



HAL
open science

Détection et identification de *Plasmodium falciparum* sur des images microscopiques

Idir Aniss Acherar

► **To cite this version:**

Idir Aniss Acherar. Détection et identification de *Plasmodium falciparum* sur des images microscopiques. Médecine humaine et pathologie. Sorbonne Université, 2024. Français. NNT : 2024SORUS019 . tel-04698420

HAL Id: tel-04698420

<https://theses.hal.science/tel-04698420>

Submitted on 16 Sep 2024

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

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



Sorbonne Université

École doctorale Pierre Louis de santé publique
Institut Pierre-Louis Epidémiologie et Santé Publique

Détection et identification de *Plasmodium falciparum* sur des images microscopiques.

Par Idir Aniss ACHERAR

Thèse de doctorat en science des données

Dirigée par le Pr. Renaud PIARROUX et encadrée par le Pr. Xavier TANNIER

Présentée et soutenue publiquement le 12 mars 2024

Devant un jury présidé par Mme Nadjia KACHENOURA, et composé de :

M.	Renaud	PIARROUX	Directeur
M.	Xavier	TANNIER	Encadrant
Mme	Sandrine	HOUZÉ	Rapporteuse
M.	Camille	KURTZ	Rapporteur
Mme	Isabelle	VILLENA	Examinatrice

Résumé

L'utilisation du microscope dans le diagnostic du paludisme est depuis longtemps un outil essentiel pour les professionnels de la santé. Cependant, il existe des défis inhérents à cette méthode, notamment en termes de subjectivité et de variabilité des résultats dus à des facteurs tels que la qualité des frottis sanguins, les variations de coloration et la présence d'artefacts. De plus, l'interprétation des images de frottis sanguins peut varier en fonction de l'expertise de l'opérateur, ce qui souligne la nécessité d'explorer des approches plus objectives et reproductibles pour le diagnostic du paludisme.

L'application future de l'apprentissage automatique et profond dans le diagnostic assisté par ordinateur du paludisme soulève une limitation majeure : ces modèles sont généralement validés sur des données de recherche, qui diffèrent considérablement des données cliniques de routine. Pour comprendre véritablement les performances de ces outils dans des contextes réels, il est impératif de les valider sur de vastes ensembles de données cliniques.

La validation sur des ensembles de données cliniques revêt une importance capitale pour évaluer la capacité des modèles d'apprentissage automatique et profond à relever les défis spécifiques de la pratique clinique quotidienne. Les données issues de la routine peuvent varier considérablement en termes de qualité des frottis, d'images, de variations de coloration et d'artefacts, ce qui peut avoir une incidence sur la performance des modèles. Il est donc essentiel de tester ces modèles sur des ensembles de données cliniques réelles afin d'évaluer leur adaptabilité et leur fiabilité dans des situations concrètes.

Dans notre première étude, nous avons créé une base de données d'images annotées, comprenant plus de 60 000 images, pour identifier les globules rouges parasités par *Plasmodium falciparum* et les composants sanguins non infectés. L'objectif était de garantir une représentation fidèle des conditions diagnostiques afin d'améliorer les performances des systèmes de classification. En contrôlant la diversité des images, nous avons pu améliorer la capacité des modèles à généraliser sur de nouveaux échantillons de patients. À l'aide d'une approche de validation croisée, nous avons évalué différentes architectures de modèles en utilisant notre base de données ainsi que celle mise en ligne par le National Institute of Health (NIH). Nous avons ensuite analysé la capacité des modèles entraînés sur chaque base de données à généraliser et à fournir des prédictions précises sur un ensemble de données cliniques indépendant comprenant des images de frottis sanguins issues de 200 patients différents.

Dans notre deuxième étude, nous nous sommes concentrés sur la mesure automatique de la parasitémie, un facteur crucial dans le diagnostic du paludisme, en particulier pour les cas causés par *Plasmodium falciparum*. Nous avons proposé une méthode automatisée pour estimer la parasitémie en analysant un nombre prédéterminé de champs de frottis sanguins minces. Cette approche automatique standardise la mesure et réduit la subjectivité liée à l'interprétation manuelle. Parallèlement, nous avons identifié les limites des méthodes manuelles traditionnellement utilisées en routine pour mesurer la parasitémie sur les frottis sanguins minces, en raison de la variabilité significative dans l'interprétation et l'analyse microscopique, pouvant entraîner des erreurs et une fiabilité réduite dans le diagnostic.

Enfin, nous avons comparé les performances de notre système proposé sur un ensemble de 102 patients. En évaluant sa corrélation avec deux autres méthodes, manuelles et semi-automatiques, nous avons démontré la validité et la fiabilité de notre approche automatique dans l'estimation de la parasitémie.

Mots-clés : Paludisme, Apprentissage Automatique, Apprentissage Profond, Diagnostic Automatique, Microscopie, *Plasmodium falciparum*, Parasitémie.

Abstract

The use of microscopy in malaria diagnosis has long been an essential tool for healthcare professionals. However, this method comes with inherent challenges, particularly in terms of subjectivity and variability in results due to factors such as the quality of blood smears, variations in staining, and the presence of artifacts. Moreover, the interpretation of blood smear images can vary depending on the operator's expertise, highlighting the need to explore more objective and reproducible approaches for malaria diagnosis.

The future application of machine learning and deep learning in computer-aided malaria diagnosis poses a significant limitation : these models are typically validated on research data, which significantly differs from routine clinical data. To truly understand the performance of these tools in real-world contexts, it is imperative to validate them on large clinical datasets.

Validation on clinical datasets is of paramount importance to assess the ability of machine learning and deep learning models to address the specific challenges of daily clinical practice. Routine clinical data can vary significantly in terms of smear quality, image characteristics, staining variations, and artifacts, all of which can impact model performance. Therefore, it is essential to test these models on real clinical datasets to evaluate their adaptability and reliability in practical scenarios.

In our first study, we created an annotated image database comprising over 60,000 images to identify *Plasmodium falciparum* infected red blood cells and uninfected blood components. The goal was to ensure a faithful representation of diagnostic conditions to enhance the performance of classification systems. By controlling image diversity, we improved the models' ability to generalize to new patient samples. Using a cross-validation approach, we evaluated different model architectures using our database and one made available by the NIH. Subsequently, we assessed the models' capacity to generalize and provide accurate predictions on an independent clinical dataset comprising blood smear images from 200 different patients.

In our second study, we focused on the automated measurement of parasitemia, a crucial factor in malaria diagnosis, particularly for cases caused by *Plasmodium falciparum*. We proposed an automated method to estimate parasitemia by analyzing a predetermined number of thin blood smear fields. This automated approach standardizes measurements and reduces the subjectivity associated with manual interpretation. Concurrently, we identified the limitations of traditional manual methods used in routine practice to measure parasitemia on thin blood smears due to significant variability in microscopic interpretation and analysis, potentially leading to errors and reduced diagnostic reliability.

Finally, we compared the performance of our proposed system on a dataset of 102 patients. By assessing its correlation with two other methods, manual and semi-automated, we demonstrated the validity and reliability of our automated approach in parasitemia estimation.

Keywords : Malaria, Machine Learning, Deep Learning, Automated Diagnosis, Microscopy, *Plasmodium falciparum*, Parasitemia.

Remerciements

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué à la réalisation de cette thèse. Leur soutien inestimable, leurs conseils judicieux et leur engagement ont joué un rôle déterminant dans l'aboutissement de ce travail.

Je tiens tout d'abord à exprimer ma profonde gratitude envers mes encadrants de thèse, les professeurs Renaud Piarroux et Xavier Tannier, pour leur soutien constant tout au long de ce parcours. Merci d'avoir accepté de me guider et de m'avoir accordé votre confiance pour mener à bien ces travaux. C'est grâce à ces échanges, aussi fréquents que précieux, ainsi qu'à vos conseils et votre disponibilité, que j'ai pu concrétiser ce travail.

Je remercie les membres de mon jury de thèse. Merci à Sandrine Houzé, Isabelle Villena, Nadja Kachenoura et enfin à Camille Kurtz pour m'avoir guidé sur le chemin de la recherche.

Mes remerciements s'étendent également à tous ceux qui ont partagé leurs connaissances et leurs perspectives, enrichissant ainsi le contenu de cette thèse. Je tiens particulièrement à remercier Marc et Ilhame, ainsi que l'ensemble du personnel du laboratoire de parasitologie-mycologie de la Pitié-Salpêtrière.

Je suis reconnaissant envers toutes les institutions, Sorbonne université, le laboratoire LIMICS ainsi que le SCAI et ses pilliers, Xavier Fresquet et Gérard Biau qui ont contribué à faire de ce projet une réalité.

Enfin, mes remerciements vont à ma famille pour leur soutien constant et leur encouragement tout au long de ce parcours académique.

Enfin et surtout, merci à Natali, de m'avoir accompagné dans cette dernière ligne droite, d'avoir été mon oreille attentive à mes doutes et interrogations, et de m'avoir soutenu dans les moments les plus difficiles.

Table des matières

Liste des figures	5
Liste des tableaux	7
1 Introduction	11
1 Le Paludisme : un problème de santé publique	12
1.1 Contexte de la maladie	12
1.2 Importance du diagnostic précoce et précis du paludisme	13
2 Techniques de diagnostic du paludisme	14
2.1 La microscopie	14
2.2 Les indicateurs de performance d'une technique diagnostique	15
2.3 Les tests de diagnostic rapide (TDR)	16
2.4 Les techniques moléculaires	18
3 Les autres techniques diagnostiques	19
3.1 La sérologie	19
3.2 La cytométrie en flux	19
4 Les défis liés aux techniques diagnostiques	20
4.1 Disponibilité et accessibilité des techniques diagnostiques dans les régions endémiques	20
4.2 Les défis liés à la surveillance épidémiologique	21
5 Besoins et perspectives pour de nouvelles approches de diagnostic	21
5.1 L'intégration de l'intelligence artificielle dans le diagnostic du paludisme	22
6 Contributions scientifiques	24
6.1 Résumés des travaux publiés dans le domaine du diagnostic du paludisme	24
6.2 Autres contributions en co-auteur	25
2 Etat de l'art	26
1 Classification automatique et diagnostic du paludisme	26
1.1 Introduction	26
1.2 Classification automatique des images pour le diagnostic du paludisme : ap- proches traditionnelles et traitement des images	26

1.3	Classification automatique d'images : Apprentissage profond pour l'identification de stades parasitaires de <i>Plasmodiums</i>	27
1.4	Focus sur le serial-killer : <i>Plasmodium falciparum</i> et mesure automatique de la parasitémie	29
1.5	Quelles solutions sont disponibles? Applications mobiles et systèmes «End-to-End» pour le diagnostic automatique du paludisme	31
3	Identification de globules rouges parasités par <i>Plasmodium falciparum</i> sur des images de frottis sanguins minces	36
1	Contexte et Objectifs	36
2	Matériel et Méthodes	37
2.1	Description du jeu de données du NIH	37
2.2	Construction d'un jeu de données CNR représentatif de la réalité diagnostique	37
2.3	Validation croisée et classification des images du jeu de données CNR	41
2.4	Evaluation rétrospective : Généralisation des modèles sur le jeu de validation .	41
3	Résultats	42
3.1	Annotation en deux étapes des images composants sanguins : FL-CNN et validation des experts	42
3.2	Validation croisée : évaluation au niveau cellulaire	42
3.3	Validation patient : évaluation au niveau patient	46
4	Mesure automatique de la parasitémie : un critère biologique important	52
1	Contexte et Objectifs	52
2	Matériel et Méthodes	53
2.1	Sélection et description des patients	53
2.2	Mesure de la parasitémie	54
2.3	Les techniques manuelles : sources de variabilité pour la mesure parasitémie .	56
2.4	La répétabilité de la parasitémie avec MALARIS : l'impact du nombre de champs	57
2.5	Comparaison de méthodes pour la mesure de la parasitémie : MALARIS Vs. cellule de Miller Vs. cytométrie en flux	58
3	Résultats	59
3.1	Les techniques manuelles : sources de variabilité pour la mesure de la parasitémie	59
3.2	La répétabilité du système MALARIS : l'impact du nombre de champs	59
3.3	Comparaison de méthodes pour l'estimation de la parasitémie sur 102 nouveaux patients	61
5	Discussion	66
1	Discussion générale	66
6	Conclusions et Perspectives	72
1	Synthèse des contributions	72
2	Perspectives de recherche à développer	73
2.1	Amélioration de la détection des cellules	73
2.2	Amélioration du jeu de données CNR : qualité et diversité	78
2.3	Optimisation des performances et déploiement du système	79
	Références	81

Annexes	87
A Annexe A	89
B Annexe B	103

Table des figures

1.1	Les espèces de <i>Plasmodium</i> qui infectent les humains [Poostchi et al., 2018b].	13
1.2	Deux différents TDR positifs pour <i>Plasmodium falciparum</i>	17
1.3	Architecture simple d'un CNN pour la classification binaire d'images de globules rouges. Le CNN est composé de 2 couches de convolutions et 2 couches de pooling, une sortie à 2 classes : infecté ou non infecté.	22
1.4	Représentation d'une image RVB : (a) Globule rouge infecté par <i>Plasmodium falciparum</i> . (b) Image pixelisée avec une taille de (MxN), où M est le nombre de colonnes et N le nombre de lignes. (c) La matrice des valeurs d'intensité correspondantes définies par trois valeurs, rouge, vert et bleu (RVB), sur une plage de 256 valeurs possibles : un mélange de ces trois valeurs produit des intensités dans l'image.	23
1.5	Activations d'un CNN à trois couches de convolution pour l'identification d'un globule rouge parasité par <i>Plasmodium falciparum</i>	24
2.1	Approches et étapes clés d'un système automatisé de classification d'images pour le diagnostic du paludisme	28
3.1	Diagramme de construction de la base d'images CNR	38
3.2	La diversité biologique des composants sanguins inclus dans le jeu de données CNR.	39
3.3	Architecture du modèle FL-CNN à 5 couches avec K = 3 classes pour la classification en tant que modèle d'annotation.	40
3.4	Boîtes à moustaches des distributions de performance pour la classification de 200 patients, comprenant 422 015 composants sanguins non infectés et 3 574 globules rouges infectés. Les triangles jaunes représentent la valeur moyenne de chaque métrique.	47
3.5	Les matrices de confusion pour la classification de 3 574 hématies infectées par <i>P. falciparum</i> et de 422 015 composants sanguins non infectés provenant de 200 patients	49
3.6	Diagramme en barres de la distribution des faux positifs pour 150 patients négatifs.	50

3.7	Deux images prédites par VGG-19-CNR et VGG-19-NIH. Boîtes vertes : vrais négatifs de composants sanguins non infectés correctement classifiés par les deux modèles. Boîtes rouges : vrais positifs d'hématies infectées par <i>P. falciparum</i> correctement interprétés par les deux modèles. Boîtes oranges : faux positifs de plaquettes et d'artefacts de coloration mal interprétés par VGG-19-NIH et correctement interprétés par VGG-19-CNR comme des composants sanguins non infectés.	51
4.1	Architecture de l'étude pour l'évaluation de la parasitémie.	54
4.2	Image d'un frottis sanguin mince infecté par <i>Plasmodium falciparum</i> à un grossissement de x1 000, sans et avec la cellule de Miller. À un grossissement de x1 000, une graduation = 1 micromètre. La taille moyenne des GR est de 7 µm (de 5 à 10).	55
4.3	MALARIS	57
4.4	Écart type relatif (RSD) pour différentes techniques. Comparaison entre MALARIS, la cellule de Miller et la technique de mesure standard. Les triangles jaunes représentent la valeur moyenne de RSD pour chaque technique	60
4.5	(a) Comparaison de l'écart type relatif (RSD) parmi les 14 patients, classés par parasitémie croissante, entre deux techniques manuelles, comprenant 10 mesures sélectionnées au hasard parmi 10 microscopistes, l'écart type relatif moyen est de 26,40 % pour la méthode assistée par la cellule de Miller et de 38,22 % pour la technique de mesure standard. (b) Distribution de l'écart type relatif pour les trois méthodes, y compris l'estimation de la parasitémie par MALARIS sur 10 images de frottis sanguins (voir le tableau détaillé en annexe A). Au-dessus de chaque histogramme, la valeur de parasitémie mesurée pour chaque patient, disposée par ordre croissant de parasitémie, est affichée pour la technique correspondante.	61
4.6	Comparaison des estimations de la parasitémie et des intervalles de confiance chez 14 patients. La cellule de Miller vs. Mesure Standard (a). Malaris vs. Cellule de Miller vs. Mesure Standard (b). Chaque disque est centré sur la valeur moyenne de l'échantillon correspondant.	62
4.7	La répétabilité du comptage automatique de la parasitémie par le système MALARIS sur 14 patients, avec une parasitémie variant de 0,7 % à 4,10 % et une valeur médiane de 2,51%. La représentation du taux d'erreur standard relatif (%) (a) pour la répétabilité des différentes combinaisons K (de 1 à 5) à partir de 10 images de frottis sanguins minces et (b) comparaison avec la cellule Miller et le comptage manuel standard. La ligne en pointillés rouges représente l'intervalle d'erreur standard de ±5 %. Les triangles jaunes représentent la valeur moyenne de chaque paire de mesures.	63
4.8	Répétabilité de l'estimation automatique de la parasitémie par le système MALARIS évaluée sur 102 patients avec une parasitémie ordonnée de 0,03 % à 27,32 %, et une valeur médiane de 3,63 %. La répétabilité a été évaluée en utilisant le taux d'erreur standard relative (RSE) (%) en analysant plusieurs combinaisons de 5 images de frottis sanguins minces, pour un total de 126 paires.	64
4.9	Comparaison des graphiques de régression pour l'estimation de la parasitémie. MALARIS (axe des x) par rapport au cellule de Miller (axe des y) (a), MALARIS (axe des x) par rapport à la cytométrie en flux (axe des y) (b), et cellule de Miller (axe des x) par rapport à la cytométrie en flux (axe des y) (c). Sous les graphiques, un accent particulier est mis sur les valeurs de parasitémie comprises entre 0 et 5 %.	65

4.10	Graphiques de Bland-Altman des techniques utilisées pour l'estimation de la parasitémie. (a) Comparaison entre MALARIS et le cellule de Miller. (b) Comparaison entre MALARIS et la cytométrie en flux. (c) Comparaison entre le cellule de Miller et la cytométrie en flux. Les graphiques affichent la différence entre les mesures en fonction de la moyenne des deux mesures, la ligne rouge représentant la différence moyenne. .	65
6.1	Résultat de segmentation des cellules par watershed sur une image de frottis sanguin mince au x500.	74
6.2	Division d'une image de frottis sanguin mince au x500 en 16 imagerie.	75
6.3	Résultat de détection des cellules par YOLOv8 pour chaque imagerie.	75
6.4	Résultat de segmentation des cellules par watershed.	76
6.5	Résultat de segmentation des cellules par YOLOv8.	77
6.6	Résultat de segmentation des cellules combinant les deux méthodes watershed et YOLOv8.	77
A.1	Échantillons de faux positifs mal classés par les trois modèles. Pour montrer le type d'erreurs de prédiction pour chaque modèle, des images de GR comprenant des artefacts, des chevauchements de plaquettes ou des précipitations de coloration (a), des images de GR comprenant des précipitations de coloration qui induisent une interprétation difficile (b), des images de GR comprenant des impuretés, des précipitations d'huile, une partie de globule blanc et un GR touchant une plaquette (c) et (d), des GR chevauchés, un globule blanc, un GR touchant une plaquette et un GR comprenant des précipitations de coloration (e).	91

Liste des tableaux

2.1	Synthèse des approches pour le diagnostic automatique du paludisme.	34
3.1	Paramètres des modèles entraînés.	44
3.2	Performance des modèles (moyenne \pm écart-type) pour la classification (5 folds) sur les deux bases de données : CNR (202 patients) vs. NIH (200 patients).	45
3.3	Les performances de la classification 200 patients au niveau en positif et négatif avec un seuil de 1 GRP infecté par <i>P. falciparum</i>	48
1	Carbon footprint per fold (5 folds of 30 epochs)	89
2	Les hyperparamètres des modèles expérimentaux et les métriques d'évaluation	89
3	Caractéristiques des ensembles de données CNR et NIH	90
4	Les mesures de parasitémie moyenne sur n champs (%) de chaque méthode pour les 14 patients	91
5	Les parasitémies (%) mesurées par le système MALARIS par champ pour un total de 10 champs	92
6	Les parasitémies (%) mesurées par 10 microscopistes sur 14 patients, la mesure est assistée par la cellule de Miller en microscopie optique	93
7	Les parasitémies (%) mesurées par 10 microscopistes sur 14 patients, la mesure est réalisée en suivant la procédure standard [WHO, 2016b]	94
8	Les parasitémies (%) mesurées par les trois techniques sur 102 patients	95
9	Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients	98
9	Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)	99
9	Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)	100
9	Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)	101
10	Évaluation de la sensibilité et de la spécificité à différents seuils de détection de la parasitémie, 4 % et 10 %.	102
11	Comparaison entre les deux approches de segmentation watershed et YOLOv8.	103
12	Scores IoU pour chaque patient.	104

Liste des abréviations

OMS	Organisation Mondiale de la Santé
WHO	World Health Organization
VP	Vrai Positif
VN	Vrai Négatif
FP	Faux Positif
FN	Faux Négatif
IA	Intelligence Artificielle
ML	Machine Learning
DL	Deep Learning
RVB	Rouge Vert Bleu
RGB	Red Green Blue
HSV	Hue Saturation Value
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
MLP	Multilayer Perceptron
SVM	Support Vector Machine
GAN	Generative Adversarial Networks
RBC	Red Blood Cells
GR	Globule Rouge
GRP	Globule Rouge Parasité
TDR	Test de Diagnostic Rapide
PCR	Polymerase Chain reaction
ADN	Acide Désoxyribonucléique
LAMP	Loop-Mediated Isothermal Amplification
IFA	Immunofluorescence Antibody
FSRCNN	Fast Super-Resolution Convolutional Neural Networks
CNR	Centre National de Référence
NRC	National Reference Center
NIH	National Institutes of Health
C-CNN	Customized-CNN
FL-CNN	First Labeling CNN

ReLU	Rectified Linear Unit
RMSProp	Root Mean Square Propagation
NA	Numerical Aperture
RESA	Ring parasite-infected Erythrocyte Surface Antigen
p/ μ L	Parasites per microlitre
RSE	Relative Standard Error
SE	Standard Error
SD	Standard Deviation
RSV	Relative Standard Variation
PARs	Parasitémies
YOLO	You Only Look Once
IOU	Intersection Over Union
BCCD	Blood Cell Count and Detection
NMS	Non-Maximum Suppression

Le paludisme, également connu sous le nom de malaria, est l'une des maladies les plus répandues et mortelles dans le monde. Il s'agit d'une infection parasitaire causée par des parasites appartenant au genre *Plasmodium*, transmis à l'homme par les piqûres de femelles moustiques infectées du genre *Anopheles*. Le paludisme est un problème de santé publique majeur, touchant principalement les régions tropicales et subtropicales, où les conditions environnementales favorisent la reproduction des moustiques vecteurs.

Chaque année, des centaines de millions de personnes sont infectées par le paludisme, et des centaines de milliers en meurent, principalement des jeunes enfants en Afrique subsaharienne. Selon l'Organisation mondiale de la santé (OMS), environ 247 millions de cas de paludisme ont été recensés en 2021, entraînant plus de 619 000 décès [WHO, 2021]. La même année, 95 % des cas de paludisme et 96 % des décès dus à la maladie ont été enregistrés dans le continent africain [Talapko et al., 2019]. Le fardeau du paludisme est disproportionné dans les pays à faible revenu, où les systèmes de santé sont souvent limités et les ressources pour la prévention, le diagnostic et le traitement sont insuffisantes.

Le paludisme présente de nombreux défis en matière de prévention, de diagnostic et de traitement. La prévention repose principalement sur l'utilisation de moustiquaires imprégnées d'insecticide, sur la pulvérisation d'insecticides à l'intérieur des habitations et sur la prise de médicaments prophylactiques pour les voyageurs se rendant dans des zones à risque. Cependant, malgré ces efforts, la transmission du paludisme persiste et de nouveaux outils de prévention sont nécessaires pour lutter contre cette maladie.

Le diagnostic du paludisme est un élément clé de la prise en charge des patients. Il permet d'identifier rapidement et avec précision les personnes infectées, de fournir un traitement approprié et de prévenir la propagation de la maladie. Cependant, le diagnostic du paludisme présente également des défis. Les méthodes de diagnostic traditionnelles, telles que l'examen microscopique du frottis sanguin et les tests de diagnostic rapides basés sur la détection d'antigènes spécifiques, ont leurs limites en termes de sensibilité, de spécificité et de disponibilité.

C'est dans ce contexte que l'intelligence artificielle (IA) et l'apprentissage automatique ont émergé comme des outils prometteurs pour améliorer le diagnostic du paludisme. En utilisant des algorithmes de reconnaissance d'images et des modèles d'apprentissage automatique, il est

possible d'identifier les globules rouges parasités sur des images de frottis sanguins.

Cette thèse se focalise sur l'application de l'apprentissage profond pour le diagnostic du paludisme. Dans cette optique, une base de données d'images de frottis sanguins a été construite afin de représenter la diversité biologique propre à la réalité diagnostique. Différentes approches et modèles d'apprentissage automatique sont explorés pour détecter les parasites du paludisme, notamment *Plasmodium falciparum*, dans les frottis sanguins minces. Une attention particulière est portée sur la sensibilité, la spécificité et les performances de généralisation des méthodes proposées. Par ailleurs, l'impact de ces approches sur un critère biologique important, à savoir l'estimation de la parasitémie, est également évalué. La parasitémie se réfère à la mesure quantitative de la proportion de globules rouges parasités dans le sang.

1 Le Paludisme : un problème de santé publique

1.1 Contexte de la maladie

La transmission du paludisme se fait par la piqûre d'un moustique infecté par le parasite. Une fois dans le corps humain, les parasites du paludisme se multiplient dans le foie, puis envahissent les globules rouges, provoquant des symptômes tels que fièvre, frissons, maux de tête, fatigue et anémie. Dans les cas graves, le paludisme peut entraîner des complications potentiellement mortelles, telles que des atteintes cérébrales, rénales et pulmonaires.

Le paludisme peut être causé par plusieurs parasites protozoaires du genre *Plasmodium*. Cinq espèces sont capables d'infecter les humains (figure 1.1) : *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* et *P. knowlesi*. Parmi elles, *P. falciparum* est l'espèce la plus mortelle provoquant la grande majorité des décès dus au paludisme [Heide et al., 2019]. La proportion des différentes espèces de *Plasmodium* peut varier selon les régions géographiques. Par exemple, *P. falciparum* est prédominant en Afrique subsaharienne, où il est responsable de la majorité des cas de paludisme grave. *P. vivax* est plus répandu dans certaines régions d'Asie et d'Amérique du Sud. *P. ovale*, *P. malariae* et *P. knowlesi* sont moins fréquents, mais peuvent également causer des infections chez les humains dans certaines zones géographiques spécifiques [Talapko et al., 2019].

