method overcame this limitation by illuminating the sample with spatially structured excitation light, revealing high-resolution information normally inaccessible. Using structured illumination patterns combined with image processing, they outperformed conventional resolution by a factor of two, delivering resolution-enhanced images on complex biological samples. This approach has opened up possibilities for high-resolution imaging in various fields, biology, and other disciplines requiring precise visualization of structures. [150] investigated the application of SIM in hematology and pathology to visualize cellular structures with improved spatial and axial resolution. Their study showed the ability of SIM to obtain detailed three-dimensional morphological images of normal and abnormal blood cells, opening new perspectives for research in hematology and pathology. SIM revealed subtle morphological details, outperforming conventional and confocal microscopy techniques for better visualization of cells. [97] explored SIM microscopy as a diagnostic tool for nephrotic diseases. Their study compared SIM to electron microscopy for visualizing podocytes,

demonstrating that SIM can provide a similar resolution to electron microscopy while enabling 3D visualization and providing greater speed and ease of use. This method presents an effective and potentially less expensive alternative to electron microscopy for diagnosing nephrotic diseases, showing detailed images and clearly distinguishing healthy from diseased tissues. [139] used SIM microscopy to visualize ultrastructural details of placental tissue, overcoming the limitations of conventional microscopy. Their method offered improved visualization of ultrastructural structures, with better resolution and increased contrast compared to deconvolution microscopy.

2.3.4 Synthesis and comparison of techniques

Unconventional techniques have opened up new lanes to surmount the constraints of conventional microscopy. Holographic, lensless, and structured illumination microscopy are notable examples of such techniques. By presenting superior visualization and analysis capabilities, they have brought about a paradigm shift in the field of microscopy. A key advantage of these innovative techniques is their capacity to surpass the inherent resolution limitations of traditional microscopy. By surpassing the diffraction limit, these methodologies can achieve remarkably improved spatial resolutions, enabling the observation of tiny cellular and subcellular structures with unprecedented precision. The second advantage is the capability of label-free imaging, which provides a distinct advantage. These techniques maintain the natural state of biological samples by eliminating the need for labels or dyes. This enables accurate observations of cellular architecture and dynamics without alteration while reducing preparation time. Quantitative phase imaging is also a distinctive feature of many unconventional microscopy techniques. It helps to measure the physical properties of cells and tissues. This quantitative information can be collected without staining or labeling, making it an invaluable tool for understanding cellular characteristics and behavior. The presented techniques exhibit exceptional flexibility in various sample types, surroundings, and imaging methods. Unconventional microscopy offers numerous advantages and is a game-changer. It gives researchers and healthcare professionals a new perspective on biomedical applications imaging, enabling them to explore the intricacies of life at the cellular and molecular levels.
Despite their potential benefits, these diagnostic methods face significant limitations restricting their everyday use. For instance, holography can be pro-

hibitively expensive and can be sensitive to vibrations and environmental changes. Structured Illumination also has challenges, such as the need for controlled lighting conditions that may only sometimes be feasible in specific diagnostic settings. Additionally, it can be disrupted by reflective surfaces and translucent materials. Furthermore, lens-free imaging reconstruction remains a complex process that requires significant processing times and may not offer the level of spatial resolution necessary to visualize fine details or structures.
Recently, even more promising perspectives have emerged in the field of microscopy, addressing certain limitations of existing techniques and offering increased stability and efficiency. One such advancement is Fourier ptychographic microscopy, which integrates computational imaging with traditional microscopy to overcome the constraints of conventional methods. FPM holds significant promise due to its compatibility with clinical environments and its cost-effectiveness. This unconventional imaging technique will be described in details in Chapter 3.

2.4 Artificial Neural Networks

Artificial Neural Networks (ANNs) are computer systems that are inspired by the functioning of the human brain. These systems are designed to learn from data, detect complex patterns, and perform various tasks, from image recognition to language translation to complex decision-making. Today, ANNs are widely used in various fields, such as computer science, biology, finance, and medicine.
In this section we will not provide a comprehensive view of this floorishing domain. We concentrate on the fondamentals and the Convolutional Neural Networks(CNN) which are the basis elements of all the neural networks models that have been used in this thesis.

2.4.1 Brief history

The history of neural networks dates back to Aristotle's proposal of "associationism" in 300 B.C. Since then, researchers have been working to create a machine that can imitate human brains. In 1943, the McCulloch-Pitts (MCP) model was introduced as one of the prototypes of artificial neural models. This model laid the theoretical foundation for formal neural networks by explaining how

biological neurons could be modeled using simple logical elements. However, the MCP model could not be learned. In the 1950s and 1960s, a neurophysiologist named Frank Rosenblatt created the perceptron. This was one of the early models of artificial neural networks that could learn. The perceptron was explicitly designed to classify various tasks by analyzing patterns and making decisions based on the input provided. However, due to its linear structure, it had limitations and needed to be improved for solving complex problems. The field of neural networks experienced an "A.I. winter" in the 1970s and 1980s due to limited progress, technical challenges, and computing power constraints. As a result, funding and academic interest were significantly reduced. During the 1990s, significant advancements in learning algorithms (ability to process non-linearity), increased availability of big data, and a boost in computing power sparked renewed interest in neural networks. As a result, more complex architectures, like Deep Neural Networks (DNNs), have been developed, allowing for more efficient problem-solving and processing of complicated data such as images, temporal sequences, and textual information. However, the full power of Neural Networks is essentially visible in the present period, when spectacular performances are obtained in many applicative fields.

2.4.2 Fundamental principle of artificial neural networks

The neuron

Artificial neural networks are designed to mimic the functioning of biological neural networks, which are responsible for implicit associations in the brain. These networks are composed of individual artificial neurons that follow the principles of biological nerve cells. To understand how a biological nerve cell

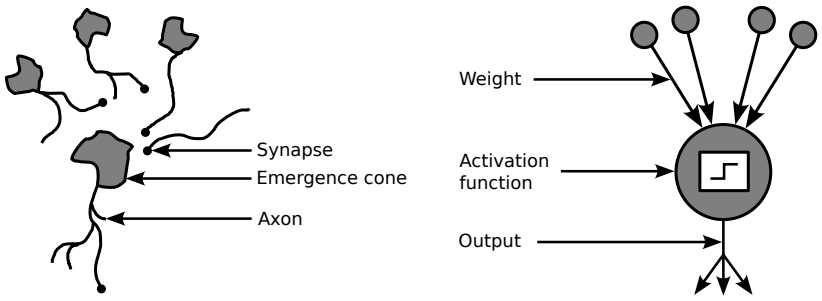


Figure 2.2: Matching biological and artificial neurons.

works, it receives input signals from other neurons through synapses. It is made up of dendrites which receive electrical signals. These signals are summed at the level of the emergence cone, and if they reach or exceed a given threshold, the neuron produces a nerve impulse. This impulse excites other neurons by crossing the axon and the terminal branches. The frequency of emission of nerve impulses by a neuron depends on the sum of excitations received. The artificial neuron is a device with several inputs and an output. It receives a variable number of inputs from upstream neurons. Each of these inputs is assigned a weight, representing the strength of the connection. Each neuron produces a single output, which results from thresholding the sum of the received inputs. This output then branches to power a variable number of downstream neurons. In comparison, as illustrated in Figure 2.2, the analogy between the biological and artificial neurons is their ability to receive weighted inputs, sum them, and produce an output if a threshold is reached. Artificial neurons use weights to adjust each input's impact, mimicking how biological synapses strengthen or weaken connections between neurons.

Multilayers perceptron

Further, similarly to biological nervous tissue, an artificial neural network is formed by an assembly of interconnected neurons. This network is generally organized into different layers of neurons. The most common architecture corresponds to multilayer perceptrons (MLP) [118], in which the neurons are organized into different layers The architecture of such networks typically includes three primary layers: input, hidden, and output. The input layer serves as the starting point, providing input signals to the following layers for processing. The hidden layers consist of sets of parallel neurons that work together to process signals, with each layer connected to the subsequent layer through weighted connections. Finally, the output layer at the network's end generates the final results or outputs based on the problem being addressed. This layered structure enables hierarchical information processing as input signals journey through the network. Specific weights are associated with connections between neurons, regulating signal transmission. These weights determine the modification and propagation of signals across the neural network, facilitating data processing and output generation. In summary, the neural network is designed to produce an output as a function of an input signal. Mathematically, the output y of a Multilayer Perceptron (MLP) with a single hidden layer comprising

m neurons can be represented mathematically as:

$$y(x) = \sum_{i=1}^{m} w_i \cdot f(w_{i0} + \sum_{j=1}^{n} w_{ij} \cdot x_j)$$

This equation illustrates the output generated by the MLP, where w_i and w_{ij} represent the network's weights, x_j denotes the input features, and $f()$ represents the activation function applied in the hidden layer to introduce nonlinearity. A structured architecture enables the MLP to approximate various complex functions by assigning suitable weights and activation functions. This relationship between the network's structure and functional approximation abilities is consistent with the principles of the universal approximation theorem. Indeed, within a feedforward neural network, a single hidden layer with adequate neurons can theoretically approximate any continuous function on a bounded and closed input space.

The learning principle

The learning process in neural networks refers to the iterative procedure through which a neural network adjusts the weights assigned to connections between neurons to improve its ability to perform a specific task [71]. The neurons' weights are generally initialized with a small random number following a Gaussian distribution. Therefore, MLPs produce random outputs at the beginning. The adjustment of the weights occurs primarily through a technique known as error backpropagation. It is essentially the mechanism by which the network learns to recognize patterns and perform tasks more accurately through exposure to data and adjustment of its internal parameters. This process involves presenting data to the network to produce an output called the forward passes. A defined loss or loss (or cost) function that measures the error between the network's outputs and the expected or actual results is backpropagated and used to adjust the weights (backward passe) and reduce this error gradually. The contribution to the error is determined at each neuron based on the weights associated with its connections. These contributions are used to update the weights incrementally, typically leveraging optimization algorithms like stochastic gradient descent (SGD) [19], Adam [60], or RMSprop[40]. The training phase of a neural network relies first on an input dataset commonly referred to as the training set, facilitating weight adjustments through backpropagation. The quality and diversity of data utilized for neural network training play a pivotal role in its

predictive accuracy. Consequently, the dataset's quality, volume, and representativeness are pivotal determinants in ensuring the neural network's efficacy and reliability during learning. Following the learning phase, the model undergoes a testing phase to evaluate its generalization capabilities on an independent dataset, distinct from the training set, known as the test set. Neural networks rely on an objective function, also known as cost or loss functions, that vary according to the task. The Mean Squared Error (MSE) cost function is commonly used for regression tasks where the output is a continuous number, but the Mean Absolute Error (MAE) can also be used. Cross-entropy or binary cross-entropy can be used for classification tasks. Other loss functions for more specific tasks, such as text generation and generative tasks, are also used. The capabilities of traditional neural networks, such as Multilayer Perceptrons, have expanded considerably in data modeling. However, processing images with such architectures is limited. The abundance of visual data in various fields necessitates more complex models that can comprehend and analyze this information. Images contain a wealth of spatial information and pixel-by-pixel relationships that MLPs cannot efficiently capture since they process each pixel separately, obscuring pixel relationships' spatial structure and significance. Object recognition, pattern detection, and understanding of visual contexts require a more comprehensive understanding of local patterns, feature hierarchy, and translation invariance, which traditional architectures have difficulty capturing due to their uniform structure. As a result, their ability to recognize objects in complex visual contexts is limited.

2.4.3 Convolution neural networks

The shortcomings of the MLPs previously described give rise to new types of neural networks. These models, called convolutional neural networks, are inspired by how the human visual cortex processes visual information. The human brain visually breaks down complex scenes into simple features, processed hierarchically across different specialized regions of the visual cortex. Early regions detect features such as edges, textures, and shapes, while more advanced regions aggregate these features to form increasingly complex visual representations. The transposition of this biological intuition into CNNs has revolutionized image processing [70]. CNNs, designed with a specialized layer structure inspired by the visual cortex, have ushered in a new era in computer vision. The first layers of CNNs act as filters, identifying basic patterns like lines, shapes

and textures, while the upper layers integrate these patterns to recognize complex visual entities. This specialized layered architecture allows CNNs to efficiently capture the spatial structure of images, significantly increasing their capabilities to perform sophisticated tasks such as classifying objects in complex scenes. CNNs are built on specific principles that allow efficient analysis of images [67], exploiting the spatial structure and local relationships between pixels. These networks are designed around specialized layers to extract essential visual features progressively.

Convolutional Layer

The convolutional layer in a CNN is based on a spatial filter, where a receptive field moves across an input image. This process involves a localized filter systematically traversing the input data, performing element-wise multiplication and summation within overlapping regions. Each spatial location undergoes this operation, generating a single value representing a distinctive feature of that segment. The principle is illustrated in Figure 2.3. In the case of multi-

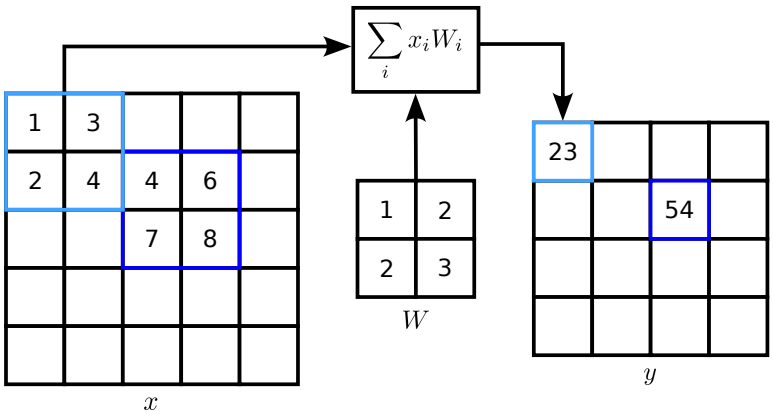


Figure 2.3: Convolution principle.

channel input, specific filters are applied separately to each channel, and the results are aggregated using a summation.

Pooling Layers

Pooling layers typically follow convolutional layers and are responsible for reducing the spatial dimensionality of the activation maps generated by convolution. Max pooling, a commonly used technique, selects the maximum value in

specific regions, reducing spatial resolution while retaining essential features. This dimensionality reduction enhances computational efficiency and network robustness.

CNN learning Principle:

A CNN processes input images through modules. Successive convolutional blocks, comprised of multiple convolutions and pooling operations, progressively extract information, ultimately generating a low-dimensional yet detailed representation called the final feature map. These convolution layers initially employ filters with learnable weights, functioning as localized windows over the input image to capture fundamental features like edges, gradients, and basic textures. As the network progresses through deeper layers, these learned filters evolve, becoming more sophisticated and abstract. They start detecting higher-level features, including intricate shapes, complex textures, and even specific components of objects. These extracted features evolve into highly specialized and task-oriented representations moving deeper into the network. The deeper layers decode increasingly complex concepts, such as object parts or specific patterns for the network's ultimate task. Subsequent layers, designed for specific tasks such as classification or segmentation, receive this distilled information as input. The effectiveness of this hierarchical feature extraction process heavily relies on the convolution block weights, which the model learns during training. The quality of the extracted features profoundly impacts the model's overall performance, making the evolution from basic edges to intricate, task-relevant features a critical aspect of CNNs. Note that the filter coefficients are learned by the backpropagation algorithm as in MLP and are therefore adapted to the task at hand in opposition to classical image processing filters the coefficients of which are fixed.

2.4.4 CNN for microscopic diagnosis

Convolutional neural networks (CNNs) have been largely explored in medical images applications [116, 31, 78]. We will limit ourselves in this section to details some of its outcomes for nonconventional optics as it corresponds to the imaging context of this thesis. Convolutional neural networks (CNNs) have been explored for unconventional optics techniques and applications, aiming to overcome some limitations. [114] demonstrate the capability of neural

networks in performing phase recovery and holographic image reconstruction from intensity-only measurements, enhancing imaging efficiency and computational speed.[55] propose a sensitive system for anthrax detection in biodefense using holographic microscopy and deep learning algorithm exploring phase images. The system achieves high sensitivity and specificity, offering rapid, label-free screening of pathogens. [123] propose using Deep Convolutional Neural Networks (DCNNs) to simplify optical systems for phase imaging. This framework aims to replace traditional optical lenses with computational algorithms, potentially leading to more cost-effective and more straightforward systems for phase imaging in biological specimens. [54] employ deep CNNs to predict the focal position in whole slide imaging, addressing autofocusing challenges without axial scanning, enhancing imaging quality, and system throughput in digital pathology and microscopy. [148] combine Digital Holographic Microscopy with a classifier based on handcrafted features extracted from phase to improve RBC segmentation accuracy and separation ability, particularly in complex cellular arrangements, offering potential applications in medical diagnostics and cell biology research. [77] employ Mask R-CNN to differentiate between thalassemic and healthy RBCs in quantitative phase images, achieving high accuracy in detection and segmentation. This model provides detailed single-cell characterization, aiding clinical decision-making in detecting abnormal RBCs. [98] delves into lensless imaging systems. These systems, employing microfluidic devices, diffusers, and CMOS sensors, demonstrate the potential of CNNs in cell identification, especially in distinguishing between cow and horse red blood cells. The article showcases the robustness and superiority of CNN-based lensless cell identification systems. [122] explore supervised machine learning for automated classification of lipid droplets in label-free, quantitative-phase images, showcasing the superiority of Convolutional Neural Networks (CNNs) in accurately classifying these structures. [143] introduce an automated RBC counting system employing Fourier Ptychographic Microscopy and a Convolutional Neural Network and achieving remarkably high accuracy in estimating the number of RBCs. [142] discusses the role of Quantitative Phase Imaging (QPI) and Fourier Ptychographic Microscopy (FPM) in label-free cell imaging. It emphasizes the potential of integrating deep learning approaches with FPM to enhance image correction, analysis throughput, and overall cell analysis in high-throughput setups. [130] discuss the limitations of 2D CNNs in processing volumetric information encoded within holograms. They propose 3D Convolutional Neural Networks to decode volumetric data efficiently, showcasing supe-

rior classification performance and enhanced depth information extraction for holographic imaging. [132] explore Fourier Ptychographic imaging combined with deep learning for accurate breast cancer histopathological image classification. This hybrid approach showcases significant strides in accurately classifying breast cancer images.

In summary, convolutional neural networks have demonstrated significant advantages in the area of non-traditional imaging. They are precious in overcoming device limitations and speeding up the acquisition process. However, the number of applications developed in this field has been limited, with a primary focus on phase modality imaging of transparent objects. Furthermore, these methods have been constrained by the size of their databases. Presently, there is no effective approach for processing bi-modal information and exploring its complementarity, nor its potential impact on studying colored objects.

2.5 Biological context

The experimental part of this thesis focusses on two specific applications related to the health field: detecting the parasite responsible for malaria from thin blood smears for parasitology and classifying white blood cells from thin blood smears for hematology.
This part, therefore, aims to present the proposed applications demonstarting the current limits of microscopy in their specific context and how FPM imaging could help to overcome these limitations.

2.5.1 Parasitology

Malaria, a parasitic disease caused by the Plasmodium genus, remains a serious threat to global health, particularly in tropical and subtropical areas. This disease is transmitted by Female Anopheles mosquitoes, impacting millions of individuals annually. The 2019 statistics underscore the magnitude of malaria, with an estimated 229 million people worldwide affected and causing nearly 409,000 fatalities. Sub-Saharan Africa bears over 90% of the global mortality burden, with multiple parasite species adding to the complexity of the disease.
Malaria is caused by several types of parasites, each with varying degrees of severity. Among all, *Plasmodium falciparum* is considered the most dangerous

form of the parasite responsible for malaria due to several characteristics inherent to this species. *Plasmodium falciparum* can lead to severe cases of malaria. Infections caused by this species can rapidly escalate to critical conditions, including cerebral malaria, severe anemia, and organ damage.
Early malaria diagnosis is fundamental in fighting against the disease as it helps prevent drug resistance and the spread of the disease. An efficient diagnostic system enables a quick response during epidemics, reducing the impact on public health. It is the cornerstone of strategies to combat malaria, facilitating prompt therapeutic measures, containing transmission, and optimizing the use of medical resources, all of which are essential in minimizing the impact of malaria on global health.

Challenges in automated malaria diagnosis

In order to diagnose malaria, two important medical questions need to be answered. The first question is related to the amount of parasites present in the body (parasitic load), while the second one deals with the species of the parasite. However, it can be difficult for medical professionals to analyze a sample when the parasitic load is extremely low. Standards set the detection limit for the ability to detect one parasitized red blood cell at 200000. This would require too much attention and observation time for a human to be feasible, with sufficient patient safety. For these reasons, determining the parasitic load is subdivided into two steps: the first seeks to determine if parasites are present in the blood. Highly sensitive techniques such as PCR or thick smear are used instead of smear observation in this case. The response is binary. In case of a positive response, the second step deals with determining the parasitic load. For this, a restricted area of the smear containing typically 5000 red blood cells is observed. It allows distinguishing the parasitic load level per slice (typically > 5%, between 2-5%, 0.5-2%, and less than 0.5%). These slices guide the choice of treatment.
Thus, a complete malaria diagnosis cannot be done using microscopic observation of the blood smear exclusively, and several complementary techniques are indeed required. This diagnosis must be made in less than 40 minutes. It should be noted that in addition to the various steps mentioned, smears must be prepared beforehand. This involves spreading blood on a microscope slide, fixing it, and staining it. Staining allows identifying and distinguishing cellular compartments of interest. It consists of methylene blue or methylene azure B

and eosin [101]. Depending on the nature of the parasite compartments (nucleus, vacuole, cytoplasm, ...), specific interactions/bonds are established. The immobilization of the dye molecules is then relatively specific. Depending on the case, some preferentially bind to proteins, others to amino acids [59]. Since their absorption properties are also distinct, distinguishing the nature of the compartments of interest by the appearance of different colors is possible. It should be noted that the compartments sought are often transparent. Thus, dyes make them visible to detectors sensitive to light intensity (eye or camera) by modifying their absorption properties.
The advent of sample digitization techniques by microscopy was supposed to simplify approaches and free up expert time for the most serious situations. However, this is not the case yet, and to date, there is no solution to diagnose malaria completely from a digitized smear. The company Noul proposes the closest solution to achieving this. It consists of a traditional microscope that integrates a motorization device for scanning large sample areas. A spreading and coloring system is also present to prepare the smear from a drop of blood. Finally, a deep-learning algorithm exploits the images (intensity only) to count the percentage of parasitized red blood cells.
The sensitivities achieved are still insufficient for reasons that are not completely established. Among the main hypotheses, this is due to the difficulty posed by microscopy in producing images with a fully controlled quality over a complete field of view. At the high resolutions that is employed (typically 200-300 nm), focus conditions are not constant within a field of view. This problem does not exist when a human observes the sample under the microscope, as the operator can adjust the microscope focus according to the red blood cell he is looking at. For a digitized sample, the focus is set globally over the field of view. Moreover, some fine information about the parasite may be deteriorated or lost during digitization. A camera has a linear response, unlike the human eye, which has a logarithmic response. The quantization step (or associated quantization noise) can then hinder the exploitation of images.
Motic proposes the second available solution with its easy-scan-go equipment. Similarly to Noul, a traditional microscope is used here again. The device integrates a stage to measure the thick smear to achieve sufficient sensitivity. In this sense, the solution is hybrid and does not rely solely on smear analysis.

FPM for malaria diagnosis

FPM is an alternative approach to address the challenge of studying a large number of blood cells in a reasonable time. Indeed, FPM provides a more favorable balance between resolution and field of view size than conventional microscopy techniques.
This technique will be described with more details in Chapter 3. We illustrate in this section the intensity and phase images, it provides, offer complementary information that helps to characterize the parasite and red blood cells precisely, as shown in Figure 2.4. Indeed, in addition to the compartments of the parasite visible in the intensity image $\mathbf{a_I}$, the phase image $\mathbf{a}_\Phi$ shows the nuclear material (N arrow) in white and the cytoplasm (C arrow). Additionally, the phase $\mathbf{b}_\Phi$ shows the hollow center of the red blood cell not visible on the intensity image $\mathbf{b_I}$.

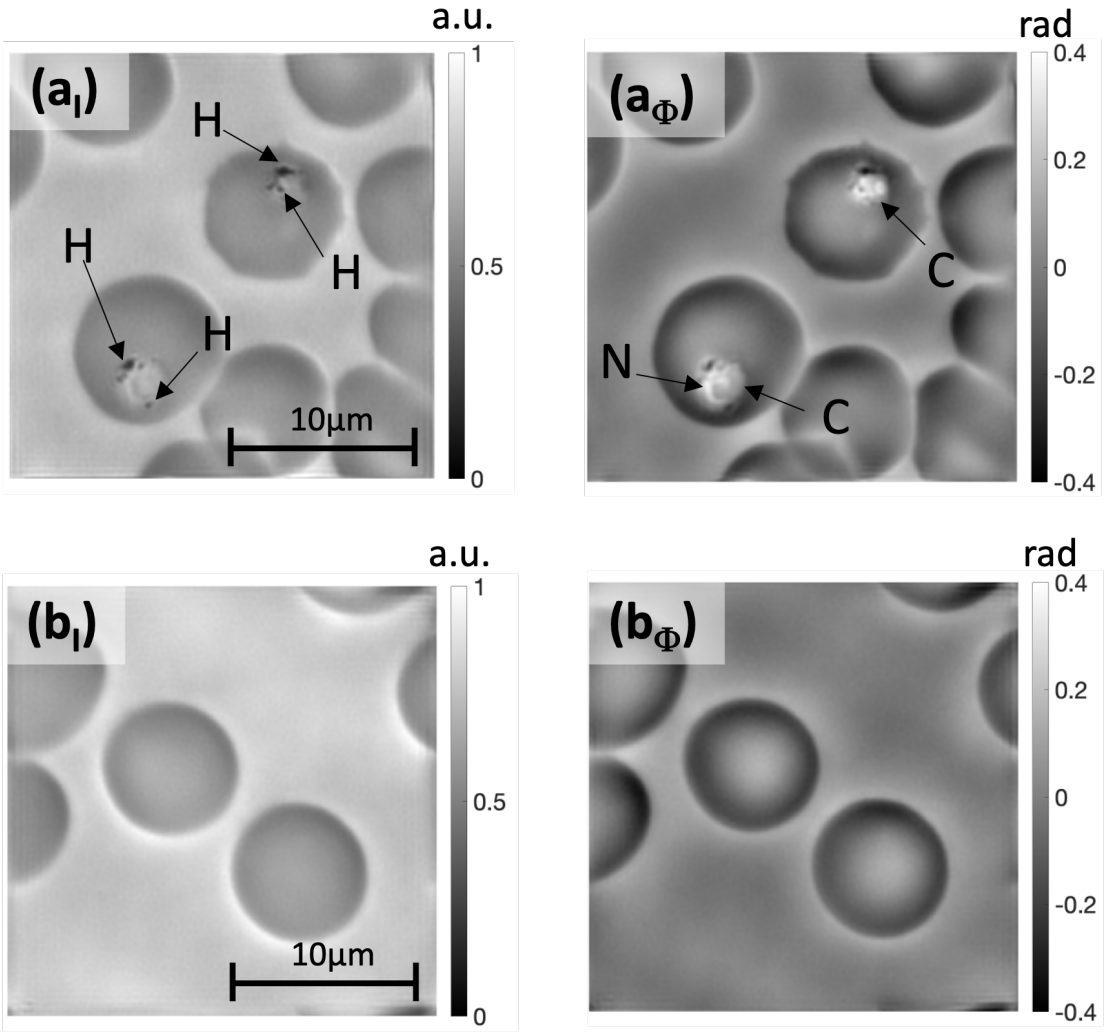


Figure 2.4: Illustrative images of stained blood smears obtained with FPM after reconstruction. ($\mathbf{a_I}$) Intensity and ($\mathbf{a_\Phi}$) phase images of two red blood cells parasitized with *Plasmodium falciparum*. In the intensity image, hemozoin pigments are indicated with the H arrow. Details of parasite structures hardly visible in the intensity image such as nuclear material N and cytoplasm C are revealed in the phase image. For comparison, intensity and phase images of healthy red blood cells are presented in ($\mathbf{b_I}$,$\mathbf{b_\Phi}$) respectively.

2.5.2 Hematology

Hematology, a specialized branch of medicine, focuses on the study of blood cells, providing a valuable window into the body's immune status and over-

all health. This discipline plays an essential role in the early diagnosis of diseases, monitoring immune responses and detecting any hematological abnormalities. At the heart of diagnostic hematology is the identification and counting of leukocytes, commonly known as white blood cells, which are of paramount importance in the evaluation of various medical conditions.
Leukocyte identification is a important process in the diagnosis of medical conditions. It involves the differentiation of different types of white blood cells present in a blood sample, each characterized by distinct morphological traits observable under a microscope. For example, neutrophils are distinguished by their multilobed nucleus and fine cytoplasmic granulations. On the other hand, Eosinophils are characterized by a bilobed nucleus and eosinophilic granulations and are involved in allergic responses and parasitic infections. Basophils have a bilobed nucleus and cytoplasmic granulations, which are major in allergic reactions. Lymphocytes, notably T and B lymphocytes have round nuclei and are essential for the immune response, while monocytes, with their kidney-shaped nucleus and agranular cytoplasm, develop into macrophages in tissues. Leukocyte count, another essential component of hematology, involves determining the concentration of white blood cells in a microliter of blood. This measurement is generally carried out using hematological devices, which allow rapid and precise counting, thus facilitating diagnosis and clinical monitoring.
White blood cell identification and counting are important in diagnosing many medical conditions. In particular, they make it possible to detect bacterial and viral infections, autoimmune diseases such as systemic lupus erythematosus, leukemia, and blood cancers, allergies, and parasitic infections. Additionally, hematological disorders and genetic diseases, such as aplastic anemia, thrombocytopenia or sickle cell disease, can be diagnosed through white blood cell analysis.

Limits of classical microscopy in white blood cell classification

Despite its long history of use in diagnostic hematology, conventional light microscopy has several intrinsic limitations that limit its ability to classify white blood cells accurately. Among these limitations, optical resolution plays a important role. This technique is subject to the constraints of the optical properties of the lenses used and the wavelength of visible light. These resolution limitations can make it difficult to distinguish the fine morphological features of different types of white blood cells, mainly when dealing with subpopulations

of cells with morphological similarities. Another major limitation of conventional optical microscopy lies in the morphological resemblance between certain classes of white blood cells. For example, neutrophils, eosinophils, and basophils can share similar morphological characteristics, making them difficult to differentiate using conventional microscopy. This morphological similarity can lead to confusion in the manual classification of cells, thus compromising diagnostic accuracy. Finally, manual identification and classification of white blood cells using conventional microscopy can be time-consuming and require skilled human resources, which can be costly in terms of time and labor. This time and cost constraint may limit the scalability of this approach and its ability to handle large sample volumes.

FPM advanatges for white blood cells classification

FPM offers a promising solution to overcome the limitations inherent in conventional optical microscopy, particularly with regard to the accurate classification of white blood cells. It provides access to the intensity and phase pair. By exploiting phase, which is not usually exploited, this technique can provide information on refractive index variations within cells, which can be used to differentiate cells with similar morphological characteristics but different optical properties.
Figure 2.5 illustrates the difference between images of a white blood cell captured using conventional microscopy and those taken using FPM. Acquiring images of white blood cells with a conventional microscope can often be affected by (a^1) focusing problems. On the other hand, FPM offers the possibility of carrying out digital focus correction after the acquisition, thus making it possible to obtain improved and homogeneous focus across all acquisitions. Furthermore, even when the focus is correct (a^1), the detail and contrast of the chromatin in the nuclei of the globules and the definition of the granules are often insufficient. However, when we observe the images obtained from the FPM (b_I) and (b_ϕ), they present a different contrast. The complementarity between these two types of images allows a better definition of cellular compartments, in particular thanks to the phase image (b_ϕ), which highlights the details of the chromatin in the nuclei of the globules.