En plus des conséquences médicales, le paludisme a également un impact socio-économique significatif [Ren, 2019]. Les régions touchées par le paludisme voient leur développement économique entravé, en raison de la morbidité et de la mortalité élevées chez les adultes en âge de travailler et les enfants en âge scolaire. La maladie a un effet dévastateur sur les familles et les communautés, entraînant une diminution de la productivité et une augmentation des dépenses de santé. La pandémie de COVID-19 a impacté les activités de prévention et de prise en charge du paludisme, ce qui a conduit à une augmentation des cas et des décès liés à cette maladie. Cependant, grâce aux efforts de l'OMS, des programmes nationaux de lutte contre le paludisme et des partenaires, le pire a été évité. Bien que les services de prévention et de prise en charge du paludisme aient en grande partie repris leurs activités, les difficultés d'approvisionnement en produits essentiels de lutte contre le paludisme et l'incertitude quant au financement de la lutte subsistent [Heuschen et al., 2021].

Les populations les plus vulnérables au paludisme sont les enfants de moins de cinq ans et les femmes enceintes. Les enfants représentent une part importante des décès liés au paludisme dans les régions touchées. Cette maladie constitue donc un obstacle majeur à l'amélioration de la santé et du bien-être des populations, ainsi qu'à la réalisation des objectifs de développement durable des Nations Unies [Simba et al., 2018].

Human Malaria					
Species	Ring	Trophozoite	Schizont	Gametocyte	
<i>P. falciparum</i>					<ul style="list-style-type: none"> Parasitised red cells (pRBCs) not enlarged. RBCs containing mature trophozoites sequestered in deep vessels. Total parasite biomass = circulating parasites + sequestered parasites.
<i>P. vivax</i>					<ul style="list-style-type: none"> Parasites prefer young red cells pRBCs enlarged. Trophozoites are amoeboid in shape. All stages present in peripheral blood.
<i>P. malariae</i>					<ul style="list-style-type: none"> Parasites prefer old red cells. pRBCs not enlarged. Trophozoites tend to have a band shape. All stages present in peripheral blood
<i>P. ovale</i>					<ul style="list-style-type: none"> pRBCs slightly enlarged and have an oval shape, with tufted ends. All stages present in peripheral blood.
<i>P. knowlesi</i>					<ul style="list-style-type: none"> pRBCs not enlarged. Trophozoites, pigment spreads inside cytoplasm, like <i>P. malanae</i>, band form may be seen Multiple invasion & high parasitaemia can be seen like <i>P. falciparum</i> All stages present in peripheral blood.

FIGURE 1.1 – Les espèces de *Plasmodium* qui infectent les humains [Poostchi et al., 2018b].

La lutte contre le paludisme repose sur plusieurs piliers, notamment la prévention des piqûres de moustiques à l'aide de moustiquaires imprégnées d'insecticide [Michalakakis and Renaud, 2009], l'utilisation de médicaments antipaludiques pour traiter les cas infectés, et le contrôle des vecteurs de la maladie. Cependant, pour mettre en œuvre ces mesures de manière efficace, il est essentiel de disposer de méthodes de diagnostic précises, rapides et accessibles.

1.2 Importance du diagnostic précoce et précis du paludisme

L'importance du diagnostic précoce et précis du paludisme ne peut être sous-estimée. Un diagnostic rapide et précis permet une prise en charge appropriée des patients, ce qui peut sauver des vies et réduire la propagation de la maladie.

Premièrement, un diagnostic précoce du paludisme permet d'initier rapidement le traitement approprié. Le paludisme non traité ou mal traité peut entraîner des complications graves, voire la mort. En identifiant rapidement la présence du parasite *Plasmodium* dans le sang, les médecins peuvent prescrire les médicaments antipaludiques appropriés pour éliminer l'infection [Griffith et al., 2007]. Un traitement précoce réduit les risques de complications et améliore les chances de guérison complète.

Deuxièmement, un diagnostic précis est essentiel pour différencier le paludisme d'autres maladies présentant des symptômes similaires. De nombreux autres problèmes de santé, tels que la dengue, la fièvre typhoïde ou la grippe, peuvent présenter des symptômes similaires à ceux du paludisme, notamment la fièvre, les maux de tête et la fatigue [Sharma et al., 2014]. En effectuant un diagnostic précis, les médecins peuvent éviter la prescription inutile de médicaments antipa-

ludiques et orienter les patients vers le traitement approprié.

Troisièmement, un diagnostic précoce et précis permet de mettre en place des mesures de prévention et de contrôle adéquates. En identifiant rapidement les cas de paludisme dans une communauté ou une région, les autorités sanitaires peuvent prendre des mesures pour prévenir la propagation de la maladie. Cela peut inclure des campagnes de pulvérisation d'insecticides pour éliminer les moustiques vecteurs, la distribution de moustiquaires imprégnées d'insecticide ou des interventions de sensibilisation pour promouvoir les comportements préventifs tels que l'utilisation de répulsifs et l'élimination des eaux stagnantes [Michalakis and Renaud, 2009].

En outre, un diagnostic précis du paludisme est crucial dans les régions où la résistance aux médicaments antipaludiques est un problème croissant. Certaines souches de *Plasmodium falciparum*, ont développé une résistance aux médicaments couramment utilisés [Plowe, 2022]. Des techniques moléculaires permettent d'identifier les souches résistantes et de prescrire des médicaments alternatifs efficaces, évitant ainsi le développement de la résistance et garantissant un traitement approprié.

Enfin, un diagnostic précis du paludisme joue un rôle clé dans les efforts mondiaux visant à éliminer la maladie. Pour éliminer le paludisme, il est essentiel d'identifier et de traiter rapidement tous les cas, y compris les porteurs asymptomatiques. En diagnostiquant de manière précise et précoce, il est possible d'identifier les porteurs asymptomatiques et de les traiter pour interrompre la transmission de la maladie.

Cependant, malgré l'importance cruciale du diagnostic du paludisme, des défis majeurs persistent. Dans de nombreuses régions, en particulier les zones rurales et les pays en développement, l'accès aux services de diagnostic de qualité est limité. Les tests diagnostiques traditionnels, tels que la microscopie, nécessitent une expertise et les équipements spécialisés nécessaires aux autres techniques disponibles (PCR, LAMP-PCR) ne sont pas toujours disponibles. De plus, la réalisation de ces tests peut prendre du temps, ce qui retarde le diagnostic et le traitement approprié.

En conclusion, le diagnostic précoce et précis du paludisme est d'une importance capitale pour sauver des vies, prévenir la propagation de la maladie et contribuer aux efforts d'élimination du parasite. Les avancées de l'intelligence artificielle offrent de nouvelles perspectives pour améliorer les outils de diagnostic et surmonter les défis actuels. Il est essentiel de continuer à investir dans la recherche et le développement de nouvelles technologies de diagnostic afin de lutter efficacement contre le paludisme et d'améliorer la santé publique dans le monde.

2 Techniques de diagnostic du paludisme

2.1 La microscopie

Principes de base de la microscopie pour le diagnostic du paludisme

La microscopie demeure une méthode fondamentale et de référence pour le diagnostic du paludisme, en dépit de l'émergence de nouvelles technologies [WHO, 2016b]. Elle repose sur l'examen microscopique de frottis sanguins colorés.

1. **Préparation de l'échantillon :** Un échantillon de sang est prélevé sur le patient. Deux types de frottis sont généralement préparés : un frottis mince et un frottis épais. Le frottis mince est utilisé pour identifier l'espèce de *Plasmodium*, tandis que le frottis épais est utilisé pour détecter la présence de parasites.
 - **Frottis épais :** Un volume important de sang (environ 10 à 20 microlitres) est déposé sur une lame de microscope, déshémoglobinisé puis coloré, généralement avec une tein-

ture Giemsa. Cette technique est très sensible pour la détection de la présence de parasites du paludisme car elle examine une plus grande quantité de sang [WHO, 2016c].

- **Frottis mince** : Une petite quantité de sang est étalée en couche fine sur une lame de microscope, fixée et colorée. Cette méthode offre une meilleure résolution des parasites, permettant une identification plus précise de l'espèce et une estimation de la parasitémie [WHO, 2016c].
- 2. **Coloration** : Les frottis sont ensuite colorés [Sathpathi et al., 2014], généralement avec une coloration de Giemsa [WHO, 2016a]. Cette coloration permet de distinguer les parasites du reste des cellules sanguines.
- 3. **Examen au microscope** : Les frottis colorés sont ensuite examinés au microscope. Les microscopistes recherchent la présence de parasites dans les globules rouges. Ils peuvent également déterminer l'espèce de *Plasmodium* et estimer la quantité de parasites présents.
- 4. **Interprétation des résultats** : Si des parasites sont détectés, le patient est diagnostiqué comme ayant le paludisme. L'espèce de *Plasmodium* identifiée peut aider à déterminer le traitement approprié. Si aucun parasite n'est détecté, le patient est généralement considéré comme n'ayant pas le paludisme. Cependant, un test négatif ne l'exclut pas complètement, surtout si le patient présente des symptômes. Dans ce cas, le test peut être répété.

Avantages et limitations de la microscopie

Avantages

- Elle est une des méthodes les plus accessibles et disponibles pour le diagnostic du paludisme, particulièrement dans les zones où la maladie est endémique Guintran et al. [2006].
- Son coût est faible par rapport aux autres méthodes.
- La microscopie peut détecter et identifier toutes les espèces de parasites du paludisme.
- Elle permet une quantification précise de la parasitémie.
- Les résultats peuvent généralement être obtenus en moins d'une heure.

Limitations

- La microscopie nécessite des compétences techniques pour la préparation des lames et l'interprétation des résultats, ce qui peut entraîner une variabilité dans la précision du diagnostic entre différents opérateurs et laboratoires [O'Meara et al., 2006].
- La qualité des résultats dépend de nombreux facteurs, dont la qualité de la préparation des lames, de la coloration et de l'interprétation [Poostchi et al., 2018b].
- La microscopie peut manquer de sensibilité pour détecter de faibles niveaux de parasitémie ou les infections mixtes par plusieurs espèces de parasites.
- Cette technique requiert un équipement microscopique qui peut ne pas être disponible ou fonctionnel dans certaines zones à ressources limitées. De plus, le maintien de l'équipement et l'approvisionnement en consommables peuvent poser des problèmes.

2.2 Les indicateurs de performance d'une technique diagnostique

Sensibilité (Sensitivity) ou Rappel (Recall) : La sensibilité mesure la capacité d'un test à identifier correctement les cas positifs. Elle est définie comme le rapport entre le nombre de vrais positifs (VP) et la somme des vrais positifs et des faux négatifs (FN).

Sensibilité (Sensitivity) ou Rappel (Recall) :

$$\text{Sensibilité (Sensitivity)} = \frac{VP}{VP + FN} \quad (1.1)$$

Spécificité (Specificity) : La spécificité mesure la capacité d'un test à identifier correctement les cas négatifs. Elle est définie comme le rapport entre le nombre de vrais négatifs (VN) et la somme des vrais négatifs et des faux positifs (FP).

$$\text{Spécificité (Specificity)} = \frac{VN}{VN + FP} \quad (1.2)$$

Exactitude (Accuracy) : L'exactitude représente la proportion globale de prédictions correctes du test, quelle que soit la classe. Elle est définie comme le rapport entre le nombre de vrais positifs et de vrais négatifs ($VP + VN$) et la somme totale des prédictions ($VP + VN + FP + FN$).

$$\text{Exactitude (Accuracy)} = \frac{VP + VN}{VP + VN + FP + FN} \quad (1.3)$$

Précision (Precision) : La précision mesure la proportion de résultats positifs correctement identifiés parmi tous les résultats positifs prédits par le test. On parle aussi de valeur prédictive positive (VPP). Elle est définie comme le rapport entre le nombre de vrais positifs (VP) et la somme des vrais positifs et des faux positifs ($VP + FP$).

$$\text{Précision (Precision)} = \frac{VP}{VP + FP} \quad (1.4)$$

F1-Score : Le score F1 est une mesure combinée de la précision et du rappel, calculée comme la moyenne harmonique de ces deux mesures. Définition : Il est défini comme le double du produit de la précision et du rappel, divisé par la somme de la précision et du rappel.

$$\text{F1-Score} = \frac{2 \times \text{Précision (Precision)} \times \text{Rappel (Recall)}}{\text{Précision (Precision)} + \text{Rappel (Recall)}} \quad (1.5)$$

Dans le contexte du diagnostic du paludisme, la sensibilité et la spécificité sont particulièrement importantes en raison de la gravité potentielle de la maladie. Un faux négatif (c'est-à-dire, un individu qui a le paludisme mais qui est testé négatif) peut entraîner un retard de traitement, ce qui peut avoir des conséquences graves, voire mortelles. D'autre part, un faux positif (c'est-à-dire, un individu qui n'a pas le paludisme mais qui est testé positif) peut entraîner un traitement inutile, ce qui peut avoir des effets secondaires et contribuer à la résistance aux médicaments. La résistance aux médicaments est un problème croissant dans le traitement du paludisme, et le diagnostic précis peut aider à surveiller et à gérer ce problème.

2.3 Les tests de diagnostic rapide (TDR)

Alors que la microscopie reste une méthode de diagnostic précieuse pour le paludisme, elle présente des défis en termes de temps, de formation et d'expertise nécessaires. Dans ce contexte, les tests de diagnostic rapide (TDR) sont devenus une alternative populaire, en particulier dans les régions où l'accès à la microscopie de qualité est limité [Mathison and Pritt, 2017]. Les TDR offrent plusieurs avantages, notamment la rapidité des résultats, la facilité d'utilisation et la possibilité d'être utilisés sur le terrain.

Principe de fonctionnement des TDR

Les tests de diagnostic rapide pour le paludisme sont des outils qui ont révolutionné le diagnostic du paludisme, en particulier dans les régions où l'accès à la microscopie de qualité est limité. Les TDR sont des tests immunochromatographiques qui détectent des antigènes spécifiques produits par les parasites du paludisme dans le sang d'une personne.

1. **Prélèvement d'échantillon** : Un échantillon de sang est prélevé sur le patient, généralement par une piqûre au bout du doigt.
2. **Application de l'échantillon** : L'échantillon de sang est appliqué sur une zone spécifique du TDR. Dans certains cas, un diluant est également ajouté.
3. **Migration de l'échantillon** : L'échantillon de sang migre le long de la bande de test par capillarité. S'il y a des antigènes de paludisme dans l'échantillon, ils se lieront aux anticorps spécifiques qui sont fixés sur la bande de test.
4. **Lecture des résultats** : Les résultats sont généralement visibles en 15 à 30 minutes. Un test est généralement considéré comme positif s'il y a deux lignes de couleur, une ligne de contrôle pour montrer que le test fonctionne correctement, et une ligne de test pour montrer la présence d'antigènes de *Plasmodium* (figure 1.2). Un test est considéré comme négatif s'il n'y a qu'une seule ligne de couleur (la ligne de contrôle).

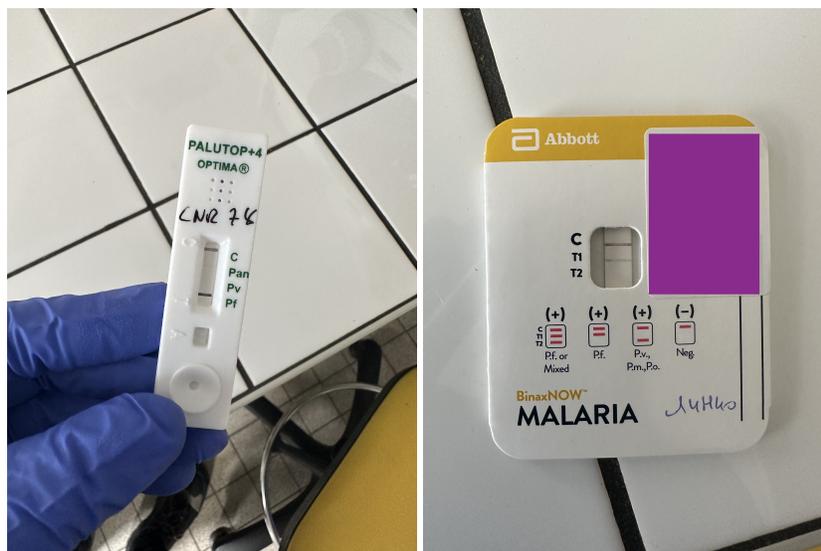


FIGURE 1.2 – Deux différents TDR positifs pour *Plasmodium falciparum*.

Avantages et limitations des TDR

Avantages des TDR

1. **Facilité d'utilisation** : Les TDR sont conçus pour être simples à utiliser, ce qui signifie qu'ils peuvent être utilisés par du personnel non spécialisé avec une formation minimale.
2. Les TDR permettent l'identification d'espèces permettant de différencier principalement une infection par *Plasmodium falciparum* des autres espèces.

3. **Rapidité** : Les TDR fournissent des résultats en 15 à 30 minutes, ce qui permet un diagnostic rapide et la possibilité de commencer le traitement immédiatement.
4. **Implémentation sur le terrain** : Les TDR ne nécessitent pas d'équipement de laboratoire spécialisé, ce qui signifie qu'ils peuvent être utilisés dans des zones éloignées ou des zones avec des ressources limitées.
5. **Stabilité** : Les TDR sont généralement stables à température ambiante, ce qui facilite leur stockage et leur transport.
6. **Coût** : Les TDR sont généralement moins chers que le reste des autres techniques.

Limitations des TDR

1. **Sensibilité** : Les TDR peuvent ne pas être aussi sensibles que la microscopie, en particulier pour les infections à faible parasitémie. Cela signifie qu'ils peuvent ne pas détecter le paludisme chez certains patients qui sont effectivement infectés [Ogunfowokan et al., 2020].
2. Une faible sensibilité pour la détection des espèces de *Plasmodium ovale* et *Plasmodium malariae* [Wongsrichanalai et al., 2007].
3. **Spécificité** : Bien que les TDR soient généralement spécifiques pour le paludisme, ils peuvent parfois donner des résultats faussement positifs en présence de certaines autres infections ou conditions médicales.
4. **Incapacité à quantifier la parasitémie** : Contrairement à la microscopie, les TDR ne permettent pas de déterminer la quantité de parasites présents. Cela peut être important pour le suivi du traitement et pour déterminer la gravité de l'infection notamment pour l'espèce *Plasmodium falciparum*, qui est la plus mortelle.
5. **Coût** : Bien que les TDR soient généralement moins chers que d'autres méthodes de diagnostic, leur coût peut toujours être prohibitif dans certaines régions à ressources limitées.

2.4 Les techniques moléculaires

Polymerase Chain Reaction (PCR)

La PCR (Polymerase Chain Reaction) est une technique qui permet d'amplifier une séquence spécifique d'ADN pour faciliter sa détection. Dans le contexte du paludisme, la PCR cible généralement des séquences spécifiques de l'ADN du parasite *Plasmodium* [Rypien et al., 2017]. La PCR peut être très sensible et spécifique, ce qui en fait une méthode de choix pour le diagnostic du paludisme, en particulier dans les cas où la charge parasitaire est faible [Pöschl et al., 2010].

Avantages et limitations des techniques moléculaires

Avantages

- **Sensibilité et spécificité élevées** : La PCR est extrêmement sensible et peut détecter de très faibles niveaux de parasites [Leski et al., 2020]. Elle est également très spécifique et, selon les cibles utilisées, elle peut permettre de distinguer entre différentes espèces de *Plasmodium*.

Limitations

- **Équipement et formation spécialisés nécessaires** : La PCR nécessite un équipement de laboratoire spécialisé et du personnel formé.
- **Coût et temps** : Elle est plus coûteuse que les tests rapides et la microscopie et prend plus de temps à être réalisée.
- **Sensibilité excessive** : De plus, la PCR peut être trop sensible, détectant des parasites dans des situations qui ne sont pas cliniquement significatives (portage simple) [Farcas et al., 2004].

LAMP (Loop-Mediated Isothermal Amplification)

La LAMP est une autre technique d'amplification de l'ADN qui a été utilisée pour le diagnostic du paludisme. Contrairement à la PCR, la LAMP se déroule à une température constante, ce qui simplifie l'équipement nécessaire. La réalisation de la LAMP est également plus rapide que pour la PCR.

Avantages

- **Rapidité et simplicité** : La LAMP est rapide, simple et peut se faire avec un équipement portable.
- **Sensibilité et spécificité élevées** : Elle est également très sensible et spécifique [Selvarajah et al., 2020].

Limitations

- **Sensibilité excessive** : Comme la PCR, la LAMP peut être trop sensible [Hopkins et al., 2013].
- **Expertise nécessaire** : Bien que la LAMP soit généralement plus simple que la PCR, elle nécessite une certaine expertise pour être réalisée correctement.
- **Coût** : Un coût élevé.

3 Les autres techniques diagnostiques

3.1 La sérologie

La sérologie est une méthode qui permet de détecter la présence d'anticorps spécifiques produits par le système immunitaire en réponse à une infection par le parasite du paludisme. Ces anticorps peuvent être détectés dans le sang des individus infectés, même après la disparition des symptômes de la maladie. Cela en fait un outil précieux pour les études épidémiologiques et de séroprévalence. Cependant, la recherche d'anticorps n'est pas adaptée au diagnostic d'accès palustre (la positivité de la sérologie est trop tardive par rapport à la clinique).

Le Test d'Immunofluorescence (IFA) est une technique utilisant des antigènes spécifiques du parasite du paludisme pour détecter et quantifier les anticorps IgG et IgM dans les échantillons de sérum. Les anticorps IgG sont généralement produits plus tard dans l'infection et peuvent persister pendant une longue période, tandis que les anticorps IgM sont généralement produits plus tôt et disparaissent plus rapidement.

En combinant la sérologie avec d'autres techniques telles que la détection d'antigènes et l'analyse de l'ADN, les chercheurs peuvent obtenir une vision plus complète de la transmission du pa-

ludisme [Oviedo et al., 2020]. La détection d'antigènes peut permettre d'identifier les infections actives, tandis que l'analyse de l'ADN permet d'identifier les différentes espèces de parasites. Ensemble, ces données contribuent à surveiller les tendances de la maladie, à identifier les zones à haut risque et à évaluer l'efficacité des interventions de lutte contre le paludisme.

3.2 La cytométrie en flux

La cytométrie en flux est une technique qui utilise la lumière laser pour analyser les caractéristiques physiques et chimiques des cellules individuelles à mesure qu'elles passent à travers un faisceau de lumière. Dans le contexte du paludisme, elle peut être utilisée pour compter le nombre d'érythrocytes (globules rouges) qui sont infectés par le parasite du paludisme [Grimberg, 2011, Janse and Van Vianen, 1994]. Cela peut fournir une mesure précise de la charge parasitaire, qui est un indicateur important de la gravité de l'infection.

Le Sysmex XN-31 est un exemple d'appareil de cytométrie en flux qui a été spécifiquement conçu pour le diagnostic du paludisme. Il utilise la fluorescence pour détecter et quantifier les parasites du paludisme dans les échantillons de sang. En plus de quantifier les parasites, il peut également déterminer l'espèce de *Plasmodium*, ce qui peut aider à orienter le traitement.

Néanmoins, la cytométrie en flux présente des limites, car une étude a révélé la possibilité de faux positifs lorsque les érythrocytes présentent une morphologie anormale. De plus, les performances de l'appareil Sysmex XN-31 restent à évaluer dans les régions endémiques de paludisme, car les tests ont été principalement réalisés dans des régions non endémiques à partir de cas importés [Khartabil et al., 2022].

4 Les défis liés aux techniques diagnostiques

4.1 Disponibilité et accessibilité des techniques diagnostiques dans les régions endémiques

Les techniques de diagnostics dépendent de nombreux facteurs que nous aborderons dans cette section. En effet, le paludisme est principalement une maladie tropicale qui touche les pays d'Afrique subsaharienne et l'Asie du Sud-Est. Il est important de replacer l'utilisation des techniques dans un contexte géographique qui permet d'évaluer la capacité et les ressources à disposition pour lutter contre cette maladie [Bronzan et al., 2008]. La disponibilité et l'accessibilité des techniques de diagnostic du paludisme dans les régions endémiques sont des facteurs essentiels pour le contrôle de la maladie [Varo et al., 2021]. Cependant, ces régions sont souvent confrontées à de nombreux défis qui rendent difficile l'accès à des tests de diagnostic précis et fiables [Guintran et al., 2006]. Tout d'abord, les infrastructures du système de santé dans de nombreuses régions endémiques, notamment en Afrique subsaharienne sont limitées, en particulier dans les zones rurales ou éloignées. Cela peut rendre difficile l'accès à des laboratoires de diagnostic et à des professionnels de santé formés.

Le coût des tests de diagnostic peut être prohibitif dans les régions à faibles ressources économiques. Bien que certains tests, comme les TDR, soient relativement abordables, d'autres, comme la PCR et la LAMP, restent des techniques très coûteuses et inaccessibles. De plus, l'approvisionnement en tests de diagnostic et leur distribution aux centres de santé peuvent être un défi similaire à celui de la distribution des médicaments, en particulier dans les zones éloignées qui peuvent principalement être les plus touchées. Certains tests nécessitent des conditions de stockage spécifiques qui peuvent être difficiles à maintenir (TDR par exemple) [Albertini et al., 2012]. L'utilisation

correcte de nombreux tests de diagnostic nécessite une formation. Par exemple, la microscopie, qui est la méthode dite de référence couramment utilisée pour le diagnostic du paludisme, nécessite une formation pour être réalisée correctement. Ce facteur important peut donner lieu à une variabilité d'interprétation diagnostique due à un manque d'expertise et de formation.

Afin de répondre à l'ensemble de ces défis freinant la lutte contre le paludisme, diverses stratégies sont en cours d'élaboration et de mise en œuvre depuis de nombreuses années. Cela passe par le développement de tests de diagnostic plus abordables, accessibles et plus faciles d'utilisation, l'amélioration de l'infrastructure et des systèmes de santé, la formation du personnel de santé notamment en microscopie, et des campagnes d'éducation pour augmenter l'acceptabilité des tests de diagnostic du paludisme.

4.2 Les défis liés à la surveillance épidémiologique

Passant de la disponibilité et de l'accessibilité des techniques de diagnostic du paludisme à la surveillance épidémiologique, il est important de noter que ces deux aspects sont étroitement liés. En effet, une surveillance épidémiologique efficace repose sur la capacité à diagnostiquer avec précision et rapidement le paludisme dans les populations à risque. Cependant, la surveillance épidémiologique du paludisme présente ses propres défis.

1. **Collecte de données** : L'obtention et la collecte de données précises et fiables est un des défis majeur dans la mise en place d'une stratégie efficace. Ce défi particulièrement complexe dans les pays et régions où les accès aux soins de santé sont limités [Bempah et al., 2020]. De plus, les systèmes de surveillance peuvent ne pas recenser tous les cas de paludisme et biaiser ainsi les estimations, en particulier si les individus pratiquent l'automédication ou si ils ne cherchent pas de soins médicaux dans des établissements officiels.
2. **Qualité des données** : La qualité des données peut varier en fonction de la variabilité des tests et techniques utilisés, de la formation du personnel de santé en termes d'examen microscopique mais aussi par rapport aux méthodes utilisées pour rapporter et remonter les données. Des données erronées ou incomplètes peuvent entraver les efforts de surveillance et de contrôle efficace [Alegana et al., 2020].
3. **Ressources** : La surveillance épidémiologique nécessite des ressources humaines et matérielles [Kiszewski et al., 2007], y compris du personnel qualifié, des systèmes de gestion et de stockage des données ainsi que du financement permettant d'assurer la qualité des infrastructures en temps réel [Mwenesi et al., 2022]. Toutes ces exigences sont difficiles à assurer étant donné l'hétérogénéité des ressources économiques des pays touchés par la maladie.
4. **Variabilité de la transmission du paludisme** : La transmission du paludisme peut varier considérablement d'une région à l'autre et au fil du temps, en fonction de facteurs tels que le climat [Zhou et al., 2005], les méthodes de contrôle du paludisme et les mouvements démographiques. cela rend la surveillance épidémiologiques d'autant plus nécessaire.
5. **Résistance aux médicaments** : La résistance aux médicaments est une problématique croissante dans le traitement du paludisme [Wicht et al., 2020]. La surveillance de la résistance aux médicaments est essentielle pour guider les politiques et stratégies de traitement, mais elle peut être difficile à réaliser en raison de la complexité des méthodes diagnostiques à mettre en oeuvre.