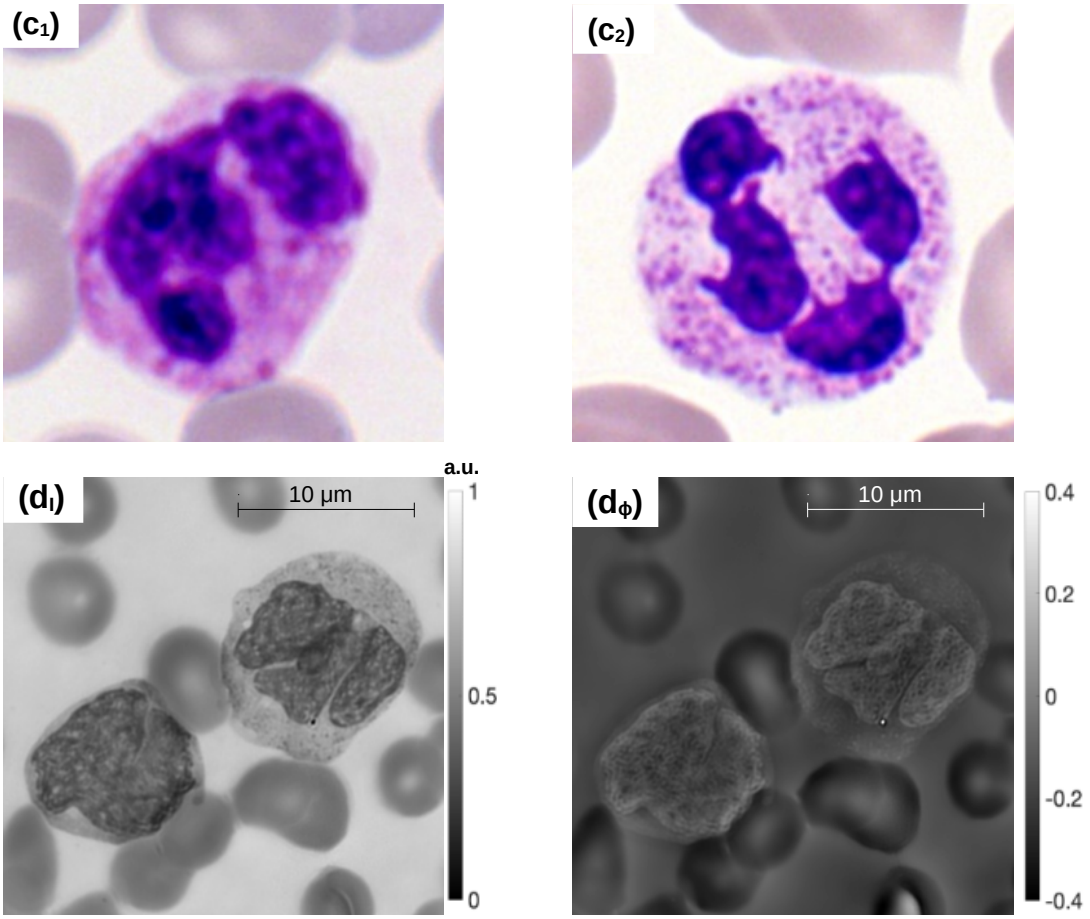


Figure 2.5: Illustrative images of stained blood smears obtained. $\mathbf{c_1}$ and $\mathbf{c_2}$ are images from *A single-cell morphological dataset of leukocytes from AML patients and non-malignant controls (AML-Cytomorphology LMU)*. ($\mathbf{c_1}$) white blood cell out of focus and ($\mathbf{c_2}$) a single white blood cell on focus. ($\mathbf{d_I}$,$\mathbf{d_\Phi}$) are respectively the intensity and the phase image of white blood cells from an FPM acquisition.

2.5.3 Conclusion

The images presented in Figure 2.4 and Figure 2.5 suggets that FPM holds great promise forimproving medical diagnosis in parasitology and hematology. In parasitology, where the early detection of malaria is critical for global public health, FPM imaging offers a precise visualization of Plasmodium parasites in blood samples, providing a valuable alternative to conventional microscopy. In hematology, FPM imaging can differentiate blood cells with similar morpho-

logical characteristics but different optical properties, improving the accuracy of leukocyte classification and facilitating the diagnosis of various medical conditions.
FPM imaging has significant advantages over conventional light microscopy, such as improved resolution and increased ability to differentiate similar cells. It is exepcted that its adoption in parasitology and hematology can lead to significant improvements in medical diagnosis, which can have positive implications for the clinical management of patients and global public health.

Chapter 3

Phase contribution for stained objects analysis

Contents

The main reasons for using bimodal microscopy lie primarily in its ability to reveal invisible structures on label-free samples. This is particularly interesting for transparent samples as it has been studied in the literature.
In this chapter, in contrast to the application cases covered in previously published works, we question the relevance of bimodal microscopy in the context of stained samples. To address this question, we consider a biological application (malaria) that requires automatic detection of parasitized red blood cells on images of stained smears. This topic is chosen because it is still unsolved. Its relatively limited complexity (two classes of objects to detect, healthy red blood cell / parasitized red blood cell) is used to evaluate the interest in exploiting phase images in addition to intensity images. For this purpose, a performance study in terms of sensitivity and specificity is conducted. The performances obtained are compared to those relying on intensity images alone. Faster R-CNN architecture is used to detect red blood cells. Training deep neural networks often requires a large dataset. We focus on FPM microscopy rather than quantitative phase microscopy (QPI) or digital holography to facilitate their acquisition. This recent technique is stable and relatively insensitive to environmental conditions (including vibrations). It is, therefore, well adapted for acquiring large sample areas with reduced experimental effort, as is necessary for dataset collection.
The chapter is organized as follows. The first section introduces the role of dyes and their possible effect on phase images. A simplified optical model is introduced and studied to identify the possible distinctions that can be expected between an intensity image and a phase image on stained smears. The second section introduces the principle of FPM microscopy. Its main characteristics and positioning are detailed compared to digital holography approaches, including QPI. In the third section, the experimental work and data exploitation are introduced. The performances achieved by exploiting intensity and phase images are discussed comparatively to those obtained from intensity images alone.

3.1 Phenomenological model of stained samples

As mentioned in section 2.5.1, dyes are used to modify the optical properties (absorption) of the targeted cellular compartments, allowing their identification and distinction based on their spectral absorption characteristics.
Regarding image analysis, parasite detection involves identifying typical mor-

phological forms within red blood cells. These forms result from contrasts in signal intensity due to the presence of dyes. Each pixel's relative variations in intensity levels are thus exploited, as are their spectral variations associated with compartment colors.
However, another optical property that could potentially be relevant for parasite detection is the variations in optical path length introduced by different cellular compartments of the parasite, as illustrated in Figure 3.1.

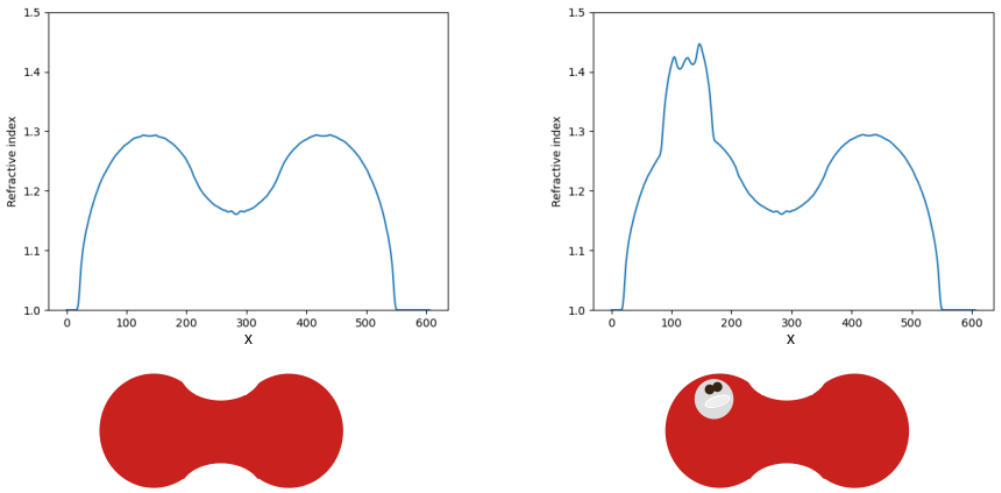


Figure 3.1: Schematic cross-section of a red blood cell (left) a red blood cell not containing a parasite, (right) a red blood cell infected with a parasite.

The left side of the figure illustrates a schematic cross-section of a red blood cell. The optical index is homogeneous throughout the volume, with no nucleus present. However, when the cell hosts a parasite, i.e., its complete cell (right), the overall volume of the cell may be altered. Additionally, specific compartments within the parasite cell may introduce an optical index different that of the red blood cell, such as its cytoplasm, nucleus, and vacuoles. This is undoubtedly the case for hemozoin, a compound resulting from the parasite's hemoglobin degradation.
By definition, the optical path length traveled by light is the curvilinear integral of the path G taken by light from point A to B.

$$L_\Gamma = \int_\Gamma n \, dl \tag{3.1}$$

Given that the optical index is affected by the presence of the parasite, its pres-

ence should be observable in the phase images recorded on the sample since :

$$\Delta\phi = \frac{2\pi L_\Gamma}{\lambda} \tag{3.2}$$

An illustration of the expected phase signal for a sample measured in transmission is presented in Figure 3.1(top). For a healthy red blood cell, the phase measured as a function of x should exhibit two lobes; however, when the cell hosts a parasite, this curve is no longer as regular. It should exhibit more or less pronounced variations depending on the thickness and nature of the parasite compartments traversed.

Phase images are, therefore, interesting as they can reveal fine variations in the optical thickness of the measured sample. Digital holography or QPI approaches are traditionally used to obtain high-resolution phase images. Their main advantage is that they can dispense with any sample labeling or staining, i.e., label-free imaging.

To the best of our knowledge, the potential benefit of characterizing samples through phase imaging on stained smears has not yet been studied. However, it seems entirely conceivable that the additional information provided by a phase image and the intensity image could make distinguishing between a healthy red blood cell and a parasitized red blood cell easier for image analysis algorithms. This represents a promising approach to explore when seeking to improve parasite detection sensitivity.

Nevertheless, ensuring that these two images are not equivalent is essential. To do this, let's analyze the impact of a dye on the phase image using a phenomenological optical model of the sample. The simplest model often used describes the sample as a collection of dipoles [1] with a dipole moment $\vec{P}(r)$, where P is the volumetric dipole moment of the sample and r is the coordinate space vector (x, y, z). The resulting total dipole moment associated with the collection of dipoles contained within an elementary volume dv centered at r then equals

$$dP = \vec{P}\,(r)\; dv \tag{3.3}$$

This dipole collection description is interesting because it allows us to relate the material's optical index to $\vec{P}$. Let us recall that a dipole is defined as a system composed of two opposite electric charges of the same magnitude, separated by a certain distance, as illustrated in Figure 3.2; $\vec{P}$ is represented by an arrow indicating the direction of the dipole moment that measures the separation of positive and negative electrical charges. In this description, the system's behav-

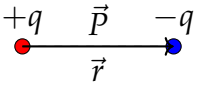


Figure 3.2: Representation of an electric dipole with a positive charge $+q$ and a negative charge $-q$ separated by a distance $\vec{r}$, forming a dipole moment $\vec{p}$.

ior influenced by an external electromagnetic field is described by the equation associated with a harmonic oscillator, see Equation 3.4 :

$$m\frac{d^2\vec{r}}{dt^2} + m\gamma\frac{d\vec{r}}{dt} + m\omega_0^2\vec{r} = \vec{F}_{\text{ext}} \tag{3.4}$$

Where m is the mass of the dipole, γ is the damping coefficient, ω_0 is the natural frequency of the dipole, and $\vec{F}_{\text{ext}}$ is the external force of the electromagnetic field. The term $m\omega_0^2\vec{r}$ represents the restoring force, assumed to be the same in all directions (proportional to the electron's movement) in the case of an isotropic oscillator. The term $m\gamma\vec{r}$ describes the energy dissipation process experienced by the dipole.
Under the effect of a monochromatic electric field at a frequency ω, this dipole experiences quasi-elastic forces and, therefore, has a natural oscillation frequency ω_0 related to a spring constant. The oscillating motion is governed by the equation 3.5. In Fourier domain this equation can also be written as follows:

$$-m\omega^2\tilde{r} + jm\omega\gamma\tilde{r} + \omega_0\tilde{r} = qE(\omega) \tag{3.5}$$

Under the hypothesis of a linear response, the optical index of the medium can be deduced from this equation and the formula of the electrical susceptibility in the form:

$$\begin{aligned} n = n' + jn'' \text{ with } n' &= \sqrt{1 + \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega)^2 + \gamma^2\omega^2}} \\ \text{and } n'' &= \sqrt{1 + \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}} \end{aligned} \tag{3.6}$$

The optical index is complex. The real part of n, n' is associated with the phase delay experienced by light propagation and the imaginary part n'' with its absorption properties. Its spectral characteristics are shown in Figure 3.3. The left panel corresponds to the imaginary part of n, while the right panel represents

its real part. The model's free parameters are w_0 and γ. They can be completely determined from the measured absorption spectral properties (resonance wavelength and linewidth).

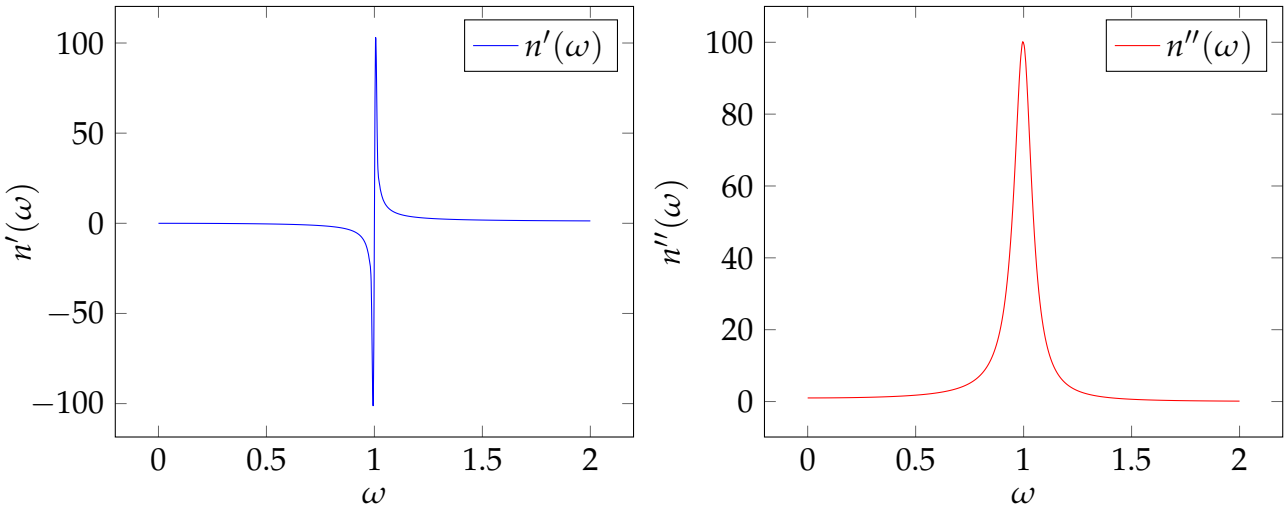


Figure 3.3: Example of Kramers-Kronig relations between the real and imaginary part of the refractive index for a specific ω_0.

It is interesting to note that near the wavelength where absorption is maximal, significant variations in phase are expected. This implies that the presence of the dye should also be observable in the phase image. However, the contrasts expected in the intensity and phase images should differ because of the different curve evolutions. Thus, although we expect to observe various compartments of the parasites in phase images attached to label-free samples, these observations should be even more evident on stained samples. This comes from the relationship between the imaginary part of the index (absorption) and its real part (phase shift). Although the invoked model is phenomenological, this property is general due to the Kramer-Krönig relations [66], which are always valid and based on a principle of causality. In the continuation of the work, we will mainly consider stained samples for the following reasons:

- Effects of dyes on phase images are expected.
- Ground truth, as defined by biologists, relies on observing stained samples. Comparing the analyses obtained by the developed microscopy and those currently used by biologists will thus be simplified.

3.2 Fourier ptychography microscopy for phase imaging

As mentioned in Chapter 2, throughout this thesis, we focus on Fourier Ptychographic Microscopy (FPM) technique. We use it to access phase images because it's easy to implement and highly tolerant to vibrations. This allows it to measure large number of samples, which is necessary for the studies conducted. The quantativity of the phase images produced remains unresolved in the literature. Although we don't directly address the question, we are more interested in identifying whether the phase images it is capable of producing can be of interest for automatic diagnosis.

3.2.1 Device and principle

FPM is a cost-effective device that uses an LED array illumination light source instead of the conventional microscope light source, a monochromatic camera, as shown in Figure 2.1. Unlike interferometry systems, these devices do not require complex optical components such as lasers and precise optical mirrors, which considerably lower costs.
The data collection procedure of FPM is a simple process that involves placing a two-dimensional sample at the microscope's focal plane and illuminating it successively with each diode of the LED matrix. This process enables the acquisition of multiple low-resolution images that are subsequently reconstructed into a high-resolution image using an iterative algorithm. FPM does not require any reference beam to generate interference since it only collects intensity images, which reduces environmental stability requirements and can be a significant advantage in clinical settings.

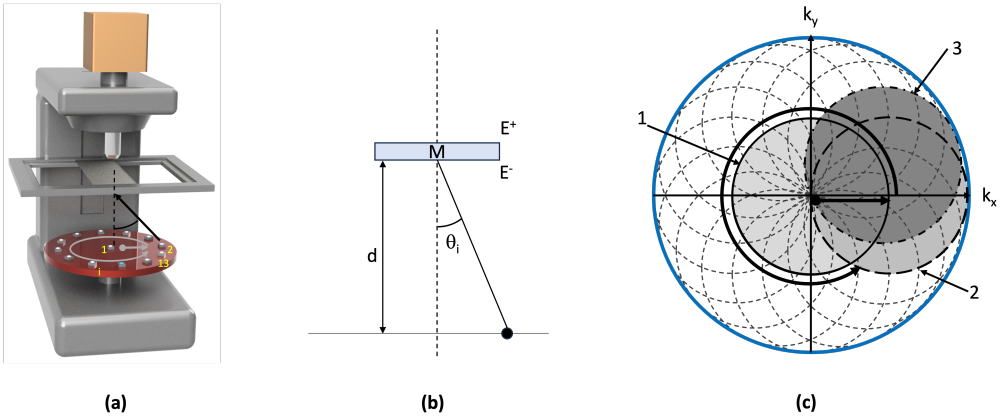


Figure 3.4: FPM principle. (**a**) Sketch of microscope configuration equipped with its LED matrix, (**b**) illustration of the angular illumination of the sample as determined with spatial position of LED i, (**c**) individual spectral regions acquired by each individual LED. The different raw images captured by the camera are assembled in Fourier Domain with a phase retrieval algorithm. The grayed regions are related to LEDs 1, 2, and 3.

In detail, let us consider an LED array consisting of a central diode placed on the optical axis and a ring array of n-1 individually controllable LEDs. The diodes in the matrix are conveniently indexed from 1 to n and are positioned in a sequential order starting from the center. This led matrix is placed at distance d (typically several cm) such that the illumination at the sample level can be considered coherent (Zernike –Van Cittert theorem) [144] and assimilable to a plane wave of wave vector $\mathbf{k_i}$. The wave vector $\mathbf{k_i}$ is defined as: $|\mathbf{k_i}| = \frac{2\pi}{\lambda}$ by a far-field approximation. Each diode i is thus assigned to a unique illumination propagation direction $\mathbf{u_i}$ of inclination θ_i. λ represents the wavelength used. Under these conditions, the electromagnetic field immediately beneath the sample that originates from the diode i is written as:

$$E_i^-(x,y) = A \cdot e^{j(k_i^x \cdot x + k_i^y \cdot y)} \tag{3.7}$$

where A, k_i^x, and k_i^y represent its complex amplitude and the projection of the vector $\mathbf{k_i}$ along the x and y axes. Furthermore, under a thin-sample approximation, the interaction between the light and the sample is modeled by a transfer matrix $M(x,y)$. For the LED i, the electromagnetic field immediately above the sample is calculated from M with $E_i^+(x,y) = M(x,y) \cdot E_i^-(x,y)$ and the electromagnetic field formed at the camera plane by the microscope objective of the

point-spread-function PSF(x,y) is:

$$E_i^{\text{cam}}(x,y) = \left[M\left(\frac{x}{G},\frac{y}{G}\right)\cdot E_i^{-}(x,y)\right] * \text{PSF}(x,y) \tag{3.8}$$

where G represents the magnification of the microscope objective and * is the convolution product, as in [39]. In the spectral domain, E^{cam} is expressed as:

$$\widehat{E}_i^{\text{cam}}(k^x,k^y) = A\cdot\widehat{M}\left(G\left(k^x-k_i^x\right),G\left(k^y-k_i^y\right)\right)\cdot\text{CTF}\left(k^x,k^y\right) \tag{3.9}$$

Where CTF denotes the coherent optical transfer function of the microscope objective (the Fourier Transform of PSF). For a lens that is aberration free, the $CTF\left(k^x,k^y\right))$ function is equal to 1 inside a disk D of radius $r = 2\pi\cdot\frac{NA}{\lambda}$ centered at $(0,0)$ and zero elsewhere. Here, NA denotes the objective's Numerical Aperture, and $\widehat{\ }$ denotes the function in the Fourier domain. The i^{th} recorded image I_i is then associated with the following equation:

$$I_i(x,y) = \left|E_i^{\text{cam}}(x,y)\right|^2 \tag{3.10}$$

This last equation corresponds to the forward problem relative to the image formation of the sample with the i^{th} LED. Examination of Equations (3.9) and (3.10) reveals that the region of the object spectrum imaged with diode i corresponds to the spectral region of M bounded by the $\left(k_i^x,k_i^y\right)$ translated disk D. For illustrative purposes, figure 3.4.c shows the zones corresponding to each LED in the spectrum.

It is interesting to note that capturing the sample from various illumination angles allows for the covering of a much broader spectral region than what is accessible by the objective lens. To utilize this information, it is necessary to know how to assemble the different images in the spectral domain appropriately (aperture synthesis mechanism) and also to retrieve the phase of the images, which is inevitably lost during measurement.

For this purpose, a reconstruction algorithm is employed. It relies on solving an inverse problem and utilizes the various acquired images. The model incorporates knowledge of the objective's transfer function and the source configuration (position of each diode in space, emission wavelength).

3.2.2 Reconstruction process

Image reconstruction through the inversion of forward models is a standard process in unconventional optics. Solving the inverse problem involves minimizing an error function using algorithms like the Gerchberg-Saxton (GS) algorithm or its variants [152, 32].
Several algorithms have been proposed to recover intensity and phase. In this thesis, we explore Extended PIE (e-PIE) [81], an improved version of PIE [115]. E-pie was chosen because it incorporates probe correction, support constraints, and enhanced convergence properties.
The e-PIE algorithm operates through iterative steps to refine the estimates of the object and illumination functions, continuing until convergence is attained. It begins by initializing the object to be reconstructed, denoted as $O_i(r)$, and the illumination function $P_i(r)$, representing the illuminating waveform. Typically, the initial object is a randomly chosen function awaiting determination, while the initial waveform is a roughly estimated support function based on the illumination region of the wavefront.
Then, the complex object undergoes multiplication by the illumination function, applying a specific shift to each region $R(i)$ knowing that each region corresponds to one measure made by the device:

$$\psi_i(r) = O_i(r)P_i(r - R(i)) \tag{3.11}$$

Next, a Fourier transform is applied to the resulting $\psi_i(r)$:

$$\tilde{\psi}_i(\nu) = \sqrt{I_i(\nu)}\frac{F[\psi_i(r)]}{|F[\psi_i(r)]|} \tag{3.12}$$

Subsequently, a corrected output field is computed:

$$\psi_i'(r) = F^{-1}[\psi_i(\nu)] \tag{3.13}$$

Finally, for iteration i, the object O_{i+1} and P_{i+1} are updated using the ePIE algorithm:

Algorithm 1 e-PIE Algorithm.

Input: Experimental measurements $\psi_i'(r)$
Output: Reconstructed object $O_{i+1}(r)$ and $P_{i+1}(r)$
Initialize object $O_i(r)$ and $P_i(r)$
Set maximum number of iterations $N_{\max}$ and convergence criterion

Loop: $i \leftarrow 1$ **to** $N_{\max}$

- Calculate $\psi_i(r) = O_i(r)P_i(r - R(i))$
- Perform Fourier transform: $\tilde{\psi}_i(\nu) = \sqrt{I_i(\nu)}\frac{F[\psi_i(r)]}{|F[\psi_i(r)]|}$
- Compute corrected output field: $\psi_i'(r) = F^{-1}[\psi_i(\nu)]$
- Update object: $O_{i+1}(r) = O_i(r) + \alpha\frac{P_i^*(r-R(i))}{|P_i(r-R(i))|^2_{\max}}(\psi_i'(r) - \psi_i(r))$
- Update probe: $P_{i+1}(r) = P_i(r) + \beta\frac{O_i^*(r+R(i))}{|O_i(r+R(i))|^2_{\max}}(\psi_i'(r) - \psi_i(r))$

  If Convergence criterion is met:

  – **Break**

In the e-PIE algorithm, α and β represent the correction coefficient used to weigh the object and illumination function corrections, respectively. This iterative process continues until convergence or until the maximum number of iterations is reached, resulting in a reconstructed phase and intensity of the object.

3.2.3 Intensity and phase images obtained on label-free and stained sample

Figure 3.5 shows FPM images obtained on a parasitized red blood cell to illustrate the impact of dyes. The top images correspond to the intensity and phase images produced on the sample under label-free conditions. Conversely, after staining the smear, the bottom images correspond to the images obtained on the same cells after staining.

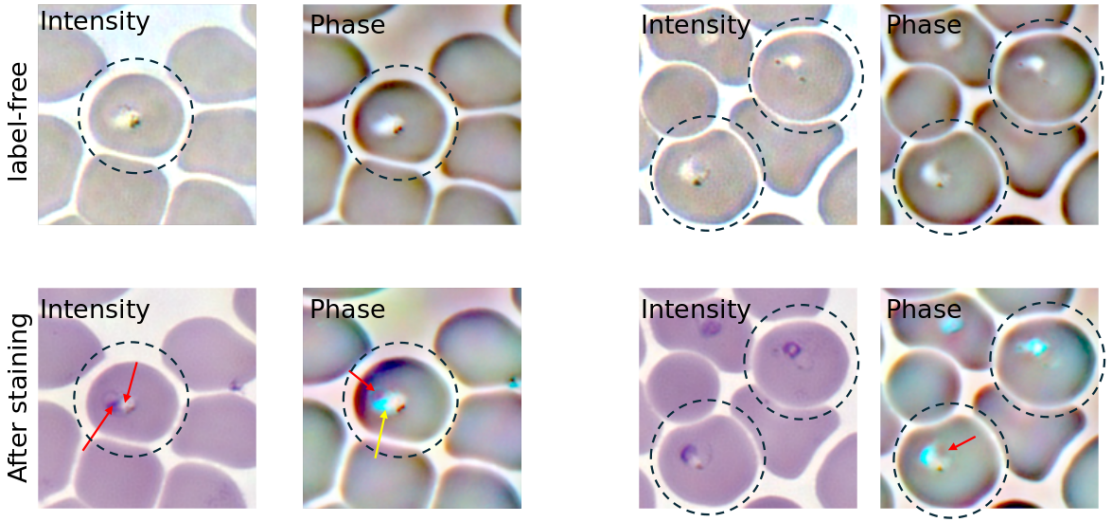


Figure 3.5: FPM images from a parasitized red blood cell. The top row displays intensity and phase images acquired under label-free conditions, while the bottom row showcases images obtained after staining the smear.

The intensity image serves as a reference, making it easy to locate the cellular compartments of the parasite due to the staining provided by MGG. Noticeably, the phase image also allows for the identification of the main compartments sought. It also presents color variations by combining images corresponding to an optical delay in the three channels (Red, Green, and Blue). The appearance of colors should thus be interpreted as variations in the light propagation time in the sample depending on the wavelength. Although the noticeable compartments in these two images are close, the arrows indicate noticeable differences. The yellow arrow points to a delicate structure not visible in the intensity image.
Examining the images taken before staining, we can still detect the parasite. However, many details about the compartmental structure are lost, and there is minimal disparity between the intensity and phase images.
Thus, as envisaged in section 3.1, staining seems to impact the phase image. It allows for obtaining images with complementary information and reveals structures at specific locations that are different from those identifiable in the intensity image. The significance of this complementarity needs to be clarified from the perspective of automatic diagnosis through statistical analysis. This question is addressed in the following section.

3.3 Automatic bimodal analysis of blood smear images

In this study, the case of *Plasmodium Falciparum* is examined to study the contribution of the phase in FPM for the stained samples. *Plasmodium Falciparum* is the causative agent of the most severe and usually fatal form of malaria. At the beginning of the infection, parasitemia may be quite low. The goal is to detect red blood cells and classify them as parasites or healthy cells. At the time of the study, the data was acquired using a single channel, green. This was due to the device's limitations, which only allowed this wavelength exploitation, and also because the green channel is particularly interesting for studying red blood cells and parasites [103].
This section will detail the dataset used for evaluation, the deep learning architecture implemented for detecting parasites in red blood cells over the complete field of view (localization and classification of red blood cells into two classes), and the results obtained.

3.3.1 Microscope configuration and dataset

The microscopic configuration used consists of an upright microscope equipped with a Plan Apochromat objective with a magnification of 10× and a numerical aperture of 0.45. The camera uses a large "1.1" CMOS sensor with a resolution of 12.34 Mpix and a pixel size of 3.45 µm (UI-3200SE from IDS). The camera has a global shutter and an optical area of 14.158 mm × 10.370 mm. The total resolution of the camera is 4104 × 3006 pixels. Similar to [126, 76], the LED array consists of a central diode placed on the optical axis and from a ring array of 12 individually controllable LEDs (from Adafruit industry). Each sample is illuminated by a single wavelength $\lambda = 525$. The data set was created from scanning thin blood smears from nine patients and was scanned using the FPM microscope setup described above. The area of good spreading is fully scanned with 16 fields of view (FOV). For each FOV acquired, 13 raw images were acquired, each corresponding to an LED with a specified illumination angle. The reconstruction of the bimodal phase intensity images is then carried out using the E-Pie algorithm. The reconstructed images are then cropped into smaller bimodal images, so that each image contains approximately 200 to 300 red blood cells. An expert labeled all fields by indicating the position of each red blood cell with a box around the cells and providing a label indicating the presence or

absence of a parasite. The final dataset contains 2216 bimodal intensity phase images of 896 $\times$ 896 pixels, showing 418389 healthy cells and 65140 infected cells. Although the database is limited in number of patients, it should be noted that the images contain parasites of different sizes and the infection levels vary between 1 and 5%, which is representative of real-world use cases in the diagnosis of malaria.

3.3.2 Faster-RCNN architecture

Numerous object detection models have been developed in the literature. These models aim to localize objects in an image and label them simultaneously. These models vary in accuracy, speed, and complexity, offering different trade-offs depending on the application's needs. They generally rely on a reasonably sized labeled database, which indicates each object localization and class label. One of the earliest region-based approaches for object detection was R-CNN [35], which successfully detected objects with high precision. However, it needed to be faster due to its multiple processing steps and required significant computational resources. Fast R-CNN [34] was introduced to overcome this limitation, which used a single convolutional pass over the entire image and managed regionalization externally. This approach improved speed and accuracy but still needed to overcome the challenge of generating regional proposals. Faster R-CNN [113] introduced the Region Proposal Network (RPN) to address this issue, integrating regionalization more effectively. However, its implementation can be challenging, and its speed depends on the number of proposals generated. Another popular object detection model is YOLO (YouOnly LookOnce). YOLOv1 and v2 are well-known for their fast and efficient single-inference passes, making them ideal for real-time object detection. However, the available versions at the moment this work was conducted may struggle with identifying small items and are less precise in complex scenes. SSD (SingleShot Multibox Detector) was introduced to address this limitation, offering high performance and speed that is capable of detecting objects of various sizes. However, it may deliver less accurate results than FasterR-CNN for smaller objects [111, 7]. Lastly, YOLO v3 boasts improved accuracy compared to YOLO v1 and v2 and is better at identifying small items. However, it requires significant computing power. In this chapter, the Faster-RCNN architecture was explored for its ability to generate regional boundaries with minimal time cost compared to older versions of RCNN and for its accuracy in detecting small

objects compared to other models such as Yolo and SDD. Its architecture is pre-

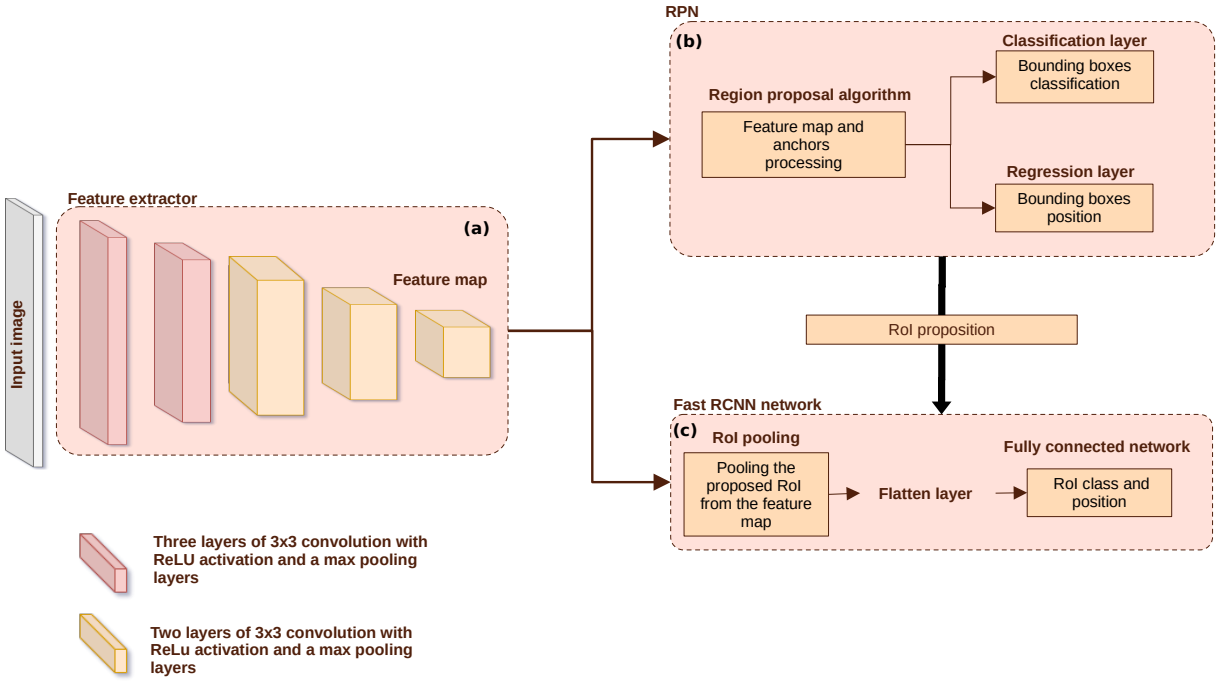


Figure 3.6: The architecture and different modules of the Faster-RCNN model: (**a**) the feature maps extraction module, (**b**) the region proposal networks modules, and (**c**) the Fast-RCNN networks module for classification.

sented in the figure, it relies on two modules: two decision networks (b) and (c) connected through a feature extractor (a). An image is given as an input to the model and processed by the feature extractor layers, represented in Figure 3.6 (a). Here, the feature extractor is based on VGG-16 architecture [124], which is a succession of 16 layers of 3x3 convolutions with ReLU activation functions and pooling layers. A feature map is then produced and transmitted simultaneously to modules (b) and (c).

Module (b) corresponds to the Regional Proportional Network (RPN). The RPN is responsible for detecting objects of interest in an image by generating a set of candidates represented by potential bounding boxes that may contain objects of interest. Formally, the RPN proposes the bounding box coordinates $\mathbf{p}$ and their associated probability $P(\text{object})$ of containing an object of interest.

Its architecture is presented in Figure 3.6; it relies on two modules: two decision networks (b) and (c) connected through a feature extractor (a). An image is given as an input to the model and processed by the feature extractor layers

Figure 3.6 (a). Here, the feature extractor is based on VGG-16 architecture [124], which is a succession of 16 layers of 3x3 convolutions with ReLU activation functions and pooling layers. A feature map is then produced and transmitted simultaneously to modules (b) and (c).
Module (b) corresponds to the Regional Proportional Network (RPN). The RPN is responsible for detecting objects of interest in an image by generating a set of candidates represented by potential bounding boxes that may contain objects of interest. Formally, the RPN proposes the bounding box coordinates $\mathbf{p}$ and their associated probability $P(\text{object})$ of containing an object of interest.
This is achieved using a set of predefined anchors $\mathbf{a}$, placed at evenly distributed positions across the feature map generated by the CNN backbone; the anchors are predefined bounding boxes of various sizes and aspect ratios (Usually, three different size and three ratios are fixed before the learning step according to the objects in the dataset specificities, this produces 9 different anchors that are placed at different positions of the feature map). The generated proposals $\mathbf{p}$ are merged from the anchors using offsets $\Delta(x, y, w, h)$ predicted by the network. The score associated with each proposition is calculated by considering the probability that the proposition matches an object of interest and the intersection over union (IoU) between the proposition and the actual bounding box.

$$\text{score}(\mathbf{p}, \mathbf{r}) = P(\text{object}) \cdot P(\text{anchor}) \cdot \text{IoU}(\mathbf{p}, \mathbf{r}) \tag{3.14}$$

$$\mathbf{p} = \mathbf{a} + \Delta(x, y, w, h) \tag{3.15}$$

Anchors $\mathbf{a}$ are defined as boxes of fixed size with different positions x_a, y_a and scales w_a, h_a deduced from the size and ratio.

$$\mathbf{a} = (x_a, y_a, w_a, h_a) \tag{3.16}$$

The $\text{score}(\mathbf{p}, \mathbf{r})$ represents the probability that the proposition $\mathbf{p}$ corresponds to an object of interest, and it is used to classify propositions. These proposals are then used to generate regions of interest (RoIs) which will be submitted to the next module for the classification (Label association) and regression task (Bounding Box coordinates adjustment). The set of bounding boxes is sent to the second branch (module (c)) with the feature map resulting from the module (a).
In module (c), the corresponding region of the bounding box $\mathbf{p}$ is pooled from

the feature map $\mathbf{f}$ by a RoI mechanism Pooling or RoIAlign, then sent to the Fast-RCNN network to predict its class $P(\text{Class}|\mathbf{p})$ and provide its final coordinates $\Delta(x, y, w, h)$. These predictions will be used to refine the RPN results.

$$\text{RoIAlign}(\mathbf{f}, \mathbf{p}) = \sum_{i,j} \mathbf{f}_{ij} \cdot \text{area}(\text{int}(\text{bin}(\mathbf{p}))_{ij}) \tag{3.17}$$

$\text{bin}(\mathbf{p})$ calculates the bins (discrete spatial locations) of the RoI.

$$P(\text{Class}|\mathbf{p}) = \text{Softmax}(\mathbf{W}_{\text{class}} \cdot \text{FC}([\text{RoIAlign}(\mathbf{f}, \mathbf{p}), \mathbf{f}, \mathbf{p}]) + \mathbf{b}_{\text{class}}) \tag{3.18}$$

FC denotes the fully connected layer, W_{class}, b_{class} denote the weights and bias of the classification layer, respectively.

$$\Delta(x, y, w, h) = \mathbf{W}_{\text{regression}} \cdot \text{FC}([\text{RoIAlign}(\mathbf{f}, \mathbf{p}), \mathbf{f}, \mathbf{p}]) + \mathbf{b}_{\text{regression}} \tag{3.19}$$

$W_{regression}, b_{regression}$ denote the weights and bias of the regression layer, respectively.

3.3.3 Results and discussion

The intensity and phase images resulting from the FPM reconstruction are explored using Faster-RCNN. The objective is to detect the red blood cells and classify them into parasites or healthy cells (Binary classification). We recall that our dataset is composed of 418389 healthy cells and 65140 infected cells More precisely, two Faster-RCNN models for *Plasmodium Falciparum* parasites detection were implemented; the models have the same number of trainable parameters:

- A classical real-valued Faster-RCNN using intensity only ($\mathcal{I}$-RV);
- A classical real-valued Faster-RCNN using intensity and phase ($\mathcal{I}/\phi$-RV);

The specifics of the architecture are detailed in Table 3.1.

Module	Layer	Input Shape	Output Shape	Kernel Size
Feature extractor	Convolution	(896, 896, 2)	(896, 896, 64)	(3, 3)
	Convolution	(896, 896, 64)	(896, 896, 64)	(3, 3)
	MaxPooling	(896, 896, 64)	(448, 448, 64)	(2, 2)
	Convolution	(448, 448, 64)	(448, 448, 128)	(3, 3)
	Convolution	(448, 448, 128)	(448, 448, 128)	(3, 3)
	MaxPooling	(448, 448, 128)	(224, 224, 128)	(2, 2)
	Convolution	(224, 224, 128)	(224, 224, 256)	(3, 3)
	Convolution	(224, 224, 256)	(224, 224, 256)	(3, 3)
	Convolution	(224, 224, 256)	(224, 224, 256)	(3, 3)
	MaxPooling	(224, 224, 256)	(112, 112, 256)	(2, 2)
	Convolution	(112, 112, 256)	(112, 112, 512)	(3, 3)
	Convolution	(112, 112, 512)	(112, 112, 512)	(3, 3)
	Convolution	(112, 112, 512)	(112, 112, 512)	(3, 3)
	MaxPooling	(112, 112, 512)	(56, 56, 512)	(2, 2)
	Convolution	(56, 56, 512)	(56, 56, 512)	(3, 3)
	Convolution	(56, 56, 512)	(56, 56, 512)	(3, 3)
	Convolution	(56, 56, 512)	(56, 56, 512)	(3, 3)
RPN	Convolution	(56, 56, 512)	(56, 56, 512)	(3, 3)
	Convolution	(56, 56, 512)	(56, 56, 9×4)	(1, 1)
	Convolution	(56, 56, 512)	(56, 56, 9)	(1, 1)
Fast RCNN	RoI proposition function	-	(1500, 4)	-
	RoI Pooling function	-	(7, 7, 512)	-
	Flattening layer	(7, 7, 512)	(25088)	-
	Fully connected	(25088)	(4096)	-
	Dropout (0.5)	(4096)	(4096)	-
	Fully connected	(4096)	(4096)	-
	Dropout (0.5)	(4096)	(4096)	-
	Fully connected	(4096)	(4×3)	-
	Fully connected	(4096)	(3)	-

Table 3.1: Parameters of Faster-RCNN architecture as shown in Figure 3.6.

Several key hyper-parameters in the Faster-RCNN architecture have been tuned. The search for the best hyper-parameters for each network was carried out in such a way as to have an optimal classification performance for each experience. A random search was used to choose the maximum number of boxes to be predicted by the RPN. Moreover, we noticed that this number greatly influences the rate of well-detected parasitized cells (Positive Rate). Indeed, if a small number of boxes is chosen, several cells are not detected, which gives a low positive rate, knowing that most cells are healthy. Contrariwise, a very high number of boxes leads to certain confusion (several boxes are predicted for one object), degrading the classification quality. For $\mathcal{I}$-RV, this parameter was fixed to 1200 boxes, while in the $\mathcal{I}/\phi$-RV model, the optimal parameter was fixed to 1500 boxes as in the original paper. Among others, the anchors used in the RPN are the most influential parameters, which have to be chosen considering the variability in blood cell sizes between patients. This was done by calculating the average size of the boxes in the images, the minimum and the maximum, and the standard deviation. This led us to choose the three sizes 30, 50 and 70, with ratios 0.5, 1 and 2 producing the 9 anchors. The total number of predictions produced by the Fast RCNN model (branch c) was set at 301 by studying the maximum number of boxes in the images used.
All models have been trained using a 5-fold cross-validation. This method is used here to establish the confidence intervals. More precisely, in this work, we used a k-split cross-validation algorithm [125], as it better estimates the model's generalization ability by performing not one but k measures of the validation metric. Its principle is simple: The data sample is divided into k smaller sample datasets. k independent evaluations of the model are then made. A single sample dataset is retained for testing the model at each evaluation, while the remaining samples are used for training.
The experiments' results are averaged, and the standard deviation is computed to estimate each evaluation metric robustly. Here, we chose a five-fold cross-validation where the original dataset was partitioned into five equitably sized subsets (roughly 443 samples per subset).
To compare the performance of the two implementations, we use two criteria, the True Positive Rate *TPR* and the True Negative Rate *TNR* of parasite detection. This is performed on the same databases and with the same protocols. *TPR* and *TNR* result from the definitions of *TP*, *TN*, *FP*, and *FN*, which were adapted slightly to our specific *Plasmodium falciparum* detection task according

to malaria specialists as follows :

$$TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \quad TNR = \frac{TN}{N} = \frac{TN}{TN + FP} \tag{3.20}$$

- True positive *TP* is the number of well-classified parasitized red blood cells;
- True negative *TN* is the number of well-classified healthy red blood cells, the extra healthy cells, and missed healthy cells;
- False positive *FP* is the sum of the wrongly classified healthy red blood cells and the extra parasite blood cells;
- False negative *FN* is the sum of the wrongly classified parasitized red blood cells and the missed parasites.

The parasite detection results for the different implementations are in Tables 5.1 and 5.2. The details of the output boxes for each model are shown in Table 3.2.

	$\mathcal{I}$-RV	$\mathcal{I}/\phi$-RV
Well-classified infected	12,116	12,320
Misclassified infected	528	433
Well-classified healthy	81,818	82,778
Misclassified healthy	551	295
Missed infected	356	186
Missed healthy	1086	601
Added infected	127	154
Added healthy	347	554

Table 3.2: Detailed results of Faster-RCNN predictions averaged to the nearest unit over the five folds.

A localization threshold of 0.7 was used to construct the table. This means that a predicted bounding box is considered correctly positioned if it overlaps with the corresponding ground truth box by at least 70%. If the predicted label matches the ground truth label, the model considers the cell in this box well-classified. Otherwise, the cell is considered misclassified. The missed cells correspond

to the boxes containing cells labeled by the specialist that the model did not predict. Lastly, the model's added boxes are the boxes that have poor overlap with ground truth or boxes that do not correspond to labeled objects. Integrating phase information positively impacts the model performance for classifying healthy and infected cells shown in Table 3.2. The $\mathcal{I}/\phi$-RV model demonstrates superior capability than the $\mathcal{I}$-RV model for this task. Specifically, the $\mathcal{I}/\phi$-RV model correctly predicted 204 more infected cells and 960 more healthy cells than the $\mathcal{I}$-RV model. Additionally, the improvements are also evident in the reduction of classification errors. The number of misclassified infected cells is reduced by 95, while the number of misclassified healthy cells is also reduced by 256. Furthermore, the model reduces the number of undetected healthy cells by 485 and the number of undetected infected cells by 170. However, it should be noted that there is a slight increase in the number of false cell detections. This increase can be attributed to the increased number of region proposals in branch (b). In the context of model performance comparison, $\mathcal{I}/\phi$-RV is more efficient for the task than $\mathcal{I}$-RV. This suggests that integrating phase information enhances the classification abilities of the Fast R-CNN module. From Table 3.3,

	$\mathcal{I}$**-RV**	$\mathcal{I}/\phi$**-RV**
TNR	99.18 ± 0.20%	99.34 ± 0.24%
TPR	93.33 ± 1.33%	95.22 ± 0.73%

Table 3.3: Detailed results of Faster-RCNN predictions averaged to the nearest unit over the five folds.

we observe a high *TNG* and a *TPR*. If the *TNG* is equivalent between the two models. The varying *TPR* is 93.33% for $\mathcal{I}$**-RV** and 95.22% for $\mathcal{I}/\phi$**-RV**. These values are promising, as the main goal in *Plasmodium Falciparum* detection is to obtain a high *TPR* to detect cases of early-stage infection, where parasite cells are extremely sparse among the total red blood cells.
To statistically validate the performance difference between these two models, we employed the McNemar test [109, 88]. The resulting p-values, lower than 0.001 for each fold, indicate that the two models do not share the same error rate. The results, therefore, illustrate the contribution of phase information to improving the quality of diagnosis of rare events. In this case study, we observe that the phase image brings complementary information to intensity, even in the case of non-transparent objects, where the phase is usually considered to

have a minor effect. It can, therefore, provide valuable information for analyzing microscopic images.

3.4 Conclusion

In the first part, this chapter presents the details of the FPM, which is considered in this thesis to be an efficient and non-costly tool for acquiring phase information in complement of intensity. In the second part, the impact of the bimodal intensity-phase images was validated in a relatively uncomplicated but unsolved malaria diagnostic case study: detecting and classifying Plasmodium falciparum parasites in red blood cells. By combining intensity and phase information, FPM significantly improves the sensitivity of infected cell detection compared to intensity alone. These first results will be further studied in the next chapter. Indeed, Chapter 4 will explore the benefits of FPM for a more complex classification task (White Blood Cell), which has several challenges, such as class imbalance and a restricted quantity of data.

Chapter 4

White blood cell classification

Contents

In the previous chapter, a proof of concept demonstrated the contribution of the joint use of phase and intensity derived from FPM for stained samples by studying the simple problem of detecting malaria parasites in red blood cells. This chapter extends these promising results to address a more complex challenge: classifying white blood cells from stained thin blood smears.
This study is part of TAMIS project. TAMIS project focuses on the reliable and efficient automated characterization of blood cells, encompassing leukocytes, schistocytes, and platelets from thin blood smears, using FPM images that allow access to enriched information. This chapter will only address the leukocyte problem.
The main goal is to accurately and automatically classify distinct leukocyte subtypes typically found in the bloodstream of healthy individuals, including neutrophils, eosinophils, basophils, lymphocytes, and monocytes.
The first part of this work aims to extend the proof of concept of the effectiveness of bimodal information in classifying colored samples to a more complex problem. This problem is considered more complex than parasite detection because it involves 5 classes that strongly resemble each other, with natural biological imbalances in the representation of these classes in the blood.
The second part proposes an innovative method for generating synthetic images to complete the under-represented classes of white blood cells, aiming to improve classification performance. This method's originality lies in first introducing the physical image formation model in a GAN model, increasing the diversity of the images generated and reducing the phenomenon of collapse mode.
A protocol for using similarities between certain classes to propose a tine-tuning approach is implemented. It offers the possibility of learning GANs for synthesizing images of classes with only a very limited number of samples.

4.1 FPM bi-modal images contribution for white blood cells classification

Given the regulatory framework of the "RGPD" and anonymization constraints, this work will focus on white blood cell classification in patients without hematological problems into five classes, namely neutrophils, lymphocytes, monocytes, eosinophils, and basophils. Although this step does not yet allow the diagnosis of hematological diseases such as leukemia, it is crucial for detect-

ing infections, autoimmune diseases, inflammatory reactions, and allergic reactions.
As Section 2.5.2 explains, specialized centers use automatic systems to carry out this step. However, due to their limited performances in many situations, manual proofreading by a specialist using a conventional microscope is often necessary, thus highlighting the need to improve this automatic classification.
This study involves the creation of a first dataset of images of white blood cells using FPM that will be described in Section 4.1.1 and the choice of classification models adapted to this task. This study mainly investigates FPM contribution by comparing the performances of three neural network models:

- a model exploiting only the intensity (I-CNN)
- a model exploiting only the phase (ϕ-CNN)
- a model exploiting the intensity-phase jointly (I-ϕ-CNN)

to study the differences and similarities in the classification results obtained by these three configurations and their respective advantages in distinguishing the different classes of white blood cells.
Furthermore, we extend our analysis by comparing these results with those obtained by a model trained using conventional images provided by a reference white blood cell classification device. Those images were acquired on the same thin blood smears of our FPM dataset.

4.1.1 White blood cell FPM dataset

Microscope configuration

The microscope setup is similar to that used in Chapter 3, where the objective lens was replaced with a $20\times$ objective with a numerical aperture of 0.75. This configuration uses three wavelengths separately to scan the samples: $\lambda = 625,\ 525,$ and 465.

The dataset

This study focuses on patients without hematologic pathology from the intensive care unit. These patients mainly have healthy white blood cells from the 5 classes mentioned earlier but may have some abnormal white blood cells that rarely appear. The blood smears were prepared with a Sysmex spreader stainer.

They were separated into two batches, each with a specific purpose: the first batch was used for training the models, and the second was used to evaluate generalization performance.
The first batch included 54 thin blood smears, while the second had 75. For each slide, 60 fields were scanned to cover the area of good spreading, ensuring a minimum of 200 white blood cells per slide, or approximately three blood cells per field. An operator carefully examined each field to extract white blood cells centered in 512×512 pixels images.
Then, for each white blood cell, a reconstruction was carried out for the three wavelengths in intensity and phase, thus making it possible to create a representation including three intensity modalities and three phase modalities, as illustrated in Figure 4.1 (a). Each modality carries unique information due to the distinct absorption properties of dyes for each wavelength. The phase images are also different, which is likely related to Kramer-Kronig relations. These different modalities provide additional information; for example, the phase in the red channel reveals the presence of vacuoles, invisible in the other phase's images, while the green channel intensity map shows a granulation density not visible in the other modalities.
In addition to these 6 modalities, the combination of the three intensity channels is also carried out to provide a color image of the specimen for annotation by a specialist (See Figure 4.1 (b)).

A hematology expert carefully annotated each image using a classification system including the five classes of white blood cells. Additional classes have been added to exclude blurry images and white blood cells that do belong to abnormal classes to exclude them, ensuring reliable annotations.
After this first labeling step, the annotations were checked and corrected if necessary by the expert, thus guaranteeing the quality and consistency of the annotated data. This meticulous annotation process ensured the quality and reliability of the dataset while eliminating problematic elements.

Table 4.1 presents the number of samples used in the training and test datasets. We notice a strong imbalance in the number of representatives, which reflects natural imbalances in the human body. The models were trained on Batch 1 using a 75% split for the training and 25% split for the validation and tested on Batch 2.

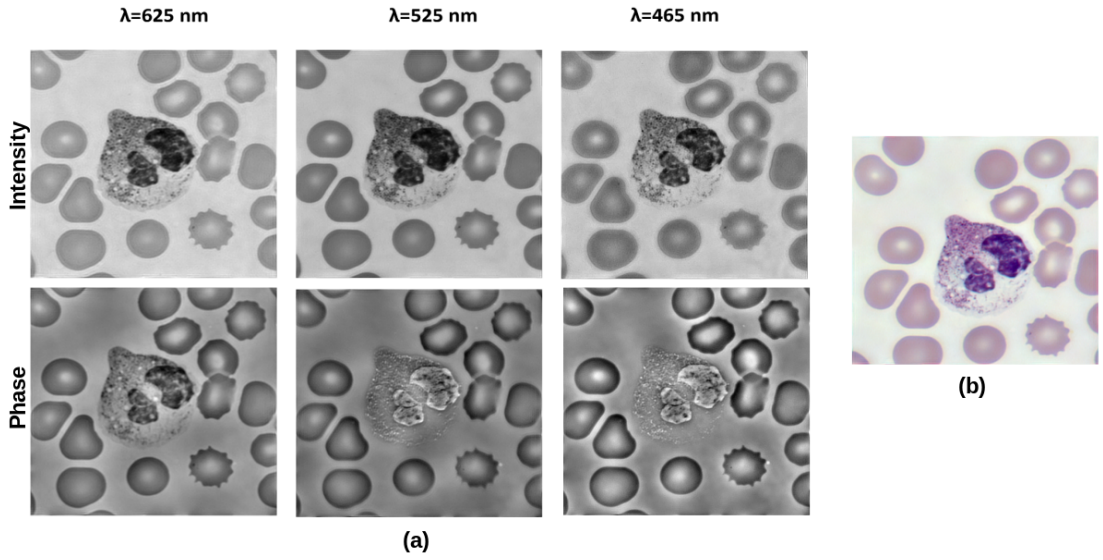


Figure 4.1: (a) Example of a neutrophil polynuclear white blood cell acquired with FPM in three wavelengths reconstructed according to 6 different modalities (Objective lens 20×0.75); (b) Combination of the three intensity images to reproduce an image similar to that seen with conventional color camera microscopy.

Cell Type	**Batch 1**	**Batch 2**
Neutrophils	8520 *(73%)*	11378 *(72%)*
Eosinophils	296 *(2%)*	294 *(2%)*
Basophils	56 *(Less than 1%)*	93 *(Less than 1%)*
Lymphocytes	2144 *[18%)*	2906 *(18%)*
Monocytes	616 *(5%)*	1023 *(7%)*
Total	11632	15694

Table 4.1: White blood cell counts in Batch 1 and Batch 2 collected in FPM white blood cells dataset.

4.1.2 Architecture choice

Initially, the architecture chosen was MobileNetV1 [50] due to its relatively large representation capacity compared to its small size, making it suitable to be embedded as part of an automatic scanning and classification system design. How-

ever, this choice was quickly abandoned because, during the learning process, the curves presented extreme oscillations, as shown in Figure 4.2.

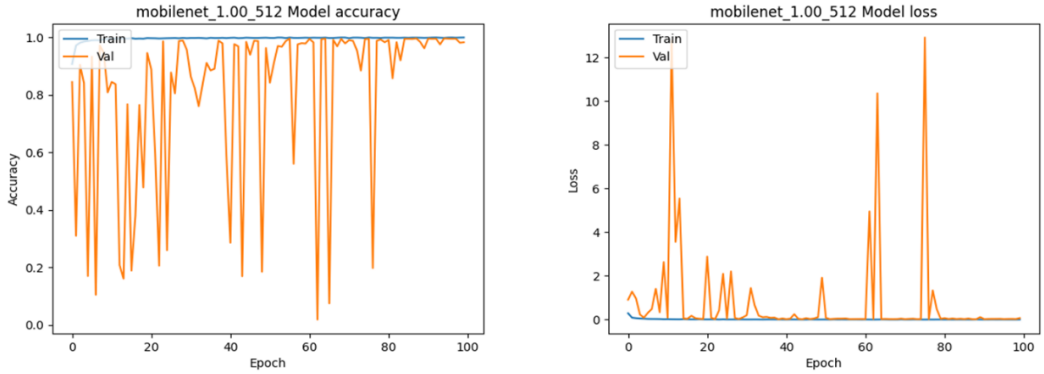


Figure 4.2: MobileNet architecture learning and accuracy curves showing high oscillations.

These oscillations may indicate overfitting due to the network's high modeling capacity, which explains its poor fit on the validation set at some points. They can also be due to numerical instability when calculating gradients. Unfortunately, this model does not make it possible to reliably choose an optimal point based on the validation set.
For this reason, VGG16 [124] architecture was tested but quickly showed signs of overfitting on the training set. To limit this overfitting, we reduced the number of free parameters of this architecture by modifying the number of neurons in the fully connected layers and introducing dropout layers for regularization. The architecture is detailed in Table 4.2.

Layer	Type	Description
Convolution layers		
1	Conv2D	64 3x3 filters, ReLU
2	Conv2D	64 3x3 filters, ReLU
3	MaxPooling2D	window 2x2
4	Conv2D	128 3x3 filters, ReLU
5	Conv2D	128 3x3 filters, ReLU
6	MaxPooling2D	window 2x2
7	Conv2D	256 3x3 filters, ReLU
8	Conv2D	256 3x3 filters, ReLU
9	Conv2D	256 3x3 filters, ReLU
10	MaxPooling2D	window 2x2
11	Conv2D	512 3x3 filters, ReLU
12	Conv2D	512 3x3 filters, ReLU
13	Conv2D	512 3x3 filters, ReLU
14	MaxPooling2D	window 2x2
15	Conv2D	512 3x3 filters, ReLU
16	Conv2D	512 3x3 filters, ReLU
17	Conv2D	512 3x3 filters, ReLU
18	MaxPooling2D	window 2x2
Fully connected layers		
19	Dense	500 units, ReLU
20	Dropout	rate = 0.5
21	Dense	200 units, ReLU
22	Dropout	rate = 0.5
23	Dense	5 units, softmax

Table 4.2: Description of the adapted layers of the VGG16 model.

In the original architecture, the fully connected layers had 4096 neurons each, representing a high complexity. This number was reduced to 500 for the first

fully connected layers and 200 for the second connected layers. Two dropout layers with 50% dropout were introduced for regularization. The last fully connected layer is set to 5 neurons, according to the classification task need. The models were trained with cross-entropy loss and a learning rate set to 10^{-4} for all experiments, with a batch-size equal to 64. The models' performances are compared on three metrics :

- **Accuracy**: This metric measures the overall correctness of the model's predictions by comparing the total number of correct predictions to the total number of samples. It provides a general assessment of the model's performance across all classes.
- **Precision**: Precision focuses on the relevance of the model's positive predictions within a specific class. It calculates the ratio of true positive predictions to all positive predictions made by the model for that class. In other words, precision quantifies how many of the predicted positive samples are actually relevant.
- **Recall**: Recall, also known as sensitivity, measures the ability of the model to identify all relevant instances of a class correctly. It calculates the ratio of true positive predictions to all actual positive instances in the dataset. Recall provides insight into how well the model captures all positive cases, regardless of the number of false positives.

in addition, two other metrics were used, **Average Precision (Avg Precision)** that measures the overall precision of a model across multiple classes. It is calculated by averaging the precision values for each class, and **Average Recall (Avg Recall)** that quantifies the overall recall of a model across multiple classes. It is calculated by averaging the recall values for each class.

4.1.3 Results and discussion

Bi-modal information contribution

The performance of the I-ϕ-CNN, I-CNN and ϕ-CNN models were evaluated for white blood cell classification. The results are presented in the two tables 4.3 and 4.4, where the standard deviations were calculated according to the Gaussian approximation, taking into account the large size of the test set and some statistical hypothesis and assumptions as explained in [109].

Model	Accuracy	Avg. Precision	Avg. Recall
I-CNN	96.19% $\pm$ 0.001	88.34%	81.44%
ϕ-CNN	94.96% $\pm$ 0.001	86.66%	83.88%
I-ϕ-CNN	98.34% $\pm$ 0.001	94.66%	93.82%

Table 4.3: Globals models performance metrics.