Pour surmonter l'ensemble de ces défis, il est nécessaire d'améliorer les systèmes de collecte de données, de renforcer la formation du personnel de santé, d'investir dans les infrastructures de

soins et de mettre en œuvre des stratégies pour surveiller et contrôler la résistance aux médicaments.

5 Besoins et perspectives pour de nouvelles approches de diagnostic

Des avancées technologiques récentes ont ouvert de nouvelles perspectives pour le diagnostic du paludisme. Les progrès technologiques permettent désormais de capturer des photographies, y compris à partir d'un microscope et d'envoyer ces images à un logiciel afin de les analyser. Ces possibilités ouvrent le champ à des procédures de diagnostic de téléexpertise permettant d'avoir un diagnostic depuis presque n'importe où dans le monde. Pour développer un tel système de diagnostic en ligne, il est nécessaire d'implémenter des méthodes d'analyse automatique des images obtenues dans des conditions réelles sur le terrain [Madhu, 2020]. Cela ouvre une voie aux approches de l'intelligence artificielle (IA) pour relever les défis et les besoins techniques spécifiques aux pays endémiques, apportant de nouveaux outils pour le contrôle du paludisme [Beck, 2022].

5.1 L'intégration de l'intelligence artificielle dans le diagnostic du paludisme

Les progrès en biologie computationnelle, associés à un meilleur accès aux infrastructures technologiques de l'information et à la puissance de calcul mobile dans de nombreux pays à revenu faible ou intermédiaire, offrent l'opportunité de déployer des systèmes de santé assistés par l'IA [Schwalbe and Wahl, 2020]. L'IA implique des méthodes d'apprentissage supervisé, qui sont un ensemble d'algorithmes appliqués à des données déjà étiquetées pour apprendre la reconnaissance de motifs à un modèle statistique de classification ou de régression. Ces modèles doivent être capables de généraliser la tâche apprise à de nouvelles données indépendantes.

L'apprentissage profond, également connu sous le nom de Deep Learning, est une branche de l'IA qui vise à reproduire la capacité d'apprentissage et de compréhension des humains par les machines. Il s'agit d'une approche computationnelle basée sur des réseaux neuronaux artificiels avec de multiples couches de traitement de l'information. Un élément central de l'apprentissage profond est l'utilisation de réseaux neuronaux artificiels, en particulier les réseaux neuronaux à convolutions (CNN). Les CNN sont principalement conçus pour l'analyse d'images et sont capables de détecter et de reconnaître des motifs visuels complexes.

Avec l'avènement de l'apprentissage profond, les CNN ont atteint des performances comparables à celle des experts dans la détection de nombreuses pathologies. Ces performances élevées reposent principalement sur l'analyse d'images médicales, telles que les scans pour le diagnostic des maladies rétinienues [De Fauw et al., 2018] ou les lames d'histopathologie pour le diagnostic des maladies cancéreuses [Zhang et al., 2019], pour en citer quelques exemples. Contrairement aux méthodes de sélection des caractéristiques, les approches d'apprentissage profond évitent généralement l'étape d'extraction des caractéristiques.

Les CNN sont des architectures de réseaux neuronaux spécifiquement conçues pour l'analyse de données structurées en grille, telles que des images (figure 1.4) ou des séquences temporelles. Les CNN ont révolutionné le domaine de la vision par ordinateur en permettant une extraction automatique de caractéristiques à partir des données brutes, ce qui a considérablement amélioré les performances de diverses tâches de traitement d'images telles que la classification, la détection d'objets et la segmentation sémantique.

Une image est un ensemble fini de valeurs numériques appelées éléments d'image ou pixels. L'image est composée d'un nombre fixe de lignes et de colonnes de pixels. Les pixels sont les plus

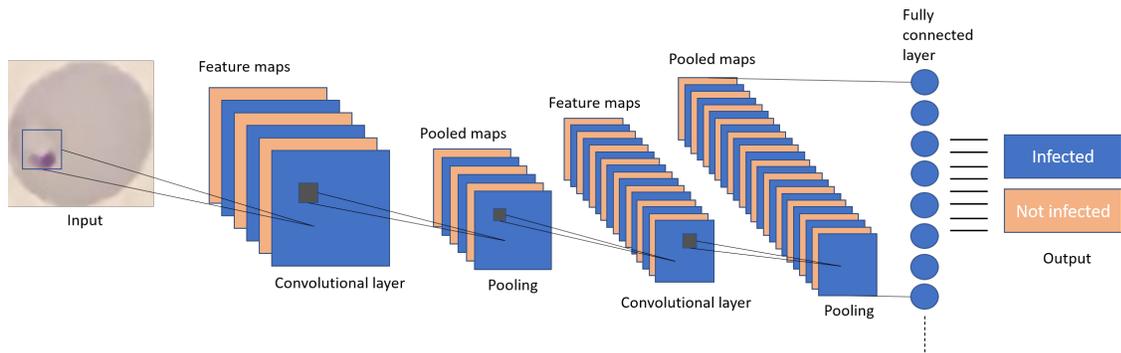


FIGURE 1.3 – Architecture simple d’un CNN pour la classification binaire d’images de globules rouges. Le CNN est composé de 2 couches de convolutions et 2 couches de pooling, une sortie à 2 classes : infecté ou non infecté.

petits éléments individuels dans une image, qui contiennent des valeurs représentant l’intensité d’une couleur donnée à un point spécifique. Habituellement, une image numérique est traitée mathématiquement en utilisant la représentation rouge, vert et bleu (RVB) qui convertit une image pixelisée en une matrice de valeurs d’intensité de rouge, vert et bleu (figure 1.4).

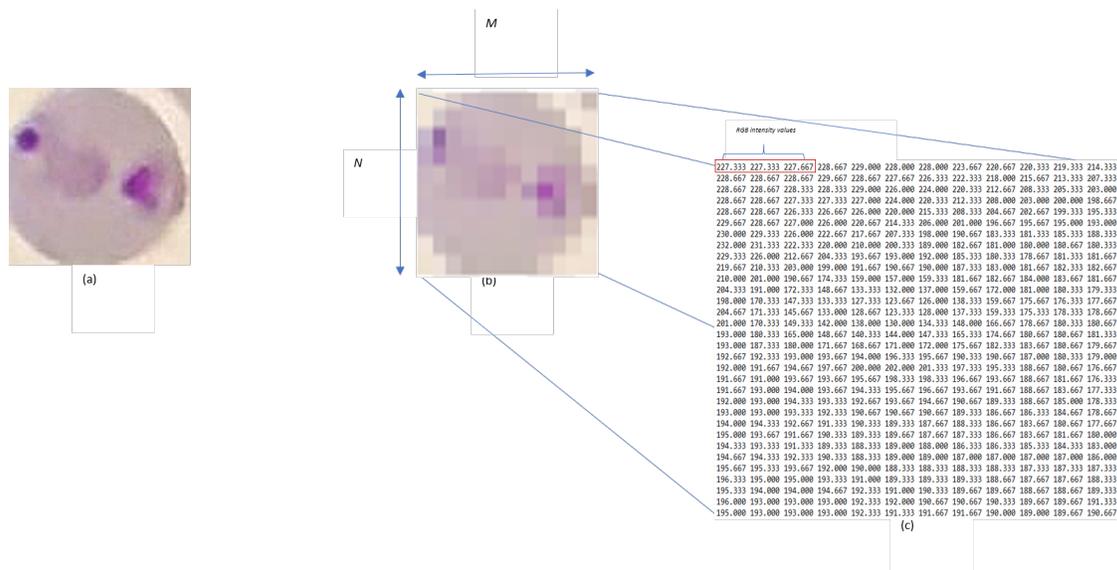


FIGURE 1.4 – Représentation d’une image RVB : (a) Globule rouge infecté par *Plasmodium falciparum*. (b) Image pixelisée avec une taille de (MxN), où M est le nombre de colonnes et N le nombre de lignes. (c) La matrice des valeurs d’intensité correspondantes définies par trois valeurs, rouge, vert et bleu (RVB), sur une plage de 256 valeurs possibles : un mélange de ces trois valeurs produit des intensités dans l’image.

Le principe fondamental des CNN est la convolution, qui est une opération mathématique entre un noyau de convolution (également appelé filtre) et une fenêtre de l’image d’entrée. Le filtre est un petit tableau de poids qui est appliqué à chaque région de l’image pour calculer des caractéristiques locales. En glissant le filtre sur l’image, des caractéristiques telles que des bords, des

textures ou des motifs spécifiques sont détectées et extraites. Cela permet au réseau de capturer des informations pertinentes à différentes échelles spatiales.

Les CNN sont généralement composés de plusieurs couches (figure 1.3). Les couches de convolution sont responsables de l'extraction des caractéristiques en appliquant des filtres sur les données d'entrée. Les couches de pooling réduisent la dimension spatiale des caractéristiques extraites en effectuant une opération de sous-échantillonnage, ce qui permet de réduire la quantité de calcul nécessaire et d'améliorer la robustesse face aux variations mineures. Les couches entièrement connectées, également appelées couches de sortie, agrègent les caractéristiques extraites par les couches précédentes et les utilisent pour la classification finale (figure 1.5).

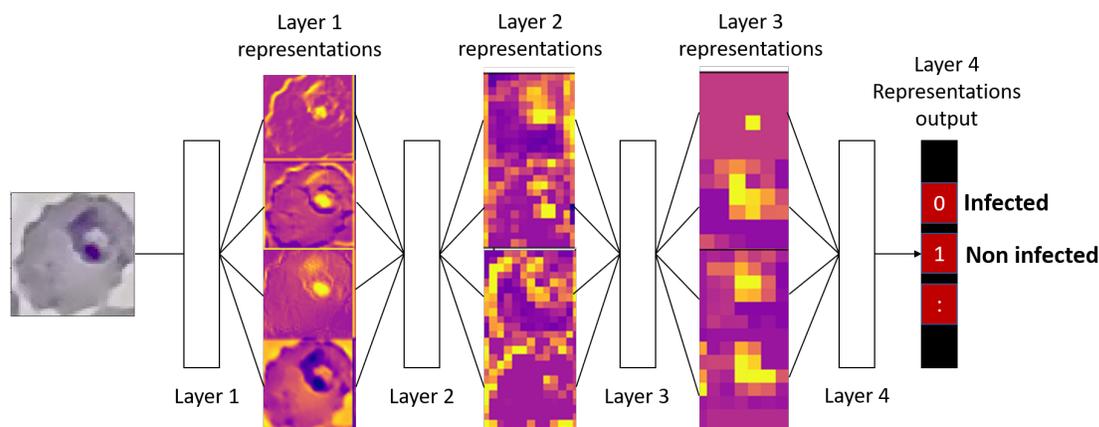


FIGURE 1.5 – Activations d'un CNN à trois couches de convolution pour l'identification d'un globe rouge parasité par *Plasmodium falciparum*.

L'apprentissage des CNN se fait par rétropropagation, une technique d'optimisation qui ajuste les poids des filtres afin de minimiser une fonction de perte entre les prédictions du modèle et les étiquettes réelles des données d'entraînement. Les gradients de cette fonction de perte sont propagés de la couche de sortie vers les couches précédentes, permettant ainsi d'ajuster les poids pour améliorer les performances du modèle. L'apprentissage est généralement réalisé sur de grands ensembles de données annotées afin d'obtenir des modèles généralisant bien sur de nouvelles données.

L'intégration de l'apprentissage profond dans le diagnostic du paludisme peut apporter plusieurs avantages. Il permet une analyse automatisée des images, réduisant ainsi le temps et l'effort nécessaires pour examiner manuellement chaque frottis sanguin. De plus, les modèles d'apprentissage profond peuvent aider à surmonter les limites de compétence et de pratique des microscopistes, en fournissant une approche standardisée et reproductible pour le diagnostic, notamment dans les régions endémiques. Ces applications visent à faciliter le diagnostic du paludisme en détectant les parasites *Plasmodium*, en identifiant les espèces et en estimant la parasitémie.

6 Contributions scientifiques

Dans le chapitre 3, le contenu s'appuie sur des travaux publiés dans le journal "Informatics in Medicine Unlocked" [Acherar et al., 2022] et dans le chapitre 4, les travaux découlent d'un second article scientifique en cours de publication dans le journal "PLOS ONE".

6.1 Résumés des travaux publiés dans le domaine du diagnostic du paludisme

Après avoir présenté les défis liés au diagnostic de la malaria, un diagnostic précis est donc essentiel pour guider le traitement et évaluer la gravité de l'infection par *Plasmodium falciparum*. Cependant, l'identification microscopique des parasites dans les globules rouges (GR) à partir de frottis sanguins minces est une tâche laborieuse, sujette à des variations et requérant une expertise spécifique. Dans ce contexte, l'utilisation de techniques d'apprentissage profond, telles que les réseaux neuronaux à convolution, offre de nouvelles perspectives pour automatiser la détection de la malaria et l'estimation de la parasitémie. Ce qui nous amène à exposer 2 questions de recherches :

QR1 Est-il crucial d'avoir un jeu de données riche avec une diversité morphologique et biologique afin d'avoir un modèle capable de différencier de manière fiable un globule rouge parasité par *Plasmodium falciparum* d'un globule rouge non infecté ou d'un autre composant sanguin ?

QR2 Grâce à un modèle d'apprentissage profond avec une sensibilité élevée, est-il possible d'estimer de manière automatique, fiable et rapide une parasitémie sur des images de frottis sanguins minces au x500 ?

Afin de répondre à ces questions de recherches, nous avons, dans notre première étude, nous avons proposé un système que l'on nommera MALARIS, pour la détection des parasites dans les GRP par *Plasmodium falciparum* et la classification des composants sanguins non infectés. Nous avons construit un ensemble de données représentatif de la réalité diagnostique en capturant des images de composants sanguins de 202 patients à un grossissement de 500x. Nous avons évalué plusieurs architectures de CNNs à travers une étude de validation croisée sur deux jeux de données incluant les données du NIH. Les modèles entraînés sur notre ensemble de données ont démontré de meilleures performances en termes de généralisation, avec des scores de précision, de sensibilité et de spécificité atteignant respectivement 99,7 %, 77,9 % et 99,8 %.

Dans notre seconde étude, nous avons étudié la variabilité de la mesure de la parasitémie en comparant trois méthodes : l'examen microscopique, l'utilisation de la microscope assistée par la cellule de Miller et notre système MALARIS présenté dans une première étude. Pour ce faire, nous avons analysé une collection de frottis sanguins minces anonymisés provenant du Centre de Référence National (CNR) pour le Paludisme à Paris, en France. Nous avons examiné l'impact du nombre d'images de champs microscopiques sur la mesure de la parasitémie et étudié la variabilité entre différents lecteurs. De plus, nous avons élargi notre étude en incluant des images supplémentaires de frottis sanguins minces de 102 patients, pour comparer les performances et la cohérence de notre système avec la microscopie assistée par la cellule de Miller et la cytométrie en flux. Les résultats ont montré de fortes corrélations entre les trois méthodes, avec des coefficients de détermination allant de 0,87 à 0,92.

6.2 Autres contributions en co-auteur

- [Normand et al., 2022] *Identification of a clonal population of Aspergillus flavus by MALDI-TOF mass spectrometry using deep learning*. Normand, AC., Chaline, A., Mohammad, N., Godmer, A., **Acherar, A.**, Huguenin, A., Ranque, S., Tannier X. & Piarroux, R.
- [Moreno-Sabater et al., 2023] *Intestinal Candida albicans overgrowth in IgA deficiency*. Moreno-Sabater A, Sterlin D, Imamovic L, Bon F, Normand AC, Gonnin C, Gazzano M, Bensalah M,

Dorgham K, Ben Salah E, **Acherar A**, Parizot C, Rigourd V, Begue H, Dalle F, Bachmeyer C, Hennequin C, Yssel H, Malphettes M, Fieschi C, Fadlallah J, Gorochov G. I.

Dans la suite de cette thèse, nous aborderons différentes sections qui mettent en avant les contributions scientifiques réalisées dans le domaine du diagnostic du paludisme. Nous commencerons par présenter dans le chapitre 2, un état de l'art approfondi pour situer nos travaux par rapport aux avancées existantes. Dans les chapitres 3 et 4, nous développerons plus en détail nos contributions, en exposant les méthodes utilisées, les résultats obtenus et leur interprétation. Enfin, nous conclurons cette thèse par une discussion générale, où nous analyserons les résultats, les comparerons avec d'autres travaux similaires et discuterons des perspectives futures dans le domaine du diagnostic du paludisme. Cette structure de thèse permettra de mettre en valeur les avancées réalisées et de fournir une vision globale des contributions scientifiques apportées.

1 Classification automatique et diagnostic du paludisme

1.1 Introduction

Pour introduire les objectifs de ce travail de thèse d'identification de globules rouges parasités par *Plasmodium falciparum* et de la mesure automatique de la parasitémie, nous devons d'abord présenter les développements récents portant sur le diagnostic automatique du paludisme à partir d'images microscopiques.

Nous examinerons les travaux réalisés sur la classification automatique d'images pour le diagnostic automatique du paludisme, ainsi que les limites qui subsistent. Nous aborderons également la manière dont nos contributions visent à répondre à ces limites.

Dans un premier temps, nous aborderons les travaux portant sur la classification automatique des images pour le diagnostic du paludisme en utilisant des techniques de traitement d'images et des approches traditionnelles de classification. Nous mettrons en évidence les limites de ces méthodes.

Nous aborderons ensuite les travaux portant sur l'apprentissage profond pour l'identification des espèces et des stades parasitaires de *Plasmodium*.

Enfin, nous examinerons les systèmes commercialisés déployés pour le diagnostic du paludisme, tels que les dispositifs d'imagerie automatisée et les applications mobiles. Nous évaluerons leur accessibilité, leur fiabilité diagnostique et discuterons des limites et des défis associés à leur utilisation à grande échelle, en particulier dans les régions endémiques où les ressources médicales peuvent être limitées.

1.2 Classification automatique des images pour le diagnostic du paludisme : approches traditionnelles et traitement des images

Pour le diagnostic automatisé du paludisme par microscopie, trois étapes sont communes à toutes les approches utilisées :

1. La préparation de l'échantillon, comprenant la réalisation d'un frottis sanguin mince ou épais ainsi que la technique de coloration utilisée;

2. L'acquisition des images à travers un dispositif de capture;
3. Un algorithme de pré-traitement des images acquises;
4. Un algorithme pour analyser et classer les images traitées.

Une architecture d'un système classique de classification automatique basé sur l'imagerie pour le diagnostic du paludisme est un système en plusieurs étapes composé de deux processus principaux : le traitement des données et la partie décisionnelle de classification (figure 2.1). Le traitement des données inclut l'acquisition d'images et des techniques de pré-traitement telles que la réduction du bruit ou la normalisation de la coloration ou du contraste des images (référence). La classification des images numériques consiste à identifier les globules rouges parasités (GRP) et non infectés et d'identifier si besoin l'espèce de *Plasmodium* et ses stades de développement.

Les approches traditionnelles pour détecter les parasites responsables du paludisme sur des frottis sanguins minces impliquent des étapes spécifiques de traitement des données, telles que l'extraction et la sélection de caractéristiques à partir d'images. Cette étape consiste à extraire des représentations et des signatures à partir des images qui peuvent aider à caractériser des objets spécifiques tels que les GR (également appelés objets d'intérêt). Les caractéristiques sont ensuite sélectionnées pour permettre la classification des GRP et des cellules non infectées, telles que la morphologie (forme, texture, etc...), la couleur et la granulométrie.

Différentes représentations des espaces de couleur peuvent être utilisées, comme la valeur de teinte saturation (HSV) [Savkare and Narote, 2015a, Nasir et al., 2012], où la couleur Lab*, où L* représente la luminosité, a* caractérise la valeur rouge/verte, et b* caractérise la valeur bleu/jaune, qui est le canal de couleur le plus prédominant pour caractériser les parasites du paludisme [Khan et al., 2014]. Les approches traditionnelles utilisent des modèles d'apprentissage automatique, tels que les machines à vecteurs de support (SVM) [Linder et al., 2014a] multiclasse [Hearst, 1998], les naïve Bayes ou les K-means mobiles [Savkare and Narote, 2015b, Abdul Nasir et al., 2012, Das et al., 2015]. Globalement, les performances de classification de ces modèles d'apprentissage automatique varient de 80 % à 99 % d'exactitude pour la détection de *Plasmodiums*. Cependant, les évaluations ont été réalisées sur un nombre relativement limité d'images, correspondant à un nombre faible d'une centaine de frottis sanguins. La plupart de ces études proposent des approches basées sur des ensembles de données unicentriques, souvent non représentatifs de la réalité diagnostique, ce qui les rend inapplicables pour un diagnostic de routine en raison de leur faible capacité de généralisation et de leur manque de robustesse. La diversité des méthodes de coloration et la qualité de la préparation des frottis sanguins microscopiques rendent complexe la classification de caractéristiques sélectionnées à l'aide d'approches de traitement d'images. Enfin, la sélection manuelle des caractéristiques d'intérêt nécessite une expertise et est souvent sujette à des biais importants, en particulier pour discriminer les espèces et les stades de développement de *Plasmodium*.

1.3 Classification automatique d'images : Apprentissage profond pour l'identification de stades parasitaires de *Plasmodiums*

Les approches traditionnelles d'apprentissage automatique pour la détection des stades parasitaires présentent de nombreuses limites. Ces approches impliquent souvent des étapes de traitement des données manuelles, telles que la sélection et l'ingénierie des caractéristiques, qui consistent à extraire des informations spécifiques des images afin d'y appliquer une classification. Cependant, ces caractéristiques peuvent être difficiles à définir de manière précise et complète, ce qui peut entraîner une perte d'informations importantes, notamment lorsque les objets analysés

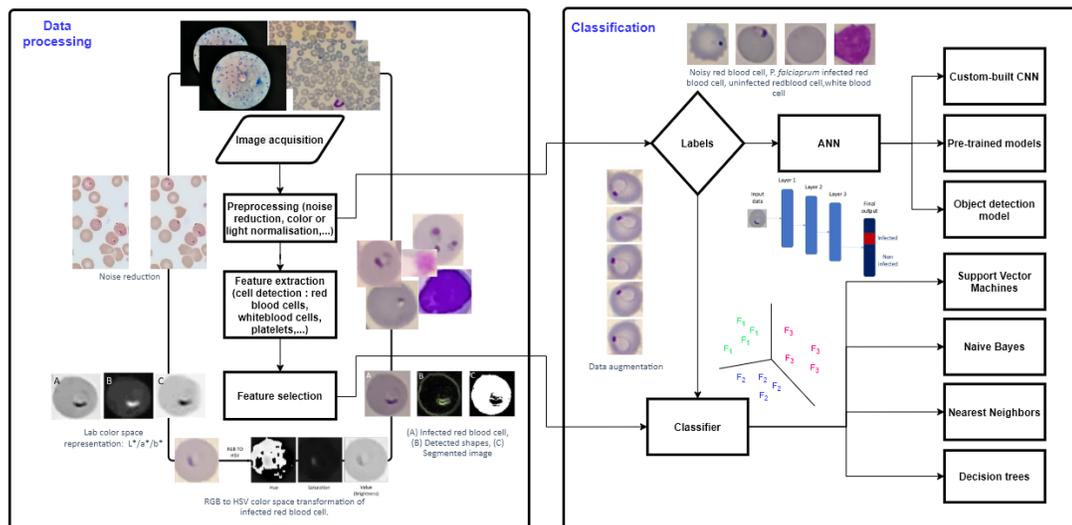


FIGURE 2.1 – Approches et étapes clés d'un système automatisé de classification d'images pour le diagnostic du paludisme

comportent une complexité et des caractéristiques souvent très proches d'un stade à un autre. De plus, la performance de ces modèles traditionnels dépend fortement de la qualité et de la représentativité des caractéristiques sélectionnées, ce qui peut limiter leur capacité à généraliser à de nouvelles données ou à des variations courantes dans les images.

En revanche, les méthodes d'apprentissage profond, en particulier les CNN, offrent une approche plus prometteuse pour la détection des stades parasitaires *Plasmodiums*. Les CNN peuvent extraire automatiquement les caractéristiques discriminantes à partir des données brutes, ce qui permet de s'affranchir des biais de sélection manuelle des caractéristiques. Cela permet aux modèles d'apprendre des représentations plus riches et plus complexes des images, capturant ainsi des informations subtiles et des détails pertinents pour la classification des différents stades parasitaires.

Peu d'études se sont intéressées à la détermination des stades de développement du *Plasmodium*. Pour démontrer l'efficacité des réseaux profonds par rapport aux approches traditionnelles d'apprentissage automatique, une première étude [Bashar, 2019] a utilisé l'application d'un CNN pour identifier les stades de développement de *Plasmodiums vivax* (gamétocytes, trophozoïtes et schizontes). Les résultats obtenus, avec une exactitude de 97,7 %, sont supérieurs de 10 % à ceux obtenus avec un modèle SVM.

D'autres travaux ont abordé l'utilisation des modèles de détection d'objet (Faster R-CNN) [Ren et al., 2015] capable de détecter et de distinguer les stades de développement de *Plasmodium vivax* à partir de frottis sanguins minces [Hung and Carpenter, 2017]. Un modèle de détection d'objets fonctionne directement sur l'image, génère des boîtes englobantes et classe les globules rouges détectés, ce qui permet de s'affranchir d'une étape de segmentation permettant de séparer les cellules des autres régions dans l'image. Les globules rouges ont été détectés et séparés des autres composants sanguins, puis envoyés au modèle AlexNet pour les classifier plus précisément en fonction de leurs stades de développement. L'exactitude du système était de 98 %, tandis que l'exactitude moyenne obtenue en comparaison avec deux annotateurs humains n'était que de 72 % pour la même tâche d'identification des stades parasitaires. Cette évaluation de l'expertise diag-

nostique des microscopistes par rapport à la prédiction du modèle est importante pour valider les performances et le potentiel d'un tel système. Une grande variabilité a été observée entre les deux opérateurs testés, suggérant une différence significative d'expertise dans l'identification des stades parasitaires. Cela suggère que 2 experts sont insuffisants afin d'évaluer de manière pertinente l'efficacité du système. La variabilité dans l'interprétation diagnostique est un facteur important, qui justifie les développements de système de diagnostic automatique du paludisme, permettant de contribuer à réduire cette variabilité. Ce qui nous a conduit à étudier cette variabilité plus en détails (voir chapitre 4). Plus récemment, plusieurs études ont exploré l'utilisation de transformateurs visuels pour l'identification des stades parasitaires de *Plasmodium vivax* [Sengar et al., 2022] afin d'explorer l'interprétabilité des modèles.