In Table 4.3, the I-ϕ-CNN model performs best in all measured metrics. Although its accuracy is close to that of the other models, it still shows 2% improvement. The fact that accuracy values are near for the models can be attributed partly to the strong representation of the neutrophil class, which all three models characterize well due to its high number of samples, contributing significantly to overall accuracy.
However, the difference becomes more notable when looking at average Precision and average Recall metrics. The I-ϕ-CNN model stands out with an average precision of 94.66% and a recall of 93.82%, demonstrating its ability to characterize the classes and effectively separate them distinctly.
The I-CNN and ϕ -CNN models exhibit similar performance levels. However, when analyzing their average Precision and average Recall, we can see differences in their ability to characterize and distinguish different classes of white blood cells.
Comparing these two models, we notice that the I-CNN model demonstrates a higher average precision, implying that it performs better in separating white blood cell classes. However, it may encounter difficulties characterizing each element correctly, as evidenced by its lower average recall. In contrast, the ϕ -CNN model exhibits a higher average recall, indicating its superior ability to classy white blood cells. Nevertheless, it may struggle with separating the different classes due to its lower average Precision.
This suggests that the contribution observed by jointly exploiting intensity and phase is not due to only one of the modalities but to the reinforcement of information by exploiting both, which shows an important interest in using information from the phase, even in the case of colored objects.

For a more detailed analysis of the performance of each model, it is necessary to examine specific metrics for each class, as presented in Table 4.4.

By analyzing the performance for each class, it becomes evident that the I-

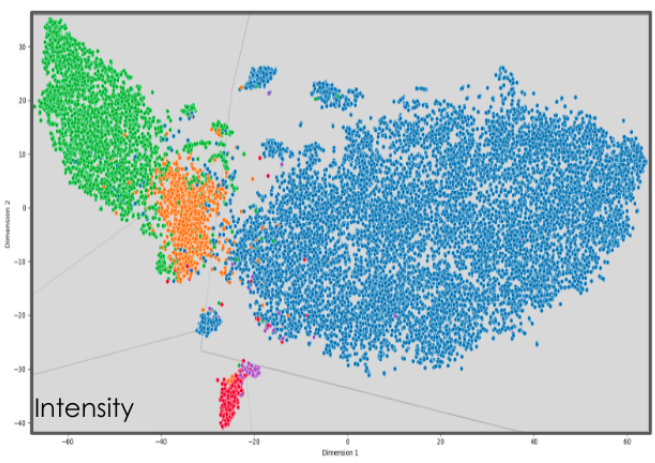
Classes	I-ϕ-CNN		I-CNN		ϕ-CNN	
	Precision	**recall**	**Precision**	**recall**	**Precision**	**recall**
Basophils	98.64% $\pm$ 0.013	81.49% $\pm$ 0.042	91.72% $\pm$ 0.037	49.91% $\pm$ 0.051	88.15% $\pm$ 0.037	72.04% $\pm$ 0.046
Éosinophiles	86.30% $\pm$ 0.018	98.63% $\pm$ 0.006	67.71% $\pm$ 0.023	92.00% $\pm$ 0.016	72.29% $\pm$ 0.025	77.21% $\pm$ 0.024
Lymphocytes	98.27% $\pm$ 0.002	95.96% $\pm$ 0.003	95.56% $\pm$ 0.004	95.14% $\pm$ 0.003	93.83% $\pm$ 0.012	90.24% $\pm$ 0.012
Monocytes	88.69% $\pm$ 0.009	93.63% $\pm$ 0.007	89.08% $\pm$ 0.010	71.06% $\pm$ 0.014	81.97% $\pm$ 0.012	81.89% $\pm$ 0.012
Neutrophils	99.54% $\pm$ 0.0006	99.43% $\pm$ 0.0007	97.99% $\pm$ 0.001	99.29% $\pm$ 0.007	97.05% $\pm$ 0.001	97.85% $\pm$ 0.001

Table 4.4: Metrics performances of each class.

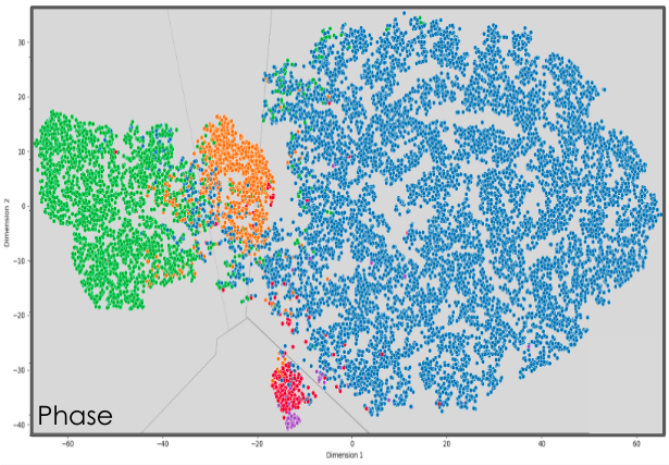
ϕ-CNN model generally has the best performance among the three evaluated models. Although the three models present similar performances for the neutrophil class, significant differences appear for the other classes where the I-ϕ-CNN model stands out for its improved performance. In particular, it handles class imbalances better than the other two models. The small amount of data in some classes, such as eosinophils (296 images) and basophils (56 images), further affects the performance of the I-CNN and ϕ-CNN models. The I-CNN model shows 49.91% Recall for basophils and 67.71% Precision for eosinophils, while the ϕ-CNN model shows 72.04% Recall for basophils and lower precision of 72.29% for eosinophils. In contrast, the I-ϕ-CNN model stands out with higher Recall for eosinophils 98.64% and higher Precision for basophils 81.49%. This shows that the joint exploitation of intensity and phase when classes are underrepresented improves performances, even if these performances must be improved to reach the levels required for clinical use.

The model also improves performance in classes with more representatives, such as monocytes and lymphocytes. In general, the I-ϕ-CNN model stands out for its ability to provide both high Precision and high Recall for all white blood cell classes. It is followed by the ϕ-CNN model, which shows slightly lower but still strong performance, demonstrating the importance of phase even for colored samples. In contrast, the I-CNN model generally presents the lowest performance among the three models evaluated, which is consistent with the weaknesses observed in conventional microscopy.

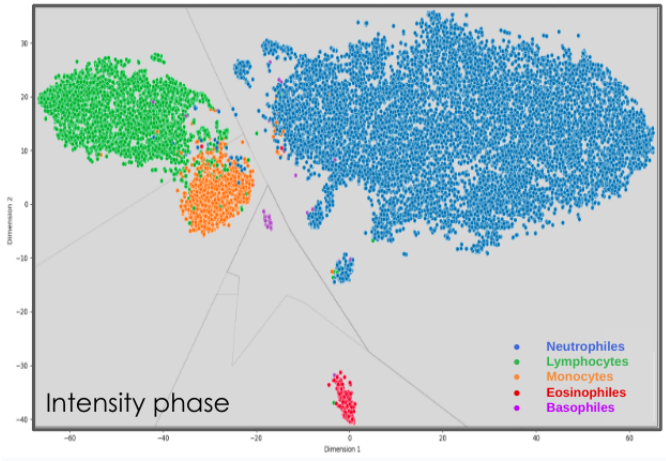
To better visualize our models' differences, a dimensionality reduction technique, t-SNE (t-distributed Stochastic Neighbor Embedding) [138], was performed on the images to visualize the feature representations extracted by the three models. t-SNE allows the data to be projected into a reduced-dimensional space while preserving the relationships (even non-linear) between the original points as much as possible. Using this technique, we can visualize the data distribution and evaluate the separability of classes in the feature space. In this case, we projected the representative vector of size 200 before the output layer provided by each model into a two-dimensional space.



(a) t-SNE I-CNN representation.



(b) t-SNE ϕ-CNN representation.



(c) t-SNE I-ϕ-CNN representation.

Figure 4.3: t-SNE representation for each model.

Figure 4.3, the t-SNE for each model is represented. Each point in the graph represents an image in the dataset, and the colors of the points indicate the corresponding class labels. The light gray lines represent the decision boundary between different data classes in this reduced-dimensional space.
By observing the three representations, we see that the main advantage of the I-ϕ-CNN model lies in its ability to better separate classes than other models. Unlike the other two models, we can see the basophil and eosinophil classes, which are far apart from each other and all the other classes as well.
Additionally, images of the monocyte class are represented as a continuation of the lymphocyte class for the I-CNN and ϕ-CNN models. This is a common confusion, for example, among young hematologists. However, in the I-ϕ-CNN model, it is possible to better separate these two classes by creating more distant groups with a few border elements.
For the I-CNN and ϕ-CNN models, the linear separators presented fail to distinguish class boundaries clearly. This indicates that the boundaries between classes are complex and cannot be summarized as a linear boundary. For the I-ϕ-CNN model, we can see that the linear separators are sufficient to separate the classes accurately. This suggests that this model might have a better ability to learn the distinctive features of each class, thus allowing it to separate them more efficiently in the latent space, unlike both I-CNN and ϕ-CNN models.

Furthermore, an analysis of the classification results was carried out with the hematologists; it involved examining all the classification errors made by the models and their respective images to identify any discernible patterns. This analysis showed that for the I-ϕ-CNN model, the border elements between monocytes and lymphocytes, often difficult to classify, are mainly due to a subclass, granular lymphocytes, a case of granule lymphocytes. This observation made the confusion understandable.
Probability analysis showed that, in this case, the model produces shared probabilities of affiliation distributed equally between the two classes of lymphocytes and monocytes.
The creation of a class of grain lymphocytes could also help resolve this confusion and would constitute an interesting solution to explore in the future.

Comparaison with existing commercial system

The following work studies the differences between FPM images and conventional images by comparing the results obtained from FPM images to those obtained by a commonly used imaging commercial system.
To carry out this comparison, the blood smears digitized in FPM were digitized using the Di-60® Cellavision, an automated machine used in a clinical routine that allows, among other things, the digitization of white blood cells.
Several differences can be noted between the two configurations. While we use a $20\times$ magnification lens with a numerical aperture of 0.75, the Di-60® uses a $100\times$ magnification lens with a numerical aperture 1.5. Indeed, thanks to the synthetic numerical aperture mechanism and the configuration of the LEDs chosen, the FPM configuration achieves a numerical aperture of 1.5. In turn, while preserving a wider field of view in one shot, our device collects a minimum of 3 white blood cells with a high resolution. For the Di-60® device, the numerical aperture 1.5 guarantees high resolution (around 206 nm according to the Nikon), but the use of high magnification, which in turn considerably reduces the size of the field of view in one shot and does not allow you to acquire only one white blood cell at a time.
Furthermore, the Di-60® configuration incorporates a color camera, enabling direct production of color images containing a single white blood cell with dimensions of 352×357 pixels. Unfortunately, the device does not allow any phase imaging or reconstruction. A first visual comparison of the same white blood cells acquired by the two methods is illustrated in Figure 4.4.

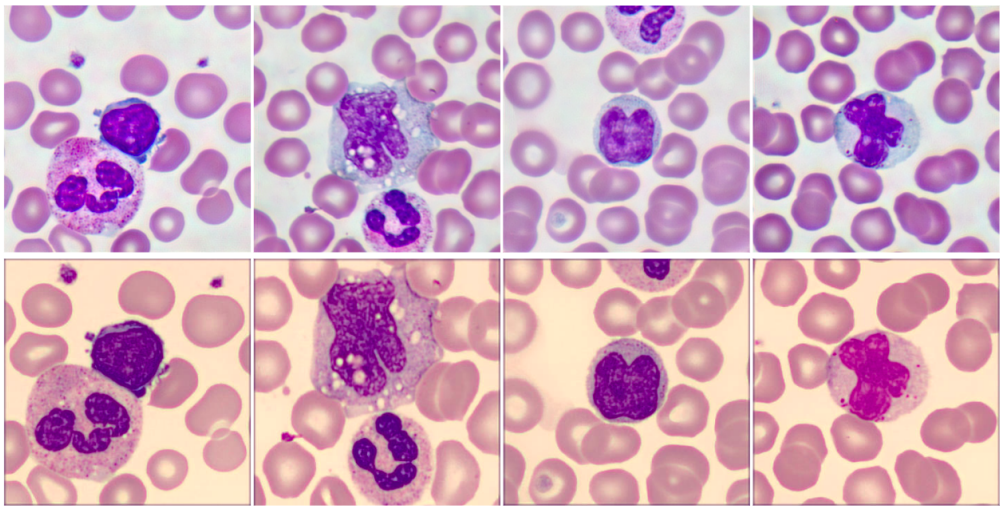


Figure 4.4: Comparison of FPM intensity images (top) and their equivalent in Di-60 (bottom) on different white blood cells.

In Figure 4.4, FPM images (top) have the same quality and level of resolution as their Di-60 equivalents (bottom). FPM demonstrates that it can obtain images of similar resolution with half the numerical aperture (0.75) and a synthetic numerical aperture mechanism. Additionally, although only the intensity is represented here, this method provides access to the sample phase.

The Di-60® dataset was acquired and labeled employing an identical protocol as that used for the FPM dataset; a Di-60-CNN model was trained on a training dataset containing the same blood smears as those used in the FPM dataset. Although it cannot be guaranteed that the same cells were captured, this discrepancy is insignificant given the minimal variability of white blood cells within a single healthy patient.
The architectural framework employed for the Di-60-CNN model mirrors the one used for the FPM dataset and is consistent with the same training procedures. Subsequently, performance metrics were computed for both models and reported in Table 4.5.

Classes	I-ϕ-CNN		Di-60-CNN	
	Precision	**recall**	**Precision**	**recall**
Basophils	98.64% $\pm$ 0,013	81.49% $\pm$ 0,042	100% $\pm$ 0,0	51.92% $\pm$ 0,06
Éosinophiles	86.30% $\pm$ 0,018	98.63% $\pm$ 0,006	67.79% $\pm$ 0,06	78.43% $\pm$ 0,05
Lymphocytes	98.27% $\pm$ 0,002	95.96% $\pm$ 0,003	95.18% $\pm$ 0,006	95.36% $\pm$ 0,006
Monocytes	88.69% $\pm$ 0,009	93.63% $\pm$ 0,007	87.28% $\pm$ 0,01	84.39% $\pm$ 0,01
Neutrophils	99.54% $\pm$ 0,0006	99.43% $\pm$ 0,0007	97.74% $\pm$ 0,002	98.98% $\pm$ 0,003

Table 4.5: Metrics performances of each class.

These results highlight the notable superiority of the I-ϕ-CNN model compared to the Di-60-CNN, particularly for the basophil, eosinophil, and monocyte classes. Confusion between monocytes and lymphocytes is significantly more pronounced in Di-60-CNN, where the model only provides 84% Precision and Recall. When examining the performance of the models for the basophils and eosinophils. The I-ϕ-CNN model shows an accuracy of 98.64% , while the Di-60-CNN model achieves a perfect precision of 100%. However, regarding Recall, the I-ϕ-CNN model achieves 81%, while the Di-60-CNN model only achieves 51.92%. For eosinophils, the I-ϕ-CNN model shows superior performance with 86.30% Precision and 98.63% Recall, while the Di-60® RV-CNN model shows 67.79% lower Precision and a recall of 78.43%.
This significant difference can be attributed to a problem od data imbalance between the classes that affects the Di-60-CNN model more, resulting in lower performance in recognizing and separating these classes.

This study shows that the quantity of information can considerably impact the performance of the models. Furthermore, this highlights the increased potential of phase and intensity joint exploitation, particularly in complex contexts, suggesting that its impact could be even more significant in pathological white blood cell classification cases.

4.2 Synthetic bi-modal image generation using GAN

The previous classification study demonstrated performance improvement by combining intensity and phase data in neural networks to classify white blood

cells. Although the model performs well in classifying the most represented classes, such as neutrophils and lymphocytes, its accuracy still needs to be improved for less represented classes. Notably, the performance gap can be mainly observed in the basophil class. To improve this class's performance, we will explore the effect of increasing the number of samples. To this end, we will explore a bimodal GAN. Given the difficulties in implementing and training classical GANs, such as mode collapse and instability, an innovative approach is explored: a physics-informed GAN, in which the direct image formation equations are inserted in the neural formalism. Combined with fine-tuning, this model can synthesize additional examples that enrich the underrepresented classes.

4.2.1 Generative Adversarial models principle

Generative modeling [36] is a field of machine learning that aims to generate new examples that closely resemble the original dataset by identifying patterns in input data. One promising approach in this field is Generative Adversarial Networks (GANs)[38], which use neural network architectures within deep learning. GANs involve a new framework for estimating generative models via an adversarial process, where two models, the generator (G) and the discriminator (D), are trained together.
The generator takes random noise vector $\vec{z}$ as input and generates synthetic data samples that mimic the data distribution $\vec{x}_{\text{fake}}$, while the discriminator uses fake and real data samples $\vec{x}_{\text{real}}$ estimates the probability that a sample came from the real data rather than G by performing binary classification, Figure 4.5.

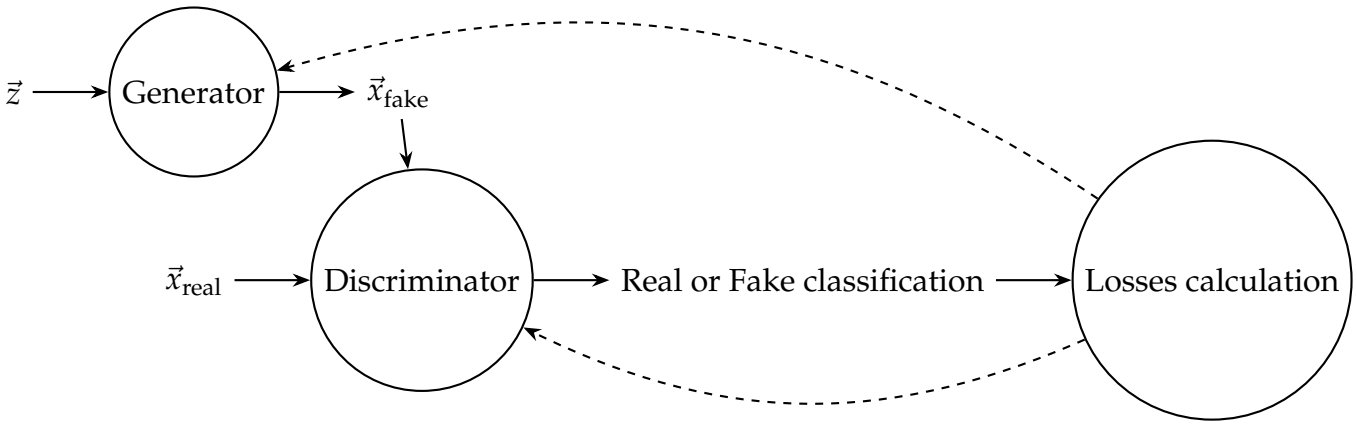


Figure 4.5: Illustration of a Generative Adversarial Network (GAN) architecture. The figure shows the components of a GAN, including the generator (G), discriminator (D), and loss calculation. Dashed arrows show backpropagation, where gradients are propagated backward through both networks during training.

The core of our approach lies in the adversarial training of the generator and discriminator performed using losses based on the classification task and propagated through the discriminator and the generator. This is achieved through the use of a minimax objective function, which we define as:

$$Loss(D,G) = \min_G \max_D \left[\mathbb{E}_{\vec{x_{real}} \sim p_{\text{data}}(\vec{x_{real}})} [\log D(\vec{x_{real}})] + \mathbb{E}_{\vec{z} \sim p_{\text{noise}}(\vec{z})} [\log(1 - D(G(\vec{z})))] \right] \tag{4.1}$$

Where $p_{\text{data}}(\vec{x_{real}})$ denotes the distribution of real data, and $p_{\text{noise}}(\vec{z})$ represents the distribution of the noise vector $\vec{z}$. ~ indicate the belonging fact. Here, D() and G() denote the discriminator's and the generator's output, respectively. $\mathbb{E}$ is the expected value of a function over a probability distribution (Expectation). The discriminator aims to maximize the objective function using gradient ascent, whereas the generator aims to minimize it using gradient descent. The generator faces convergence and learning issues due to the discriminator's gradient ascent. This happens because when the generator creates a sample, it is often classified as fake by the discriminator. The generator wants to learn from the gradients to fool the discriminator, but they tend to be relatively flat, mak-

ing it harder for the model to learn. The generator's objective function is for this modified as follows:

$$\mathcal{L}_G = \max_G \mathbb{E}_{\vec{z} \sim p_{\text{noise}}(\vec{z})} [\log D(G(\vec{z}))] \tag{4.2}$$

with this formulation, the generator aims to maximize the likelihood of the discriminator being wrong rather than minimizing the likelihood of the discriminator being correct. The interplay between the generator and discriminator drives the iterative training process of GANs, continuously improving the generator's ability to produce increasingly realistic outputs. As the competition between the generator and discriminator intensifies, the GAN converges towards an equilibrium, where the generator produces synthetic samples indistinguishable from real data.
It is important to emphasize that for this interplay to work optimally, the generator and the discriminator must have equivalent complexity. Indeed, if the generator is simple compared to the discriminator, it risks easily fooling it. Conversely, if the discriminator is very efficient, the generator will need help to improve and will produce images of limited quality. Adjusting the balance between the two components constitutes a major challenge in designing effective GAN models.

The loss formulation presented in equations 4.1 and 4.2 represents the classical approach commonly used in GAN training. Other cost functions are also used to improve the stability of GAN models, facilitate convergence, and increase diversity. The most widely used cost function is Wasserstein's loss [9]; the primary distinction of this loss lies in the fact that the discriminator no longer performs classification but instead produces a single value using an embedding in latent space to describe the images.
Unlike traditional GANs, where the discriminator outputs a probability indicating whether the input is real or fake, the Wasserstein GAN (WGAN) framework uses the discriminator to approximate the Wasserstein distance between the data distribution and the model distribution. This distance measures how far two probability distributions are from each other. The Weierstrass loss can be formulated as follows for the generator:

$$\mathcal{L}_G = -\mathbb{E}_{z \sim p_{\text{noise}}} [D(G(z))] \tag{4.3}$$

And for the discriminator:

$$\mathcal{L}_D = -\mathbb{E}_{x\sim p_{\text{data}}}[D(x)] + \mathbb{E}_{z\sim p_{\text{noise}}}[D(G(z))] \tag{4.4}$$

Other regularization techniques are also commonly used to stabilize the training of GANs and improve their convergence. Among them, gradient regularization and weight normalization are particularly widespread [72]. The Weierstrass loss and the regularization technique will be employed in some of the implemented architecture to promote diversity and stability.

Once the GAN training is done, the trained GAN models are evaluated according to a qualitative criterion through visual observation of generated images and the FID score metric that provides the quantitative evaluation. The Frechet Inception Distance (FID) is defined by using the Inception V3 model pre-trained on the ImageNet dataset to produce an embedding for each image, then calculated as :

$$\text{FID} = \|\mu_r - \mu_g\|_2^2 + \text{Tr}(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{1/2})$$

g and r are the real and synthetic embeddings assumed to be two multivariate normal distributions, μ_g and μ_r are the magnitudes of the vector g and r. Tr is the trace of the matrix, Σ_g and Σ_r are the covariance matrix of the vectors. The lowest the FID scores, the higher the similarity between the two groups of images.

4.2.2 Classical GAN approach

FPM dataset has the particularity of containing bimodal intensity and phase images. The GAN model must respect this property by generating an intensity and a phase image simultaneously. To this perspective, we propose a first classic approach. it architecture is illustrated in Figure 4.6.

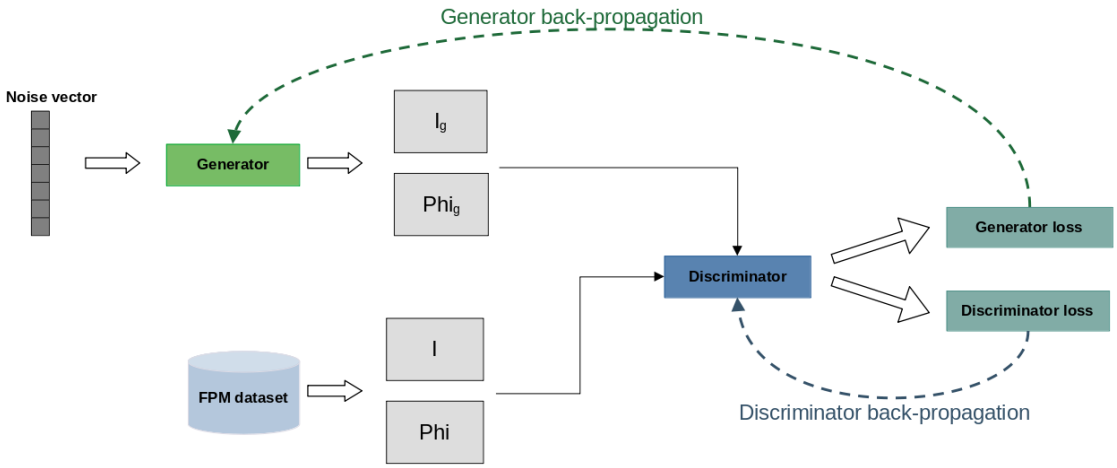


Figure 4.6: Classical GAN approaches applied to intensity and phase bimodal data.

The generator takes random noise as input and produces two channels simultaneously representing intensity and phase. The first channel is associated with the object's intensity, while the second corresponds to the phase image. The discriminator takes two channels representing intensity and phase generated by the generator or extracted from the FPM dataset as input. Then, using neural architecture, the discriminator extracts relevant features from the input images, which are then used to classify the examples as real or synthetic. The classification task is then used to compute losses and to perform backpropagation for both the generator and discriminator.
In this context, the main challenge is to design a discriminator that adequately interprets image intensity and phase information of images. In our architecture, the discriminator is designed to take an image represented by two channels as input: one channel for the intensity and another for the phase. Using a convolutional neural network (CNN) architecture, the model must learn to extract relevant features from these two channels to adequately encode essential features in both intensity and phase and carry out a classification that considers both. It is important to note that intensity and phase are two types of information coming from different representation domains (absorption and optical path). Despite their correlation, they have different value ranges and distributions, even after normalization. This divergence in representation spaces can make the task of generating bimodal images (intensity and phase) particularly complex.

Let us recall that our goal is to generate intensity and phase images for the basophil class. Due to the limited number of samples available for this class, we will initiate the GAN training process using a similar class: the neutrophils. Subsequently, we will refine the model using a fine-tuning approach targeting the basophil class. Note that the proposed approaches only focus on a single wavelength (green channel).

Neutrophil GAN

A collection of 8520 bimodal images (comprising both intensity and phase) of 256×256 pixels, with neutrophil blood cells positioned at the center used to train the model.
The choice of generator and discriminator architectures was made empirically. This empirical research also focused on some hyper-parameters, such as batch-size and learning rates. The objective was to optimize the stability of the model by guaranteeing appropriate complexity for the optimization of the adversarial game and to optimize the model's performance. The generator architecture was chosen inspired by the architecture of StyleGAN2 [57]. It first introduces a mapping network, detailed in Table 4.6, that aims to transform the initial random noise vector by projecting it into a specific feature space, allowing more meaningful control over the generated images, which facilitates the manipulation of particular attributes of the generated images, leading to more diverse and realistic results. The generator uses the resulting vector in a series of generator blocks to generate the final image incrementally. Each generator block uses a combination of convolution layers, batch normalization, and leaky ReLUs to generate increasingly detailed features. Finally, a convolution layer with tanh activation produces the final image with Table 4.7.
The architecture of the generator is complex. This choice was made during the architecture research phase, where smaller generators (fewer layers and fewer filters, no mapping network...) were used, but the generated images needed more resolution and details, and only one of the channels was optimized (either the intensity or the phase).

Layer (Type)	**Output Shape**	**Parameters number**
Input (InputLayer)	(None, 128)	0
Dense (Dense)	(None, 128)	16512
LeakyReLU (LeakyReLU)	(None, 128)	0
Dense (Dense)	(None, 128)	16512
LeakyReLU (LeakyReLU)	(None, 128)	0
Dense (Dense)	(None, 128)	16512
LeakyReLU (LeakyReLU)	(None, 128)	0
Dense (Dense)	(None, 128)	16512
LeakyReLU (LeakyReLU)	(None, 128)	0

Table 4.6: Mapping network architecture.

Layer (Type)	Output Shape	Parameters number
Input (InputLayer)	(None, 128)	0
Dense (Dense)	(None, 131072)	16777216
BatchNormalization (BatchNormalization)	(None, 131072)	524288
LeakyReLU (LeakyReLU)	(None, 131072)	0
Reshape (Reshape)	(None, 16, 16, 512)	0
Conv2D (Conv2D)	(None, 16, 16, 256)	3276800
BatchNormalization (BatchNormalization)	(None, 16, 16, 256)	1024
LeakyReLU (LeakyReLU)	(None, 16, 16, 256)	0
UpSampling2D (UpSampling2D)	(None, 32, 32, 256)	0
Conv2D (Conv2D)	(None, 32, 32, 128)	819200
BatchNormalization (BatchNormalization)	(None, 32, 32, 128)	512
LeakyReLU (LeakyReLU)	(None, 32, 32, 128)	0
UpSampling2D (UpSampling2D)	(None, 64, 64, 128)	0
Conv2D (Conv2D)	(None, 64, 64, 64)	204800
BatchNormalization (BatchNormalization)	(None, 64, 64, 64)	256
LeakyReLU (LeakyReLU)	(None, 64, 64, 64)	0
UpSampling2D (UpSampling2D)	(None, 128, 128, 64)	0
Conv2D (Conv2D)	(None, 128, 128, 64)	102400
BatchNormalization (BatchNormalization)	(None, 128, 128, 64)	256
LeakyReLU (LeakyReLU)	(None, 128, 128, 64)	0
UpSampling2D (UpSampling2D)	(None, 256, 256, 64)	0
Conv2D (Conv2D)	(None, 256, 256, 2)	3200

Table 4.7: Generator detailed architecture for neutrophils generation using classique GAN approach.

This generator was trained along with the discriminator described in Table 4.8 using batch-size fixed to 4, learning rate fixed to 2×10^{-4} and 4×10^{-4} for the discriminator and the generator respectively. The complete model was trained for 1000 iterations, each generating 10000 images. The best model selection was based on the FID score calculated separately for intensity and phase. The strat-

egy adopted was to choose the model offering the best balance between the minimization of the FID for the intensity and that of the FID for the phase. In practice, it turned out that the model that minimizes the FID for intensity also minimizes the FID for phase. The model was trained using a Weierstrass loss, as the initial version was observed to produce no variability.

Layer (Type)	**Output Shape**	**Parameters number**
InputLayer (InputLayer)	(None, 128, 128, 2)	0
Conv2D (Conv2D)	(None, 64, 64, 128)	2432
LeakyReLU (LeakyReLU)	(None, 64, 64, 128)	0
Conv2D (Conv2D)	(None, 32, 32, 128)	147584
BatchNormalization (BatchNorm)	(None, 32, 32, 128)	512
LeakyReLU (LeakyReLU)	(None, 32, 32, 128)	0
Conv2D (Conv2D)	(None, 16, 16, 256)	295168
BatchNormalization (BatchNorm)	(None, 16, 16, 256)	1024
LeakyReLU (LeakyReLU)	(None, 16, 16, 256)	0
Conv2D (Conv2D)	(None, 8, 8, 256)	590080
BatchNormalization (BatchNorm)	(None, 8, 8, 256)	1024
LeakyReLU (LeakyReLU)	(None, 8, 8, 256)	0
GlobalAveragePooling2D	(None, 256)	0
Dense (Dense)	(None, 1)	257

Table 4.8: Discriminator detailed architecture for neutrophils generation using classique GAN approach.

Examples of images produced by this model in intensity and phase were compared to real images to assess the generation quality in Figure 4.7.