1.4 Focus sur le serial-killer : *Plasmodium falciparum* et mesure automatique de la parasitémie

À ce jour, l'état de l'art montre une prédominance des études portant sur la détection du *Plasmodium* principalement concentrées sur les cas d'infection par *Plasmodium falciparum*. L'utilisation de CNN a été introduite pour la première fois pour l'identification d'érythrocytes infectés et non infectés par *Plasmodium falciparum* par Liang et al. en 2016 [Liang et al., 2016]. De bonnes performances ont été obtenues (exactitude de 97,37 %) après l'entraînement d'une architecture de CNN (from scratch) sur plus de 27 500 images. En comparaison, le CNN pré-entraîné, AlexNet [Krizhevsky et al., 2017], qui a une architecture plus complexe et est connue pour ses bonnes performances, s'est révélée être moins performante pour la même tâche de classification (exactitude de 91,99 %).

Néanmoins, d'autres études ont démontré l'efficacité des modèles pré-entraînés pour l'identification des parasites ou l'extraction des caractéristiques. Des réseaux pré-entraînés (AlexNet [Krizhevsky et al., 2017], VGG [Simonyan and Zisserman, 2015a] et GoogLeNet [Liu et al., 2016a]) ont été évalués dans une étude pour analyser des frottis épais pour la détection de la forme en anneau de *Plasmodium falciparum* [Mehanian et al., 2017b]. Les CNN ont été entraînés sur des frottis de 109 patients (78 positifs et 31 négatifs), et ont obtenu une exactitude de 90 %. Bien que cette étude soit la seule à avoir analysé des frottis épais sur un nombre relativement important de patients et à avoir obtenu des résultats prometteurs, les performances restent faibles pour des applications cliniques en routine. D'autres études ont comparé les performances des systèmes basés sur des modèles de détection d'objets à celles d'un panel d'experts afin de démontrer leur potentiel. Les résultats montrent que l'algorithme a une performance de niveau expert pour détecter les globules rouges infectés par *Plasmodium falciparum*, avec une précision moyenne de 88,5 % et une sensibilité de 99,5 %. Ces performances sont comparables à celles des microscopistes expérimentés [Kuo et al., 2020].

Le Transfer learning ou l'utilisation de modèles pré-entraînés améliorés se sont généralisés de manière croissante pour l'identification de *Plasmodium falciparum* [Rajaraman et al., 2018]. Dans cette étude, d'autres modèles CNN sont utilisés, ResNet-50 [He et al., 2015], a offert les meilleures performances (exactitude de 98,6 %) par rapport à d'autres modèles pré-entraînés comme DenseNet-121 [Huang et al., 2017] et Xception [Chollet, 2017]. L'entraînement a été réalisé sur l'ensemble de données du National Institute of Health (NIH) sur le paludisme (<https://lhncbc.nlm.nih.gov/publication/pub9932>) composé de plus de 27 000 images de globules rouges infectés par *Plasmodium falciparum* et non infectés de manière équilibrée obtenues à partir de frottis sanguins minces. Cependant, lors de l'évaluation des modèles sur différentes lames provenant de patients individuels, une baisse de performance a été observée en raison des variations de coloration sur

les lames (exactitude de 95,9 %). Il s'agit là d'un paramètre important à prendre en compte, car les lames de frottis sanguins peuvent présenter des colorations hétérogènes qui affectent facilement les résultats de la classification durant l'étape d'évaluation. Pour se rapprocher le plus possible des conditions diagnostiques réelles, les modèles ont été entraînés sur des images hétérogènes, en tenant compte de la variabilité des résultats de coloration uniquement et non de la diversité morphologique des images en termes de bruit ou d'artefacts. Les techniques d'augmentation artificielle des données permettent d'enrichir les données disponibles, d'apprendre de nouvelles représentations et de rendre un modèle plus efficace et plus généralisable. Une étude a proposé des améliorations de la base de données NIH [Rajaraman et al., 2018] afin de proposer de meilleurs scores de classification avec une exactitude de 97,64 % [Rahman et al., 2019]. Cependant, l'augmentation artificielle des images et l'application d'opérations de transformations, telles que les effets miroir, les rotations et les translations, n'a que légèrement amélioré les performances (exactitude de 97,89 %). Cela suggère qu'un jeu de données bien que diversifié et augmenté en termes de colorations et de transformations, peut présenter des limites importantes en matière de généralisation.

Les modèles génératifs ont également le potentiel de créer de nouvelles images ou de modifier des images existantes. Ils sont devenus incontournables ces dernières années pour différentes applications. Cette spécificité de modèles génératifs également appelée réseaux adversaires génératifs ou GAN en anglais, a été abordé dans de nombreux travaux pour générer de nouvelles images réalisées de cellules sanguines [Bailo et al., 2019]. L'approche implique l'utilisation de modèles génératifs, tels que le pix2pixHD [Wang et al., 2018], pour traiter des masques synthétiques d'images de cellules sanguines créés par le modèle FCN-8s [Long et al., 2015]. Un système capable de générer de nouvelles représentations de cellules sanguines serait intéressant pour générer des données souvent difficiles à acquérir en raison des contraintes de rareté tels que pour des espèces de *Plasmodiums* moins représentés avec une prévalence plus faible allant de 5 à 20 % (ovale ou malariae). Cela peut permettre de combler le déséquilibre des classes lors de l'entraînement des modèles. D'autres études ont également exploré l'utilisation de transformateurs visuels sur le jeu de données du NIH pour la détection de *Plasmodium falciparum* [Islam et al., 2022].

Les approches d'apprentissage profond et autres approches basées sur l'apprentissage automatique offrent la possibilité de mesurer la parasitémie de manière automatique sur des cas de paludisme à *Plasmodium falciparum* de manière automatique. Un système de mesure automatique de la parasitémie présente plusieurs avantages comme l'amélioration des mesures plus précises et plus fiables de la parasitémie, qui reste un critère de gravité important de l'infection. En outre, sa rapidité permet d'avoir des mesures en un temps beaucoup plus court par rapport à une évaluation manuelle. L'utilisation de CNN a été abordée pour détecter et compter le nombre de globules blancs et mesurer la parasitémie sur des images de frottis sanguins épais, où la concordance de quantification avec la microscopie experte a montré des coefficients de corrélation faibles allant de 0,59 à 0,86 [Torres et al., 2018]. Certaines études ont montré qu'il peut y avoir une faible corrélation entre les résultats de la microscopie manuelle et les mesures estimés par les modèles profonds avec un coefficient de corrélation ($R^2=0,55$) [Manescu et al., 2020] ou une surestimation du modèle pour des charges parasitaires <1000 p/ μ L [Mehanian et al., 2017a].

Parmi les quelques tentatives consacrées à l'évaluation de la parasitémie à partir de frottis sanguins minces avec d'autres techniques, des approches traditionnelles et des réseaux neuronaux ont été utilisées, tels que les SVM et les perceptrons multicouches (MLP), pour mesurer la parasitémie sur des caractéristiques pré-extraites de globules rouges à partir d'images de frottis sanguins minces [Le et al., 2008]. De fortes corrélations ont été observées ($R^2=0,97$) entre les résultats manuels et semi-automatiques; cependant, la taille de l'échantillon reste faible, avec un total de 2

500 globules rouges analysés sur 31 patients (19 patients atteints de paludisme et 12 témoins non infectés) [Linder et al., 2014b]. D'autres études ont abordé l'utilisation d'un réseau neuronal artificiel (ANN) pour la classification de caractéristiques pré-extraites comparée à la cytométrie en flux commerciale [Poostchi et al., 2018a]. L'évaluation réalisée sur 2 souris avec un suivi de 2 semaines a également montré une surestimation de la parasitémie due à la présence de faux positifs induits par des artefacts de coloration. La taille de l'échantillon reste réduite et ne permet pas une évaluation précise de la corrélation entre les 2 techniques.

Ces études démontrent et confirment le potentiel de l'apprentissage automatique pour le diagnostic automatique du paludisme. L'utilisation croissante des CNN a permis d'identifier les limites et les perspectives d'améliorations en termes de données d'apprentissage. En effet, assurer la reproductibilité des résultats reste un défi majeur, qui nécessitera des évaluations multiples à grande échelle en prenant en compte une diversité morphologique proche de la réalité diagnostique ou encore de comparer les performances avec des opérateurs et experts humains.

1.5 Quelles solutions sont disponibles ? Applications mobiles et systèmes «End-to-End» pour le diagnostic automatique du paludisme

Associées à des modèles d'apprentissage automatique, les systèmes automatisés développés se déclinent en deux types : les applications en ligne et les systèmes « end-to-end ». Parmi les travaux publiés, des équipes ont proposé des preuves de concepts d'applications Android [Zhao et al., 2020], combinant des modèles de détection d'objets et modèles pré-entraînés, tels que SSD300 [Liu et al., 2016b] et VGG-16. L'application permet de classer des images individuelles de globules rouges infectés par *Plasmodium vivax* et *Plasmodium falciparum* sur des frottis sanguins minces au x1 000. L'application a été évaluée pour la détection des globules rouges (exactitude de 90,4 %) et la classification en infectés et non infectés (exactitude de 96,5 %). Elle permet également d'avoir une estimation de la charge parasitaire mais cette étape n'a pas été évaluée avec d'autres techniques afin de démontrer sa fiabilité. Les auteurs ont également utilisé des modèles d'augmentation de résolution (FSRCNN) [Dong et al., 2016] pour améliorer la qualité des images. Les modèles entraînés reposent principalement sur les jeux de données accessibles du NIH (<https://lhncbc.nlm.nih.gov/publication/pub9932>) [Rajaraman et al., 2018] pour la détection de *Plasmodium falciparum* et le jeu de données de l'Institut Broad (<https://bbbc.broadinstitute.org/BBBC041>) pour la détection de *Plasmodium vivax*. Le jeu de données de *Plasmodium vivax* comprennent 1 364 images de frottis sanguins infectés et non infectés, avec 80 000 globules rouges individuellement annotés, et un taux de globules rouges infectés de 5 %. L'application offre également la possibilité de charger d'autres modèles de classification. [Yu et al., 2020] ont également développé Malaria Screener, une application disponible sous Android pour le diagnostic du paludisme à partir de frottis minces et épais capturés grâce à un smartphone. Les performances d'évaluation atteignent une exactitude allant de 96,9 % à 98,6 % pour l'analyse de frottis sanguins épais [Yang et al., 2019] et minces [Yu et al., 2020], respectivement.

Les systèmes «End-to-End» pour le diagnostic du paludisme se distinguent des solutions présentées précédemment. Ces applications permettent de s'affranchir de la tâche d'analyse au microscope, mais la préparation manuelle de l'échantillon reste une étape manuelle et nécessaire. Cependant, il est important de noter que la qualité de la préparation de l'échantillon peut avoir un impact significatif sur les performances et la qualité des modèles de classification d'images. Des erreurs de coloration ou d'autres problèmes liés à la préparation de l'échantillon peuvent entraîner des résultats moins fiables. Par conséquent, il est essentiel de standardiser et de contrôler rigoureusement cette étape afin d'obtenir des performances optimales des systèmes automatisés de

diagnostic du paludisme. C'est pour cela que les systèmes «End-to-End» sont considérés comme plus performants dans le domaine du diagnostic du paludisme. En intégrant l'ensemble du processus, de l'acquisition de l'image à l'analyse automatisée, ces systèmes réduisent la dépendance aux étapes manuelles et aux variations associées à la préparation des échantillons. Ils offrent ainsi une approche plus cohérente et standardisée, ce qui améliore la fiabilité et la précision des résultats. Ce type de solutions se compose généralement d'un système optique et d'un dispositif de capture pour numériser les lames microscopiques, contrairement aux applications pour smartphones qui dépendent fortement de la qualité de leur équipement de capture mais aussi des conditions dans lesquelles les images sont prises (stabilité, source de lumière variable, mauvais contraste, etc...). Un système visant à standardiser et normaliser l'acquisition d'images facilitera l'analyse des frottis sanguins et l'identification précise des parasites. Pour répondre à ces défis, différents travaux ont été réalisés comme les scanners de lames basés sur le focus stacking pour l'analyse des frottis sanguins minces [Gopakumar et al., 2017]. Les globules rouges infectés par *P. falciparum* ont été identifiés en analysant les régions candidates et les meilleures images de mise au point des objets d'intérêt. Le scanner de frottis portable a montré une sensibilité de 97 % et une spécificité de 98,5 %. Les stratégies de mise au point automatique permettent de réduire le taux d'erreurs du modèle CNN passant de 1051 à 912 faux positifs. Cependant, l'estimation de la parasitémie reste imprécise en raison de la surestimation importante de globules rouges infectés. Divers systèmes sont ou sont en cours de commercialisation sur le marché. MotiC digital® propose un automate d'analyse des frottis sanguins, avec une sensibilité de 91,1 % et une spécificité de 73,6 % pour la détection de *Plasmodium falciparum* (<https://motidigitalpathology.com>). D'autres systèmes sont accessibles sur le marché, notamment CellaVision DM96, qui est principalement utilisé dans les services d'hématologie. Cette automate permet d'analyser environ 2 000 hématies par lame, une évaluation sur 220 lames, dont 98 lames de patients infectés et 122 lames négatives, a permis de mettre en évidence des performances faibles pour la détection de globules rouges infectés par *P. falciparum* (14 lames) et *P. vivax* (84 lames). Les scores de sensibilité (23,5 %) et de spécificité (81,1 %) résultent de taux de faux négatifs et positifs élevés [Yoon et al., 2019]. Un autre système prometteur, entièrement automatisé proposé par Noul®, miLab™, capable d'analyser environ 400 à 500 images de frottis sanguins minces en 15 minutes. La plateforme intègre la préparation de l'échantillon y compris la coloration et l'analyse des lames pour la détection et la quantification pour les espèces de *P. falciparum* et *P. vivax*. Une évaluation sur 1 869 patients a démontré des scores de sensibilité et de spécificité de 95 %.

Même si des progrès ont été réalisés dans le domaine du diagnostic automatisé du paludisme ces dernières années, les applications présentées montrent que certaines barrières dans la conception d'un système de diagnostic automatisé du paludisme persistent et n'ont été que partiellement résolues :

1. **Manque de validation sur un grand nombre de patients** : Beaucoup de ces études sont réalisées sur un nombre limité de patients, ce qui limite leur généralisabilité. Pour confirmer l'efficacité et l'efficacité de ces systèmes, il est nécessaire de mener des études sur de plus grands échantillons de population, incluant une diversité de contexte.
2. **Manque de diversité des données** : Les modèles d'apprentissage profond sont très bons lors des processus d'entraînement et moins bons lors des validations sur de nouveaux jeux de données. Si les données d'entraînement ne reflètent pas la diversité des conditions réelles de diagnostic, le modèle risque de ne pas être performant dans des conditions réelles. Par exemple, si un modèle est formé uniquement sur des images ne présentant pas de bruit ou d'artefacts, il risque de ne pas être performant sur des images plus diverses et variables, ce

qui est courant dans les régions à faibles ressources.

3. **Utilisation de bases de données limitées** : De nombreuses études utilisent des bases de données publiques (NIH, Broad Institute, etc...) qui sont souvent de bonne qualité et ne présentent pas beaucoup de bruit. Cela peut conduire à une surestimation de la performance du modèle lorsqu'il est appliqué à des images réelles qui peuvent présenter du bruit et d'autres artefacts ou encore d'autres objets d'intérêts non pris en compte.
4. **Un déficit notable de travaux portant sur la mesure automatique de la parasitémie pour les cas de *P. falciparum*** : La plupart des études se sont concentrées sur la détection des parasites, mais peu ont abordé la question de la mesure automatique de la parasitémie, une information cruciale pour le diagnostic et le suivi de la maladie.
5. **Coût élevé des automates et des systèmes «End-to-End»** : Bien que ces systèmes puissent améliorer l'efficacité du diagnostic de manière significative, leur coût élevé peut être un obstacle majeur, en particulier dans les régions à faibles ressources où surviennent la majorité des cas de paludisme. En plus du coût initial d'achat de l'équipement, il peut y avoir des coûts supplémentaires pour la formation, la maintenance et la mise à jour du système.
6. **Complexité et nécessité de formation** : L'utilisation de ces systèmes peut nécessiter une formation supplémentaire pour les professionnels de la santé ou les utilisateurs finaux, ce qui peut être un défi dans les régions où les ressources sont limitées.

TABLE 2.1 – Synthèse des approches pour le diagnostic automatique du paludisme.

Catégorie	Avantages	Inconvénients	Performances	Références
Approches traditionnelles et traitement d'images	Permet la détection précise des globules rouges parasités et non infectés.	Nécessite une expertise pour la sélection manuelle des caractéristiques.	Variabilité des performances (80% à 99% d'exactitude).	Savkare and Narote [2015a], Nasir et al. [2012], Khan et al. [2014], Linder et al. [2014a], Hearst [1998], Savkare and Narote [2015b], Abdul Nasir et al. [2012], Das et al. [2015]
Apprentissage profond pour l'identification de stades parasitaires de <i>Plasmodium</i>	Extraction automatique de caractéristiques.	Difficulté de définition précise des caractéristiques.	Supériorité des réseaux profonds avec des exactitudes allant jusqu'à 97,7%.	Bashar [2019], Ren et al. [2015], Hung and Carpenter [2017], Sengar et al. [2022]
Mesure automatique de la parasitémie pour <i>Plasmodium falciparum</i>	Amélioration de la mesure de la parasitémie.	Limitation sur la précision de l'estimation de la parasitémie.	Performances variables avec des sensibilités et spécificités atteignant 95%.	Liang et al. [2016], Krizhevsky et al. [2017], Simonyan and Zisserman [2015a], Liu et al. [2016a], Mehanian et al. [2017b], Kuo et al. [2020], Rajaraman et al. [2018], He et al. [2015], Huang et al. [2017], Chollet [2017], Rahman et al. [2019], Bailo et al. [2019], Wang et al. [2018], Long et al. [2015], Islam et al. [2022]
Applications mobiles et systèmes «End-to-End»	Accessibilité accrue et facilité d'utilisation.	La qualité de la préparation de l'échantillon impacte les performances.	Exactitudes allant de 90,4% à 98,6% pour différentes applications.	Zhao et al. [2020], Liu et al. [2016b], Dong et al. [2016], Rajaraman et al. [2018], Yu et al. [2020], Yang et al. [2019], Gopakumar et al. [2017], Yoon et al. [2019]

Cette thèse vise à combler plusieurs lacunes significatives dans le domaine du diagnostic du paludisme assisté par l'IA. Nous allons nous attaquer au manque de généralisation des systèmes en évaluant notre approche sur des données diverses, y compris celles du NIH. Nous allons également aborder la question sous-explorée de la mesure automatique de la parasitémie, qui est essentielle pour le diagnostic et le suivi précis du paludisme. En utilisant la microscopie optique et la capture par smartphone, nous visons à rendre notre système à la fois abordable et accessible, même dans les régions à ressources limitées. Ainsi, ce travail aspire à faire progresser la capacité de l'apprentissage automatique à améliorer le diagnostic du paludisme dans les contextes où ils sont les plus nécessaires.

Identification de globules rouges parasités par *Plasmodium falciparum* sur des images de frottis sanguins minces

1 Contexte et Objectifs

Jusqu'à présent, toutes les études montrent des performances satisfaisantes sur des échantillons majoritairement infectés, ce qui ne permet pas vraiment d'évaluer la spécificité des systèmes proposés. De plus, l'évaluation des systèmes publiés jusqu'à présent montre qu'il existe encore plusieurs limites, notamment le sur-apprentissage sur le jeu de données ainsi que l'absence d'une évaluation externe indépendante des modèles sur un ensemble de données supplémentaire de patients. Seule une telle évaluation permet de reproduire les différentes méthodes et de les comparer à l'expertise humaine. De plus, le jeu de données rendu accessible par le NIH, acquis avec un grossissement de 1 000x (objectif d'immersion d'huile 100x et oculaire 10x), est principalement utilisé dans les travaux scientifiques car il facilite la détection des éléments parasitaires. Cependant, une telle résolution entraîne un inconvénient important, notamment à travers le nombre de globules rouges examinés à partir d'un seul champ qui reste faible (environ 200 globules rouges par champ). Il est recommandé, dans la pratique actuelle, d'examiner au moins 200 champs avant de conclure un diagnostic. Pour faciliter la transition vers un système d'analyse automatique, il est important de réduire le nombre de champs capturés à analyser et donc de réduire le grossissement pour maintenir le même nombre de globules rouges à analyser par le système.

Dans ce chapitre nous décrivons les contributions apportées par nos travaux afin de répondre aux limites de l'état de l'art décrites précédemment. Pour cela, nous allons décrire le système semi-automatisé développé pour la détection de globules rouges infectés par *P. falciparum* à partir d'images de frottis sanguins minces capturées au grossissement de x500 (objectif d'immersion d'huile x50 et oculaire x10). Nous avons dans un premier temps, construit un jeu de données composé d'images collectées dans des conditions diagnostiques de routine en utilisant une méthodologie originale permettant de renforcer de manière considérable la diversité des classes. Dans un second temps, nous avons comparé différentes architectures de réseaux de neurones à convolution et leur performance sur le jeu de données construit localement ainsi que sur le jeu de données du NIH. L'objectif de cette étude était de comparer les performances en termes de généralisation

entre les modèles pré-entraînés sur notre jeu de données CNR et ceux pré-entraînés sur le jeu de données du NIH à travers une étape de validation sur un échantillon de patients indépendant.

Plus précisément :

1. Nous avons construit un jeu de données basé sur des images de cellules sanguines infectées et non infectées par *P. falciparum*, augmentées par des artefacts et d'autres composants sanguins afin de rapprocher de la réalité diagnostique.
2. Le grossissement au x500 pour le jeu de données CNR est inférieur à celui du NIH (x1 000), ce qui permet d'analyser 4 fois plus d'images de cellules par champ microscopique, permettant ainsi d'analyser une quantité d'hématies plus importante par patient.
3. Nous avons effectué une étude de validation croisée avec et sans modèles pré-entraînés pour la classification d'images de globules infectés par *P. falciparum* et des images de composants sanguins non infectés.
4. Nous démontrons que les modèles entraînés sur notre jeu de données CNR avec un grossissement plus faible présentent de meilleures performances en termes de généralisation au niveau cellulaire mais aussi au niveau patient par rapport au jeu de données NIH.

2 Matériel et Méthodes

2.1 Description du jeu de données du NIH

Le jeu de données sur le paludisme du NIH [Rajaraman et al., 2018] contient 27 558 images individuelles de globules rouges, déjà segmentées et annotées. La moitié de ces cellules sont non infectées, tandis que l'autre moitié est infectée par *P. falciparum* (soit 13 779 images de globules rouges pour chaque classe). Les images acquises (<https://lhncbc.nlm.nih.gov/publication/pub9932>) proviennent de frottis sanguins minces colorés au Giemsa de 50 patients sains et 150 patients infectés par *P. falciparum*, collectés et capturés à l'Hôpital Médical de Chittagong au Bangladesh avec une magnification de x1 000. Les images ont été capturées à l'aide de deux différents microscopes optiques d'Olympus et Motif. L'annotation des images a été réalisée à l'Unité de Recherche de Médecine Tropicale Mahidol-Oxford à Bangkok, Thaïlande.

2.2 Construction d'un jeu de données CNR représentatif de la réalité diagnostique

Nous décrivons ici comment nous avons construit notre jeu de données d'images de cellules sanguines infectées par *P. falciparum* et non infectées à partir des images de frottis sanguins minces décrites ci-dessus. La figure 3.1.a et 3.1.b illustrent le processus global de construction de cette base d'images.

Tout d'abord, nous avons collecté des données du laboratoire de parasitologie du CNR pour le paludisme de la Pitié-Salpêtrière et du laboratoire d'hématologie de l'Hôpital Saint Antoine, en France. Au total, les données de 101 patients infectés par *P. falciparum* et de 101 patients sains ont été collectées de 2019 à 2021. Les frottis sanguins colorés au Giemsa ont été réalisés par 4 différents microscopistes et capturés en utilisant un appareil photo de smartphone fixé à un microscope optique au format JPG RGB. Pour ces expériences, nous avons utilisé un microscope optique Olympus BX51 pour capturer les frottis à un grossissement de X500. Un total de 1 250 images ont été capturées sous un microscope optique. Deux smartphones différents ont été utilisés pour acquérir les images : la caméra arrière de 12 mégapixels d'un iPhone 7 et un Google Pixel 6 avec un système de caméra arrière double composé de lentilles grand angle de 50 et 12 mégapixels.

Dans un second temps, une étape de segmentation des objets d'intérêt a été réalisée en utilisant l'algorithme de watershed [Beucher, 1992], qui permet de détecter chaque composant sanguin dans l'image du frottis sanguin. L'algorithme de watershed est une technique de segmentation d'images qui fonctionne en visualisant l'image en niveaux de gris comme un paysage topographique. Les pixels les plus sombres, qui correspondent aux cellules et autres composants sanguins, sont interprétés comme des vallées, tandis que les pixels plus clairs sont interprétés comme des sommets. L'algorithme "inonde" ensuite ce paysage à partir des vallées, en construisant des barrages pour empêcher les eaux de différentes vallées de se mélanger. Ces barrages forment les lignes de partage des eaux, qui correspondent aux contours des composants sanguins que nous voulons segmenter. Ainsi, chaque composant sanguin individuel est identifié et séparé des autres. Une étape finale consisté à utiliser un réseau de neurones à convolution (master) afin d'alléger la tâche d'annotation faites par les experts humains afin de différencier les GRP des autres composants sanguins.

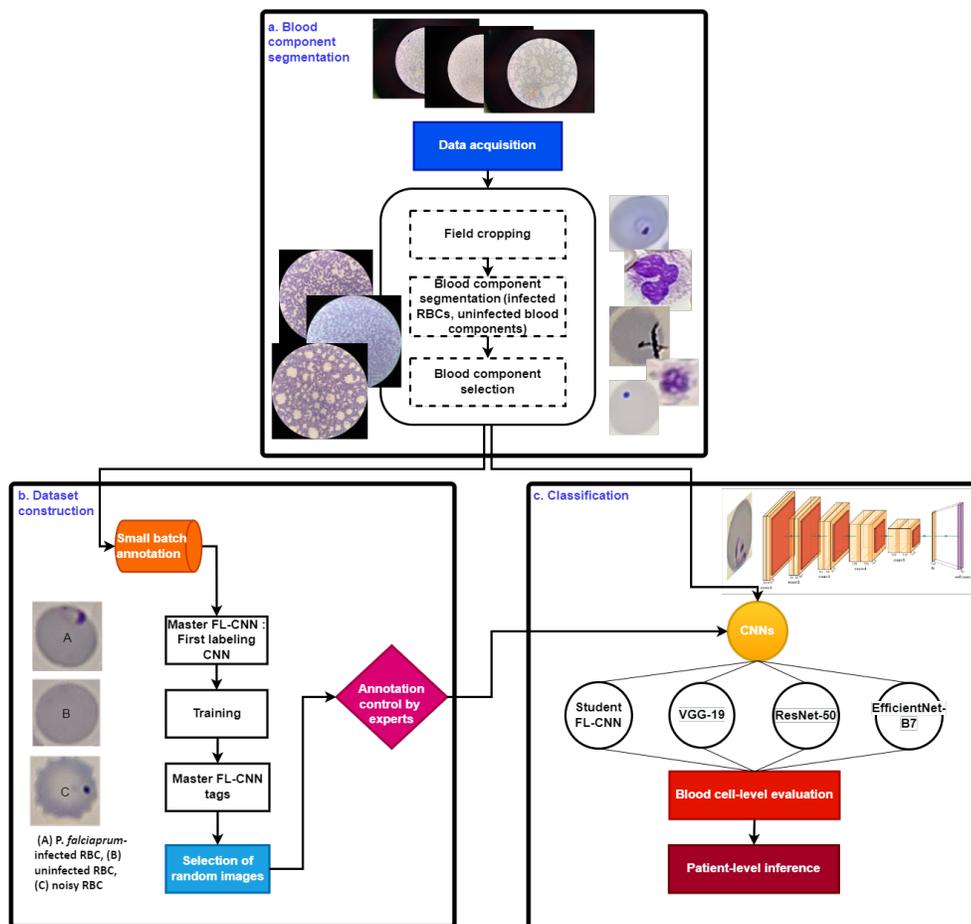


FIGURE 3.1 – Diagramme de construction de la base d'images CNR

Sélection et annotation des images de composants sanguins

Suite à l'étape de segmentation, les images incluent différentes formes de composants sanguins. Les cellules sanguines infectées contiennent des parasites *P. falciparum* avec des morphologies différentes; les parasites peuvent être en périphérie, ou plusieurs parasites peuvent également être présents à l'intérieur du globule rouge. D'autres images contiennent différents types d'objets, y compris des globules blancs, des plaquettes, des cellules sanguines avec des artefacts de coloration ou des tâches et des globules rouges avec des corps de Howell-Jolly. Les cellules non infectées étant protéiformes, celles-ci nécessitent un tri supplémentaire crucial afin de séparer les cellules parfaitement saines de celles comportant un objet biologique ou une impureté, pour cela nous utilisons un premier modèle permettant de traiter cette granularité. La diversité biologique des objets colorés segmentés est illustrée par la figure 3.2.

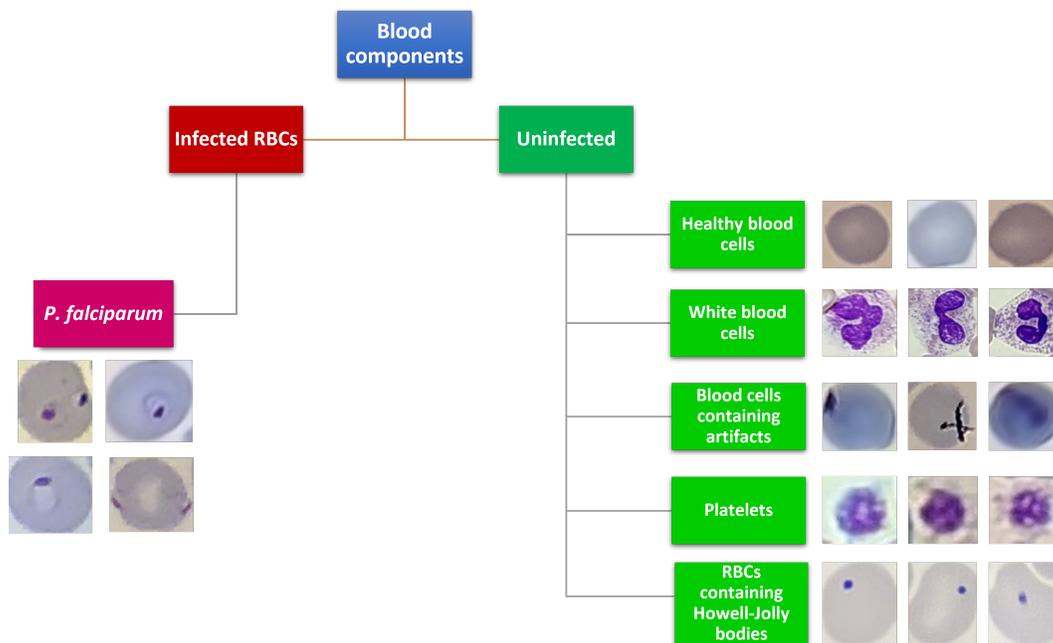


FIGURE 3.2 – La diversité biologique des composants sanguins inclus dans le jeu de données CNR.

FL-CNN : Un premier modèle d'annotation automatique des images segmentées

FL-CNN (first labeling CNN) est un premier réseau de neurones à convolution, entraîné sur un petit ensemble d'images. L'architecture du modèle repose sur 5 couches de convolution et deux couches entièrement connectées avec une entrée de taille $50 \times 50 \times 3$ pixels (voir figure 3.3). Les couches de convolutions sont caractérisées par des filtres de taille 3×3 avec un pas d'un pixel suivi d'une fonction d'activation unité linéaire (ReLU). La couche de max-pooling comprend une fenêtre de 2×2 , et une couche de « batch normalisation ». La première couche dense entièrement connectée à 512 neurones, et la deuxième couche dense est suivi d'un classifieur softmax. L'optimiseur utilisé est RMSProp.

Ce processus consistant à entraîner un premier modèle sur un petit jeu de données afin d'avoir un modèle d'annotation. La stratégie de « Noisy Student » [Xie et al., 2020] a été appliquée afin d'ob-

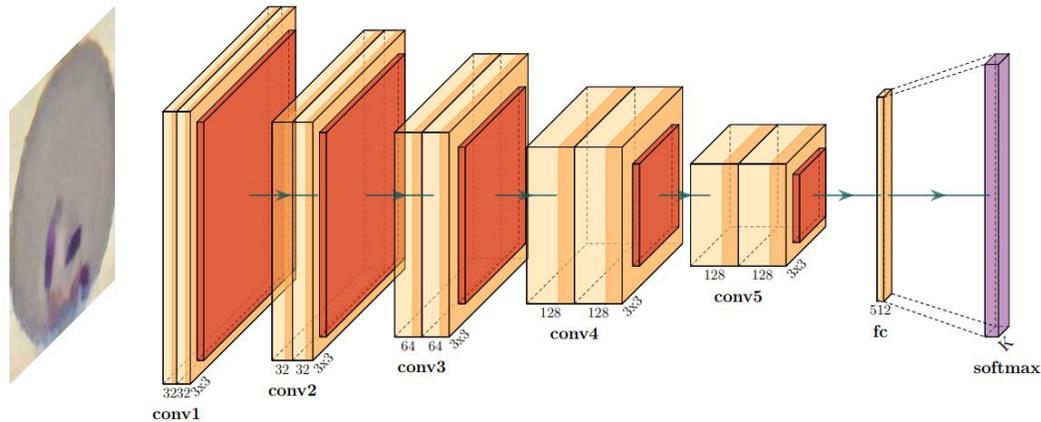


FIGURE 3.3 – Architecture du modèle FL-CNN à 5 couches avec $K = 3$ classes pour la classification en tant que modèle d’annotation.

tenir ce premier modèle d’annotation automatique. Le modèle FL-CNN (master) a été entraîné sur 951 images annotées manuellement pour une classification des objets sanguins en 3 classes : 214 images de globules rouges infectés par *P. falciparum*, 610 images de globules rouges sains, et 127 images d’autres composants sanguins, comprenant non seulement des globules rouges potentiellement infectés mais difficilement interprétables en raison de leur faible qualité, mais aussi des globules blancs, des plaquettes et des artefacts ou tâches de colorant. La classification multi-classe du modèle FL-CNN permet d’obtenir une pré-annotation satisfaisante des composants sanguins d’intérêt, avec seulement quelques données étiquetées. Les hyperparamètres des modèles expérimentaux sont présentés dans le tableau 2. Pour toutes les expériences, les images ont été normalisées à une résolution de 50×50 pixels. Les modèles ont été entraînés et évalués sur Ubuntu 20.04.3 LTS avec un processeur graphique (GPU) Nvidia TITAN Xp, Python 3.7.7, Keras 2.4.3 avec TensorFlow 2.5.0 et CUDA 11.2.

La validation humaine des images pré-annotées par le modèle permet de reclasser l’ensemble des images en 2 classes :

1. Une classe négative comprenant des images ne présentant aucune tâche ou autre objet à l’intérieur de leurs membranes, celles-ci ont été systématiquement validées comme composants sanguins non infectées, elles sont les plus représentées et les plus simples à confirmer. Les images de composants sanguins non infectées incluant la présence de tâches, plaquettes, corps de Howell-Jolly ou des globules blancs) ont également été sélectionnés et soumis à une validation par des experts humains afin d’enrichir la classe négative.
2. Une classe positive comportant des globules rouges infectés par *P. falciparum*.

Quatre experts, médecins biologistes et techniciens du laboratoire de parasitologie mycologie de l’Hôpital Pitié-Salpêtrière, ont participé à ce processus d’annotation. Ce processus a conduit à l’obtention d’un jeu données CNR diversifié biologiquement et morphologiquement composé de cellules issues de patients infectés et non infectés.

Préparation d'un jeu de données de validation indépendant

Nous avons également sélectionné un panel de 200 nouveaux patients aléatoires pour évaluer la qualité des modèles entraînés. Nous avons collecté 50 patients infectés par *P. falciparum* et 150 patients négatifs non infectés ayant été testés pour un diagnostic du paludisme durant la période de fin 2021 et début 2022. Parmi les patients non infectés, certaines images contiennent des corps de Howell-Jolly et d'autres artefacts ou impuretés pouvant complexifier la tâche de classification des globules rouges parasités pour le système.

Pour chaque patient, nous avons confirmé l'infection des globules rouges parasités par *P. falciparum* en confirmant la présence de parasites au microscope optique en réexaminant directement la lame concernée. Concernant les globules rouges positifs douteux, nous avons contrôlé leur infection en optimisant les paramètres microscopiques durant la capture de l'image en faisant augmenter la résolution ou en calibrant la profondeur de mise au point.

Les expériences menées précédemment nous ont permis jusqu'ici d'obtenir un jeu de données CNR riche et diversifié représentatif des conditions réelles d'un examen microscopique. Celui-ci va nous permettre d'entraîner différentes architectures de modèles profonds à travers une validation croisée.

2.3 Validation croisée et classification des images du jeu de données CNR

Nous décrivons ici les modèles utilisés pour classer les composants sanguins en globules rouges infectés par *P. falciparum* et en composants sanguins non infectés. Nous avons comparé les performances des quatre architectures pour la classification des images de composants sanguins : un modèle à convolution C-CNN (Student) basé sur l'architecture du modèle utilisé pour l'annotation FL-CNN avec une sortie binaire ($k=2$), ainsi que trois architectures de modèles pré-entraînés, VGG-19 [Simonyan and Zisserman, 2015b], ResNet-50 [He et al., 2015] et EfficientNet-B7 [Tan and Le, 2020]. Nous avons également comparé les performances obtenues par l'entraînement sur les ensembles de données CNR et NIH pour évaluer leur efficacité et la généralisation des modèles pour de nouveaux patients. Les blocs de « feature extraction » de VGG-19, ResNet-50 et EfficientNet-B7 sont suivis de deux couches denses entièrement connectées composées de 128 neurones et d'une fonction d'activation ReLu. Nous avons utilisé l'optimiseur Adam. Notons que la taille des images a également été normalisée pour l'entraînement de ces modèles : 50x50 pixels pour les jeux de données CNR et NIH.

Une validation croisée en 5-folds ou sous-ensembles a été appliquée sur les ensembles de données CNR et NIH. Les critères de performance traditionnels ont été utilisés volet, l'exactitude ou accuracy (proportion d'images correctement classées), la précision (ou la valeur prédictive positive), la sensibilité (ou le rappel), la spécificité et le F1-score (moyenne harmonique de la précision et du rappel), pour la classe de globules rouges infectés. Le meilleur modèle sélectionné est celui qui offre le meilleur F1-score parmi les 5 folds. Une estimation des empreintes carbone pour chaque système a été fournie, en utilisant l'outil CarbonTracker [Anthony et al., 2020].

2.4 Evaluation rétrospective : Généralisation des modèles sur le jeu de validation

Nous avons validé la performance de notre meilleur modèle au niveau des patients en utilisant le jeu de données de validation incluant 200 patients supplémentaires : 150 patients non infectés et 50 patients infectés. Pour chaque patient, 2 images de frottis sanguins minces sont capturés pour l'analyse à l'aide de deux sources de microscopes optiques, les microscopes Olympus BX51 et Leitz Laborlux S. Nous avons comparé les performances des meilleurs modèles parmi les 5 folds

de la validation croisée sur les ensembles de données NIH et NRC en termes de sensibilité et de spécificité de deux manières :

- Classification au niveau cellulaire des images composants sanguins pour chaque patient
- Classification des patients au niveau individuel en 2 classes : infecté (avec présence d'au moins 1 globule rouge infecté) ou non infecté.

Nota bene : Notez que la distribution des patients dans l'ensemble de validation, avec 50 patients positifs et 150 patients négatifs, rend l'utilisation de mesures telles que la précision ou le rappel non pertinentes pour la classification des patients individuels comme infectés ou non infectés. En effet, ces mesures dépendent de la distribution des patients. Dans notre cas, la proportion de patients négatifs est trois fois plus élevée que celle des patients positifs, ce qui peut biaiser ces mesures. Par conséquent, il est important de prendre en compte cette distribution lors de l'évaluation de la performance des modèles de classification.

La suite de ce chapitre sera consacrée à l'illustration des résultats obtenus pour chaque méthodologie.

3 Résultats

3.1 Annotation en deux étapes des images composants sanguins : FL-CNN et validation des experts

La segmentation de 1 250 images microscopiques provenant de 202 lames de sang (c'est-à-dire, 101 patients infectés par *P. falciparum* et 101 patients non infectés) a conduit à l'extraction de 1 360 615 images d'objets colorés. Parmi les 101 patients positifs, le modèle FL-CNN à trois classes a identifié 53 101 images de globules rouges infectés par *P. falciparum*, 1 065 295 images de globules rouges sains et 65 004 images d'autres composants sanguins non infectés. Pour les 101 patients non infectés, le modèle a identifié 26 006 images de globules rouges infectés par *P. falciparum* (faux positifs de préannotation), 156 066 images de globules sanguins sains et 1 143 images incluant d'autres composants sanguins non infectés, tels que des plaquettes ou des globules blancs. L'annotation finale fusionnée des patients infectés et non infectés déterminée par le modèle FL-CNN comprend 79 107 images de globules rouges infectés par *P. falciparum*, 1 215 361 images de globules rouges sains et 66 147 images d'autres composants sanguins, y compris des globules blancs, des plaquettes, des artefacts et des globules sanguins parasités douteux.

Après validation par les quatre experts en diagnostic du paludisme, nous avons sélectionné un ensemble de données final à deux classes composé de 68 485 images de cellules avec des proportions : 14 752 images de globules rouges contenant des parasites *P. falciparum* et 53 733 images de composants sanguins non infectés. Parmi les globules rouges non infectés, 32 904 provenaient de patients négatifs et 20 829 provenaient de patients positifs.

Nota bene : Notez que de nombreuses images négatives de pré-annotation ont été ignorées dans l'ensemble de données final. Cette sous-échantillonnage est un moyen courant de rééquilibrage d'une distribution de classes naturellement et massivement déséquilibrée en apprentissage automatique lorsqu'il s'agit de grands ensembles de données [Beckmann et al., 2015].

3.2 Validation croisée : évaluation au niveau cellulaire

Le Tableau 3.2 présente les performances globales de classification des quatre modèles (Student C-CNN, VGG-19, ResNet-50 et EfficientNet-B7) sur nos ensembles de données CNR et NIH avec les paramètres définis dans le tableau 3.1, en moyenne sur les 5 partitions. VGG-19, ResNet-50 et

EfficientNet-B7 ont obtenu les meilleurs résultats. Les trois modèles affichent des scores d'exactitude approchant les 99 %. Cependant, une légère supériorité peut être observée pour VGG-19 (exactitude de $99,7 \% \pm 0,4 \%$) et ResNet-50 (exactitude de $99,4 \% \pm 0,2 \%$), tandis que le modèle (Student) C-CNN présente les performances les plus faibles (exactitude de $98 \% \pm 0,1 \%$). Ces performances indiquent une capacité élevée à distinguer les globules rouges infectés par *P. falciparum* des globules sains et d'autres composants sanguins.

Ces résultats suggèrent que VGG-19, ResNet-50 et EfficientNet-B7 sont des modèles très performants pour identifier des globules rouges parasités à partir d'images microscopiques au x500. Cependant, il convient de prendre en compte d'autres facteurs comme le fait que la validation croisée ait été faite sur l'ensemble des images et que les images du jeu de test peuvent provenir du même patient. Il faut également considérer que les erreurs de prédiction faites par les différents modèles ont été vérifiées et certaines images ont été re-labellisées. Il s'agit ici d'un apprentissage final après de multiples mise à jour de labels pour certaines images notamment négatives qui peuvent prêter à confusion lors d'une validation humaine.

TABLE 3.1 – Paramètres des modèles entraînés.

Modèle	Nb de Convolutions	Taille des Filtres	Nb de MaxPooling	Couches Dense	Nb de Neurones par Couche Dense	Poids initialisés	Fonction d'activation finale
C-CNN	5	3x3 pour toutes les convolutions	5	3	512, nombre de classes=2	Non	Softmax
VGG-19	16	3x3 pour toutes les convolutions	5	3	128, nombre de classes=2	Imagenet	Softmax
ResNet-50	53	1x1, 3x3 (blocs convolutifs variés)	Variable	3	128, nombre de classes=2	Imagenet	Softmax
EfficientNet-B7	179	Variable (3x3 principalement, ajustée)	Variable	3	128, nombre de classes=2	Imagenet	Softmax

Model	Accuracy	Precision		Sensitivity		Specificity		F1-Score	
		NIH	CNR	NIH	CNR	NIH	CNR	NIH	CNR
C-CNN	0.980 ± 0.001	0.953 ± 0.004	0.945 ± 0.010	0.961 ± 0.013	0.938 ± 0.004	0.985 ± 0.004	0.952 ± 0.004	0.953 ± 0.005	0.952 ± 0.004
VGG-19	0.997 ± 0.003	0.985 ± 0.001	0.971 ± 0.011	0.974 ± 0.018	0.984 ± 0.006	0.992 ± 0.003	0.986 ± 0.010	0.971 ± 0.014	0.986 ± 0.012
ResNet-50	0.994 ± 0.002	0.982 ± 0.001	0.991 ± 0.004	0.985 ± 0.005	0.984 ± 0.008	0.997 ± 0.001	0.980 ± 0.008	0.988 ± 0.005	0.980 ± 0.009
EfficientNet-B7	0.993 ± 0.004	0.988 ± 0.003	0.986 ± 0.010	0.980 ± 0.011	0.981 ± 0.008	0.996 ± 0.002	0.989 ± 0.003	0.983 ± 0.009	0.985 ± 0.004

TABLE 3.2 – Performance des modèles (moyenne \pm écart-type) pour la classification (5 folds) sur les deux bases de données : CNR (202 patients) vs. NIH (200 patients).

L’empreinte carbone des modèles entraînés est un aspect important à considérer, car elle reflète l’impact environnemental de leur développement. Selon les résultats obtenus à l’aide de CarbonTracker, le modèle C-CNN présente la plus petite empreinte carbone parmi les modèles évalués, avec une consommation d’énergie de 3,52 kWh pour chaque fold d’entraînement. Cela équivaut à environ 1 035 g eqCO₂ (équivalent CO₂).

Une observation intéressante est que le modèle EfficientNet-B7, malgré des performances similaires aux modèles VGG-19 et ResNet-50, affiche une empreinte carbone beaucoup plus faible. Pour chaque jeu d’entraînement, l’EfficientNet-B7 consomme 7 kWh d’énergie, ce qui correspond à environ 2,7 kg eqCO₂. En comparaison, VGG-19 nécessite 11 kWh par jeu d’entraînement (environ 3,2 kg eqCO₂) et ResNet-50 nécessite 25 kWh par jeu d’entraînement (environ 7,5 kg eqCO₂). Ces résultats mettent en évidence la capacité de l’EfficientNet-B7 à atteindre de bonnes performances tout en minimisant l’impact environnemental.

Il est important de noter que ces chiffres reflètent l’empreinte carbone pendant la phase d’entraînement des modèles et ne tiennent pas compte de la phase d’inférence, qui est généralement plus courte et moins énergivore. Ces informations soulignent l’importance de prendre en compte l’empreinte carbone lors du développement de modèles d’apprentissage automatique. En choisissant des modèles plus efficaces sur le plan énergétique, comme l’EfficientNet-B7, il est possible de réduire l’impact environnemental tout en maintenant de bonnes performances de diagnostic.

3.3 Validation patient : évaluation au niveau patient

Dans cette section, nous évaluons les performances du meilleur modèle résultant de l’étape de validation croisée à la fois au niveau cellulaire et au niveau individuel par patient. Nous présentons les résultats de deux manières différentes :

1. Évaluation au niveau cellulaire : Nous effectuons une classification de toutes les images de cellules provenant des 200 patients. Les résultats de cette évaluation sont présentés dans la figure 3.4)
2. Évaluation au niveau individuel par patient : Nous effectuons une classification individuelle des patients et présentons les résultats dans le (Tableau 3.3).

Parmi les 50 patients infectés, la vérité terrain comprend 3 574 globules rouges infectés par *P. falciparum* et 104 200 composants sanguins non infectés. Dans cette étape de validation, nous excluons le modèle ResNet-50 qui a été entraîné sur le jeu de données NIH. Bien que ce modèle ait obtenu des scores de classification très élevés lors de l’étude de validation croisée, il présente des performances considérablement inférieures lors de l’inférence au niveau du patient, notamment un grand nombre d’images de cellules sanguines mal classées, en particulier des faux positifs. Cette baisse de performance peut s’expliquer par le fait que les images du jeu de données NIH contiennent des objets d’intérêt individuels segmentés sur un fond noir. Nous excluons également le modèle C-CNN (Student) afin de garantir une comparaison optimale se concentrant sur les modèles pré-entraînés. Ces exclusions nous permettent de mieux évaluer et comparer les performances des autres modèles dans cette étape d’évaluation au niveau cellulaire et au niveau individuel par patient.

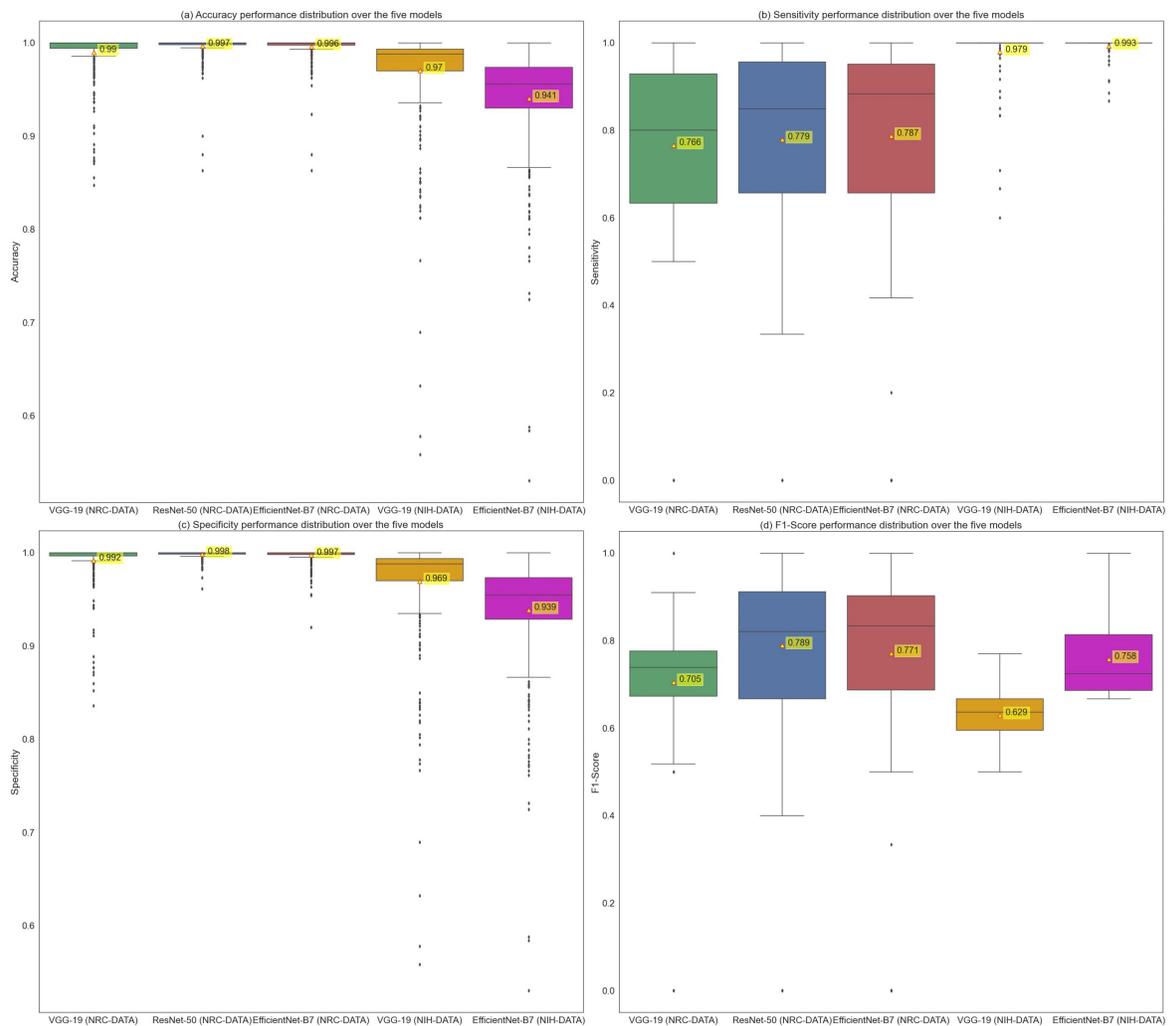


FIGURE 3.4 – Boîtes à moustaches des distributions de performance pour la classification de 200 patients, comprenant 422 015 composants sanguins non infectés et 3 574 globules rouges infectés. Les triangles jaunes représentent la valeur moyenne de chaque métrique.

Model	Positive patients ($n = 50$)		Negative patients ($n = 150$)		Sensitivity	Specificity
	TP	FN	TN	FP		
VGG-19-CNR	49	1	84	66	0.98	0.56
ResNet-50-CNR	50	0	76	74	1	0.50
EfficientNet-B7-CNR	49	1	73	77	0.98	0.49
VGG-19-NIH	50	0	3	147	1	0.02
EfficientNet-B7-NIH	50	0	0	150	1	0

TABLE 3.3 – Les performances de la classification 200 patients au niveau en positif et négatif avec un seuil de 1 GRP infecté par *P. falciparum*.

La figure 3.4 présente la distribution des performances de classification des cellules sanguines pour plusieurs modèles entraînés sur les ensembles de données CNR et NIH. Les modèles incluent VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR entraînés sur le jeu de données CNR, ainsi que VGG-19-NIH et EfficientNet-B7-NIH entraînés sur le jeu de données NIH.

D’après les boîtes à moustaches dans la figure 3.4. Il est clair que les modèles VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR ont obtenu les meilleures performances de classification des cellules sanguines. Ces modèles affichent une exactitude moyenne de 99,4 %, une sensibilité de 77,3 % et une spécificité de 99,6 %. VGG-19-NIH et EfficientNet-B7-NIH ont quant à eux obtenu une exactitude moyenne de 95,5 %, une sensibilité de 98,6 % et une spécificité de 95,4 % (voir figure 3.4.a et 3.4.b).

De plus, les matrices de confusion de la figure 3.5 révèlent une surestimation des hématies infectées par *P. falciparum* par les modèles VGG-19-NIH et EfficientNet-B7-NIH (Figure 3.7), qui ont mal classifié respectivement 3 902 et 24 039 images de composants sanguins non infectés. Nos modèles présentent un nombre total inférieur de faux positifs (860 faux positifs pour VGG-19-CNR, 581 faux positifs pour ResNet-50-CNR et 812 faux positifs pour EfficientNet-B7-CNR) (Figure 3.5.a, 3.5.b et 3.5.c).