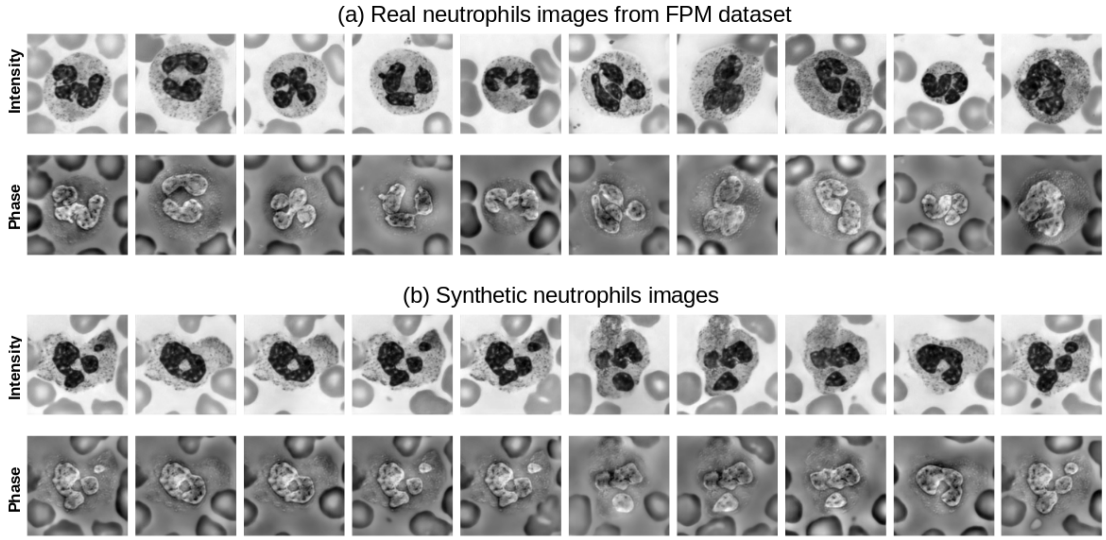


Figure 4.7: Illustration of different neutrophils. (a) shows images of neutrophils that were randomly extracted from the FPM database in intensity and phase; (b) shows images of neutrophils that were randomly extracted from a set of 10,000 images generated by the GAN model.

In Figure 4.7, we observe a random sample of neutrophils exhibiting high variability of multilobed nuclei, coated in cytoplasm containing fine granules, shown in the first two rows in intensity and phase. The second row represents neutrophil images generated by the randomly extracted GAN model. These generated images have several positive aspects. First, the resolution of the images generated is quite satisfactory, allowing details such as granules to be observed. Furthermore, the overall structure of the images shows a multilobed organization with fine granulations consistent with the expected data characteristics. However, despite these positive aspects, certain weaknesses persist. We indeed observe a certain repetition in the generated images, testifying to a lack of variability in the results. Indeed, the images of neutrophils presented in columns 2 and 3 are identical, as are those in columns 4 and 5, which are very similar. The same goes for the three neutrophils in columns 6, 7 and 8. Of the ten examples shown here, only six images are unique. Even after adding the Weierstrass loss in addition to the mapping network, the problem persists. Although this loss helps introduce some diversity into the generated samples, it fails to resolve it

completely.

This model faces a challenge called mode collapse, in which the generator tends to produce a limited set of outputs, often repeating the same patterns or features across different samples. Mode collapse occurs when the generator discovers a strategy to effectively fool the discriminator by producing a set of very realistic but little varied samples. In such situations, the discriminator may need to be more confident in distinguishing real samples from fakes, leading to generator stagnation. The latter then finds itself unable to explore the complete set of possible data modes, limiting itself to a restricted subset. More formally, the generator fails to capture the diversity inherent in the data distribution, harming the generated samples' quality and variety.
This problem is well known in the literature of GAN models, for which several studies have proposed solutions such as the Wasserstein loss or even approaches based on regularization[119, 13, 100, 72]. Despite these advances, collapse mode can still occur under certain conditions. In the case of the studied bimodal GAN model, the fact that it generates two distinct outputs (intensity and phase) can aggravate the problem of mode collapse due to the complexity of the task. Generating two distinct outputs simultaneously can place additional constraints on the generator network, making it more challenging to explore the full diversity of the data distribution. This problem has also been observed in other applications seeking to produce multiple outputs, such as magnetic resonance imaging (MRI) [68], where complex tasks can make limiting collapse mode particularly difficult [120, 121, 62, 20].

Basophils GAN

Despite the persistent collapse mode when training GANs on the neutrophil class, an attempt was made to synthesize basophil images to improve the basophil class. Unlike the neutrophil class, which had a larger dataset, the basophil dataset only included 54 images for training. Our attempts to generate basophils from scratch completely failed. The similarity between basophils and neutrophils was explored through a fine-tuning process to address this challenge. These two types of white blood cells are characterized by cytoplasmic granules visible in their cytoplasms. Additionally, basophils and neutrophils have a lobed nucleus, although that of neutrophils. We trained the neutrophil GAN and used fine-tuning approaches to adjust the parameters of the first

model so that we could generate basophils using the same architecture. Compared to real images, the images produced are represented in Figure 4.8.

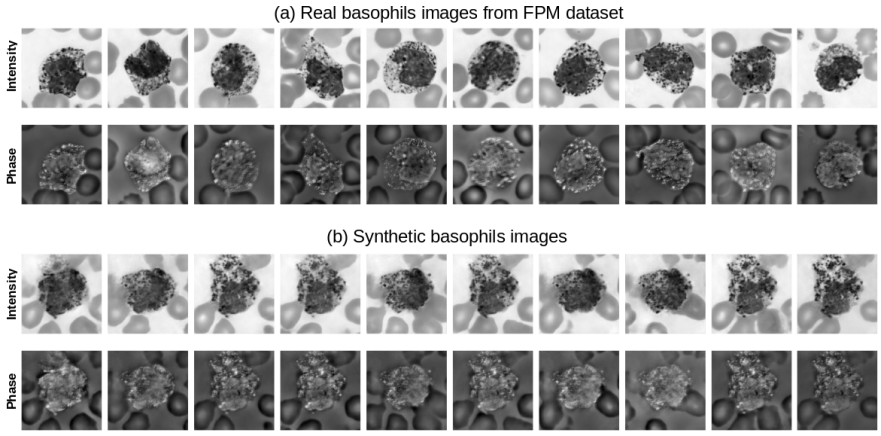


Figure 4.8: Illustration of different basophils. (a) shows images of basophils that were randomly extracted from the FPM database in intensity and phase; (b) shows images of basophils that were randomly extracted from a set of 1000 images generated by the GAN model.

We see that the images remain generally faithful to the representation of real images of basophils. However, we also see that the outlines of the nuclei are not as well defined as in real images of basophils. However, it is also evident that the model is more affected by a severe collapse mode. The images represented have minimal variability, and only two main forms are produced. It can be observed that the basophils generated in columns 3, 4, 6, 9 and 10 are the same, as are the images in columns 2, 5, 7 and 8. Only 3 different patterns are observed over the 10 representations.

A classification-based approach is employed to assess the generative model of basophils and the quality of the generated images beyond a visual inspection. The performance of a model trained exclusively on a dataset containing 54 real basophils is compared to that of a model trained on a mixture of these same real basophils and 1000 basophils generated by a GAN. For each model, 75% of the data was used for training and 25% for validation. The test dataset contains 93 images of real basophils in both cases. The results are summarized in Table 4.9. The classification results show no improvement despite the augmentation protocol. The models with data augmentation even get the worst performances on

Classes	Model without augmentation		Model with augmentation	
	Precision	**Recall**	**Precision**	**Recall**
Basophils	94.00%	99.66%	92.59%	99.52%
Eosinophils	98.31%	96.70%	97.87%	95.50%
Lymphocytes	99.07%	98.38%	98.70%	97.35%
Monocytes	96.73%	97.24%	92.25%	96.47 %
Neutrophils	99.69%	99.00%	99.70%	99.78%

Table 4.9: Performance metrics for each class.

the other classes. This may be caused by the augmentation that added many elements that were too similar, leading the model to overfit these specific training elements. As a result, the model lost its ability to generalize effectively over the test base. In agreement with the state of the art, the complexity of our problem, mainly due to the bimodal images, does not allow us to produce effective GAN models. This type of approach does not provide any performance improvement for under-represented classes.

4.2.3 Physics-informed GAN approach

We performed multiple experiments to adjust and optimize the architecture and hyper-parameters with the aim of escaping the mode collapse phenomenon. This this task was a significant challenge, given the number of possible combinations of the architectures and proposals for different possible cost functions. The results obtained with conventional GAN for producing joint intensity-phase bimodal images remain unsatisfactory, which is in line with the known challenges associated with multimodal GANs, as reported in the literature.
In our case, bimodal images are obtained through a reconstruction process from a physical model. Given this particularity, we propose an innovative GAN approach.
This new approach aims to maintain a consistent link between intensity and phase images throughout the generation process. To do this, we rely on physics-inspired neural networks (PINNs for Physics-Inspired Neural Networks). The idea is to develop a GAN architecture that integrates these physical principles

to better capture the relationship between the two image modalities. This approach is detailed below.

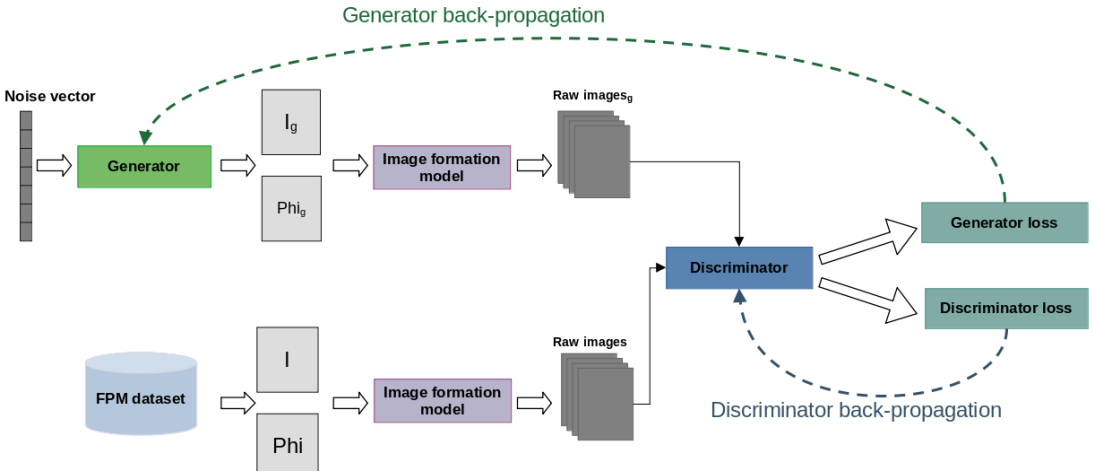


Figure 4.9: Physics-informed GAN approaches applied to intensity and phase bimodal data.

The main concept is shown in Figure 4.9. The process begins with the generator, which takes random noise as input and produces two distinct channels simultaneously representing the intensity and phase of the sample. The first channel encodes the intensity of the simulated object, while the second channel corresponds to the spatial distribution of the phase, as in classical GAN. This generated intensity and phase data is then used to feed an imaging model, which uses the equations detailed in Chapter 3 to produce a stack of raw images that simulate what the FPM microscope would produce as raw images for a specific object under a defined configuration.
The discriminator then comes into play, taking as input multi-channel data representing the raw images produced by the imaging models. These images are constructed either from intensity and phase data produced by the generator or from intensities and phases extracted from the FPM dataset. Using its neural architecture, the discriminator extracts relevant features from input images and then uses them to classify examples as real (from the real dataset) or synthetic (generated by the generator). The classification task is then used to calculate losses and perform backpropagation, thus making it possible to adjust the generator and discriminator parameters to improve the quality of the synthetic images produced. This approach shifts the learning paradigm of the GAN.

Instead of focusing on recognizing the real intensity and phase of the synthetics, this model focuses on optimization from the stack of raw images, establishing a link between the intensity and the phase generated, which could be assimilated into a regularization. Indeed, the images produced by this physical model are only intensity images and are represented in the same space with the same range of values, which simplifies feature extraction.
We first trained a neutrophil GAN with this model, which we will use as an initialization for basophil GAN later in this chapter.

Neutrophil GAN

The model was optimized using the same architecture's search method as in the previous approach with the classical GAN on neutrophils. The generator architecture also starts with a mapping network followed by a CNN architecture as detailed in Table 4.10.
This differs from the classic GAN approach; the discriminator does not process directly the generator output. In our case, it is preceded by an imaging model that generates a stack of 13 LEDs containing 128×128 images corresponding to what the FPM microscope would capture of the observed object. The discriminator takes this stack of images as input, as input and classifies them as real or synthetic using CNN architecture. The precise details of the discriminator architecture are provided in Table 4.11.

Layer (Type)	Output size	Parameters number
Input (Input)	(latent_dim,)	-
Dense (Dense)	128	16512
LeakyReLU (Activation)	128	0
Dense (Dense)	128	16512
LeakyReLU (Activation)	128	0
Dense (Dense)	128	16512
LeakyReLU (Activation)	128	0
Dense (Dense)	128	16512
LeakyReLU (Activation)	128	0
Reshape (Reshape)	(16, 16, 512)	0
Conv2D (Convolution2D)	(16, 16, 256)	1179648
BatchNormalization (BatchNorm)	(16, 16, 256)	1024
LeakyReLU (Activation)	(16, 16, 256)	0
UpSampling2D (UpSampling2D)	(32, 32, 256)	0
Conv2D (Convolution2D)	(32, 32, 128)	294912
BatchNormalization (BatchNorm)	(32, 32, 128)	512
LeakyReLU (Activation)	(32, 32, 128)	0
UpSampling2D (UpSampling2D)	(64, 64, 128)	0
Conv2D (Convolution2D)	(64, 64, 64)	73728
BatchNormalization (BatchNorm)	(64, 64, 64)	256
LeakyReLU (Activation)	(64, 64, 64)	0
UpSampling2D (UpSampling2D)	(128, 128, 64)	0
Conv2D (Convolution2D)	(128, 128, 32)	18432
BatchNormalization(BatchNorm)	(128, 128, 32)	128
LeakyReLU (Activation)	(128, 128, 32)	0
UpSampling2D (UpSampling2D)	(256, 256, 32)	0
Conv2D (Convolution2D)	(256, 256, 2)	576
Activation (Activation)	(256, 256, 2)	0

Table 4.10: Physics-inspired generator starting with a mapping network followed by a CNN architecture.

Layer (Type)	Output	Output size
InputLayer (InputLayer)	-	(128, 128, 13)
Conv2D (Conv2D)	(None, 64, 64, 32)	832
LeakyReLU (LeakyReLU)	(None, 64, 64, 32)	0
Conv2D (Conv2D)	(None, 32, 32, 64)	51264
BatchNormalization(BatchNorm)	(None, 32, 32, 64)	256
LeakyReLU (LeakyReLU)	(None, 32, 32, 64)	0
Conv2D (Conv2D)	(None, 16, 16, 128)	204928
BatchNormalization(BatchNorm)	(None, 16, 16, 128)	512
LeakyReLU (LeakyReLU)	(None, 16, 16, 128)	0
Dense (Dense)	(None, 1)	8193

Table 4.11: Detailed architecture of discriminator of the physics-inspired GAN.

The physics informed GAN was trained using batch-size fixed to 2, learning rate fixed to 4×10^{-4} for both the discriminator and the generator. The complete model was trained for 1000 iterations, and each iteration generated 10000 images. The best model was chosen according to the minimum FID score on phase and intensity. This model generated the images shown in Figure 4.10.

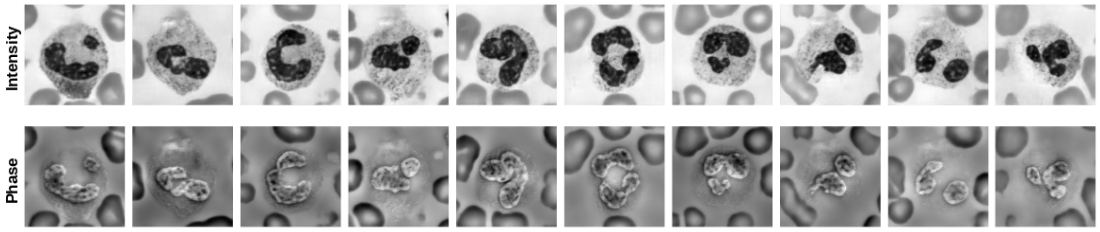


Figure 4.10: Generated neutrophils random sample from physique-informed GAN.

The images produced by the physics-informed model exhibit an overall multi-lobed structure with fine granulations, thus approximating the morphological characteristics observed in images of real neutrophils (see Figure 4.7). Furthermore, these images exhibit remarkable diversity and a striking resemblance to

authentic images, marking a marked improvement over previous results obtained with conventional GAN.

For a quantitative evaluation of the performance of classical GAN and the physics-informed GAN, the FID score was calculated over 8520 samples for real and generated images using both models. The results obtained are summarized in Table 4.12

Model	**Intensity (FID)**	**Phase (FID)**
Classical model	87.84	63.43
Physics informed model	84.90	45.69

Table 4.12: Quantitative comparison between the classical GAN and physic-informed GAN on neutrophils.

The physics-informed model displays lower FID scores for both intensity and phase, thus reflecting a substantial improvement in image generation quality compared to the classical model. Furthermore, this significant improvement in fidelity in phase reproduction could be attributed to better consistency between the generated intensity and phase. This could be justified by directly incorporating images combining intensity and phase into the discrimination process rather than extracting features separately, as is classical in classical approaches. The improvements observed with the physics-informed model can be attributed to several features inherent to its architecture. First, by integrating physical knowledge about the relationship between image intensity and phase, the model is better equipped to regularize image generation by ensuring consistency between these two aspects. Additionally, the physics-informed model's discriminator is designed to consider multiple representations of the images, allowing a more accurate assessment of their realism and consistency with real images. This approach facilitates learning more subtle and complex data features, thereby improving differentiation between real and synthetic images.

Basophil GAN

A GAN model for basophil augmentation was trained to evaluate the model's effectiveness in increasing minority classes. This model is trained using the physics-inspired architecture used for neutrophils. This architecture exploits

the structural similarity between basophils and neutrophils, similar to the approach adopted in the case of classic GAN; the neutrophil GAN model was used as an initialization for basophils GAN, and this model was then fine-tuned using the 54 available images. The best model was chosen according to the FID score. Illustrations of randomly selected images generated by this model are shown in Figure 4.11.

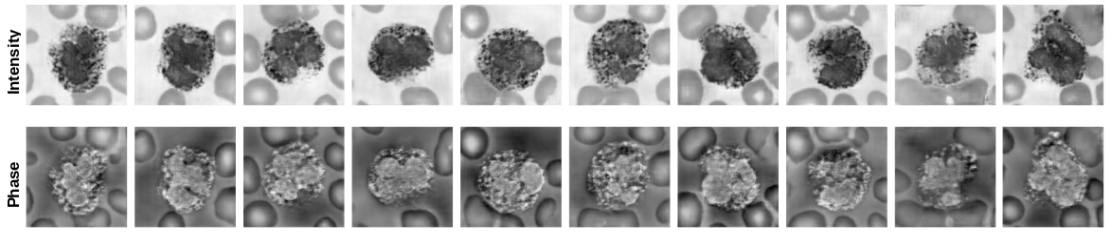


Figure 4.11: Generated basophils random samples from physics-informed GAN.

Physics-informed GAN demonstrated promising performance in basophil image generation. The generated images present appreciable diversity and resolution, which allows a faithful representation of the morphological characteristics of basophils. More precisely, we observe that the nuclei contours are more defined and the granules are clearly defined. Furthermore, this GAN effectively reproduces the original distribution of basophils, strengthening its ability to generate synthetic data consistent with real samples. The fine-tuning strategy allowed us to learn from a sample as small as 54 representatives to generate basophils.

Subsequently, a second evaluation of the basophil model was performed using a classification task. For each model, 75% of the data was used for training and 25% for validation. The tests were conducted only on real data, with an identical test set for both models. We performed 50 initializations for the model without GAN augmentation and 50 initializations for the model with GAN augmentation, where in each iteration, 100 images were randomly selected from the GAN model. The exploratory architecture adopted was that of a MobileNet, chosen for its superior performance in five-class classification.
The results show that augmenting basophil data using a GAN significantly impacts classification accuracy. The model with augmentation shows signifi-

Classes	Model without augmentation		Model with augmentation	
	Precision	**Recall**	**Precision**	**Recall**
Basophils	94.00% $\pm$ 0.024%	99.66% $\pm$ 0.009%	99.37% $\pm$ 0.002%	99.78% $\pm$ 0.002%
Eosinophils	98.31% $\pm$ 0.013%	96.70% $\pm$ 0.013%	96.62% $\pm$ 0.026%	97.73% $\pm$ 0.01%
Lymphocytes	99.07% $\pm$ 0.003%	98.38% $\pm$ 0.002%	99.24% $\pm$ 0.005%	98.98% $\pm$ 0.005%
Monocytes	96.73% $\pm$ 0.013%	97.24% $\pm$ 0.016%	97.52% $\pm$ 0.010%	97.96% $\pm$ 0.015%
Neutrophils	99.69% $\pm$ 0.001%	99.00% $\pm$ 0.001%	99.86% $\pm$ 0.0007%	99.85% $\pm$ 0.001%

Table 4.13: Performance metrics for each class.

cantly higher accuracy for basophils compared to the model without augmentation. Specifically, the model without augmentation achieves an accuracy of 94.00% with a standard deviation of 0.024, while the model with augmentation achieves a remarkably improved accuracy of 99.37% with a standard deviation of 0.002. This substantial increase in accuracy, of approximately 5%, indicates that adding augmented data has significantly enhanced the model's ability to classify basophils correctly.
The physics-informed GAN approach stands out for its ability to improve the variability of the GAN while limiting the collapse mode phenomenon. This innovative methodology generates images that complement underrepresented classes, such as basophils, in this study, significantly improving classification performance. This combination highlights the importance of physics-informed GANs combined with the fine-tuning approach to limit the challenges of blood cell classification.

4.3 Conclusion

The integration of FPM phase and intensity information demonstrated remarkable potential in classifying white blood cells from thin-stained blood smears, which was the proof of concept presented Chapter 3. Specifically, by exploiting joint intensity phase information, the models demonstrated superior performance in accurately distinguishing distinct leukocyte subtypes. This improvement is particularly notable for underrepresented classes.
Comparative analysis with conventional imaging methods highlights the advantages of FPM, Indeed, the integration of phase imaging alongside intensity

imaging has proven to be essential, improving diagnostic capabilities and outperforming traditional methods like the Di-60-CNN model.
WBC classification faces the question of class imbalance and a limited amount of data. Generative Adversarial Networks (GANs) are generally used to address these issues. Generating FPM images requires synthesizing bimodal images. Our trials to use classical GAN to this end came up against the problem of mode collapse, limiting the variability and usefulness of the generated samples. We, therefore, introduced an innovative approach, a Physics-Informed GAN, to synthesize new FPM bimodal images, improving the diversity of generated images and limiting the impact of the mode collapse problem encountered with classical GANs. This corresponds to an essential advance in the field thanks to an efficient regularization obtained by introducing the image formation model in the loss functions. This model is associated with an innovative strategy for fine-tuning GANs. This combination demonstrated its ability to complete minority classes even when only a few samples were available for learning the GAN, as in the basophil class.
This approach constitutes a first step that must be expanded to encompass the generation of three wavelengths at once. Addressing challenges such as maintaining balance and ensuring coherence among the different wavelengths still needs to be discussed.
An article is currently being prepared for the publication of this work.

Chapter 5

System optimization

Contents

In previous chapters, the value of FPM microscopy for blood cell classification has been explored, demonstrating its potential for two distinct applications. By exploiting bimodality, FPM significantly improved model performance, highlighting a synergy between intensity and phase. However, to move to an operational system, an overall optimization of the workflow is necessary.
This chapter, therefore, proposes some studies for optimizing the system along two main axes. First, the focus is put on optimizing the bimodel information exploitation in DNN. Although bimodal information revealed a significant gain in performance, how to fully exploit its potential in image processing models remains. Second, optimizing the microscope's optical configuration, particularly the lens choice, is discussed. The objective is to find the best compromise between the size of the field of view, the resolution and the classification performance. This approach combines experimental data and an optical image formation model to predict the observation of samples and studies the model's performance on these observations.
This approach to optimizing the system by modeling from experimental data is illustrated through a specific application: the detection of malaria. However, the principles and methods presented apply to any microscopy application requiring improved field of view and detection performance.

5.1 Exploitation of the bi-modal FPM information through complex-valued neural networks

FPM provides access to bi-modal data by reconstructing both intensity and phase information.
The most classic approach in neural networks considers these two images as independent information channels. Indeed, they represent two distinct physical parameters resulting from the light-matter interaction. Typically, one image encodes the absorption parameters (intensity) while the other represents the optical path (phase). In this context, it is common to apply classical real-valued convolutions to independently extract the relevant information from each image and then merge them across the different layers of the network.
However, the acquired bi-modal images can also be considered as the components of a complex-valued matrix called a transfer matrix. This matrix is closely linked to the complex electromagnetic field of the observed sample. With this in mind, using complex-valued convolutions becomes a fair processing option.

Inspired by complex filtering in physics, which has proven to reveal details (for example, in dark-field imaging) [144], the integration of complex-valued convolutions in neural networks makes it possible to simultaneously extract information from both modalities thanks to their complex representation. This approach could provide an advantage to the feature extraction stage and improve performance. Complex-valued convolution falls under the domain of complex-valued neural networks.

5.1.1 State of the art of complex-valued neural networks

Complex-valued neural networks (CVNNs) are a type of neural network architecture that processes data and parameters using complex numbers, representing amplitude and phase information. Unlike traditional neural networks that only handle real numbers, CVNNs leverage complex numbers to encode additional information about the data.
This section aims to provide a comprehensive overview of CVNNs found in the literature. It explores the theoretical foundations, practical applications, and limitations of CVNNs

Theoretical Foundations of Complex Neural Networks

The literature has extensively examined the theoretical foundations of CVNNs, covering both implementation and theoretical aspects. For instance, [48, 47] aimed to showcase the primary benefits of CVNNs, namely their capacity to control both phase rotation and amplitude. This feature enables a reduction in the degrees of freedom required for learning when compared to real-valued networks. As a result, CVNNs are particularly well-suited for applications that involve wave information processing.
Furthermore, [21] presents proof of the approximation theorem for complex-valued neural networks. Similarly, [135] investigates the blocks comprising real-valued neural networks and their counterparts in CVNNs. Mathematical and empirical studies have been conducted on back-propagation, cost function selection, and activation functions for this type of network in [65, 28, 16].

Applications of Complex Neural Networks

CVNNs have been implemented in various applications across diverse fields, showcasing their versatility and efficacy in addressing a wide range of complex

problems. We will highlight some of these applications and what they have demonstrated.

Regarding MRI data, [140] studied CVNNs for MRI fingerprinting, showing that these networks outperform real-valued models in classifying tissue parameters. Additionally, [149] presented a comprehensive complex-valued convolutional neural network for synthetic aperture radar (SAR) target classification. By avoiding the complex pooling operation and preventing overfitting, their model improved the average accuracy compared to equivalent real-valued models.

Moreover, [14] explored the effectiveness of CVNNs in classifying non-circular complex data, indicating that CVNNs had better accuracy and were more resistant to overfitting compared to real-valued models.

In the field of hyperspectral image processing, [4] proposed a complex neural network architecture to improve the spatial resolution of images. Based on frequency analysis, their approach overcame spectral distortions and improved the quality of hyperspectral images. For medical ultrasound imaging, [74] developed an innovative method using complex ultrasound signals with complex weights to improve image quality. Their fully complex neural network model demonstrated significant improvement in the resolution and contrast of ultrasound images.

In the context of radar target classification in interferometric synthetic aperture (InSAR) data, [127] developed a complex adaptive convolutional neural network capable of accurately distinguishing different terrain shapes. Their model demonstrated a remarkable ability to identify slopes, plains, and even geological features such as small volcanoes.

In these fields, where phase information is obtained directly from a measure, CVNNs exhibit smaller errors and better generalization abilities than their real-valued counterparts. However, CVNNs have yet to be applied to classify bimodal intensity-phase images produced by FPM microscopy, where phase results from computational reconstruction rather than a direct measure.

Moreover, while CVNNs have been primarily utilized in tasks such as regression and classification, they have yet to be employed in detection tasks. This highlights the potential for further exploration in leveraging CVNNs for such applications.

5.1.2 Complex-valued layers

In complex-valued convolutional neural networks, real-valued layers need to be substituted with complex-valued layers. This is done by representing complex data using its real and imaginary parts, which enables the application of this type of filtering using standard neural network implementations.
However, this filtering changes the network's operations, including activation functions and other aspects. The following section details the differences in implementing real versus complex convolution in CNN.

Convolution layers

One of the key operations in CNNs is convolution. Real-valued convolution is used to extract meaningful features from the input data while preserving the spatial relationships and patterns present in the data.

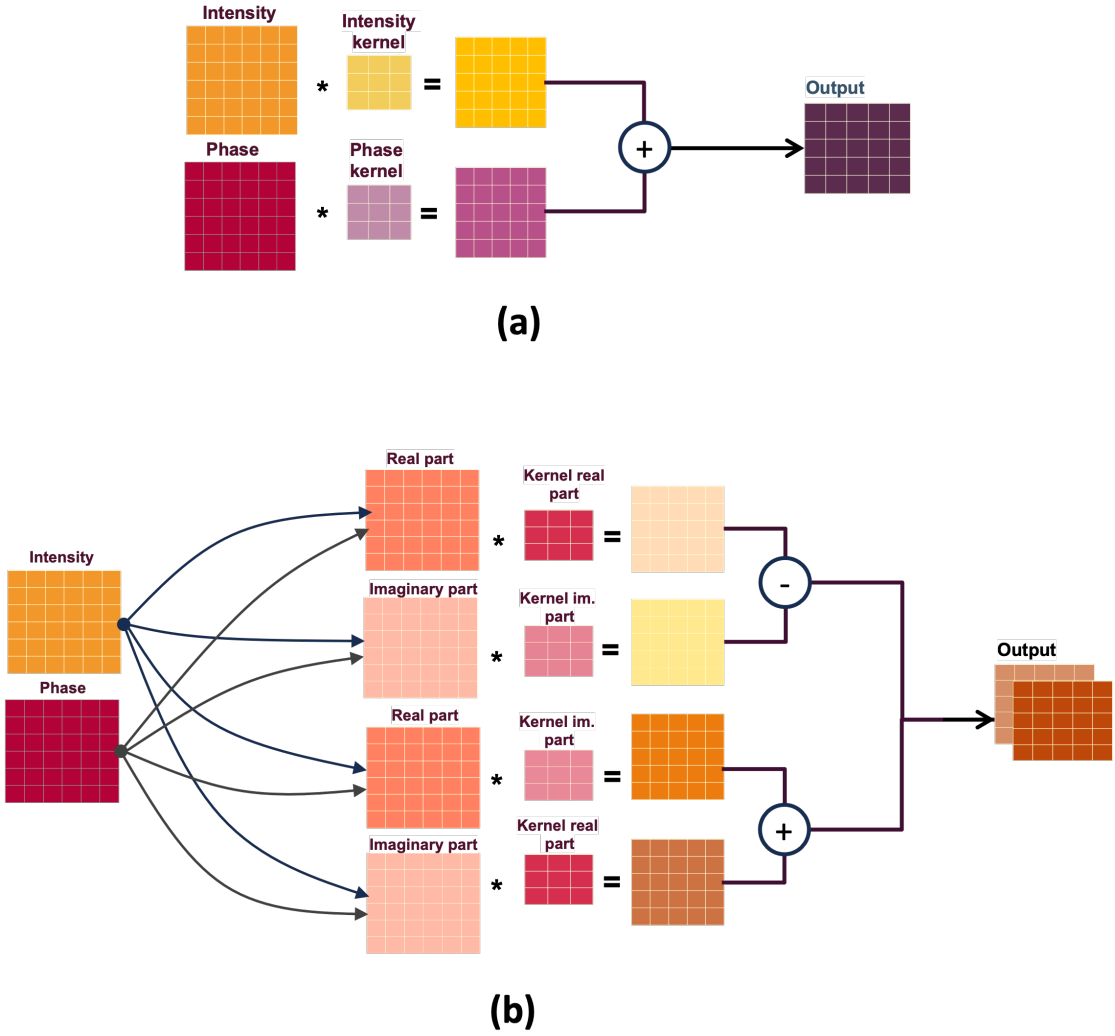


Figure 5.1: Schematic representation of the filtering operation in a real-valued implementation (**a**) versus a complex-valued one (**b**).