Ces résultats mettent en évidence les performances supérieures des modèles entraînés sur l’ensemble de données CNR (VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR) par rapport aux modèles entraînés sur l’ensemble de données NIH (VGG-19-NIH et EfficientNet-B7-NIH). Les modèles CNR affichent une exactitude plus élevée et un nombre de faux positifs inférieur. Cependant, il est important de noter que les modèles NIH ont montré une surestimation des globules rouges infectés.

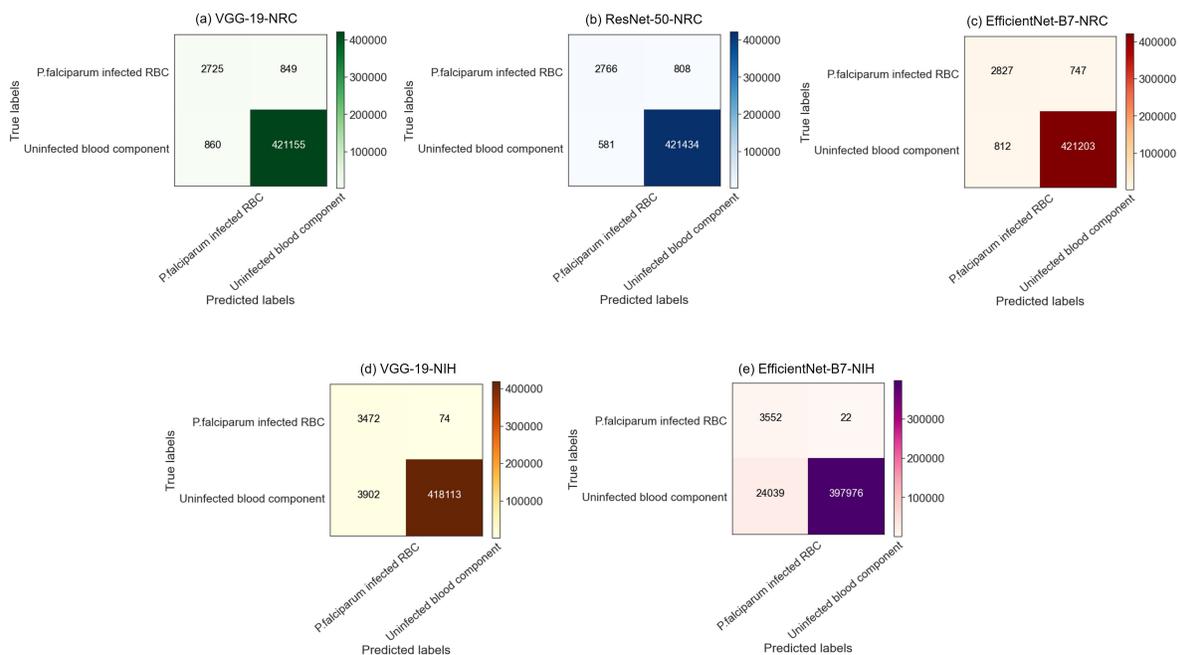


FIGURE 3.5 – Les matrices de confusion pour la classification de 3 574 hématies infectées par *P. falciparum* et de 422 015 composants sanguins non infectés provenant de 200 patients

Dans cette évaluation stricte, où un seul faux positif par patient est considéré comme une mau-

vaise classification, les scores de sensibilité restent élevés, allant de 98 % à 100 %). Cependant, la spécificité est faible, avec une moyenne de 53 % pour les modèles CNR et 2 % pour les modèles NIH (Tableau 3.3). Cela signifie que les modèles ont tendance à produire un grand nombre de faux positifs, ce qui conduit à une faible spécificité.

La figure 3.6 illustre ces résultats en montrant la distribution des faux positifs parmi les patients négatifs. Avec le modèle VGG-19-NIH, on constate que 84 % des patients négatifs ont entre 10 et plus de 100 faux positifs. Pour EfficientNet-B7-NIH, ce chiffre atteint 98,6 %, ce qui est très élevé. En revanche, les modèles VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR présentent des taux d'erreur plus faibles. Environ 0 ou 1 faux positif(s) sont observés pour 150 patients négatifs, ce qui correspond à des taux d'erreur de 77,3 % pour VGG-19-CNR et 73,3 % pour ResNet-50-CNR et EfficientNet-B7-CNR, respectivement.

Ces résultats soulignent la difficulté des modèles NIH à maintenir une spécificité élevée, conduisant à une surestimation des globules rouges infectés et à un grand nombre de faux positifs. En revanche, les modèles CNR présentent une spécificité plus élevée, avec un taux d'erreur nettement plus faible. Il est important de prendre en compte ces résultats dans l'interprétation des performances des modèles, en particulier lorsqu'il s'agit de la classification des patients. La spécificité joue un rôle crucial dans la détermination précise de l'infection par *P. falciparum*, et les modèles CNR semblent mieux adaptés pour cette tâche.

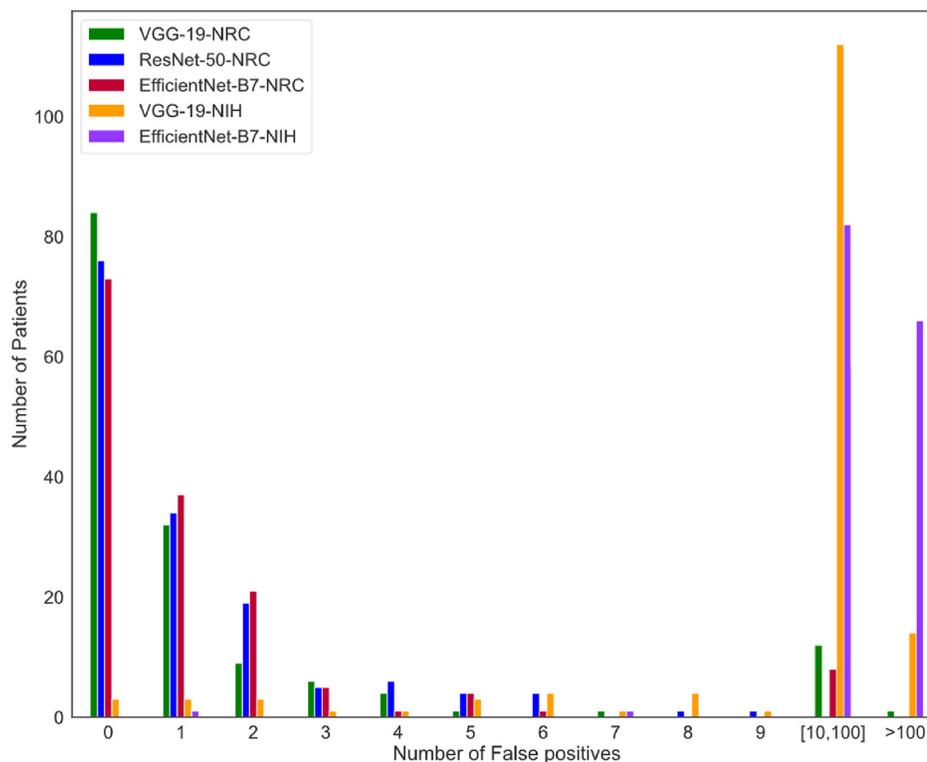


FIGURE 3.6 – Diagramme en barres de la distribution des faux positifs pour 150 patients négatifs.

Après avoir examiné les performances des modèles CNR dans la classification des cellules sanguines infectées par *P. falciparum*, il est maintenant pertinent de passer au prochain chapitre, qui se concentre spécifiquement sur l'estimation de la parasitémie. La parasitémie fait référence à la

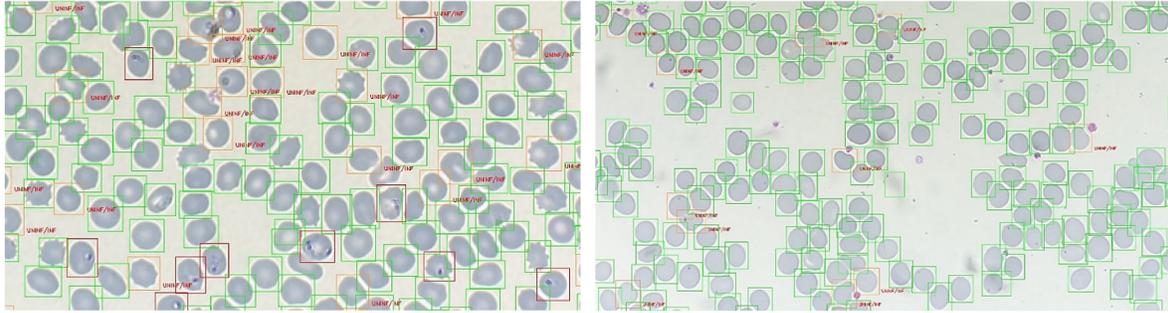


FIGURE 3.7 – Deux images prédites par VGG-19-CNR et VGG-19-NIH. Boîtes vertes : vrais négatifs de composants sanguins non infectés correctement classifiés par les deux modèles. Boîtes rouges : vrais positifs d’hématies infectées par *P. falciparum* correctement interprétés par les deux modèles. Boîtes oranges : faux positifs de plaquettes et d’artefacts de coloration mal interprétés par VGG-19-NIH et correctement interprétés par VGG-19-CNR comme des composants sanguins non infectés.

quantification de la charge parasitaire dans le sang des patients atteints de paludisme, et elle est un indicateur essentiel pour évaluer la sévérité de l’infection et guider les décisions de traitement.

Dans le prochain chapitre, nous explorerons les approches et les techniques utilisées pour estimer la parasitémie à l’aide de différentes techniques incluant les modèles présentés dans ce chapitre. Nous examinerons les différentes méthodes de quantification appliquées en laboratoire et les considérations liées à l’utilisation de ces modèles pour fournir des estimations précises de la parasitémie. En somme, ce prochain chapitre nous permettra d’explorer en détail l’utilisation des modèles CNR, en particulier EfficientNet-B7-CNR, pour l’estimation précise de la parasitémie dans le contexte du paludisme. Nous verrons comment cette approche peut potentiellement améliorer la fiabilité de l’estimation de la parasitémie.

Mesure automatique de la parasitémie : un critère biologique important

1 Contexte et Objectifs

Pour le diagnostic du paludisme dans les pays non endémiques, la stratégie recommandée est le frottis sanguin épais ou les techniques de biologie moléculaire. Il est important de noter qu'il peut exister des zones non endémiques ou des régions au sein de pays endémiques où la microscopie ou les TDR sont utilisés. En pratique, étant donné que la plupart des laboratoires ne disposent pas de l'expertise nécessaire pour réaliser des frottis sanguins épais, les techniques de biologie moléculaire, telles que la LAMP, sont couramment utilisées. En tant que deuxième étape, en cas de résultats positifs, un frottis sanguin mince est systématiquement réalisé. Le diagnostic du paludisme ne se limite pas à identifier la présence du parasite; il doit également répondre à des questions telles que : quelle espèce de *Plasmodium* est impliquée, quels stades de parasites sont présents dans le sang et quelle est la parasitémie en cas de paludisme à *P. falciparum*? Malheureusement, dans le contexte du diagnostic de routine en urgence du paludisme, certaines techniques de diagnostic, telles que les TDR ou les méthodes moléculaires, peuvent ne pas fournir des réponses complètes à toutes ces questions. Cependant, à travers un frottis mince, il est possible de déterminer la parasitémie, d'identifier les espèces et de détecter les stades du parasite. L'examen microscopique du parasite reste un aspect crucial du diagnostic du paludisme.

Une estimation précise de la parasitémie nécessite l'examen d'un frottis sanguin mince, ce qui revêt une importance particulière étant donné que la parasitémie constitue un critère significatif pour déterminer la gravité d'une attaque de paludisme. Par exemple, une parasitémie supérieure à 4 % chez les adultes et à 10 % chez les enfants de moins de 5 ans est considérée comme un critère d'infection sévère à *P. falciparum* [Bouchaud et al., 2020]. Généralement, la parasitémie est estimée manuellement à l'aide de techniques de microscopie standard, telles que la mesure sur frottis sanguin épais et la quantification de la couche leucocytaire [WHO, 2016b]. Cependant, ce processus est chronophage et il existe une variabilité importante entre les microscopistes [O'Meara et al., 2006]. De plus, il nécessite la présence d'un microscopiste expert disponible 24 heures sur 24, 7 jours sur 7 pour établir le diagnostic.

La fiabilité du diagnostic dépend fortement de l'expertise humaine, laquelle peut varier considérablement d'un microscopiste à l'autre. Il est donc essentiel de maintenir les compétences du personnel et d'assurer un contrôle de qualité rigoureux des résultats produits, en particulier dans les zones endémiques où les ressources sont limitées et où le contrôle de qualité peut être insuffisant. Bien qu'il ait été démontré, dans l'état de l'art, que l'utilisation de systèmes basés sur l'apprentissage profond (DL) peut être prometteuse pour estimer précisément la parasitémie, aucune étude n'a évalué de manière satisfaisante ces systèmes par rapport à différentes techniques manuelles et semi-automatiques sur un grand échantillon d'images de frottis sanguins minces, dans un contexte de soins courants aux patients.

Dans ce chapitre, nous nous concentrons sur l'évaluation de notre système MALARIS développé précédemment, qui vise à répondre aux besoins spécifiques liés à l'estimation de la parasitémie dans le diagnostic du paludisme. Nous avons entrepris plusieurs étapes pour évaluer les capacités de notre système et les comparer à des méthodes manuelles traditionnelles. Plus précisément :

1. Tout d'abord, nous avons réalisé une étude de reproductibilité pour examiner l'évolution de la variation de la parasitémie en fonction du nombre de champs examinés. Cette étude nous a permis d'évaluer la précision et la cohérence des mesures de la parasitémie obtenues par MALARIS avec différentes combinaisons d'images de champs.
2. Nous avons mené une expérience visant à mettre en évidence la variabilité entre différents lecteurs dans l'estimation manuelle de la parasitémie. Cette comparaison nous a permis d'évaluer la fiabilité et la cohérence des mesures obtenues par notre système par rapport à des experts humains.
3. Nous avons évalué les performances de notre système MALARIS en les comparant à celles de deux techniques manuelles et semi-automatiques. Cette comparaison nous permet de déterminer si notre système parvient à atteindre des performances similaires à celles des méthodes comparées.

La figure 4.1 illustre le déroulement des différentes expériences effectuées dans ce chapitre.

2 Matériel et Méthodes

2.1 Sélection et description des patients

Au sein du département clinique de Parasitologie-Mycologie, Centre National de Référence (CNR) pour le paludisme à Paris, France, nous avons sélectionné de manière aléatoire parmi 456 échantillons, 116 frottis sanguins minces colorés au Giemsa (RAL, France) [WHO, 2016a] provenant de patients positifs pour l'espèce *P. falciparum*. Les lames, réalisées entre janvier 2019 et août 2022, ont été anonymisées pour l'étude. Les mesures de parasitémie expertes variaient de 0,15 % à 26,2 %. Les images ont été capturées au format JPG RGB, sur des microscopes optiques Olympus BX51 et BX41 avec un grossissement de 500x. Les microscopes Olympus BX51 et BX41 sont tous deux équipés du système optique infini UIS2, qui réduit la distorsion et améliore la qualité de l'image. Pour le BX41, nous avons utilisé une lentille Olympus UPlanSApo 50x, et pour le BX51, nous avons utilisé une lentille UPlanFl 50x/0.50P. Ce sont toutes deux des lentilles à immersion dans l'huile avec une ouverture numérique (NA) accrue. Le smartphone était fixé à un oculaire WHN 10x/22 à l'aide d'un adaptateur Orion 05306 SteadyPix Pro. Pour évaluer la reproductibilité de l'estimation de la parasitémie, un panel de 14 frottis sanguins a été utilisé pour comparer deux techniques de mesure manuelle, la technique standard de l'OMS et la technique assistée par la

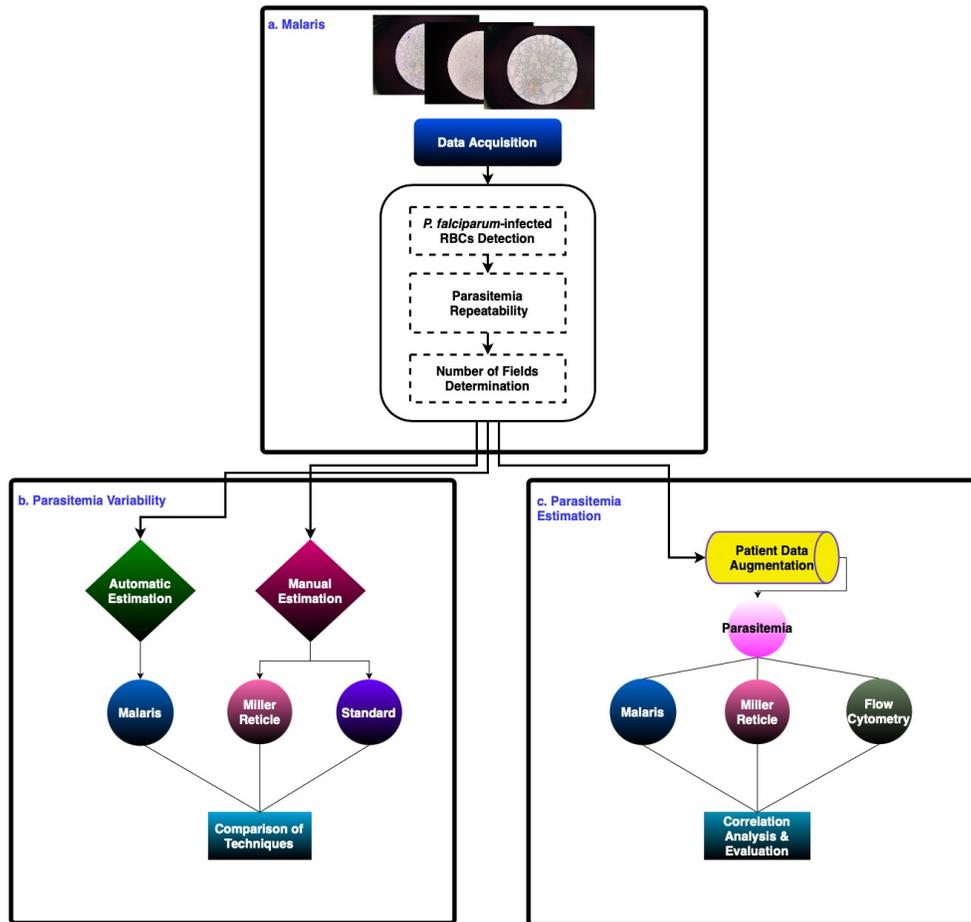


FIGURE 4.1 – Architecture de l'étude pour l'évaluation de la parasitémie.

cellule de Miller. Ensuite, 102 frottis sanguins ont été utilisés pour comparer les performances de notre système basé sur l'apprentissage profond (DL) avec à la fois le cellule Miller et une méthode automatisée de comptage par cytométrie en flux. Au total, 1160 images ont été capturées à l'aide de deux smartphones différents : l'appareil photo arrière de 12 mégapixels d'un iPhone 7 et un Google Pixel 6 doté d'un système de double appareil photo arrière. Le Pixel 6 comprenait un capteur de 50 mégapixels et un capteur de 12 mégapixels pour les prises de vue grand angle. Toutes les expériences décrites ont été réalisées entre 2019 et 2022. Une fois les données sélectionnés, plusieurs techniques manuelles et automatiques ont été utilisées pour l'estimation de la parasitémie. Bien que nous ayons employé deux moyens d'acquisition distincts pour collecter nos données, notre analyse n'a pas séparément évalué l'influence de chaque appareil photo sur la performance des résultats, laissant ainsi une voie ouverte pour des investigations futures sur l'impact potentiel des spécificités de chaque caméra sur la précision de nos modèles.

2.2 Mesure de la parasitémie

La microscopie optique : une technique de référence pour la mesure de la parasitémie

La parasitémie (formule 4.1) est calculée en divisant le nombre total de globules rouges infectés (à l'exclusion des gamétocytes) par le nombre total de globules rouges comptés ou estimés dans les champs microscopiques examinés. Pour estimer la parasitémie sur des frottis sanguins minces observés au microscope optique, nous avons suivi une procédure standard adaptée aux directives de l'OMS [WHO, 2016b]. À un grossissement de x1 000, nous avons compté manuellement tous les GR infectés dans le numérateur, tandis que le dénominateur impliquait le décompte du nombre total de GR dans un champ représentatif (Fig 2a) pour estimer le dénominateur global. Le processus a ensuite été répété en comptant les GR infectés dans chaque champ suivant. L'OMS recommande d'examiner 5 000 GR, ce qui correspond approximativement à 20 champs, pour estimer la parasitémie. Le dénominateur, le nombre total de GR examinés, a été estimé en multipliant le nombre de champs par le nombre de GR dans un champ représentatif.

$$\text{Parasitémie (\%)} = \frac{\text{Nombre de globules rouges parasités}}{\text{Nombre total de globules rouges}} \times 100 \quad (4.1)$$

La microscopie optique assistée par la cellule de Miller

La cellule de Miller (Figure 4.2) est une méthode de comptage assistée par une cellule utilisant le rapport de deux carrés de tailles différentes pour compter un grand nombre de cellules, notamment lorsque la répartition d'un champ à l'autre est très variable. La cellule de Miller était initialement utilisée en application hématologique, en particulier pour le comptage des réticulocytes [Nazi, 1986], où il a été démontré que celle-ci permettait de réduire l'imprécision lors du comptage manuel. Concernant la mesure de la parasitémie pour le diagnostic du paludisme, la procédure consiste à compter le nombre total de GR dans le petit carré, estimer leur nombre dans le grand carré en utilisant un ratio de 1:10, puis compter le nombre total de GR infectés dans le grand carré. Cela permet d'avoir une estimation plus précise du dénominateur dans chaque champ examiné (grand carré). Pour estimer la parasitémie sur des frottis sanguins minces observés au microscope optique, nous avons suivi une procédure opérationnelle standard.

Mesure de la parasitémie par cytométrie en flux : analyse des GR du patient avec un anticorps polyclonal anti-RESA et le colorant SYBR Green I.

Une troisième technique pour l'estimation la parasitémie à l'aide de la cytométrie en flux a été utilisée, le protocole expérimentale est le suivant : les globules rouges (GR) du patient (1×10^7) ont été lavés avec $200 \mu\text{L}$ de tampon TBH contenant $0,5 \times \text{HBSS}$ (Life Technologies Limited, Paisley, Royaume-Uni), $0,5 \times \text{VERSOL}^{\circledR}$ (AGUETTANT, Lyon, France) et 50 mM de Tris HCl (pH 7,0). Ensuite, les cellules ont été traitées avec 1% de glutaraldéhyde dans du PBS pendant 1 minute à température ambiante, puis lavées quatre fois. Les cellules ont ensuite été incubées à $37 \text{ }^{\circ}\text{C}$ avec agitation pendant 30 minutes avec un anticorps polyclonal anti-peptide répétitif de RESA à une concentration de $25 \mu\text{g/mL}$, développé dans notre laboratoire, puis lavées trois fois avec $200 \mu\text{L}$ de PBS. Les cellules ont ensuite été resuspendues dans $200 \mu\text{L}$ d'AlbuMaxTM II à 1% (Life Technologies NZ Ltd, Auckland, États-Unis), et 0,1% de TRITON X100 dans du PBS contenant $10 \mu\text{g/mL}$ d'anticorps anti-rabbit IgG-(H+L) fragment conjugué à Pacific BlueTM (INVITROGEN, Oregon, États-Unis) et 4x SYBRTM Green I (INVITROGEN, Carlsbad, États-Unis). Après 30 minutes d'incubation à $37 \text{ }^{\circ}\text{C}$ avec agitation, les cellules ont été lavées trois fois avec $200 \mu\text{L}$ de PBS avant d'être soumises à un

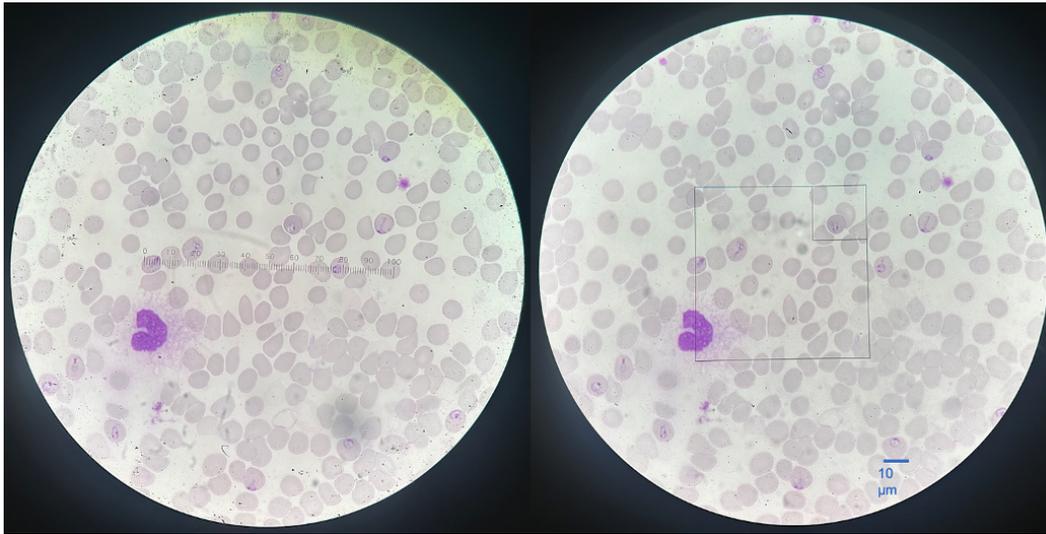


FIGURE 4.2 – Image d'un frottis sanguin mince infecté par *Plasmodium falciparum* à un grossissement de x1 000, sans et avec la cellule de Miller. À un grossissement de x1 000, une graduation = 1 micromètre. La taille moyenne des GR est de 7 µm (de 5 à 10).

test de cytométrie en flux à l'aide du CytoFlex (Beckman Coulter, Pasadena, États-Unis). Les GR ont été sélectionnés en fonction de leurs caractéristiques de diffusion avant et latérale, et 50 000 événements ont été acquis. Les GR infectés par *Plasmodium falciparum* étaient positifs pour RESA et SYBR, tandis que les GR non infectés étaient négatifs pour RESA et SYBR. Les événements ont été analysés à l'aide du logiciel CytExpert.

Mesure automatique de la parasitémie grâce au système MALARIS

Pour cette étude, nous avons nommé le système abordé dans le chapitre précédent : MALARIS, dans lequel nous avons comparé les performances de différentes architectures de CNN, telles que VGG-19, ResNet-50 et EfficientNet-B7. Leur performance et leur capacité de généralisation ont été évaluées sur 200 patients indépendants lors d'une validation croisée. Les résultats ont montré que notre système était capable de détecter les cellules infectées par *Plasmodium falciparum* dans des frottis sanguins minces avec une exactitude de 99,7 %, une sensibilité de 77,9 % et une spécificité de 99,8 %. Nous avons également pu étendre notre étude pour estimer automatiquement la parasitémie, ce qui sera abordé dans la suite de ce travail. Pour cette étude, nous avons sélectionné le modèle de classification de la série ResNet-50-NRC qui présentait des performances optimales pour distinguer les images de globules rouges infectés par *Plasmodium falciparum* de celles des composants sanguins non infectés (77,9 % de sensibilité et 99,8 % de spécificité au niveau cellulaire, et 100 % de sensibilité et 50 % de spécificité au niveau patient).