Figure 5.1 (a) illustrates the principle of real-valued convolution when applied to our case; each component of the input image, here intensity and phase, is processed independently by its corresponding kernels. The intensity component is convoluted with intensity kernels, and the phase component is convoluted with phase kernels. In this way, distinct characteristics are extracted for each channel individually.
However, after the individual convolutions, the resulting filtered images are mixed to form the final output, where the corresponding pixel values of the

different filtered images are added together to create a single output channel. In this manner, the neural network aggregates both information adaptively, depending on the task.
In other words, the two physical processes in the system are treated as separate channels during the convolution process, allowing specific features to be extracted independently from each channel. This independence ensures that the unique information from both processes is preserved and adaptively merged. As a result, the final output benefits from the complementary nature of the two independent physical processes.
While real-valued convolutions operate on real numbers, complex convolutions consider inputs as complex numbers. Complex convolutions use complex numbers as coefficients in the filters to jointly extract features from amplitude and phase.
Since standard deep-learning libraries can natively only handle real numbers, some adaptation in the implementation of the convolution block of the network is required, as depicted in Figure 5.1 (b). Complex numbers are represented as pairs of real numbers (real and imaginary parts), so the operation may be adapted to the existing implementations (Tensorflow, PyTorch) as described below. The idea is to introduce the complex domain operations in the formalism usually exploited for real-valued convolution. The changes will, therefore, concern the representation of the filter weights and the input by using a bi-modal (real, imaginary) representation and redefining the convolution operation.
More precisely, one can define the complex convolution using only real values while considering the particularities of the four arithmetic operations. The complex-valued input M is written as $M = (M_{\Re} + iM_{\Im})$. The complex convolution filter K is given by $K = (K_{\Re} + iK_{\Im})$. This is equivalent to defining two real-valued filters, where one represents the real part and the other the imaginary part of the complex filter (See Figure 5.1). The convolution of M by K is written:

$$\begin{aligned} K * M &= (M_{\Re} + iM_{\Im}) * (K_{\Re} + iK_{\Im}) \\ &= (M_{\Re} * K_{\Re} - M_{\Im} * K_{\Im}) + i(M_{\Re} * K_{\Im} + M_{\Im} * K_{\Re}) \end{aligned} \tag{5.1}$$

A tensorial representation can be given by:

$$\begin{bmatrix} \Re(M * K) \\ \Im(M * K) \end{bmatrix} = \begin{bmatrix} K_{\Re} & -K_{\Im} \\ K_{\Im} & K_{\Re} \end{bmatrix} \begin{bmatrix} M_{\Re} \\ M_{\Im} \end{bmatrix} \tag{5.2}$$

Activation fonction

Similarly, activation functions that introduce non-linearity must be adapted to complex-valued data. The usual functions are inappropriate due to their reliance on real number operations, specifically the maximum function (max). One such example is the rectified linear unit (ReLU). The max function cannot be straightforwardly extended to complex numbers. The other activation functions are also unsuitable for complex numbers; Sigmoid or Tanh have singularities for some complex numbers values $(i\pi(2N-1))$. As a result, alternative activation functions have been developed that are more suitable for complex-valued data.
Several choices have been made in the literature to design activation functions adapted to learning from complex data [28]. One of them, defined as the extended version of $ReLU$ for complex-valued data, called $\mathbf{C}ReLU$, has been proposed and tested in [135], which reports that $\mathbf{C}ReLU$ is giving the best performances in comparison to other functions like Sigmoid that seems unstable.
The complex activation $\mathbf{C}ReLU$ applies The complex activation $\mathbf{C}ReLU$ applies separate $ReLU$ on both the real and the imaginary part of a neuron, i.e.:

$$\mathbf{C}ReLU(z) = ReLU(R(z)) + iReLU(I(z)) \tag{5.3}$$

Where $R(.)$ and $I(.)$ are the real and the imaginary parts, respectively.
This activation function will be used in our complex-valued architecture. This function allows the cancellation separately (or jointly, when both real and imaginary parts are negative) of the real and imaginary parts when they are negative and preserves only the most relevant extracted information.

Pooling layers

Adapting the pooling operation to suit CVNNs represents another transformation that must be made. In CNNs, the pooling operation is crucial in retaining salient features by downsampling the feature maps. Typically, max pooling is the preferred approach in CNNs due to its ability to ensure numerical stability during training and facilitate a reduction in feature map dimensions. However, this operation poses challenges when applied to complex-valued data, as it is not inherently defined in this domain.
A plausible solution involves applying max pooling separately to the real and imaginary channels of the complex-valued data. While this approach maintains

numerical stability, it disregards the mathematical correlation between the real and imaginary components, potentially leading to suboptimal performance in feature extraction. Alternatively, an average pooling operation emerges as a viable solution for CVNNs, as it is well-defined within the complex domain, thereby preserving the integrity of complex-valued data and ensuring efficient feature extraction processes. Therefore, in our CVNN implementations, we opt for average pooling as a suitable alternative to max pooling, ensuring compatibility with complex-valued data while maintaining effective feature extraction capabilities.

5.1.3 Experimental Evaluation

Dataset

The dataset referenced in this chapter is the same as the one utilized in Chapter 3, Section 3.3.1. The dataset comprises thin blood smears obtained from nine patients and scanned through the FPM microscope, producing 2216 bi-modal intensity-phase images, each measuring 896×896 pixels. In total, the dataset contains 418389 healthy cells and 65140 infected cells.

Implementation Design

The study of the most adapted formalism for the joint exploration of intensity and phase images derived from FPM reconstruction involves an analysis utilizing both real and complex-valued representations within the framework of Faster R-CNN; this architecture has been described in Chapter 3, Section 3.3.2 for real-valued formalism. The complex-valued Faster-RCNN is obtained from the real-valued architecture by replacing the convolutions in the shared branch (Module (a)) with complex-valued convolutions, changing the activation function to CReLU and average pooling as described in Section 5.1.2. To ensure a fair comparison between the models, the number of filters for each convolution in the complex-valued version was fixed to half its number in the real-valued version. This was done because each filter in the complex-valued version includes real and imaginary representant parts. The objective remained the same. First, localize the red blood cells within these images, and second, classify them into distinct categories, namely parasites or healthy cells.
Two Faster R-CNN models were implemented for *Plasmodium Falciparum* parasite detection.

- A classical real-valued Faster-RCNN using intensity and phase ($\mathcal{I}/\phi$-RV);
- A complex-valued convolution Faster-RCNN using intensity and phase ($\mathcal{I}/\phi$-CV).

A hyper-parameter search was carried out for complex-valued architectures in the same way as for real-valued architecture described in Chapter 3. The anchor's size and ratios remain unchanged. We also used the same number of the total predictions produced by the Fast RCNN model (branch c). The maximum number of boxes to be predicted by the RPN was fixed at 1500 boxes.

Data pretreatment

The data must be normalized to ensure fast and efficient learning in the chosen architecture. Here, a normalization adapted to each representation was carried out.
In the case of the real-valued network, each channel has been normalized using a min-max normalization applied to each channel separately to ensure the information can be mixed and extracted consistently as described Chapter 3. In the case of complex networks, the min-max normalization is performed on the module of the data to guarantee the preservation of the relationship between the module and the phase and, thus, preserve its physical meaning.

5.1.4 Results and Discussion

The performance evaluation of the two models involves a comprehensive comparison based on two key performance indicators: the True Positive Rate (TPR) and the True Negative Rate (TNR). These indicators have been slightly adjusted to accommodate the use case. For a detailed explanation of these performance metrics, please refer to Section 3.3.3 of Chapter 3.
The parasite detecting results for the two implementations are provided in Tables 5.1 and 5.2.
The details of the output boxes for each model are shown in Table 5.1. A localization IoU threshold of 0.7 was used to construct the table. This means that a predicted bounding box is considered correctly positioned if it overlaps with the corresponding ground truth box by at least 70%. If the predicted label matches the ground truth label, the model considers the cell in this box well-classified. Otherwise, the cell is considered misclassified. The missed cells regroup the

	$\mathcal{I}/\phi$-RV	$\mathcal{I}/\phi$-CV
Well-classified infected	12,320	12,670
Misclassified infected	433	324
Well-classified healthy	82,778	82,965
Misclassified healthy	295	372
Missed infected	186	46
Missed healthy	601	169
Added infected	154	119
Added healthy	554	620

Table 5.1: Detailed results of Faster-RCNN predictions averaged to the nearest unit over the five folds.

boxes containing cells the model did not predict. Lastly, the model's added boxes are the boxes that have poor overlap with ground truth or boxes that do not correspond to actual objects.

When comparing the two models based on the performances given in Table 5.1, the $\mathcal{I}/\phi$-CV appears to be the most efficient for the task. This model achieves the highest number of accurately classified infected (350 more infected cells are detected) and healthy cells (187 more healthy cells are detected) while having the lowest count of misclassified infected cells. These results strongly suggest that incorporating the complex feature extractor improves the classification capabilities of the Fast-RCNN module (classification branch). Additionally, this model has the lowest count of missed cells in both healthy and infected categories, indicating a significant enhancement in the capabilities of the RPN module (detection branch). This suggests that this model has a greater capacity to find infected cells and distinguish them from healthy blood cells, which will allow it to better meet the high sensitivity requirements obligatory for the diagnosis of malaria.

From the performances presented in Table 5.2, it is evident that the *True Negative Rate* (*TNG*) scores indicate robust and accurate detection of healthy red blood cells for both models. Meanwhile, the *True Positive Rate* (*TPR*) are 95.22% for $\mathcal{I}/\phi$**-RV** and 97.15% $\boldsymbol{\mathcal{I}/\phi}$**-CV**. These findings are promising, especially in the context of *Plasmodium Falciparum* detection, where the primary objective is to achieve a high *TPR* for the early identification of infection cases, given the

	$\mathcal{I}/\phi$-RV	$\mathcal{I}/\phi$-CV
TNR	99.34 ± 0.24%	99.10 ± 0.07%
TPR	95.22 ± 0.73%	97.15 ± 0.30%

Table 5.2: Experiments evaluation metrics.

sparse presence of parasite cells among the total red blood cell population.
The comparison between the complex and real implementations for intensity-phase images underscores the relevance of the complex-valued formalism. Notably, a discernible improvement of approximately 2% in *TPR* with the complex formalism compared to the real one. Additionally, the narrower confidence intervals provided by the complex formalism ($\mathcal{I}/\phi$-CV versus $\mathcal{I}/\phi$-RV) further accentuate the superior generalization capabilities of complex-valued neural networks, as evidenced by the mean *TNG* and *TPR* values.
To statistically validate this performance disparity between the real and complex-valued models, we employed the McNemar test [109, 88]. The resulting p-values, lower than 0.001 for each fold, unequivocally reject the null hypothesis, indicating that the two models do not share the same error rate. This substantiates the superior performance of the model employing the complex formalism, with significantly different marginal proportions, thus affirming the validity of the performance outcomes.
Furthermore, the results reveal that even a partial implementation of the complex formalism, confined to the initial stages of the model to derive the feature map utilized in subsequent branches, yields performance enhancements.

Another interesting aspect of the complex CNN implementation is its capacity for faster convergence. This phenomenon is illustrated in Figure 5.2, where the total loss for each epoch of each model is shown. Notably, the real-valued model attains its lowest loss value (0.43) on the validation dataset after 94 epochs. In this contrast, the complex-valued model reaches its optimal validation point with a lower loss value (0.34) in just 35 epochs, as denoted by the cross markers in Figure 5.2.

It is essential to reconsider the underlying motivations driving the utilization of intensity and phase images in the proposed models. On the one hand, the real-valued convolution approach explores parameters associated with light-matter interactions, such as absorption and optical path. On the other hand,

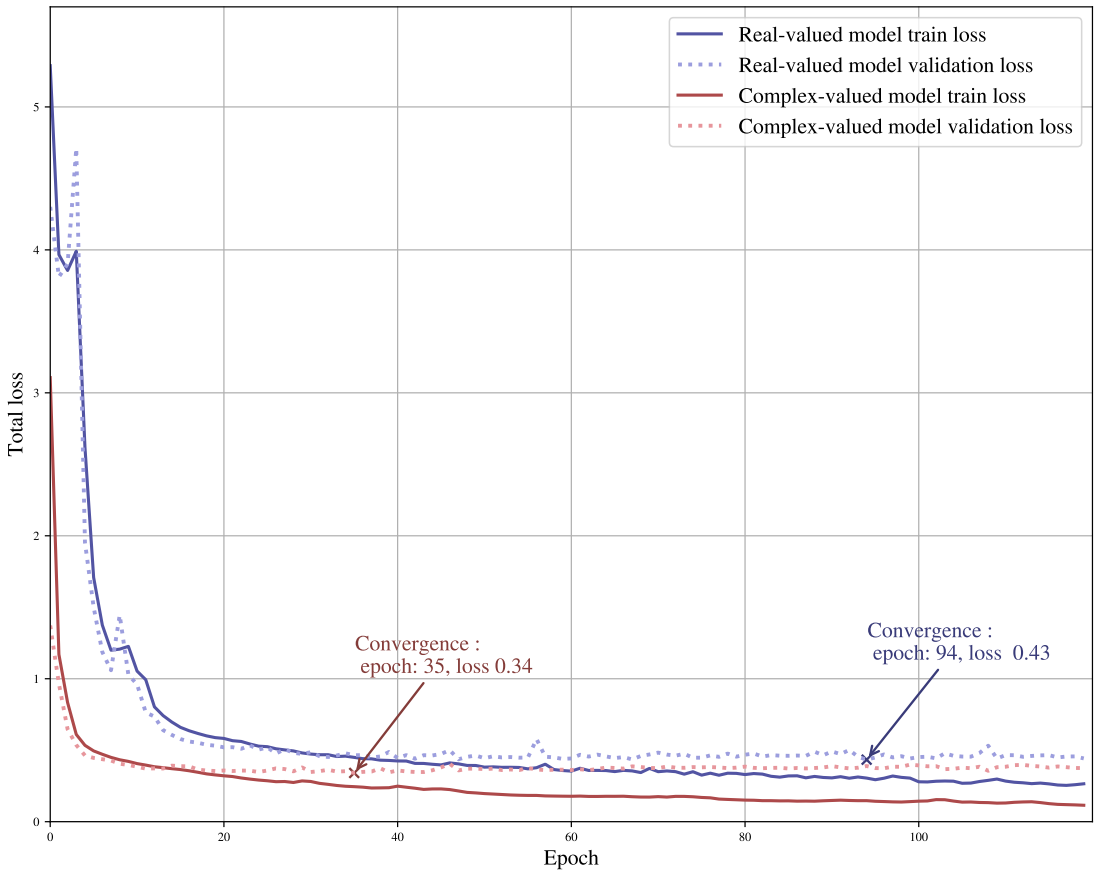


Figure 5.2: Train and validation loss curves for complex-valued and real-valued models.

the complex-valued implementation filters the image associated with matrix transfer, a process linked to the electromagnetic (EM) field emanating from the sample.

From a physical standpoint, both complex and real-valued filtering methodologies for extracting relevant features from intensity and phase data are justifiable. The performance improvement observed with the complex model can be explained by the effectiveness of the complex formalism, which includes both input coding and complex filtering. This interpretation finds support in the reduction of confidence intervals observed with complex filtering in Table 5.2, indicating greater precision, reliability, and tighter generalization capabilities. As previously cited, these findings align with established practices across various application domains.

Adopting the complex implementation is well-justified when dealing with bi-modal data. Furthermore, as highlighted by previous studies, this implementation enriches the network's representative capacity, consequently enhancing the performance of our detection application.
Note that, in previous studies, both intensity and phase result from the acquisition device (such as MRI or radar), and improvement is observed with a complex implementation. This improvement remains valid for our case, where the phase results from a computational reconstruction process.
This section shows that complex-valued formalism is more effective than its real-valued counterpart in processing bi-modal data acquired through FPM microscopy to detect malaria in red blood cells. This comparative analysis highlights the significant advantages offered by complex-valued neural networks, particularly their higher performance and strong generalization abilities. The promising results obtained demonstrate the viability and effectiveness of complex-valued neural networks, especially in situations where multi-modal data is prevalent.
Moreover, limiting the complex-based implementation to the convolution part of a DNN is sufficient to benefit from the bi-modal, complex nature of the information. This point is interesting, as it will allow generalization to many deep network models that use convolution in their first stage for feature extraction. Therefore, this work should easily be extended to other QPI imaging techniques and applications, relying on various deep network architectures.
In addition, this part also offers a first step towards optimizing the system by adapting the neural networks approach to the bi-modal nature of the data produced by the microscope, which maximizes the efficiency of the whole system.

5.2 Fourier Ptychographic microscope optimization for malaria classification: objective lens choice

In this part, we focus on optimizing the microscope setup by studying the neural network's performance for different numerical apertures. This approach aims to determine the minimum viable numerical aperture for the specific application of malaria diagnosis. Our study is limited to the binary classification of red blood cells into infected and healthy cells, thus providing valuable insights for system optimization in this specific context.

5.2.1 Objective lens optimization pipeline

Different neural networks must be trained to study the performance as a function of the image resolution in the FPM framework. To train those classifiers, different datasets at various resolutions must be created. Usually, this requires FPM data acquisition for each objective lens considered and their intensity-phase maps reconstruction. To avoid this costly process, we designed a specific protocol to derive these databases from an initial database relying on a physical image formation model. The workflow is described in figure 5.3. The process begins with acquiring a dataset of red blood cells collected from 9 infected patients. The field of view of the good spreading area was captured using the FPM device described in section 3.3.1, resulting in reconstructed intensity and phase images for the entire field of view with a synthetic aperture of 0.9 (NA = 0.45). Subsequently, red blood cells were randomly selected and extracted to generate individual images of size 64 $\times$64. This dataset includes 12,000 cells, split evenly between 6,000 healthy and 6,000 parasite-infected cells. This dataset is then combined with the image formation and reconstruction models to infer the lower-resolution datasets.

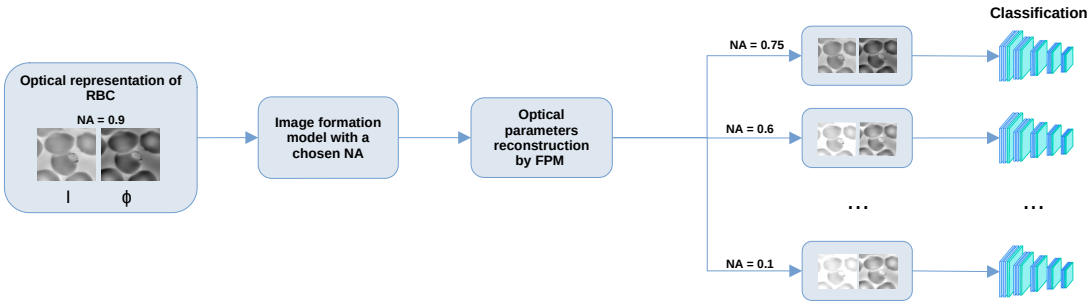


Figure 5.3: Workflow: intensity and phase maps reconstruction for different NA from experimental data and image formation model and their usage in the different classification task.

For each NA value (NA between 0.1 and 0.45), the images of the 6000 healthy and parasitized red blood cells are downgraded in resolution using the image formation model involved in the FPM reconstruction algorithm. This corresponds to complex image filtering in Fourier space. The filter shape consists of a disk centered at (0,0) with a radius $r = 2\pi NA/\lambda$. A dataset containing intensity and phase images at the corresponding resolution is generated for each NA. This principle is illustrated in Figure 5.4 for some values of NA, showing the fil-

ter's radius reduction over NAs and the modification of intensity and phase images.

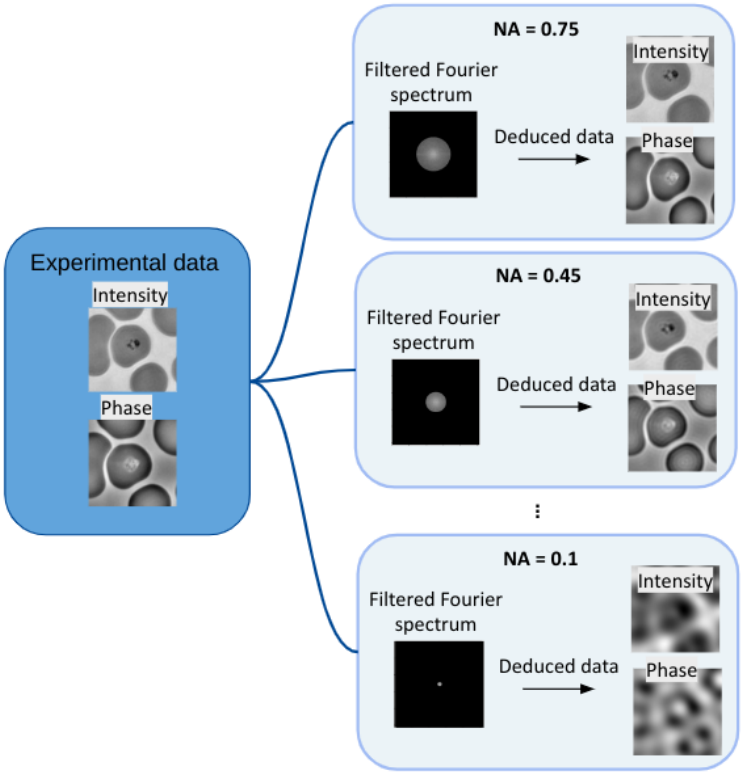


Figure 5.4: Datasets generation examples.

5.2.2 Experimental protocol

The datasets produced according to the protocol described above are used to train the two neural models ($\mathcal{I}$-RV-CNN and $\mathcal{I}/\phi$-CV-CNN) sharing the same architecture, namely nine convolutional layers with ReLU activation and three max-pooling layers, followed by two hidden layers and an output layer. The first model exclusively utilizes intensity data, while the second processes intensity and phase information. Notably, for the second model, we opt for a complex-valued architecture. Notably, for the second model, we opt for a complex-valued architecture, which is efficient for processing bi-modal data in Section 5.1.1 by replacing the real-valued convolution, the activation function, and the pooling with their equivalent in complex-valued as details in Section 5.1.1. Each model was trained using five-fold cross-validation, and the hyperparameters,

such as learning rate and batch size, were adjusted to ensure optimal convergence for each model at each numerical aperture.

5.2.3 Results and discussion

Models performances are summarized in Figure 5.5.

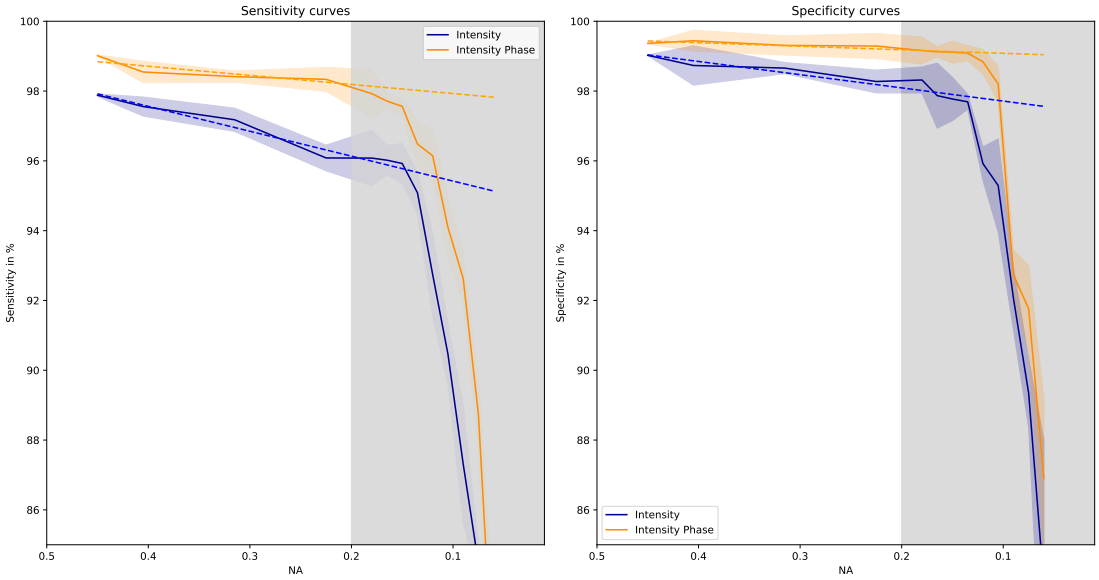


Figure 5.5: Sensitivity and specificity performances for each NA.

The left-hand figure corresponds to sensitivity, and the right-hand figure corresponds to specificity. The blue curves are obtained by exploiting intensity and phase images jointly $\mathcal{I}/\phi$-CV-CNN. The yellow curves are relative to the exploitation of intensity images alone ($\mathcal{I}$-RV-CNN model). The curves in continuous lines correspond in each point to the average obtained on the five folds of cross-valuation. The hallow around corresponds to the zone of the standard deviation of more or less the standard deviation. The dotted curves represent the tangent to each curve.
Several interesting elements can be noted. Firstly, and in coherence with the results obtained in Section 3.3.3, classification performance is systematically better when both intensity and phase images are used in a complex-valued CNN. Secondly, the evolution of the curves is different in the two cases (cf. dotted curves). In particular, it can be seen that performance for the $\mathcal{I}/\phi$-CV model is relatively stable over a wide range of NAs (down to NA=0.2), and the tan-

gent slope is near zero. This is not the case for the I-RV network, which exhibits a deterioration with decreasing NA; the slope value equals 7.34 and 3.82 for sensitivity and specificity, respectively. Finally, when NA falls below 0.15, performance drops sharply for both models.
For automatic malaria diagnosis, the results suggest that an objective lens with a numerical aperture as low as 0.2 (accessible with a 4x magnification objective) could be a practicable option. With such a lens, the associated field of view is approximately 600 times larger than that of a 100x lens, which is the standard for expert visual examination of blood smears. For $\boldsymbol{NA = 0.2}$ we estimate sensitivity and specificity at 98.1% and 99.3% , respectively. These figures could be further improved with a more extensive learning dataset.

5.3 Conclusion

This chapter was devoted to presenting tw studies aiming to optimize a system based on deep learning and an FPM microscope for specific tasks related to malaria diagnosis. Incorporating phase information in addition to intensity has significantly improved the performance of classification tasks, which confirms the proof of concept of Chapter 3. With complex-valued neural networks (CVNN) outperforming their real-valued counterparts, a first optimization at the level of exploitation has been introduced. This superiority was demonstrated in malaria detection, where CVNNs exhibited significantly higher sensitivity and specificity than real-valued models. In addition, the objective lens optimization study, based on the performance of the neural networks, has made it possible to put forward an approach to overcome as much as possible the compromise between the resolution and the size of the field of view in the diagnosis of malaria through the classification task. The results suggest that using objective lenses with a numerical aperture (NA) as low as 0.2 may be sufficient to achieve high performance in terms of sensitivity and specificity. This digital opening value considerably reduces the acquisition times required for scanning objects of interest and reduces the resolution constraints necessary in the studied problem. Validation of these findings using an experimental dataset and generalization over other applications is now needed.

Chapter 6

Conclusion and perspectives

6.1 Conclusions

As part of multidisciplinary research, this thesis explored the application of an unconventional microscopy method, FPM, in two medical applications facing the limits of classical microscopy. FPM is a low-cost, stable, quantitative phase imaging method suitable for routine clinical use. It allows obtaining both intensity and phase images using adequate reconstruction models. The main objective of this thesis was to show how the enriched information provided by the FPM framework can help to improve the diagnosis and monitoring of hematological and parasitological pathologies. Towards two specific use-cases (parasite detection for malaria diagnosis and white blood cell analysis), we explored different aspects related to the design of FPM-based detection systems in blood smears.

In Chapter 3, we studied the interest of the phase for the analysis of stained objects. Some theoretical hypotheses have been put forward and validated on the limited use-case of parasite detection (2-classes classification). The results demonstrate that coupling intensity to phase is more efficient than using intensity alone.

In chapter 4, this issue is studied on a more complicated application, namely for automatically classifying white blood cells. The results show that the model combining intensity and phase outperforms those exploiting a single modality in terms of precision and recall where an enhancement of up to 30% is observed, thus better classifying different types of white blood cells and allowing better separation of classes. In addition, the images provided by these methods of-

fer better performance than those of current automata, with an enhancement of performance of 30% over some classes. The work also addressed other limiting problems in certain biomedical applications, such as the problem of class imbalance, which is frequently encountered in biomedical applications. After unsuccessful attempts to implement a classical GAN model able to generate good quality bi-modal FPM images, we proposed an innovative approach, a physics-inspired GAN. Using the FPM image formation model coupled with adversarial generative models, this physics-informed GAN approach was able to generate images close to reality with a high diversity that made it possible to respond to the mode collapse problem. Moreover, this physics-informed GAN approach, coupled with a fine-tuning procedure, could generate synthetic images for a class initially containing very few representatives. Our results showed that integrating synthetic data significantly improved the classification performance, particularly for minority classes, without degrading the performance of non-augmented classes.
Chapter 5 is dedicated to some attempts to optimize the previous FPM framework towards an efficient diagnostic tool. First, we implemented complex-valued convolutional neural layers in a Faster RCNN to detect parasite red cells. The results show improved problem detection performance, bringing the sensitivity to 97.15%. It also enabled faster convergence and better generalization with intricate confidence intervals compared to real-valued models. The second part of chapter 5 was dedicated to optimizing the choice of the microscope objective lens by emphasizing the balance between the resolution and the speed of data acquisition, i.e., size of the field of view. The results demonstrated that FPM images allow resistance to low resolution and an increase in the field of view size. It is suggested that a numerical aperture as low as 0.2 may be sufficient for high performance while reducing acquisition time.

6.2 Perspectives

Different perspectives of this work can be envisaged in continuation of the contributions already obtained in this thesis.

6.2.1 Towards an FPM-based system for digital pathology

In continuation of chapter 5, several issues have to be tackled to comfort our first results, such as:

- **Optimization of the acquisition device:** Additional experiments are necessary to validate the results of optimizing the microscopic configuration in chapter 5.
- Considering more complicated biological tasks per the biologist's needs. The experimental databases that we considered in this thesis were of limited size and diversity. It is necessary to perform other experiments on databases acquired following the biologist's needs. For instance, for malaria diagnosis, it is crucial to expand tasks to more complex levels, such as recognizing infection stages and classifying parasite types, in order to design a complete automated diagnosis system. Furthermore, dilution tests should be performed to understand the sensitivity limits better and confirm FPM's advantages. For white blood cell classification, the FPM framework (including GAN implementation) can be tested further for classifying malignant cells and diagnosing blood pathologies such as leukemia.