2.3 Les techniques manuelles : sources de variabilité pour la mesure parasitémie

Nous avons d'abord examiné la répétabilité de deux techniques manuelles d'estimation de la parasitémie : la mesure standard et une mesure assistée par la cellule de Miller. La validation de la méthode utilisant la cellule de Miller pour estimer la parasitémie a été réalisée au laboratoire de parasitologie CNR du Paludisme à la Pitié-Salpêtrière en France. Pour confirmer son efficacité,

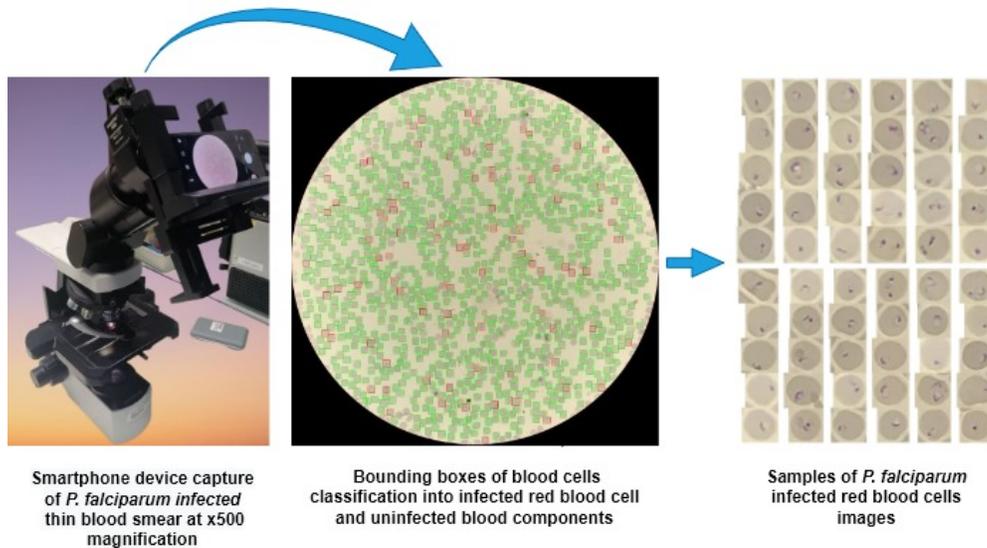


FIGURE 4.3 – MALARIS

nous avons sélectionné 14 patients infectés, avec un frottis par patient, et la parasitémie de chaque frottis a été estimée par 10 microscopistes indépendants, comprenant des internes, des médecins biologistes et des techniciens de laboratoire faisant partie du laboratoire de parasitologie. Pour chaque microscopiste, nous avons évalué la variation de son estimation par rapport à la moyenne du groupe en calculant l'écart type relatif ou relative standard deviation (RSD) (équation 4.2). Cela nous a permis d'analyser la variabilité et la répétabilité de chaque technique pour chaque patient, en tenant compte des différences entre les opérateurs. La mesure standard correspond à une estimation de la parasitémie basée sur 20 champs à un grossissement de x1 000, tandis que la technique assistée par la cellule de Miller correspond à une estimation basée sur 50 champs.

$$\text{RSD (\%)} = \frac{\text{SD}}{\bar{X}} \times 100 \quad (4.2)$$

L'écart type relatif ou RSD (exprimé en pourcentage), exprimé en pourcentage, est défini comme l'écart type (SD) des mesures obtenues par différentes paires de champs ou de microscopistes, divisé par la moyenne des mesures de parasitémie (\bar{X}). L'écart type relatif sert de mesure du degré de variabilité entre différentes mesures de parasitémie. Un écart type relatif plus faible indique un meilleur accord et une plus grande cohérence entre les mesures de parasitémie, tandis qu'un écart type relatif plus élevé suggère une plus grande variabilité et une cohérence plus faible.

2.4 La répétabilité de la parasitémie avec MALARIS : l'impact du nombre de champs

Ensuite, nous avons examiné la variabilité et la répétabilité de l'estimation de la parasitémie à l'aide de MALARIS. Nous avons utilisé un panel de 14 patients, à partir desquels nous avons obtenu 10 images de champs microscopiques par patient. Pour évaluer les variations potentielles résultant de la sélection de différents champs (répétabilité), nous avons réalisé plusieurs comparaisons entre deux combinaisons utilisant deux ensembles distincts de n champs. Avec nos 10 champs par patient, nous avons comparé les valeurs de K de 1 à 5. Pour chaque patient, nous avons un total

de 45 paires d'images, 630 paires pour 2 champs, 2 100 paires pour 3 champs, 1 575 paires pour 4 champs et 126 paires pour 5 champs.

Nous avons évalué la répétabilité et l'impact du nombre de champs sur chaque estimation de la parasitémie en utilisant l'erreur standard relative (RSE) traditionnelle entre deux estimations sur tous les patients (équation 4.3). Cela donne une indication du nombre optimal de champs à choisir pour obtenir une estimation fiable de la parasitémie tout en réduisant le nombre de champs à capturer.

$$\text{RSE (\%)} = \frac{\text{SE}}{\bar{X}} \times 100 \quad (4.3)$$

L'erreur standard relative est exprimée en pourcentage, SE pour standard error est l'erreur standard des différences entre les deux combinaisons de mesures de la parasitémie, et \bar{X} est la moyenne des deux combinaisons de mesures. L'erreur standard relative mesure la précision de la différence entre les deux combinaisons de champs. Une RSE faible indique une plus grande précision et moins de variabilité dans la différence entre les deux combinaisons de mesures. Une RSE importante indique une moins grande précision et plus de variabilité.

2.5 Comparaison de méthodes pour la mesure de la parasitémie : MALARIS Vs. cellule de Miller Vs. cytométrie en flux

S'appuyant sur l'évaluation de la répétabilité et de la variabilité de la mesure de la parasitémie à l'aide de techniques manuelles et du système MALARIS, nous avons réalisé une comparaison entre notre système MALARIS et deux autres méthodes, incluant la cellule de Miller et la cytométrie en flux. Pour évaluer la corrélation de la mesure de la parasitémie, nous avons analysé 102 patients supplémentaires et calculé le coefficient de corrélation (R^2) pour chaque paire de méthodes. L'objectif principal de cette comparaison était de mettre en évidence la mesure de la parasitémie estimée par notre système MALARIS en tant que mesure fiable et concordante par rapport aux deux autres techniques. En évaluant la corrélation entre les estimations de la parasitémie, nous souhaitons déterminer si les mesures automatiques fournies par notre système étaient en accord avec celles obtenues à l'aide des méthodes manuelles de la cellule de Miller et de la cytométrie en flux.

Pour chaque patient, le système MALARIS a analysé 5 images de champs, ce qui représente un total de plus de 5 000 cellules. D'autre part, la méthode de la cellule de Miller s'appuyait sur au moins une mesure correspondant à 50 champs examinés, soit 3 000 érythrocytes, et la cytométrie en flux fournissait une seule mesure par patient, ce qui correspond à l'analyse de 50 000 érythrocytes. De nouveaux experts et lecteurs ont été impliqués dans cette expérience pour diversifier l'interprétation et les compétences humaines par rapport aux expériences précédentes. Cette approche nous a permis d'approfondir notre compréhension des variations potentielles résultant de différentes méthodes d'estimation de la parasitémie. De plus, nous avons effectué des analyses supplémentaires pour évaluer la performance de sensibilité et de spécificité du système MALARIS aux valeurs de seuil de détection ou limit of detection (LOD) de 4 % et 10 %, en utilisant la rétine de Miller et la cytométrie en flux comme méthodes de référence. Ces analyses comparent les résultats du système MALARIS à ceux de la technique la cellule de Miller et de la cytométrie en flux. La sensibilité mesure la capacité du système à identifier correctement les patients atteints de paludisme dont les niveaux de parasitémie sont supérieurs au seuil de détection spécifié, tandis que la spécificité mesure la capacité du système à identifier correctement les patients dont les niveaux

de parasitémie sont inférieurs au seuil de détection. De plus, nous avons évalué les performances des méthodes de référence elles-mêmes.

Plus précisément :

1. Les vrais positifs (VP) : représentent le nombre de cas où à la fois le système MALARIS et la cellule de Miller/ cytométrie en flux ont correctement identifié le patient comme ayant des niveaux de parasitémie supérieurs au seuil de détection spécifié.
2. Les faux négatifs (FN) : représentent le nombre de cas où le système MALARIS a obtenu un résultat négatif (inférieur au seuil de détection), mais la cellule de Miller/cytométrie en flux a obtenu un résultat positif (supérieur au seuil de détection).
3. Les vrais négatifs (VN) : représentent le nombre de cas où à la fois le système MALARIS et la cellule de Miller/cytométrie en flux ont correctement identifié le patient comme n'ayant pas de niveaux de parasitémie supérieurs au seuil de détection spécifié.
4. Les faux positifs (FP) : représentent le nombre de cas où le système MALARIS a obtenu un résultat positif (supérieur au seuil de détection), mais la cellule de Miller/cytométrie en flux a obtenu un résultat négatif (inférieur au seuil de détection).

3 Résultats

3.1 Les techniques manuelles : sources de variabilité pour la mesure de la parasitémie

L'estimation de la parasitémie chez 14 patients à l'aide des trois méthodes différentes a conduit à une analyse comparative entre les techniques manuelles et notre système MALARIS. Sur la figure 4.5.a, on peut observer une variabilité interhumaine plus significative avec la procédure standard de l'OMS (écart type relatif (RSD) moyen de 38,22 % et médiane de 27,22 %) basée sur 10 mesures sélectionnées de manière aléatoire parmi les 10 microscopistes (voir le tableau détaillé en annexe 7), contrairement à la cellule de Miller (écart type relatif moyen de 26,40 % et médiane de 18,54 %) (voir le tableau détaillé en annexe 6). Sur la figure 4.5.b, le système MALARIS présente un écart type relatif plus faible (moyenne de 11,21 % et médiane de 10,98 %). Une variation significative est observée dans la mesure standard, avec un écart type relatif dépassant 100 % pour le patient P3, de même pour le patient P5 avec la cellule de Miller. Ces variations sont également représentées dans les boxplots de la figure 4.4, mettant en évidence les différences notables entre les méthodes.

Les intervalles de confiance présentés dans la figure 4.6 fournissent les estimations individuelles pour chaque patient, offrant ainsi des informations sur la variabilité intra-patient pour chaque technique. Ces intervalles illustrent à la fois les variations entre différentes techniques et les variations observées entre les patients individuels. Les intervalles de confiance pour les techniques manuelles telles que la procédure standard et celle assistée par la cellule de Miller présentent une variabilité significative à mesure que la parasitémie augmente. En revanche, ils montrent des estimations moyennes proches pour les trois techniques, à l'exception d'un patient (P12) où le système MALARIS sous-estime significativement la parasitémie.

Ces observations soulignent l'avantage de l'utilisation de la cellule de Miller pour des estimations plus fiables de la parasitémie. Ainsi, nous avons utilisé cette méthode comme référence pour évaluer les performances de notre système MALARIS dans la prochaine section de ce chapitre.

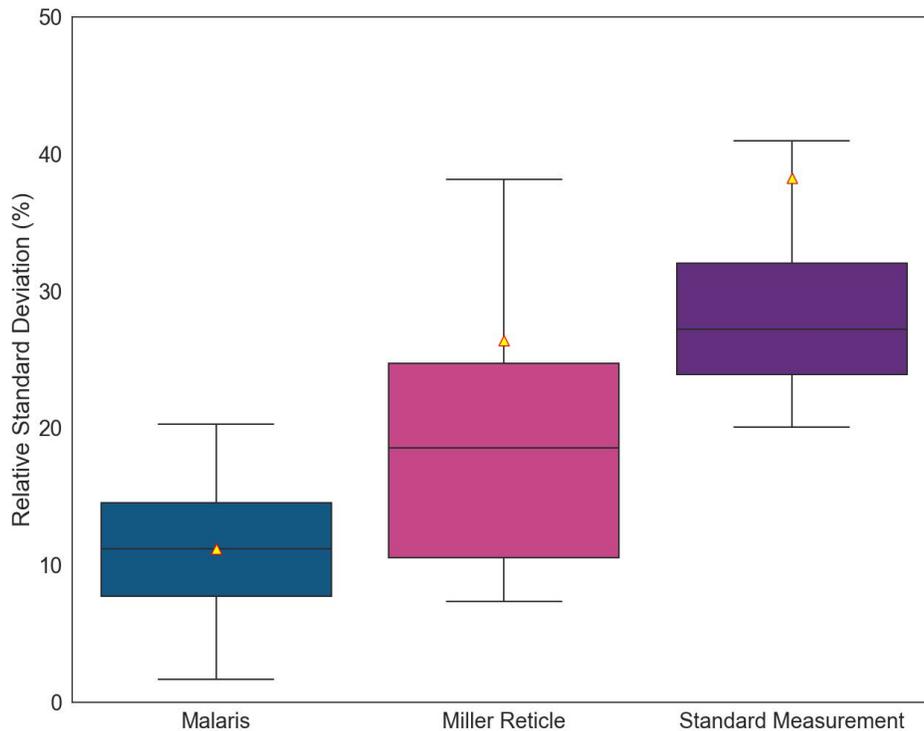


FIGURE 4.4 – Écart type relatif (RSD) pour différentes techniques. Comparaison entre MALARIS, la cellule de Miller et la technique de mesure standard. Les triangles jaunes représentent la valeur moyenne de RSD pour chaque technique

3.2 La répétabilité du système MALARIS : l'impact du nombre de champs

Dans notre étude, nous avons évalué la répétabilité de l'estimation de la parasitémie par le système MALARIS sur les 14 patients. Nous avons examiné l'impact du nombre de champs utilisés sur la quantification des globules rouges infectés par *Plasmodium falciparum* pour chaque patient.

Pour chaque combinaison d'images de champs, le système MALARIS a détecté et classé en moyenne 1 000 images de cellules sanguines par image de champ. Les diagrammes en boîtes de la figure 4.7.a représentent la distribution de la RSE en fonction de la parasitémie moyenne calculée pour chaque combinaison de 1 à 5 images de champs. En utilisant 5 champs, nous avons observé une variation moyenne de 2,12 %, avec un maximum inférieur à 10 %. Cela indique que l'utilisation de 5 images de champs à un grossissement de x500 réduit l'incertitude de mesure à un niveau inférieur à celui des approches traditionnelles.

Il est important de noter que le processus d'acquisition et d'analyse d'une image d'un champ au x500 prend en moyenne 80 à 90 secondes pour plus de 1 000 images de cellules sanguines. Cette efficacité temporelle représente un avantage significatif par rapport aux méthodes traditionnelles, qui nécessitent un temps plus important (30 minutes en moyenne) pour effectuer les mêmes mesures.

En outre, la figure 4.7.b compare les résultats obtenus avec l'utilisation de 5 champs par le système MALARIS (représenté par le diagramme en boîte bleu) avec ceux obtenus par les lecteurs (par paire), y compris la cellule de Miller (représentée par le diagramme en boîte rose) et la mesure standard (représentée par le diagramme en boîte violette). L'erreur standard RSE mesure la différence

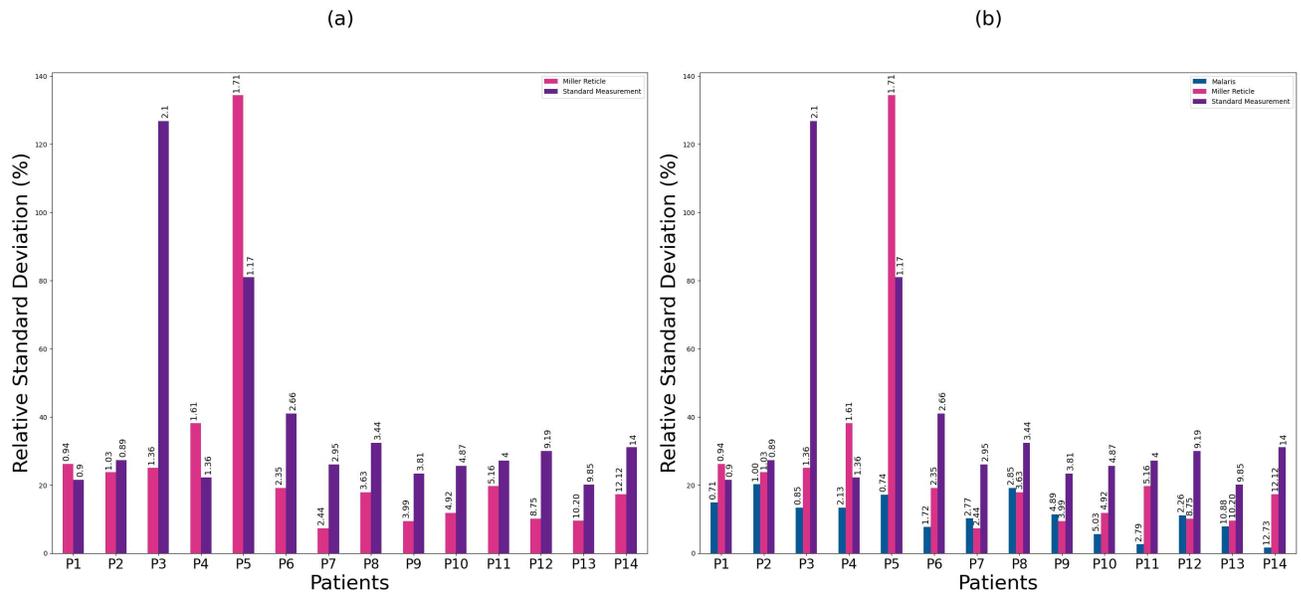


FIGURE 4.5 – (a) Comparaison de l'écart type relatif (RSD) parmi les 14 patients, classés par parasitémie croissante, entre deux techniques manuelles, comprenant 10 mesures sélectionnées au hasard parmi 10 microscopistes, l'écart type relatif moyen est de 26,40 % pour la méthode assistée par la cellule de Miller et de 38,22 % pour la technique de mesure standard. (b) Distribution de l'écart type relatif pour les trois méthodes, y compris l'estimation de la parasitémie par MALARIS sur 10 images de frottis sanguins (voir le tableau détaillé en annexe A). Au-dessus de chaque histogramme, la valeur de parasitémie mesurée pour chaque patient, disposée par ordre croissant de parasitémie, est affichée pour la technique correspondante.

entre deux combinaisons de mesures de la parasitémie. Les résultats montrent naturellement, une très faible RSE pour le système MALARIS par rapport aux autres méthodes, avec une RSE moyenne de 2,12 %, allant de 0,70 % à 4,10 %. La cellule de Miller présente également une faible RSE, avec une valeur moyenne de 13,15 % et une variation de 6,05 % à 26,01 %, mais légèrement plus élevée que celle de notre système. En revanche, la mesure standard, qui est la méthode la plus variable, montre une variation élevée avec une RSE moyenne de 23,98 % et une variation allant de 10,04 % à 48,06 %.

L'utilisation de plusieurs champs combinée à l'efficacité temporelle du système permet d'obtenir des mesures précises et reproductibles de la parasitémie. Cela offre des avantages significatifs par rapport à un examen microscopique standard nécessitant une interprétation humaine.

Après avoir évalué la répétabilité du système MALARIS dans l'estimation de la parasitémie sur un échantillon restreint de patients, nous élargirons notre étude à une plus grande cohorte de 102 nouveaux patients dans la section suivante. Cette extension nous permettra de comparer les performances de MALARIS avec d'autres méthodes d'estimation de la parasitémie. En analysant la concordance entre les mesures obtenues par notre système et celles des méthodes conventionnelles, nous pourrions évaluer la fiabilité et l'efficacité de MALARIS dans l'estimation précise de la parasitémie.

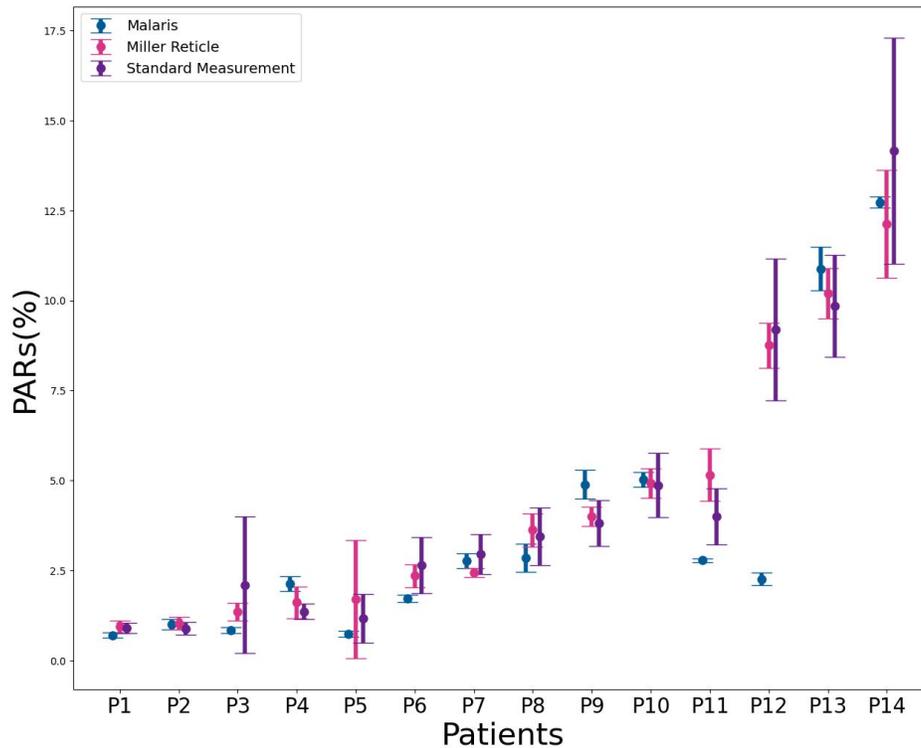


FIGURE 4.6 – Comparaison des estimations de la parasitémie et des intervalles de confiance chez 14 patients. La cellule de Miller vs. Mesure Standard (a). Malaris vs. Cellule de Miller vs. Mesure Standard (b). Chaque disque est centré sur la valeur moyenne de l'échantillon correspondant.

3.3 Comparaison de méthodes pour l'estimation de la parasitémie sur 102 nouveaux patients

Tout d'abord, nous avons évalué si 5 000 images de cellules étaient suffisantes pour notre système, comme cela avait été démontré précédemment sur 14 patients. Nous avons confirmé ces résultats en étudiant la variabilité, mais cette fois-ci sur un échantillon plus important de 102 patients. Pour cela, nous avons examiné la variabilité des combinaisons de 5 champs pour ces 102 patients en calculant l'erreur RSE pour les 126 paires possibles de K=5 champs microscopiques (figure 4.8). Les résultats ont révélé une RSE moyenne de 4,94 % et une médiane de 3,89 %. La RSE minimale observée était de 0,4 %, tandis que le maximum atteignait 22,86 %. Ces résultats confirment les résultats obtenus lors de l'analyse des 14 patients, où la RSE moyenne était inférieure à 10 %, précisément à 9,64 %. Nos résultats démontrent que, quelle que soit la combinaison des 5 images, les mesures estimées sont assez proches dans l'ensemble.

Pour évaluer davantage les méthodes d'estimation de la parasitémie, nous avons comparé les résultats obtenus pour chaque technique (voir le tableau détaillé en annexe 8) sur 102 nouveaux patients et calculé leurs valeurs R^2 pour estimer la concordance entre les résultats obtenus par chaque méthode. Notre système MALARIS a montré un fort coefficient de corrélation ($R^2=0,87$) avec la parasitémie obtenue à l'aide de la cellule de Miller (figure 4.9.a). De même, nous avons également observé une forte corrélation avec la cytométrie en flux ($R^2=0,89$), comme le montre la figure 4.9.b. De plus, nous avons noté une forte corrélation ($R^2=0,92$) entre la méthode de la cellule

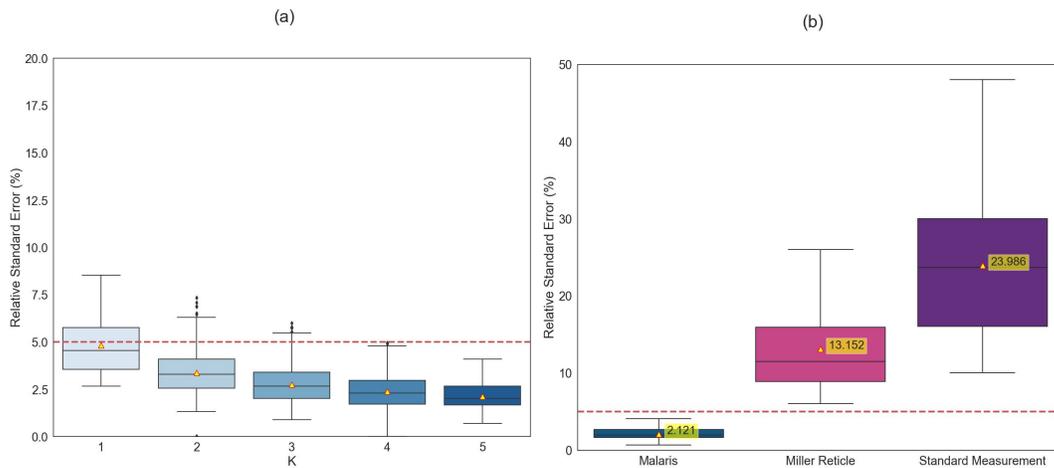


FIGURE 4.7 – La répétabilité du comptage automatique de la parasitémie par le système MALARIS sur 14 patients, avec une parasitémie variant de 0,7 % à 4,10 % et une valeur médiane de 2,51%. La représentation du taux d’erreur standard relatif (%) (a) pour la répétabilité des différentes combinaisons K (de 1 à 5) à partir de 10 images de frottis sanguins minces et (b) comparaison avec la cellule Miller et le comptage manuel standard. La ligne en pointillés rouges représente l’intervalle d’erreur standard de ± 5 %. Les triangles jaunes représentent la valeur moyenne de chaque paire de mesures.

de Miller et les mesures de cytométrie en flux (figure 4.9.c).

Concernant les seuils de détection de la parasitémie à 4 et 10 %, MALARIS avec la cellule de Miller ainsi que MALARIS avec la cytométrie en flux démontrent une sensibilité élevée. Plus précisément, MALARIS avec la cellule de Miller présentent une sensibilité de 84,1 %, tandis que MALARIS couplé à la cytométrie en flux démontrent une sensibilité plus élevée de 93,2 %. Lorsque le seuil de détection est de 10 %, les valeurs de sensibilité pour les deux méthodes restent relativement élevées. MALARIS avec la cellule de Miller atteignent une sensibilité de 83,3 %, et la paire de méthodes, MALARIS et cytométrie en flux maintient une sensibilité de 91,7 %. À un seuil de détection de 4 %, MALARIS et la cellule de Miller affichent une spécificité légèrement supérieure (86,2 %) par rapport à MALARIS avec la cytométrie en flux (77,6 %). Pour un seuil à 10 %, MALARIS et la cellule de Miller présentent une spécificité élevée de 97,8 %. MALARIS avec la cytométrie en flux, à 10 %, maintient une spécificité de 93,3 % (voir le tableau détaillé en annexe 10). De plus, notre évaluation s’est penchée sur les performances des méthodes de référence elles-mêmes. Lorsque nous avons comparé la cellule de Miller à la cytométrie en flux, quelques tendances intéressantes sont apparues. À un seuil à 4 %, la cellule de Miller associée à la cytométrie en flux a montré une excellente sensibilité à 97,8 %, avec une spécificité de 82,5 %. Cela met en évidence la robustesse de cette approche combinée pour identifier correctement les patients présentant des niveaux de parasitémie supérieurs à 4 % tout en maintenant un taux de faux positifs relativement faible. Pour un seuil à 10 %, la combinaison de la cellule de Miller et de la cytométrie en flux a maintenu une sensibilité élevée à 91,7 % et une spécificité élevée à 93,3 %.