6.2.2 Physics-Informed GANs (PI-GANs)

In the present study, we only considered one wavelength, so we had only to produce two modalities from one wavelength, one for intensity and one for phase, and the PI-GAN was implemented in this case. A further study should consider the three available wavelengths and design a synthetic image formation model in that case in order to benefit from the complementarity of the three channels. Other approaches for synthetic image production are proposed in the literature as an alternative to GAN, such as diffusion models. Implementing such models with the physics-inspired scheme is also an interesting extension of the present work. Let us quote that using PI-GAN could find interest in other contexts where images are computationally obtained via an image formation model, notably in radar approaches, often faced with limited data problems.

6.2.3 Label-free imaging

In all the datasets considered in this work, the images resulted from a coloration, and it was an interesting outcome of our work to show that intensity and phase

were complementary. They could be used efficiently in a DNN.
Accessing a label-free method presents several benefits: it simplifies the preparation process, reduces the necessity for handling potentially harmful reagents, and preserves the natural chemical state of the observed cell.
In the absence of staining (label-free mode), absorption information is limited but can be supplemented by phase information characterizing the optical pathway, which carries a differential signal depending on the nature and texture of the cellular compartments.
So, future investigations should evaluate the interest in FPM approaches in this label-free context, where phase information is even more crucial than in colored images.

6.2.4 Diffraction tomography using FPM

Fourier ptychography relies on the thin sample assumption, limiting its suitability to thin samples. However, medical diagnoses often involve thicker objects like muscle tissue, which do not meet this assumption.
Examining samples with reconstructions at different planes would be relevant in this context. This entails capturing samples using the FPM device or digitally reconstructing intensity and phase by varying the distance of the sample from the microscope. This will produce multiple pairs of bimodal images describing the object. The next step involves developing new methods to process these different intensity and phase images and extract pertinent information on the object.

Bibliography

[1] *Ondes électromagneétiques en radioeélectricité et en optique*. Enseignement de la physique,. Masson, Paris, Milan, Barcelone, 1989. OCLC: 491636923. 85

[2] E. Abbe. Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für Mikroskopische Anatomie*, 9(1):413–468. 50, 53

[3] Hasan Berkay Abdioglu, Yagmur Isik, Merve Sevgi, Ufuk Gorkem Kirabali, Caner Karaca, Yunus Emre Mert, Defne Dedehayir, Berke Tuna Bostanci, Yasemin Basbinar, and Huseyin Uvet. Biomechanical analysis of breast cancer cells: A comparative study of invasive and non-invasive cell lines using digital holographic microscopy, 2023. 59

[4] Nour Aburaed, Mohammed Alkhatib, Stephen Marshall, Jaime Zabalza, and Hussain Al-Ahmad. Complex-valued neural network for hyperspectral single image super resolution. In Nick J. Barnett, Aoife A. Gowen, and Haida Liang, editors, *Hyperspectral Imaging and Applications II*, page 15. SPIE. 146

[5] Alfonso Blazquez Castro and Juan Carlos Stockert. *Fluorescence Microscopy In Life Sciences*. Bentham Science Publishers. OCLC: 1027570314. 55

[6] Khaled Almezhghwi and Sertan Serte. Improved classification of white blood cells with the generative adversarial network and deep convolutional neural network. *Computational Intelligence and Neuroscience*, 2020, 2020. 188

[7] Adel Ammar, Anis Koubaa, Mohanned Ahmed, Abdulrahman Saad, and Bilel Benjdira. Aerial images processing for car detection using convolutional neural networks: Comparison between faster r-cnn and yolov3. *arXiv preprint arXiv:1910.07234*, 2019. 95

[8] A. Anand, V. K. Chhaniwal, N. R. Patel, and B. Javidi. Automatic identification of malaria-infected RBC with digital holographic microscopy using correlation algorithms. *IEEE Photonics Journal*, 4(5):1456–1464, 2012. 58

[9] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein GAN. 123

[10] Danilo Avola, Luigi Cinque, Alessio Fagioli, Gianluca Foresti, and Alessio Mecca. Ultrasound medical imaging techniques: A survey. *ACM Computing Surveys*, 54(3):1–38. 49

[11] D Axelrod. Cell-substrate contacts illuminated by total internal reflection fluorescence. *The Journal of cell biology*, 89(1):141–145. 51

[12] Neeraj Baghel, Upendra Verma, and Kapil Kumar Nagwanshi. WBCsnet: type identification of white blood cells using convolutional neural network. 81(29):42131–42147. 188

[13] Duhyeon Bang and Hyunjung Shim. MGGAN: Solving mode collapse using manifold-guided training. pages 2347–2356. 131

[14] J. A. Barrachina, C. Ren, C. Morisseau, G. Vieillard, and J.-P. Ovarlez. Complex-valued vs. real-valued neural networks for classification perspectives: An example on non-circular data. In *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2990–2994. IEEE. 146

[15] Kevin Barrera, Anna Merino, Angel Molina, and José Rodellar. Automatic generation of artificial images of leukocytes and leukemic cells using generative adversarial networks (syntheticcellgan). 229:107314. 188

[16] Joshua Bassey, Lijun Qian, and Xianfang Li. A survey of complex-valued neural networks. 145

[17] Frank M. Bengel, Takahiro Higuchi, Mehrbod S. Javadi, and Riikka Lautamäki. Cardiac positron emission tomography. *Journal of the American College of Cardiology*, 54(1):1–15. 49

[18] Giuseppe Biondi-Zoccai, Mariangela Peruzzi, Simona Mastrangeli, and Giacomo Frati. Introduction to clinical diagnosis. In Giuseppe Biondi-Zoccai, editor, *Diagnostic Meta-Analysis*, pages 3–9. Springer International Publishing. 48

[19] Léon Bottou. Stochastic gradient descent tricks. In Grégoire Montavon, Geneviève B. Orr, and Klaus-Robert Müller, editors, *Neural Networks: Tricks of the Trade*, volume 7700, pages 421–436. Springer Berlin Heidelberg. 66

[20] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale GAN training for high fidelity natural image synthesis. 131

[21] A. Caragea, D. G. Lee, J. Maly, G. Pfander, and F. Voigtlaender. Quantitative approximation results for complex-valued neural networks. 145

[22] Robert D. Cardiff, Claramae H. Miller, and Robert J. Munn. Manual hematoxylin and eosin staining of mouse tissue sections. *Cold Spring Harbor Protocols*, 2014(6):pdb.prot073411. 55

[23] W. Chen. Clinical applications of PET in brain tumors. *Journal of Nuclear Medicine*, 48(9):1468–1481. 49

[24] César Cheuque, Marvin Querales, Roberto León, Rodrigo Salas, and Romina Torres. An efficient multi-level convolutional neural network approach for white blood cells classification. 12(2):248. 188

[25] Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, and The National Academies of Sciences, Engineering, and Medicine. *Improving Diagnosis in Health Care*. National Academies Press. 48

[26] Ramona Corman, Willem Boutu, Anna Campalans, Pablo Radicella, Joana Duarte, Maria Kholodtsova, Laure Bally-Cuif, Nicolas Dray, Fabrice Harms, Guillaume Dovillaire, Samuel Bucourt, and Hamed Merdji. Lensless microscopy platform for single cell and tissue visualization. *Biomedical Optics Express*, 11(5):2806, 2020. 60

[27] Etienne De Harven. Electron microscopy of cancer cells: A review. *Medical Clinics of North America*, 50(3):887–900. 56

[28] Jesper Sören Dramsch, Mikael Lüthje, and Anders Nymark Christensen. Complex-valued neural networks for machine learning on non-stationary physical data. 146:104643. 145, 150

[29] Kai Moritz Eder, Anne Marzi, Alvaro Barroso, Steffi Ketelhut, Bjorn Kemper, and Jurgen Schnekenburger. Label-free digital holographic microscopy for in vitro cytotoxic effect quantification of organic nanoparticles. *Cells*, 11(4):644, 2022. 59

[30] R.F. Egerton. *Physical Principles of Electron Microscopy*. Springer International Publishing. 56

[31] Andre Esteva, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, February 2017. 69

[32] J. R. Fienup. Phase retrieval algorithms: a comparison. *Applied Optics*, 21(15):2758. 91

[33] for the Alzheimer's Disease Neuroimaging Initiative, Avinash Chandra, George Dervenoulas, and Marios Politis. Magnetic resonance imaging in alzheimer's disease and mild cognitive impairment. *Journal of Neurology*, 266(6):1293–1302. 49

[34] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015. 95

[35] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014. 95

[36] Harshvardhan Gm, Mahendra Kumar Gourisaria, Manjusha Pandey, and Siddharth Swarup Rautaray. A comprehensive survey and analysis of generative models in machine learning. 38:100285. 121

[37] Cynthia S. Goldsmith and Sara E. Miller. Modern uses of electron microscopy for detection of viruses. *Clinical Microbiology Reviews*, 22(4):552–563. 56

[38] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. 121

[39] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company publishers, 2005. 90

[40] Alex Graves. Generating sequences with recurrent neural networks. 66

[41] Alon Greenbaum, Yibo Zhang, Alborz Feizi, Ping-Luen Chung, Wei Luo, Shivani R. Kandukuri, and Aydogan Ozcan. Wide-field computational imaging of pathology slides using lens-free on-chip microscopy. *Science Translational Medicine*, 6(267), 2014. 60

[42] MAY GRUNWALD-CORANTE. May grunwald-giemsa. 55

[43] M. G. L. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy: SHORT COMMUNICATION. *Journal of Microscopy*, 198(2):82–87, 2000. 61

[44] C A Hartanto, S Kurniawan, D Arianto, and A M Arymurthy. DCGAN-generated synthetic images effect on white blood cell classification. 1077(1):012033. 188

[45] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. 191

[46] K Herholz and W Heiss. Positron emission tomography in clinical neurology. *Molecular Imaging & Biology*, 6(4):239–269. 49

[47] Akira Hirose. Recent progress in applications of complex-valued neural networks. In Leszek Rutkowski, Rafał Scherer, Ryszard Tadeusiewicz, Lotfi A. Zadeh, and Jacek M. Zurada, editors, *Artifical Intelligence and Soft Computing*, volume 6114, pages 42–46. Springer Berlin Heidelberg. 145

[48] Akira Hirose. Complex-valued neural networks: The merits and their origins. In *2009 International Joint Conference on Neural Networks*, pages 1237–1244, 2009. 145

[49] Robert Hooke, editor. *Micrographia*. Leipzig. OCLC: 1137398899. 50

[50] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. MobileNets: Efficient convolutional neural networks for mobile vision applications. 110

[51] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. 195

[52] Xiwei Huang, Yangbo Li, Xuefeng Xu, Renjie Wang, Jiangfan Yao, Wentao Han, Maoyu Wei, Jin Chen, Weipeng Xuan, and Lingling Sun. High-precision lensless microscope on a chip based on in-line holographic imaging. *Sensors*, 21(3):720, 2021. 60

[53] Toshiyuki Ishiwata, Fumio Hasegawa, Masaki Michishita, Norihiko Sasaki, Naoshi Ishikawa, Kaiyo Takubo, Yoko Matsuda, Tomio Arai, and Junko Aida. Electron microscopic analysis of different cell types in human pancreatic cancer spheres. *Oncology Letters*. 56

[54] Shaowei Jiang, Jun Liao, Zichao Bian, Kaikai Guo, Yongbing Zhang, and Guoan Zheng. Transform- and multi-domain deep learning for single-frame rapid autofocusing in whole slide imaging. *Biomedical Optics Express*, 9(4):1601. 70

[55] YoungJu Jo, Sangjin Park, JaeHwang Jung, Jonghee Yoon, Hosung Joo, Min-hyeok Kim, Suk-Jo Kang, Myung Chul Choi, Sang Yup Lee, and YongKeun Park. Holographic deep learning for rapid optical screening of anthrax spores. *Science Advances*, 3(8):e1700606. 70

[56] Terry Jones and David Townsend. History and future technical innovation in positron emission tomography. *Journal of Medical Imaging*, 4(1):011013. 49

[57] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of StyleGAN. 126, 190

[58] Amir Khodavirdipour, Mahboobeh Mehregan, Ali Rajabi, and Yasoub Shiri. Microscopy and its application in microbiology and medicine from light to quantum microscopy: A mini review. *Avicenna Journal of Clinical Microbiology and Infection*, 6(4):133–137. 56

[59] John A. Kiernan. *Histological and histochemical methods: theory and practice*. Scion, 4. ed., repr edition. 73

[60] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. 66

[61] M. Knoll and E. Ruska. Das elektronenmikroskop. *Zeitschrift for Physik*, 78(5):318–339. 51

[62] Naveen Kodali, Jacob Abernethy, James Hays, and Zsolt Kira. On convergence and stability of GANs. 131

[63] August Köhler. New method of illumination for photomicrographical purposes. *Journal of the Royal Microscopical Society*, 14:261–262, 1894. 52

[64] George Kolokolnikov and Andrey Samorodov. Comparative study of data augmentation strategies for white blood cells classification. In *2019 25th Conference of Open Innovations Association (FRUCT)*, pages 168–175, 2019. 188

[65] Vladyslav Kotsovsky, Anatoliy Batyuk, and Maksym Yurchenko. New approaches in the learning of complex-valued neural networks. In *2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP)*, pages 50–54. IEEE. 145

[66] Hendrik Anthony Kramers. La diffusion de la lumiere par les atomes. In *Atti Cong. Intern. Fisica (Transactions of Volta Centenary Congress) Como*, volume 2, pages 545–557, 1927. 87

[67] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. 68

[68] Gihyun Kwon, Chihye Han, and Dae-shik Kim. Generation of 3d brain MRI using auto-encoding generative adversarial networks. In Dinggang Shen, Tianming Liu, Terry M. Peters, Lawrence H. Staib, Caroline Essert, Sean Zhou, Pew-Thian Yap, and Ali Khan, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, volume 11766, pages 118–126. Springer International Publishing. 131

[69] P. C. Lauterbur. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature*, 242(5394):190–191. 48

[70] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324. 67

[71] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444. 66

[72] Minhyeok Lee and Junhee Seok. Regularization methods for generative adversarial networks: An overview of recent studies. 124, 131

[73] Myungjun Lee, Oguzhan Yaglidere, and Aydogan Ozcan. Field-portable reflection and transmission microscopy based on lensless holography. *Biomedical Optics Express*, 2(9):2721, 2011. 60

[74] Zhenyu Lei, Shangce Gao, Hideyuki Hasegawa, Zhiming Zhang, MengChu Zhou, and Khaled Sedraoui. Fully complex-valued gated recurrent neural network for ultrasound imaging. pages 1–14. 146

[75] Timothy G. Leighton. What is ultrasound? *Progress in Biophysics and Molecular Biology*, 93(1):3–83. 49

[76] 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. 94

[77] Yang-Hsien Lin, Ken Y.-K. Liao, and Kung-Bin Sung. Automatic detection and characterization of quantitative phase images of thalassemic red blood cells using a mask region-based convolutional neural network. *Journal of Biomedical Optics*, 25(11). 70

[78] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. Van Der Laak, Bram Van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, December 2017. 69

[79] Luciano Lombardi and Attilio Orazi. Electron microscopy in an oncologic institution. diagnostic usefulness in surgical pathology. *Tumori Journal*, 74(5):531–535. 56

[80] Li Ma, Renjun Shuai, Xuming Ran, Wenjia Liu, and Chao Ye. Combining dc-gan with resnet for blood cell image classification. *Medical & biological engineering & computing*, 58:1251–1264, 2020. 188

[81] Andrew M Maiden and John M Rodenburg. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, 109(10):1256–1262, 2009. 91

[82] Jyoti Mangal, Rashi Monga, Sandeep R. Mathur, Amit K. Dinda, Joby Joseph, Sarita Ahlawat, and Kedar Khare. Unsupervised organization of cervical cells using bright-field and single-shot digital holographic microscopy. *Journal of Biophotonics*, 12(8):e201800409, 2019. 59

[83] Christopher J. Mann, Lingfeng Yu, Chun-Min Lo, and Myung K. Kim. High-resolution quantitative phase-contrast microscopy by digital holography. *Optics Express*, 13(22):8693, 2005. 58

[84] Pierre Marquet, Benjamin Rappaz, Pierre J. Magistretti, Etienne Cuche, Yves Emery, Tristan Colomb, and Christian Depeursinge. Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy. *Optics Letters*, 30(5):468, 2005. 58

[85] Christian Matek, Simone Schwarz, Carsten Marr, and Karsten Spiekermann. A single-cell morphological dataset of leukocytes from AML patients and non-malignant controls. 184

[86] Christian Matek, Simone Schwarz, Karsten Spiekermann, and Carsten Marr. Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. 1(11):538–544. 184

[87] Christian Matek, Simone Schwarz, Karsten Spiekermann, and Carsten Marr. Human-level recognition of blast cells in acute myeloid leukemia with convolutional neural networks. 02 2019. 36, 187

[88] Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, 1947. 102, 154

[89] M. Minsky. Memoir on inventing the confocal scanning microscope. *Scanning*, 10(4):128–138. 51

[90] A. Molder, M. Sebesta, M. Gustafsson, L. Gisselson, A. Gjorloff Wingren, and K. Alm. Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography. *Journal of Microscopy*, 232(2):240–247, 2008. 58

[91] Christopher Moore and Susan B. Promes. Ultrasound in pregnancy. *Emergency Medicine Clinics of North America*, 22(3):697–722. 49

[92] Elizabeth A Morris. Breast cancer imaging with MRI. *Radiologic Clinics of North America*, 40(3):443–466. 49

[93] George F. Murphy. The clinical use of electron microscopy in dermatology. *Archives of Dermatology*, 122(9):1074. 56

[94] Gillian Murphy, Masoom Haider, Sangeet Ghai, and Boraiah Sreeharsha. The expanding role of MRI in prostate cancer. *American Journal of Roentgenology*, 201(6):1229–1238. 49

[95] Paul G. Newman and Grace S. Rozycki. THE HISTORY OF ULTRASOUND. *Surgical Clinics of North America*, 78(2):179–195. 49

[96] A. Nordberg. PET studies and cholinergic therapy in alzheimer's disease. *Revue Neurologique*, 155 Suppl 4:S53–63. 49

[97] Jonathan Nylk, James M. Pullman, Elaine C. Campbell, Frank J. Gunn-Moore, Michael B. Prystowsky, and Kishan Dholakia. Structured illumination microscopy as a diagnostic tool for nephrotic disease. In Anita Mahadevan-Jansen, Tuan Vo-Dinh, and Warren S. Grundfest, editors, *SPIE BiOS*, page 1005417, 2017. 61

[98] Timothy O'Connor, Christopher Hawxhurst, Leslie M. Shor, and Bahram Javidi. Red blood cell classification in lensless single random phase encoding using convolutional neural networks. *Optics Express*, 28(22):33504. 70

[99] Daniel C. Pease. An electron microscopic study of red bone marrow. *Blood*, 11(6):501–526. 56

[100] Sen Pei, Richard Yi Da Xu, Shiming Xiang, and Gaofeng Meng. Alleviating mode collapse in GAN via diversity penalty module. 131

[101] Eric Piaton, Monique Fabre, Isabelle Goubin-Versini, Marie-Françoise Bretz-Grenier, Monique Courtade-Saïdi, Serge Vincent, Geneviève Belleannée, Françoise Thivolet, Jean Boutonnat, Hervé Debaque, Jocelyne Fleury-Feith, Philippe Vielh, Béatrix Cochand-Priollet, Caroline Egelé, Jean-Pierre Bellocq, and Jean-François Michiels. Recommandations techniques et règles de bonne pratique pour la coloration de may-grünwald-giemsa : revue de la littérature et apport de l'assurance qualité. 35(4):294–305. 73

[102] Nandor K. Pinter and Joseph V. Fritz. Neuroimaging for the neurologist. *Neurologic Clinics*, 38(1):1–35. 49

[103] Mahdieh Poostchi, Kamolrat Silamut, Richard J. Maude, Stefan Jaeger, and George Thoma. Image analysis and machine learning for detecting malaria. *Translational Research*, 194:36–55, 2018. In-Depth Review: Diagnostic Medical Imaging. 94

[104] Douglas Queen and Keith Gordon Harding. Importance of imaging to wound care practice. *International Wound Journal*, 20(2):235–237. 48

[105] Ernst-Wilhelm Radue, Matthias Weigel, Roland Wiest, and Horst Urbach. Introduction to magnetic resonance imaging for neurologists. *CONTINUUM: Lifelong Learning in Neurology*, 22(5):1379–1398. 49

[106] Logesh Rajamani, Bhavya Basetti, Rajeswaran Rangasami, Anupama Chandrasekharan, and Rajoo Ramachandran. Magnetic resonance imaging in the evaluation of the pathologies affecting large intracranial arteries. *Journal of Neurosciences in Rural Practice*, 14:35–40. 49

[107] Benjamin Rappaz, Alexander Barbul, Annick Hoffmann, Daniel Boss, Rafi Korenstein, Christian Depeursinge, Pierre J. Magistretti, and Pierre Marquet. Spatial analysis of erythrocyte membrane fluctuations by digital holographic microscopy. *Blood Cells, Molecules, and Diseases*, 42(3):228–232, 2009. 58

[108] Benjamin Rappaz, Pierre Marquet, Etienne Cuche, Yves Emery, Christian Depeursinge, and Pierre J. Magistretti. Measurement of the integral refractive index and dynamic cell morphometry of living cells with digital holographic microscopy. *Optics Express*, 13(23):9361, 2005. 58

[109] Sebastian Raschka. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv preprint arXiv:1811.12808*, 2018. 102, 113, 154

[110] 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. 53

[111] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016. 95

[112] John Reid. Medical ultrasonics: Diagnostic applications of ultrasound. *Proceedings of the IRE*, 47(11):1963–1967. 49

[113] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 28. Curran Associates, Inc., 2015. 95

[114] 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. 69

[115] J. M. Rodenburg and H. M. L. Faulkner. A phase retrieval algorithm for shifting illumination. *Applied Physics Letters*, 85(20):4795–4797. 91

[116] 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. 69

[117] Reena M Roy and PM Ameer. Identification of white blood cells for the diagnosis of acute myeloid leukemia. *International Journal of Imaging Systems and Technology*, 32(4):1307–1317, 2022. 188

[118] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536. 65

[119] Muhammad Muneeb Saad, Ruairi O'Reilly, and Mubashir Husain Rehmani. A survey on training challenges in generative adversarial networks for biomedical image analysis. 131

[120] Mehdi S. M. Sajjadi, Olivier Bachem, Mario Lucic, Olivier Bousquet, and Sylvain Gelly. Assessing generative models via precision and recall. 131

[121] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training GANs. 131

[122] Luke Sheneman, Gregory Stephanopoulos, and Andreas E. Vasdekis. Deep learning classification of lipid droplets in quantitative phase images. *PLOS ONE*, 16(4):e0249196. 70

[123] Igor A. Shevkunov, Meenakshisundaram Kandhavelu, and Karen Eguiazarian. CNN-assisted quantitative phase microscopy for biological cell imaging. In Emmanuel Beaurepaire, Adela Ben-Yakar, and YongKeun Park, editors, *Advances in Microscopic Imaging IV*, page 40. SPIE. 70

[124] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. 2014. 96, 97, 111

[125] M. Stone. Cross-validatory choice and assessment of statistical predictions. 36(2):111–133. 100

[126] 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. Number: 1 Publisher: Nature Publishing Group. 94

[127] Yuki Sunaga, Ryo Natsuaki, and Akira Hirose. Land form classification and similar land-shape discovery by using complex-valued convolutional neural networks. 57(10):7907–7917. 146

[128] Y F Tai. Applications of positron emission tomography (PET) in neurology. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(5):669–676. 49

[129] Lei Tang and Xiaohong Joe Zhou. Diffusion MRI of cancer: From low to high b-values. *Journal of Magnetic Resonance Imaging*, 49(1):23–40. 49

[130] Dániel Terbe, László Orzó, and Ákos Zarándy. Classification of holograms with 3d-CNN. *Sensors*, 22(21):8366. 70

[131] John A. Terzakis, Eugene Santagada, Ascension Hernandez, and Metin Taskin. Scanning electron microscopy of peripheral blood smears: Comparison of normal blood with some common leukemias. *Ultrastructural Pathology*, 29(1):19–28. 56

[132] Leena Thomas and M. K. Sheeja. Fourier ptychographic and deep learning using breast cancer histopathological image classification. *Journal of Biophotonics*, 16(10):e202300194. 71

[133] Lei Tian and Laura Waller. 3d intensity and phase imaging from light field measurements in an led array microscope. *Optica*, 2(2):104–111, Feb 2015. 40

[134] M Titford. The long history of hematoxylin. *Biotechnic & Histochemistry*, 80(2):73–78. 55

[135] Chiheb Trabelsi, Olexa Bilaniuk, Ying Zhang, Dmitriy Serdyuk, Sandeep Subramanian, João Felipe Santos, Soroush Mehri, Negar Rostamzadeh, Yoshua Bengio, and Christopher J Pal. Deep complex networks. 145, 150

[136] Matthias Ugele, Markus Weniger, Maria Leidenberger, Yiwei Huang, Michael Bassler, Oliver Friedrich, Barbara Kappes, Oliver Hayden, and Lukas Richter. Label-free, high-throughput detection of *P. falciparum* infection in sphered erythrocytes with digital holographic microscopy. *Lab on a Chip*, 18(12):1704–1712, 2018. 58

[137] Matthias Ugele, Markus Weniger, Manfred Stanzel, Michael Bassler, Stefan W. Krause, Oliver Friedrich, Oliver Hayden, and Lukas Richter. Label-free high-throughput leukemia detection by holographic microscopy. *Advanced Science*, 5(12):1800761. 58

[138] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008. 116

[139] Luis E. Villegas-Hernández, Mona Nystad, Florian Ströhl, Purusotam Basnet, Ganesh Acharya, and Balpreet S. Ahluwalia. Visualizing ultrastructural details of placental tissue with super-resolution structured illumination microscopy. *Placenta*, 97:42–45, 2020. 62

[140] Patrick Virtue, Stella X. Yu, and Michael Lustig. Better than real: Complex-valued neural nets for MRI fingerprinting. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3953–3957. IEEE. 146

[141] Henry N. Wagner. A brief history of positron emission tomography (PET). *Seminars in Nuclear Medicine*, 28(3):213–220. 49

[142] Devin L. Wakefield, Richard Graham, Kevin Wong, Songli Wang, Christopher Hale, and Chung-Chieh Yu. Cellular analysis using label-free parallel array microscopy with fourier ptychography. *Biomedical Optics Express*, 13(3):1312. 70

[143] Shushan Wang, Tingfa Xu, Jizhou Zhang, Xin Wang, Yiwen Chen, and Jinhua Zhang. Automatic counting system of red blood cells based on fourier ptychographic microscopy. In Qilian Liang, Wei Wang, Xin Liu, Zhenyu Na, Xiaoxia Li, and Baoju Zhang, editors, *Communications, Signal Processing, and Systems*, volume 654, pages 891–898. Springer Singapore. 70

[144] Emil Wolf. *Introduction to the Theory of Coherence and Polarization of Light*. Cambridge university press, 2007. 89, 145

[145] Chansuda Wongsrichanalai, Mazie J Barcus, Sinuon Muth, Awalludin Sutamihardja, and Walther H Wernsdorfer. A review of malaria diagnostic tools: microscopy and rapid diagnostic test (rdt). *Defining and Defeating the Intolerable Burden of Malaria III: Progress and Perspectives: Supplement to Volume 77 (6) of American Journal of Tropical Medicine and Hygiene*, 2007. 56

[146] Saining Xie, Ross Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. 185

[147] Fannuo Xu, Zipei Wu, Chao Tan, Yizheng Liao, Zhiping Wang, Keru Chen, and An Pan. Fourier ptychographic microscopy 10 years on: A review. 13(4):324. 32, 41

[148] Faliu Yi, Inkyu Moon, and Bahram Javidi. Automated red blood cells extraction from holographic images using fully convolutional neural networks. *Biomedical Optics Express*, 8(10):4466. 70

[149] Lingjuan Yu, Yuehong Hu, Xiaochun Xie, Yun Lin, and Wen Hong. Complex-valued full convolutional neural network for SAR target classification. 17(10):1752–1756. 146

[150] Tieqiao Zhang, Samantha Osborn, Chloe Brandow, Denis Dwyre, Ralph Green, Stephen Lane, and Sebastian Wachsmann-Hogiu. Structured illumination-based super-resolution optical microscopy for hemato- and cyto-pathology applications. *Analytical Cellular Pathology*, 36(1):27–35, 2013. 61

[151] Xiaohui Zhang, Imran Khimji, Umut Atakan Gurkan, Hooman Safaee, Paolo Nicolas Catalano, Hasan Onur Keles, Emre Kayaalp, and Utkan Demirci. Lensless imaging for simultaneous microfluidic sperm monitoring and sorting. *Lab on a Chip*, 11(15):2535, 2011. 60

[152] Guo zhen Yang, Bi zhen Dong, Ben yuan Gu, Jie yao Zhuang, and Okan K. Ersoy. Gerchberg–saxton and yang–gu algorithms for phase retrieval in a nonunitary transform system: a comparison. *Appl. Opt.*, 33(2):209–218, Jan 1994. 91

[153] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang. Wide-field, high-resolution Fourier ptychographic microscopy. *Nature Photonics*, 7(9):739–745, September 2013. 40

[154] Warren R Zipfel, Rebecca M Williams, and Watt W Webb. Nonlinear magic: multiphoton microscopy in the biosciences. *Nature Biotechnology*, 21(11):1369–1377. 51

Appendix A

White Blood Cells classification using LMU dataset

Contents

It is worth noting that data accessibility for TAMIS project was constrained, with availability only during the concluding phase of the thesis. Nonetheless, the project adeptly capitalizes on an existing repository of classical imaging data, notably the A Single-cell Morphological Dataset of Leukocytes from AML Patients and Non-malignant Controls (AML-Cytomorphology_LMU) dataset [85], recognized as one of the foremost datasets. The following describes the preparatory work carried out while waiting for the availability of TAMIS data. This work focused on identifying the inherent limitations and challenges described in the contemporary scientific literature related to the classification of white blood cells. The study proposed a method to overcome one of the main limitations: the data imbalance within the classes. The methodology involves data preparation and GAN architecture with a fine-tuning strategy to achieve high-quality synthetic images in small-size classes. The work presented in this chapter greatly influenced the work of the chapter 4 and generated several ideas that were subsequently reused.

A.1 AML Cytomorphology LMU dataset

The AML Cytomorphology LMU dataset [85] was published in 2019 to showcase its capabilities in classifying white blood cells [86]. The authors noted a lack of effective automatic cytomorphological examination devices, which classify white blood cells and diagnose blood-related pathologies like acute myeloid leukemia (AML). They attributed this to the inadequacy of available databases in cytomorphology, which are often limited in terms of patient numbers or images. Consequently, the generalizability and reliability of the models developed are limited. To address this issue, the authors digitized a database containing over 18,000 images of white blood cells divided into 15 classes, making it one of the largest open databases of its kind.

A.1.1 Image Acquisition and images annotation

The dataset comprises 18,365 segmented white blood cell images ($\mathbf{400} \times \mathbf{400} \times \mathbf{3}$) annotated by hematology experts. These images were acquired from blood smears taken from 100 patients diagnosed with acute myeloid leukemia (AML) at the University Hospital of Munich between 2014 and 2017, as well as from 100 patients without signs of hematological malignancy. This dataset was ob-

tained using an M8 digital scanner (Precipoint GmbH, Freising, Germany) with an optical magnification of **100**× with oil immersion (with a resolution of 14.14 pixels per micrometer). This device uses a Köhler illumination.
The scanning process involved selecting an area of interest within the monolayer region of the blood smear, approximately 20 mm^2 in size, and scanning it at high magnification. The resulting digitized data consisted of multiresolution pyramidal images, each approximately 1 GB. These images were then annotated by trained examiners experienced in routine cytomorphological diagnostics at Munich University Hospital. Annotation was carried out single-cell, with approximately 100 cells differentiated in each smear.
The dataset was labeled according to 15 classes corresponding to cells normally present in the blood: eosinophils, lymphocytes, monocytes, basophils, neutrophils, and pathological cells indicative of conditions such as myeloblasts. Additionally, morphological classes containing fewer than 10 images were merged with neighboring classes to ensure adequate training samples. Table A.1 shows the distribution of images in the different classes: the distribution of cells among classes is highly imbalanced. Upon examination of the figures, it can be noted that the monocytes, myeloblast, lymphocytes, and neutrophils classes have significantly higher counts than the other classes.

A.1.2 Class imbalance in AML Cytomorphology LMU Dataset

Despite its volume, the database collected presents a strong imbalance of classes, limiting the performance of the neural network models that could be trained on it. These models give deficient performances for classes with low numbers of images because they have not seen enough representatives for these classes and, therefore, cannot characterize them.
The article's authors faced this problem and proposed a class rebalancing protocol. The data of each class was increased by generating additional images from existing ones. Random transformations were applied for each image in the dataset, including rotations from 0 to 359 degrees and horizontal and vertical flips, to produce 10,000 images per class.
The classification results for the 15 classes using a ResNext [146] model with 5-fold cross-validation strategies are presented in Table A.2:
The table illustrates that the most represented classes of white blood cells, including segmented neutrophils, typical lymphocytes, and monocytes, are identified with precision and sensitivity exceeding 90%, which is encouraging.

Cell Type	Number of Cells
Neutrophil (segmented)	8,484
Neutrophil (band)	109
Lymphocyte (typical)	3,937
Lymphocyte (atypical)	11
Monocyte	1,789
Eosinophil	424
Basophil	79
Myeloblast	3,268
Promyelocyte	70
Promyelocyte (bilobed)	18
Myelocyte	42
Metamyelocyte	15
Monoblast	26
Erythroblast	78
Smudge cell	15

Table A.1: Number of cells for each cell type in the dataset.

However, the algorithm's performance may be compromised for less common classes such as basophils or promyelocytes due to the limited number of training images available, despite employing augmentation strategies. For less represented classes, such as metamyelocytes and monoblasts, the performance is lower in precision and sensitivity, and the strategy employed does not improve.

A.2 State of the art of data imbalance in white blood cells classification

The class imbalance problem is not limited to the previously studied database. Indeed, it constitutes a fundamental challenge in classifying white blood cells, closely related to the natural distribution of these cells in humans. The white blood cell population shows a strong disparity between classes, with some classes

Cell Class	Precision (%)	Sensitivity (%)
Neutrophil (Segmented)	99.0	96.0
Neutrophil (Band)	25.0	59.0
Lymphocyte (Typical)	96.0	95.0
Lymphocyte (Atypical)	20.0	7.0
Monocyte	90.0	90.0
Eosinophil	95.0	95.0
Basophil	48.0	82.0
Myeloblast	94.0	94.0
Promyelocyte	63.0	54.0
Promyelocyte (Bilobed)	45.0	41.0
Myelocyte	46.0	43.0
Metamyelocyte	7.0	13.0
Monoblast	52.0	58.0
Erythroblast	75.0	87.0
Smudge Cell	53.0	77.0

Table A.2: Performance metrics of AML Cytomorphology LMU from the original paper [87].

being significantly more represented than others. This disparity is even more pronounced in cases of hematological pathologies, where certain classes of cells can become extremely rare, being present only in diseased individuals. Therefore, collecting representative images of these minority classes becomes particularly arduous, aggravating data imbalance and further complicating the classification task.
When data is imbalanced between classes, machine learning models may tend to favor majority classes over minority classes. This can cause the model's performance to decline in classifying less-represented classes, as it may struggle to learn to distinguish these classes from the majority classes effectively. Therefore, it is importante to consider and address class imbalance to achieve robust and accurate classification models in the medical domain, especially in white blood cell classification.

The solution has long been to exclude the problematic classes in several approaches. Several approaches [24, 12] have avoided basophils class, which are represented at less than 1% in the human blood. This is not acceptable when it comes to developing automatic diagnostic methods.
Other approaches, such as oversampling (increasing the number of instances in the minority class by randomly duplicating existing instances) and undersampling (reducing the number of instances in the majority class to match the size of the minority class), have been explored in previous work [117, 64]. While these approaches have shown some interest in increasing performance in underrepresented classes, they also have several limitations. Oversampling can often lead to overfitting on very weakly represented classes and can introduce bias. As for undersampling, it is often the cause of performance degradation on the most represented classes because it excludes an important number of this class samples, thus limiting variability.
More promising approaches have been proposed to overcome this challenge. Several works [80, 6] explored using synthetic data based on Generative Adversarial neural Networks (GANs) to compensate the lack of representent in minority classes. This type of model shows good performance. GANs allowed in the case where all classes have a deficient number of representatives to observe an improvement in classification performance by introducing synthetic GAN data [6, 44]. In other more complex approaches, a DC-GAN network coupled with a residual neural network (ResNet) has been used to deal with imbalanced data, emphasizing a new loss function. The experimental results demonstrated this model's effectiveness for classifying white blood cell images for certain underrepresented classes [80]. However, classes presented my few examples, such as basophils, were excluded from this study because the number of samples does not allow learning a GAN.
Other more successful works propose augmentation methods with a combination of GAN models. An approach called SyntheticCellGAN (SCG) [15] was designed with two GANs to automatically generate artificial images of white blood cells with realistic morphological characteristics. The generated synthetic images significantly improved performance and reduced the impact of underrepresented classes. However, this approach combines different literature datasets to provide sufficient samples to train the GANs. Its implementation is limited by the time required to collect the actual images and the limited number of samples for some rare cell classes.
In conclusion, this literature review shows that synthetic data generation meth-

ods utilizing GANs represent a promising approach for rebalancing imbalanced classes. However, it is worth noting that these methods often require substantial amounts of data to learn the underlying distributions effectively. Consequently, there is a clear need to develop methodologies capable of learning to generate classes, especially when the number of samples for certain classes is exceedingly low.
The methodology proposed in this chapter introduces an approach for generating synthetic images to address the under-representation of certain classes, mainly focusing on high-quality outputs. Its innovative aspect lies in its adaptive transfer learning process, which enables the model to effectively generate images for moderately under-represented classes and tackle extreme imbalance cases, even with limited samples (less than 50 representatives).

A.3 Methodology

In this study, we focus on classifying white blood cells in the context of a healthy patient. This is done to position the work in the same framework as TAMIS project. We consider five classes of white blood cells: Basophils, Eosinophils, Monocytes, Lymphocytes, and Neutrophils. For this purpose, a subset was extracted from the AML Cytomorphology LMU dataset. The distribution of the cells within the five classes is listed in Table A.3.

Cell Type	**Number**
Basophils	79
Eosinophils	424
Lymphocytes	3948
Monocytes	1789
Neutrophils	8484

Table A.3: The distribution of the five classes of white blood cells extracted from the AML Cytomorphology LMU dataset.

In Table A.3, it is apparent that the distribution of cells among the five classes is highly imbalanced. Lymphocytes, Monocytes, and Neutrophils have significantly higher counts compared to the classes of Basophils and Eosinophils.

Our aim is to design a GAN-based strategy to increase the number of samples in the under-represented classes. For this, an approach based on training a GAN for each minority class is proposed.
After evaluating various architectures, we implemented StyleGAN2 [57]. This choice was driven by the fact that other architectures yielded images with lower resolutions and lacked the desired level of detail.

A.3.1 StyleGAN architecture

StyleGAN2 [57] is a generative model developed by NVIDIA in 2019, representing the second version of the original StyleGAN model, aiming to produce even more realistic and detailed images. Similarly to usual GANs, it consists of a generator and a discriminator. The general functioning of GANs was introduced in Chapter 4
Its particularity lies on several principles introduced within the generator that enable significant improvements in the quality and variability of the produced images. The architecture of this generator is illustrated in Figure A.1.

The input layer of the generator receives a random noise vector. This noise vector is the seed from which the architecture generates the image. The noise vector passes through a mapping network. This network consists of multiple fully connected layers. It projects the noise vector into a more structured latent space. It also aims to increase the variability of the output images by increasing the differences between the input white noise vectors by passing through numerous layers. This transformation yields a style vector that guides image generation. Once the style vector is obtained, the generator creates the image using a synthesis network. This network is designed to progressively generate the image, starting from low-detailed images and gradually adding finer details. Each resolution level in the synthesis network comprises synthesis and normalization blocks. The synthesis block employs convolutional layers to generate image features at a specific resolution and upsampling layers to increase the spatial resolution of the image progressively. Modulation and demodulation layers are also introduced in the original StyleGAN2 architecture. Modulation scales each input feature map of a convolutional layer based on the incoming style vector, achieved by multiplying the convolution weights by the corresponding scale factors from the style vector. Demodulation adjusts the output feature maps

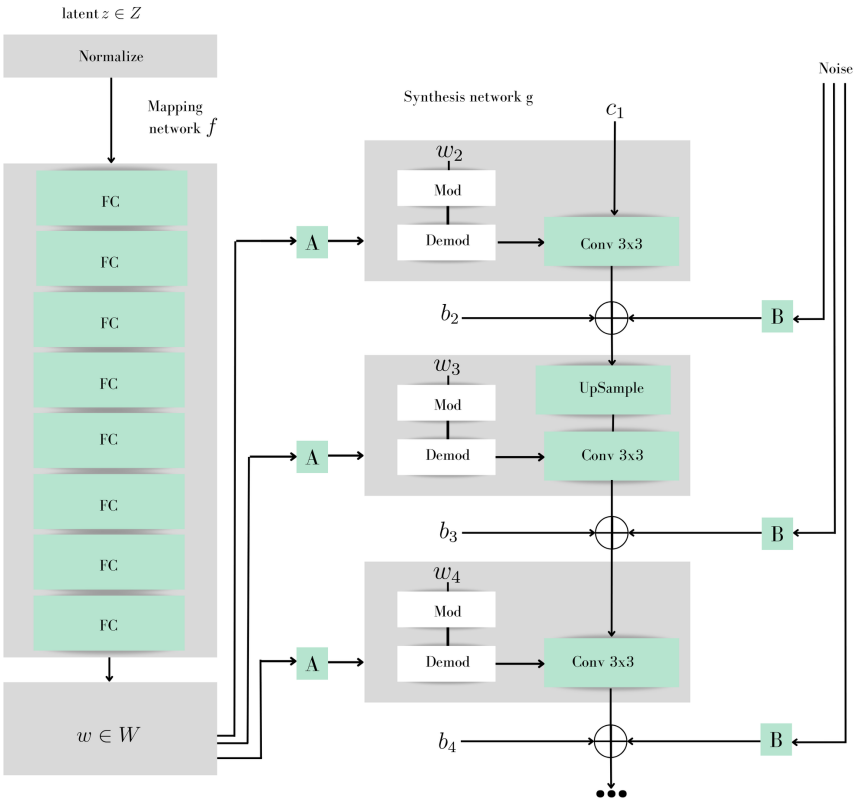


Figure A.1: Illustration of the styleGAN2 generator.

of the convolutional layer by normalizing them according to their L2 norms, ensuring a unit standard deviation for the output activations. White noise is also added at each generator stage to increase diversity and improve learning. Finally, the output layer generates the final image using the outputs of the last block using a convolution to produce the desired number of channels.
This architecture has been adapted to allow $\mathbf{400 \times 400}$ images to be generated by adjusting the convolution filter size in the architecture.
For the discriminator, we used ResNet [45] architecture.
The implemented StyleGAN2 was trained using a combination of several cost functions. The generator is based on the usual adversarial loss and perceptual loss, a mean square error applied to the embeddings of the images extracted using a pre-trained model (Here inception V3). This loss was used to improve the details of the images produced. A regularization loss on the weights was applied to penalize large weights in the network and allow better convergence. This loss introduces a penalization coefficient γ that helps control the strength

of the penalization. The greater the γ coefficient, the stronger the regularization. The discriminator is trained using adversarial loss.

A.4 GAN Implementation and Experimentation

GAN data augmentation was performed for basophils and eosinophils following different strategies, as explained below.

A.4.1 Metrics

The trained GAN models are evaluated using a qualitative criterion through visual observation of generated images and a blind evaluation by a hematologist. The FID score metric provides the quantitative evaluation. The Frechet Inception Distance (FID) is defined by using the Inception V3 model pre-trained on the ImageNet dataset to produce an embedding for each image, then calculated as :

$$\text{FID} = \|\mu_r - \mu_g\|_2^2 + \text{Tr}(\mathbf{\Sigma}_r + \mathbf{\Sigma}_g - \mathbf{2}(\mathbf{\Sigma}_r\mathbf{\Sigma}_g)^{\mathbf{1/2}})$$

Where g and r are the real and synthetic embeddings assumed to be two multivariate normal distributions, μ_g and μ_r are the magnitudes of the vector g and r. Tr is the trace of the matrix, $\mathbf{\Sigma}_g$ and $\mathbf{\Sigma}_r$ are the covariance matrix of the vectors. Low FID scores indicate that the two groups of images are similar or have similar statistics.

A.4.2 Data pre-processing and hyperparameters tuning

The number of images used for training the StyleGAN2 is 351 for the eosinophils and 64 for the basophils. Including three 90-degree rotations and two mirrors in addition to the original image allows the number of images to be multiplied by a factor, bringing the sample's number to 384 for basophils and 2106 for eosinophils.
Data normalization was applied to all images to ensure the pixel values were between -1 and 1 (i.e., 0 to 255 before normalization).
All the models were trained to generate **500000** images and evaluated each **20000** image. The best model was chosen using the lowest FID.
Two hyperparameters were important for ensuring the convergence of the models. The first is the batch-size, and the second is the γ coefficient. A batch-size

of 4 was chosen for training all the GAN because higher values lead to model divergence. The regularization coefficient γ was initially set to 8 to impose higher regularization, stabilizing the gradient values during the early stages of model training. As the model progressed and became more stable, this coefficient gradually decreased by dividing it several times by 2. Lowering γ allows the model to capture finer details more easily, facilitating learning of intricate features as the training progresses.

A.4.3 Eosinophils generation

For eosinophils, the StyleGAN was trained from scratch. The best model was chosen based on its FID score, achieving a score of 16.63 for the FID indicator. The generated images are compared to the real basophils images in figure A.2.

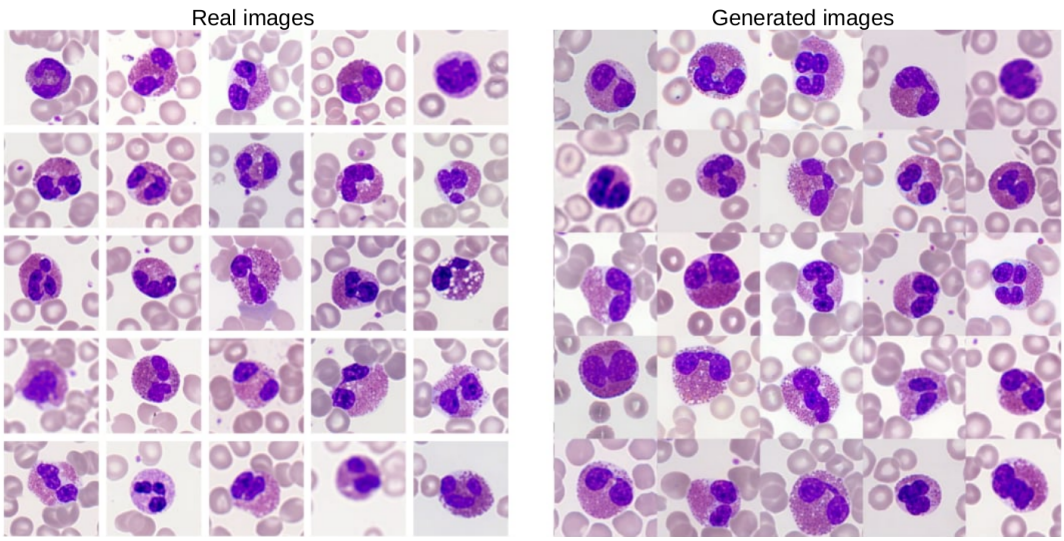


Figure A.2: Comparaison of generated and real eosinophils images.

it can be seen that the model produced high-quality images. The resemblance to real eosinophils is striking in many instances, indicating high fidelity in the generated samples. The generated eosinophils exhibit bilobed shapes with prominent pink to orange granules, often obscuring the nucleus. The model manages to avoid generating a lot of out-of-focus images even if some images out of focus exist in the training sample. It thus responds to the natural description of this class in humans.

A.4.4 Basophils generation

Unlike the eosinophils class, the basophils dataset encompasses only 64 images for the training. In our first attempt to train basophils from scratch, the FID score was stacked around 117, and the generated images were visually very bad.
To solve this problem, the similarity between basophils and neutrophils was explored in a fine-tuning process. Indeed, basophils and neutrophils share some morphological characteristics, such as multi-lobed nuclei and granular cytoplasm. First, we trained a neutrophil GAN using 80% of the available images of these classes. After achieving a satisfactory FID score of 6.02, we employed fine-tuning approaches to adjust the first model's weights to generate basophils. The basophils model was chosen based on its FID score, achieving a score of 29.58 for the FID indicator. The generated images are compared to the real basophils images in figure A.3 The model generated images of medium

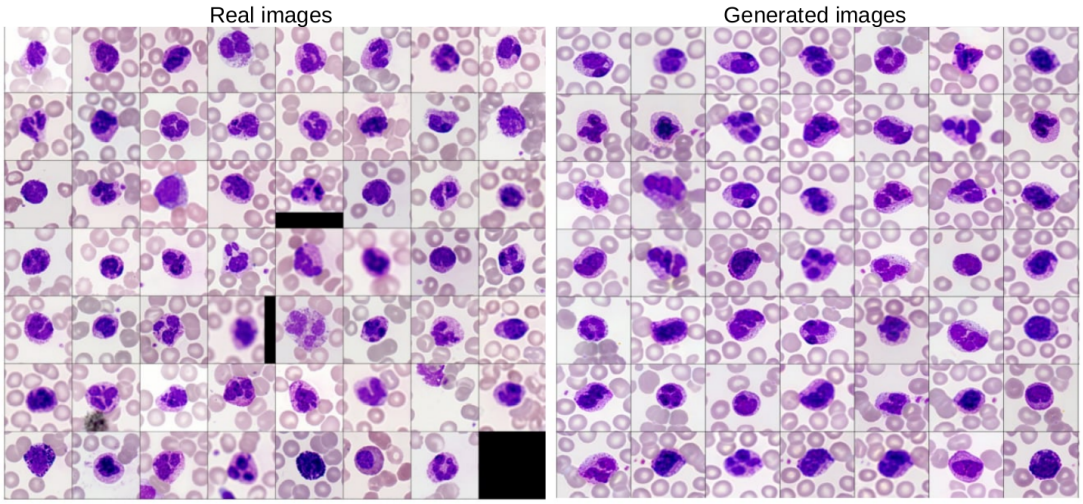


Figure A.3: Comparaison of generated and rael basophils images.

quality for basophils. The generated basophils display characteristic features such as irregularly shaped nuclei and abundant dark blue to purple granules in their cytoplasm. The model seems satisfactory when compared to the real images. However, the generated images have numerous out-of-focus images, several images without blur, and sometimes insufficient grain resolution. These imperfections are due to the limited number of images in the training set. Additionally, it is important to note that several real images exhibit out-of-focus

and blur. The quality of the GAN largely depends on the quality of the training data, which becomes even more critical when the number of representatives is reduced. Despite the rather medium visual quality of the synthetic basophil images, we performed a classification task using this augmented data to see how much it could improve the classification performance.

A.5 Classification using GAN augmented images

To study the impact of enriching the eosinophil and basophil classes, classification performances before and after this enrichment were compared. A classification into five classes using the MobileNetV1 [51] model was performed in two configurations.
For the first model, only real data was used, a classic learning approach where 80% of the data from each class was used for training and 20% for testing was used.
For the first model, only real data was used, in which 80% of the data from each class was used for training, and the remaining 20% was used for testing. For the second approach, the model was trained using synthetic images to complete the underrepresented classes. This model was then tested on the same dataset as the initial model. The model's training was carried out with 1000 examples of each class per epoch. When the class is sufficiently represented, the 1000 examples are drawn randomly at each epoch to cover all the diversity of the class (dynamic sampling). For underrepresented classes (basophils and eosinophils), a sufficient number of samples is generated using StyleGAN and added to the real data to reach 1000 images.
Figure A.4 represents the two confusion matrices before the enrichment (left) and after the enrichment (right). The confusion matrix of the model trained on the imbalanced dataset shows varying performance. Among the classes benefiting from good performance, neutrophils and lymphocytes stand out. However, the model performs poorly for the least represented class. Basophils were correctly classified in only 5 of 15 cases, corresponding to an accuracy rate of 33.3%. This is certainly due to the underrepresentation of this class in the dataset. The results in the second matrix demonstrate that improvement is evident in the case of the basophil class, where an increase in accuracy of 50% was observed. It is also remarkable that the introduction of GAN did not degrade the performance of most classes.

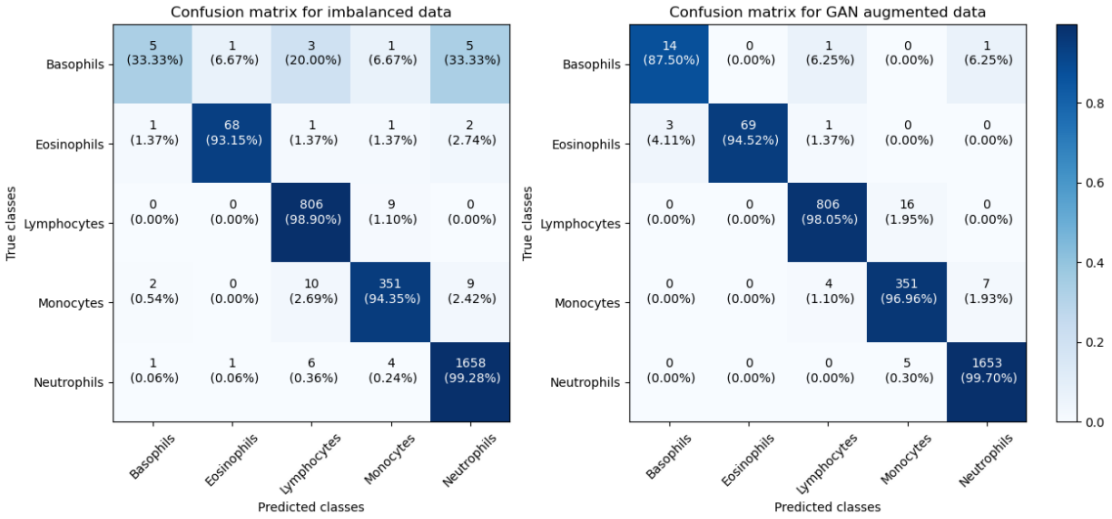


Figure A.4: Comparaison of confusion matrixes before and after GAN data augmentation.

A.6 Conclusion

Incorporating synthetic images into the dataset significantly improves classification performance, especially for minority classes. This improvement demonstrates the effectiveness of data augmentation using GANs to compensate for class imbalance and improve the generalization of classification models without degrading performance in non-augmented classes. The accuracy of the basophil class increased significantly from 33.3% before data augmentation, to 87% after adding the synthetic images. Therefore, the proposal of the fine-tuning strategy proved effective for learning to generate images of classes with less than 100 samples by exploiting the similarities between classes.

However, it is essential to emphasize that the success of this approach depends on the quality of the synthetic images generated by the GAN. The proximity of synthetic data distribution to real data is important for significantly improving classification performance. It was noted during the experiments that images generated with a high FID do not allow an improvement in classification to be observed, or even worse when the image generation has many defects and is far from the distribution of the real data, severe performance degradation can be observed.

Finally, the approach must be extended to other pathological classes in ULM dataset, which have been excluded to simplify. It should be noted that, unfortu-

nately, some classes only have around ten images, which could challenge learning the GAN even with the fine-tuning strategy. In this case, other approaches should be investigated, notably less data-consuming as diffusion models.

Appendix B

Publications

Conferences and journals

H. Hassini, B. Dorizzi, M. Thellier, J. Klossa, and Y. Gottesman, *Investigating the joint amplitude and phase imaging of stained samples in automatic diagnosis*, *Sensors*, volume 23, 2023.

H. Hassini, B. Dorizzi, J. Klossa, and Y. Gottesman, *Optimization of Fourier ptychographic microscope using phase images for malaria detection*, in *Proc. SPIE Photonics Europe*, 12996-10 (Presented at SPIE Unconventional Optical Imaging: April 8, 2024; Published: soon)

H. Hassini, L. Bouchama, M. Sextius, J. Klossa, C. Fleury, F. Baran-Marszak, F. Delhommeau, B. Dorizzi, Y. Gottesman, V. Leymarie, *The TAMIS project: Hematology Cytology and Artificial Intelligence: Revolution is in the Light Bulb*, *Proceedings of the International Society for Laboratory Hematology Conference*, (Accepted)

Presentations

H. Hassini, A. Baroni, B. Dorizzi, Y. Gottesman, J. Klossa, and M. Thellier, *Contribution of deep networks with complex extractor on augmented microscopy data for the detection of malaria*, in *Journées Imagerie Optique Non Conventionnelle - 17ème édition, GDR ISIS*, Paris, France, May 2022.

H. Hassini, A. Baroni, B. Dorizzi, Y. Gottesman, J. Klossa, and M. Thellier,

Contribution of complex extractor deep networks on augmented microscopy data for malaria diagnosis, in *2nd Biomedical Engineering (BME) Symposium*, Paris, France, 2021.

H. Hassini, A. Baroni, B. Dorizzi, Y. Gottesman, J. Klossa, and M. Thellier, *Contribution of complex extractor deep networks on augmented microscopy data for malaria diagnosis*, in *7ème édition de la journée sur la « Co-conception : capteurs hybrides et algorithmes pour des systèmes innovants, GDR Ondes*, Paris, France, May 2021.

H. Hassini, L. Bouchama, M. Sextius, J. Klossa, C. Fleury, F. Baran-Marszak, F. Delhommeau, B. Dorizzi, Y. Gottesman, V. Leymarie, *The TAMIS project: Hematology Cytology and Artificial Intelligence: Revolution is in the Light Bulb*, in *SFH - 43ème Congrès Annuel de la Société Française d'Hématologie*, Paris, France, March 2024

Posters

H. Hassini, L. Bouchama, M. Sextius, J. Klossa, C. Fleury, F. Baran-Marszak, F. Delhommeau, B. Dorizzi, Y. Gottesman, V. Leymarie, *The TAMIS project: Hematology Cytology and Artificial Intelligence: Revolution is in the Light Bulb*, in *SFH - 43ème Congrès Annuel de la Société Française d'Hématologie*, Paris, France, March 2024

DOCTORAL
SCHOOL

Titre: Analyse automatique des images des frottis sanguins : intérêt de la phase en microscopie ptychographique de Fourier

Mots clés: Microscopie ptychographique de Fourier, Réseaux de neurones profonds, Réseaux antagonistes génératifs, Hématophatologie

Résumé: La pathologie numérique constitue aujourd'hui un outil fondamental pour le diagnostic médical, exploitant les avancées technologiques en matière de numérisation pour transformer les échantillons biologiques en données numériques, facilitant ainsi leur visualisation et leur analyse. Cependant, ces méthodes, souvent basées sur la microscopie conventionnelle, rencontrent des limitations qui entravent parfois leur efficacité.

Dans ce contexte, des méthodes d'imagerie non conventionnelles telles que la microscopie ptychographique de Fourier (FPM) offrent des perspectives prometteuses pour surmonter ces limitations. En effet, la FPM offre un accès à la phase en complément de l'intensité et permet d'examiner un large champ de vision à haute résolution à un coût de conception raisonnable.

Cette thèse explore le potentiel de la microscopie ptychographique de Fourier dans l'analyse des frottis sanguins minces. Plusieurs résultats ont été obtenus grâce à une approche multidisciplinaire intégrant l'apprentissage en profondeur et la microscopie.

Nous nous concentrons d'abord sur le problème limité de la détection des parasites pour le diagnostic du paludisme. L'exploitation conjointe de l'intensité et de la phase permet d'améliorer les performances d'un détecteur de réseau neuronal profond. À cette fin, un CNN à valeurs complexes a été introduit dans l'architecture Faster-RCNN pour une extraction efficace des caractéristiques.

Ensuite, nous examinons une application plus complexe, à savoir la classification des globules blancs, où les avantages de l'exploitation conjointe de l'intensité et de la phase ont également été confirmés. Nous nous intéressons également au problème du déséquilibre des classes rencontré dans cette tâche, nous proposons un nouveau modèle GAN informé par la physique dédié à la génération d'images d'intensité et de phase. Ce modèle évite le problème de mode collapse rencontré avec l'implémentation habituelle des GAN.

Enfin, nous considérons l'optimisation de la conception du microscope FPM. À cette fin, nous explorons des stratégies combinant simulations, réseaux neuronaux et modélisation de la formation d'images. Nous démontrons que la FPM peut utiliser des résolutions faibles sans compromettre significativement les performances.

Cette thèse souligne l'intérêt d'adapter l'apprentissage automatique en lien avec les principes de la microscopie et met en évidence le potentiel de la microscopie ptychographique de Fourier pour les futurs systèmes de diagnostic automatisés.

Title: Automatic analysis of blood smears images: contribution of phase modality in Fourier Ptychographic Microscopy

Keywords: Fourier ptychographic microscopy, Deep neural networks, Génaratifs adversarial models, Hematopathology

Abstract: Digital pathology presents today a fundamental tool for medical diagnosis, exploiting technological advances in digitalization to transform biological samples into digital data, thus facilitating their visualization and analysis. However, these methods, often based on conventional microscopy, encounter limitations that sometimes hinder their effectiveness.
From this perspective, unconventional imaging methods such as Fourier ptychographic microscopy offer promising prospects for overcoming these limitations. Indeed, FPM offers access to the phase in complement of the intensity and allows examining a large Field of View at a high resolution at a reasonable design cost.
This thesis explores Fourier ptychographic microscopy (FPM) 's potential in thin blood smear analysis. Several results have been obtained thanks to a multidisciplinary approach integrating deep learning and microscopy. We have first focused our attention on the problem of limited complexity of parasite detection for malaria diagnosis. The joint exploitation of intensity and phase is shown to improve the performance of a deep network detector. To this end, a complex-valued CNN has been introduced in Faster-RCNN architecture for efficient feature extraction.
Secondly, we have considered a more complex application, namely the classification of white blood cells, where the benefits of joint exploitation of intensity and phase were also confirmed. Furthermore, to reduce the imbalance of classes encountered in this task, we propose a novel physics-informed GAN model dedicated to generating intensity and phase images. This model avoids the mode collapse problem faced with usual GAN implementation.
Finally, we have considered optimizing the FPM microscope design. To this end, we explore strategies combining simulations, neural networks, and image formation modeling. We demonstrate that FPM can use low resolutions without significantly compromising performance.
This thesis underscores the interest in tailoring machine learning in connection to microscopy principles and highlights the potential of Fourier ptychographic microscopy for future automated diagnosis systems.

Institut Polytechnique de Paris
91120 Palaiseau, France