De plus, les graphiques de Bland-Altman dans la figure 4.10 fournissent une représentation visuelle de l’accord entre les trois techniques de mesure de la parasitémie. Sur chaque graphique, la ligne rouge, représentant la différence moyenne centrée autour de 0. Cela indique un bon niveau d’accord entre les deux techniques comparées. L’alignement constant des lignes rouges autour de

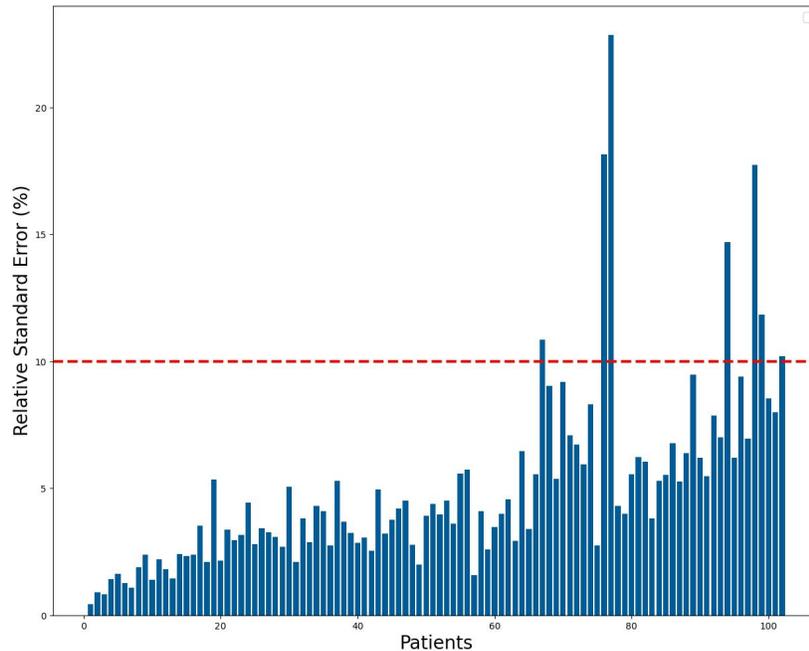


FIGURE 4.8 – Répétabilité de l’estimation automatique de la parasitémie par le système MALARIS évaluée sur 102 patients avec une parasitémie ordonnée de 0,03 % à 27,32 %, et une valeur médiane de 3,63 %. La répétabilité a été évaluée en utilisant le taux d’erreur standard relative (RSE) (%) en analysant plusieurs combinaisons de 5 images de frottis sanguins minces, pour un total de 126 paires.

0 sur les trois graphiques démontre la cohérence et la concordance des estimations de parasitémie obtenues à l’aide de différentes techniques.

L’accord et la cohérence observés dans les estimations de parasitémie entre les différentes techniques soulignent une certaine fiabilité ainsi que l’exactitude des mesures obtenues, indépendamment de la méthode utilisée.

En combinant ces différentes étapes d’évaluation, nous sommes en mesure de fournir une analyse complète de l’efficacité et de la pertinence d’un système tel que MALARIS pour l’estimation de la parasitémie. Ces résultats contribueront à notre compréhension de l’apport potentiel de notre système dans le diagnostic et le suivi du paludisme, en offrant une alternative précise et automatisée aux méthodes existantes.

Cependant, il convient de souligner que notre travail ne s’arrête pas là. La discussion qui suit analysera en détail ces résultats, leurs implications et leurs limitations. Nous aborderons également les perspectives futures pour améliorer et optimiser nos méthodes, afin de fournir des outils plus précis et fiables pour le diagnostic et la surveillance du paludisme.

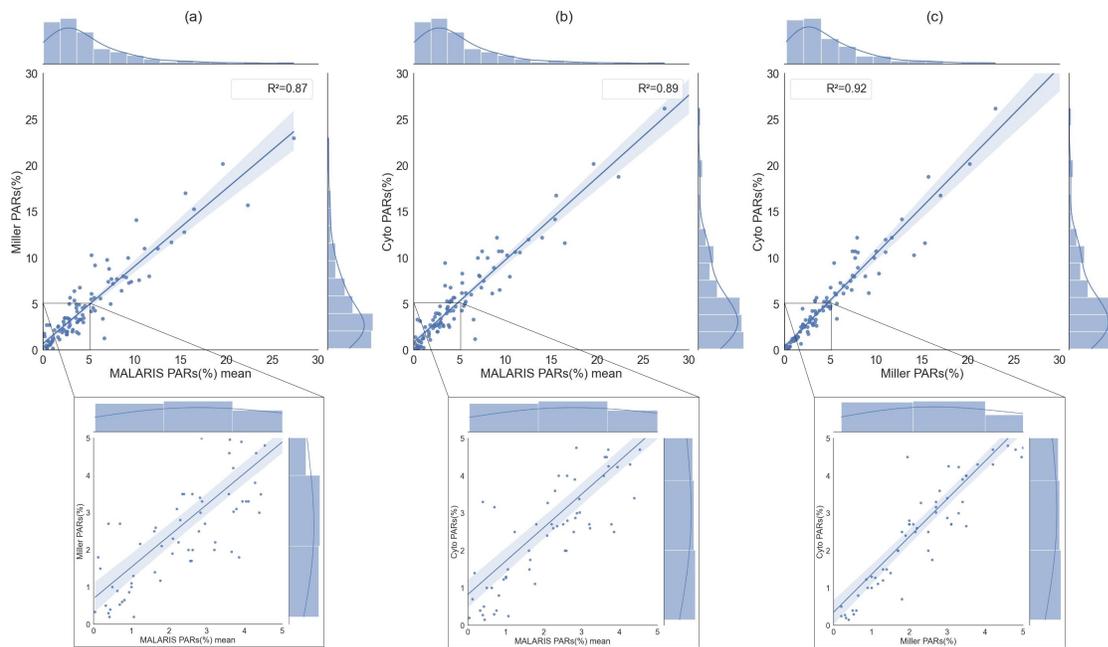


FIGURE 4.9 – Comparaison des graphiques de régression pour l’estimation de la parasitémie. MALARIS (axe des x) par rapport au cellule de Miller (axe des y) (a), MALARIS (axe des x) par rapport à la cytométrie en flux (axe des y) (b), et cellule de Miller (axe des x) par rapport à la cytométrie en flux (axe des y) (c). Sous les graphiques, un accent particulier est mis sur les valeurs de parasitémie comprises entre 0 et 5 %.

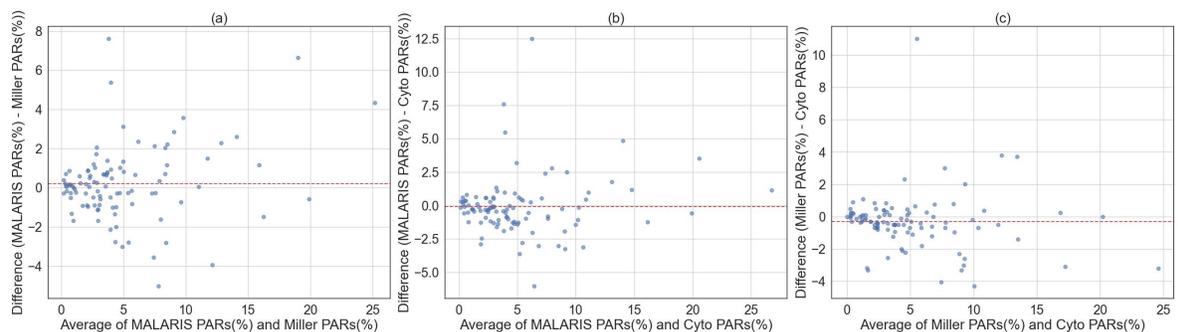


FIGURE 4.10 – Graphiques de Bland-Altman des techniques utilisées pour l’estimation de la parasitémie. (a) Comparaison entre MALARIS et le cellule de Miller. (b) Comparaison entre MALARIS et la cytométrie en flux. (c) Comparaison entre le cellule de Miller et la cytométrie en flux. Les graphiques affichent la différence entre les mesures en fonction de la moyenne des deux mesures, la ligne rouge représentant la différence moyenne.

1 Discussion générale

Dans les chapitres présentés, nous avons entrepris une étude approfondie pour évaluer différentes méthodes et modèles dans le but d'identifier de manière précise les globules rouges parasités par *Plasmodium falciparum* et d'estimer la parasitémie dans le contexte du paludisme d'importation. Pour cela, nous avons construit une base de données inédite d'images comprenant une grande diversité d'objets sanguins présents dans des frottis minces. Nous avons comparé la capacité de généralisation de nos modèles, entraînés à la fois sur notre base de données et sur des images provenant du NIH. Les résultats ont démontré une meilleure généralisation de nos modèles, leur permettant ainsi d'identifier de manière plus précise les globules rouges parasités.

Par la suite, nous avons sélectionné notre meilleur modèle pour l'utiliser à travers un système appelé MALARIS, qui permet une estimation automatique de la parasitémie. Nous avons évalué la répétabilité de notre système et l'impact du nombre de champs nécessaires pour obtenir une mesure fiable de la parasitémie. De plus, nous avons analysé la fiabilité de ces mesures en comparant les estimations du système MALARIS avec d'autres techniques couramment utilisées en laboratoire pour la mesure de la parasitémie. Cette comparaison a permis de mettre en évidence la concordance et la fiabilité des mesures fournies par notre système MALARIS.

La présente thèse a pour objectif d'étudier et d'évaluer différentes méthodes et modèles pour l'identification de globules rouges parasités par *Plasmodium falciparum* et d'estimer de manière fiable la parasitémie qui est un facteur clinique important dans un contexte de paludisme d'importation. Dans cette discussion, nous analyserons les principales contributions de notre travail, les résultats obtenus ainsi que les implications et les limitations de nos travaux.

Les résultats de notre travail mettent en évidence plusieurs contributions importantes. Tout d'abord, nous avons construit un ensemble de données CNR représentatif d'un contexte de réalité diagnostique. Cet ensemble de données est composé d'images de globules rouges infectés par *Plasmodium falciparum* et d'images de composants sanguins non infectés, provenant d'un échantillon de patients important (150 patients négatifs et 50 patients positifs). Cette constitution minutieuse de l'ensemble de données inspiré de la stratégie « noisy student » [Xie et al., 2020] permet d'obtenir des résultats plus fiables et pertinents dans un contexte clinique réel. Nous avons mené

des expériences sur deux ensembles de données différents incluant l'ensemble de données NIH et celui construit dans ce travail (voir le tableau en annexe 3). Notre ensemble de données a été constitué à partir d'images acquises à partir de frottis sanguins minces avec un grossissement de $\times 500$, incluant une plus grande diversité biologique et morphologique pour améliorer les exigences de généralisation du modèle. Le deuxième ensemble de données utilisait des images individuelles acquises à partir de frottis sanguins minces avec un grossissement de $\times 1\ 000$ et a été mis à disposition par le NIH. Dans les résultats de l'étude de validation croisée présentés dans le tableau 3.2, ResNet-50-CNR a obtenu les meilleurs scores de performance, avec une sensibilité de $98,5 \pm 0,5$ %, une spécificité de $99,7 \pm 0,1$ % et un F1-score de $98,8 \pm 0,5$ %. Pour le jeu de données NIH, VGG-19-NIH a donné les meilleurs résultats avec une sensibilité de $98,4 \pm 0,6$ %, une spécificité de $98,6 \pm 1$ % et un F1-score de $98,6 \pm 1,2$ %. Dans l'ensemble, nous pouvons constater que les modèles entraînés avec notre base de données CNR donnent de meilleurs résultats que les modèles entraînés avec la base de données du NIH.

Dans nos expériences de validation des modèles sur de nouveaux patients au niveau cellulaire, en dehors du modèle ResNet-50 entraîné avec l'ensemble de données NIH qui a donné de mauvais résultats, les cinq autres modèles ont également montré de bons scores de précision atteignant ou dépassant 99 % pour VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR, et 96,9 % de précision pour VGG-19-NIH et 93,9 % pour EfficientNet-B7-NIH. En travaillant avec l'ensemble de données NIH, plusieurs études se sont concentrées sur l'amélioration continue de différentes architectures de réseaux en utilisant des méthodes d'ensemble [Rajaraman et al., 2019] et ont montré une précision élevée allant de 97,77 % à 99,5 % pour la classification au niveau cellulaire [Allassaf and Yacin Sikkandar, 2022]. Cependant, les méthodes d'ensemble peuvent être très coûteuses en termes de ressources de calcul et prendre beaucoup de temps pour l'analyse de plusieurs globules rouges par champ. De plus, aucune étude n'a clairement testé ses modèles finaux sur un ensemble de données indépendant séparé pour examiner la généralisation de leurs systèmes.

Nous avons montré ici que les performances des modèles peuvent varier considérablement lors de l'inférence sur de nouveaux patients. En augmentant la diversité des données morphologiquement mais aussi biologiquement (Figure 3.2), en particulier pour la classe qui peut présenter beaucoup de variabilité telle que les artefacts ou des cellules qui se chevauchent, dans notre cas la classe des composants sanguins non infectés, nous avons réduit le taux de faux positifs (77,3 % des 150 patients négatifs ont environ 0 ou 1 faux positif). En effet, cela réduit le nombre de faux positifs (581 faux positifs pour ResNet-50-CNR), tout en augmentant la spécificité (99,8 %), mais réduit également la sensibilité (77,9 %). Cependant, même si la sensibilité semblait insuffisante dans cette étape d'inférence sur des patients au niveau cellulaire, cela peut être corrigé en augmentant le nombre de champs à analyser et, par conséquent, le nombre de globules rouges. Le modèle ResNet-50-CNR a pu identifier les 50 patients positifs lorsque l'examen portait sur environ 2 000 globules rouges. Plus le nombre de globules rouges analysés est élevé, plus le risque de manquer un diagnostic positif au niveau du patient est faible.

Les modèles NIH montrent une limite claire dans la classification infecté/non infecté par patient, avec un nombre élevé de faux positifs (27 941 images de composants sanguins non infectés mal classifiées au total), 84 % des patients négatifs ont entre 10 et plus de 100 faux positifs pour VGG-19-NIH et 98,6 % pour EfficientNet-B7-NIH. La présence d'une forme particulière ou d'une tache à l'intérieur d'un globule rouge donne un faux positif, probablement en raison du manque d'images représentatives de composants sanguins non infectés, tels que les corps de Howell-Jolly et les artefacts, les taches de colorant ou les cellules sanguines superposées, y compris les plaquettes, qui sont courantes dans les frottis sanguins colorés (Figure A.1 en annexe A).

Dans les résultats présentés dans le tableau 3.3 pour la classification individuelle des patients,

nos modèles VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR ont montré une spécificité plus élevée en moyenne (53 %) que les modèles NIH (2 % de spécificité). Comme recommandé par l’OMS, la réalisation d’un frottis sanguin mince est une technique de confirmation pour identifier l’espèce ou le stade du parasite et d’estimer la parasitémie; le frottis épais est la technique utilisée en premier lieu pour identifier la présence ou l’absence de parasites [WHO, 2016c]. Ainsi, la spécificité devient moins significative si seuls les frottis minces potentiellement positifs sont analysés par les modèles. De plus, nos résultats montrent un taux de faux positifs très faible par rapport aux modèles NIH, nous pouvons donc envisager de résoudre ce scénario avec un processus de contrôle et de validation manuelle par un expert, comme nous l’avons introduit. Une approche similaire a été introduite par CellaVision DM96 [Yoon et al., 2019], la sensibilité est renforcée de 33,3 % à 85,7 % en introduisant une validation humaine manuelle. Ils ont également montré que la majorité des faux positifs (57,1 %) correspondaient à des cellules sanguines contenant des corps de Howell-Jolly, un composant sanguin non infecté que nous avons également pris en compte dans notre ensemble de données CNR (Figure 3.2). Cette constatation confirme notre hypothèse selon laquelle un diagnostic entièrement automatisé de la malaria par apprentissage profond nécessitera un entraînement renforcé des modèles pour atteindre les meilleures performances attendues.

Dans ce travail d’identification de globules rouges parasités, nous avons démontré que, même si notre ensemble de données était constitué d’images avec un grossissement au x500, nos résultats ont montré de meilleures performances par rapport à l’ensemble de données de malaria NIH construit avec une résolution plus élevée de x1 000. Nous avons également souligné l’importance d’une inférence à travers une étape de validation sur des patients au niveau cellulaire en évaluant nos modèles CNR et NIH sur un ensemble de données complètement indépendant de 200 patients acquis avec différents microscopes et produits par différents opérateurs. Parmi la profusion d’éléments présents dans le sang, la reconnaissance précise d’un parasite à l’intérieur d’un globule rouge est un défi majeur. De plus, les nombreux artefacts liés à la coloration et aux modes de capture d’images sont autant d’éléments qui compliquent encore cette tâche, en particulier pour les modèles NIH. Nos résultats indiquent que pour atteindre cet objectif, il est nécessaire d’obtenir autant d’images que possible dans les différentes conditions rencontrées pour les modèles d’apprentissage en profondeur, en particulier des globules rouges contenant des artefacts et des éléments intra-érythrocytaires tels que les corps de Howell-Jolly.

L’estimation de la parasitémie vient naturellement compléter les contributions mentionnées précédemment. En effet, les scores de sensibilité élevés de nos modèles nous encouragent à aborder cette problématique très peu explorée dans l’état de l’art. Ce travail est le premier à évaluer et à comparer les performances d’un système basé sur des approches d’apprentissage profond pour la mesure de la parasitémie sur des infections à *Plasmodium falciparum* en utilisant à la fois des méthodes manuelles et semi-automatiques. Pour cela, nous avons réalisé une analyse rétrospective. Malgré le rôle crucial de la parasitémie dans l’évaluation de la gravité de la maladie et du risque de mortalité, peu d’études ont évalué les performances des systèmes proposés afin d’estimer cette mesure sur un échantillon de données patients important.

Nous avons réalisé des expériences en évaluant trois techniques différentes. La première était le système d’aide au diagnostic, MALARIS. La deuxième technique utilisait la cellule de Miller validée en interne dans notre laboratoire, et une troisième technique basée sur des marqueurs d’ADN, utilisant la cytométrie en flux. Dans nos travaux, nous avons étudié l’impact du nombre de champs sur la variation du calcul de la parasitémie en prenant 10 images de champ de 14 patients. Nos résultats ont démontré qu’une faible variation de l’erreur standard relative moyenne (RSE) (<10 %) pour 5 images de champ permettait d’obtenir une estimation de la parasitémie (Figure 4.7.b), ce qui équivaut à plus de 5 000 images de globules rouges analysées par notre système. Ces résultats

s'alignent sur les recommandations de l'OMS [WHO, 2016c].

De plus, nous avons mis en évidence la variabilité inter-humaine en microscopie manuelle pour la mesure de la parasitémie. Cette variabilité était significative (avec un écart type relatif moyen de 38,22 % en microscopie standard). La cellule de Miller présentait un écart type relatif moyen de 26,40 %, tandis que notre système MALARIS montrait une variation plus faible de 11,21 %. Cela démontre qu'un système automatisé de calcul de la parasitémie peut réduire significativement la variabilité et les erreurs de comptage lors du diagnostic (Figure 4.5.b). Pour démontrer l'efficacité de notre modèle, nous l'avons comparé à une autre méthode semi-automatique utilisant la cytométrie en flux. MALARIS a montré une forte corrélation ($R^2=0,89$).

Cependant, lors de la comparaison entre MALARIS et la cellule de Miller, nous avons observé une corrélation légèrement inférieure avec un coefficient de détermination ($R^2=0,87$) (figure 4.9.a). Cette diminution de corrélation peut s'expliquer par les différences de nombre d'erythrocytes analysés entre les approches. La cytométrie en flux est une méthode plus sensible d'analyse en solution qui mesure une parasitémie sur un nombre de cellules beaucoup plus important, tandis que la cellule de Miller analyse des champs microscopiques statiques. Cette divergence dans les principes de mesure peut influencer les résultats et contribuer à la corrélation légèrement inférieure observée. Néanmoins, il est important de noter que malgré cette légère différence, les résultats de MALARIS et de la cytométrie en flux restent hautement corrélés, soulignant le potentiel de notre système à fournir des estimations précises de la parasitémie.

Ainsi, les résultats de notre travail mettent en évidence plusieurs contributions importantes. Tout d'abord, nous avons construit un ensemble de données CNR représentatif d'un contexte de réalité diagnostique, en incorporant une plus grande diversité biologique et morphologique pour améliorer les exigences de généralisation du modèle. Cette diversité a été obtenue en renforçant les images avec du bruit, des artefacts et d'autres caractéristiques spécifiques, tels que les corps de Howell-Jolly. Cette approche a démontré son efficacité en améliorant la spécificité du système, réduisant ainsi le taux de faux positifs et assurant une distinction précise entre les globules rouges infectés et les composants sanguins non infectés. Cependant, il est important de noter que l'introduction de cette diversité peut également entraîner une diminution de la sensibilité du système, en particulier lors de la validation des patients au niveau cellulaire où un nombre limité de globules rouges est analysé. Des ajustements du protocole d'évaluation peuvent être nécessaires pour trouver un équilibre optimal entre spécificité et sensibilité.

Par ailleurs, nos résultats soulignent l'importance de la spécificité dans ce contexte. Les modèles d'apprentissage profond développés dans le cadre de cette thèse ont démontré des performances remarquables en termes de spécificité, avec des scores élevés de $99,7 \pm 0,1$ % pour ResNet-50-CNR et de $98,6 \pm 1$ % pour VGG-19-NIH, par exemple. Une spécificité élevée est cruciale pour minimiser les faux positifs et garantir des résultats fiables et pertinents dans un environnement clinique réel. En effet, des diagnostics erronés pourraient avoir des conséquences néfastes sur la prise en charge clinique des patients. Ainsi, notre approche axée sur la spécificité offre un avantage significatif pour améliorer la fiabilité du système d'identification des globules rouges parasités.

De plus, notre système présente certains avantages pratiques par rapport à d'autres méthodes couramment utilisées. Par exemple, notre approche basée sur l'apprentissage profond offre une vitesse d'analyse plus rapide par rapport à la méthode manuelle de Miller et à la cytométrie en flux. La capacité de notre système à analyser rapidement et efficacement un grand nombre d'images de globules rouges permettrait un gain de temps considérable, ce qui serait particulièrement bénéfique dans des contextes où des résultats rapides sont nécessaires, tels que les interventions sur le terrain ou les régions à ressources limitées.

Notre système présente également un avantage économique par rapport à la cytométrie en

flux, qui est une technique plus coûteuse et nécessite des équipements spécialisés. En utilisant des microscopes standard et une approche d'apprentissage profond, notre système offre une alternative plus abordable et plus accessible pour le diagnostic et la surveillance du paludisme. Cela pourrait avoir des implications significatives dans les régions où les ressources financières et matérielles sont limitées, en permettant une utilisation plus étendue de la technologie et une amélioration de l'accès aux soins de santé.

Malgré ces avancées, certaines limites doivent être prises en compte. Tout d'abord, bien que nos modèles aient montré de bonnes performances dans la classification des globules rouges parasités par *Plasmodium falciparum*, leur généralisation sur d'autres images pourrait se traduire par une baisse des performances. Des études supplémentaires et des ensembles de données spécifiques seraient nécessaires pour évaluer la capacité des modèles à détecter et classifier des images d'autres colorations par exemple. De plus, bien que notre approche d'augmentation de la diversité ait amélioré la spécificité, elle a également entraîné une diminution de la sensibilité. Des ajustements du protocole d'évaluation et du nombre de globules rouges analysés peuvent être nécessaires pour obtenir des performances diagnostiques équilibrées.

En effet, il convient de noter que notre étude est une étude monocentrique portant sur des cas de paludisme d'importation, avec des données provenant d'un seul centre médical acquises dans des conditions différentes. Cela peut présenter certaines limitations en termes de généralisation des résultats à d'autres établissements de santé et populations. Pour une évaluation plus complète et robuste de notre système, il serait important d'étendre l'étude à un cadre multicentrique, en incluant des données couvrant une plus grande diversité de patients et de conditions cliniques.

L'évaluation multicentrique permettrait d'examiner la performance du système dans des environnements cliniques variés, avec différentes pratiques de laboratoire, des opérateurs de microscopes différents et des variations potentielles dans les protocoles de coloration et d'acquisition d'images. Cela permettrait également de mieux évaluer la généralisation et la robustesse du système dans des conditions réelles et d'identifier d'éventuelles sources de variation ou de biais spécifiques à chaque centre.

En réalisant une étude multicentrique, il serait possible de recueillir des données provenant de différentes régions géographiques, présentant des variations dans la prévalence des différentes espèces de parasites du paludisme. Cela permettrait de tester la capacité des modèles à détecter et classifier efficacement les parasites d'autres espèces, ce qui est essentiel pour une application plus large et mondiale de notre système.

Par conséquent, l'évaluation multicentrique constitue une étape importante pour valider la performance et l'applicabilité de notre système dans des contextes plus diversifiés et réalistes. Cela contribuerait à renforcer la confiance dans les résultats obtenus dans ce travail et à fournir des informations précieuses pour une éventuelle implémentation clinique à plus grande échelle.

Dans l'ensemble, notre système combine des performances diagnostiques élevées, une vitesse d'analyse rapide et une approche économiquement avantageuse, ce qui en fait une solution prometteuse pour le diagnostic du paludisme dans des contextes variés, en particulier sur le terrain et dans les régions où les ressources sont limitées. Cependant, des études supplémentaires et une validation clinique approfondie seraient nécessaires pour confirmer ces avantages et évaluer la faisabilité de déploiement à plus grande échelle.

En conclusion, notre recherche a abouti à des avancées significatives dans l'identification de globules rouges parasités par *Plasmodium falciparum* et l'estimation de la parasitémie. Nous avons démontré l'importance de la spécificité dans un contexte clinique, ainsi que les avantages de renforcer les images avec de la diversité pour améliorer la généralisation du système. Cependant, certaines limites doivent être prises en compte et adressées dans les futures recherches afin d'améliorer

rer encore les performances et l'applicabilité clinique du système développé.

Conclusions et Perspectives

Au cours de ce doctorat, qui a débuté fin 2020, nous avons concentré nos efforts sur l'amélioration de l'état de l'art et, dans une certaine mesure, contribué au changement de paradigme en cours dans le domaine du diagnostic automatique du paludisme à partir d'images microscopiques de frottis sanguins. Les modèles profonds pré-entraînés ont eu un impact sur le domaine au-delà de toute attente, au point que les modèles de classifications sont devenus de véritables atouts ouvrant la voie à la prochaine génération de systèmes d'aide au diagnostic en imagerie. Ce paradigme a fortement soutenu nos réponses aux deux questions de recherche :

QR1 Est-il crucial d'avoir un jeu de données riche avec une diversité morphologique et biologique afin d'avoir un modèle capable de différencier de manière fiable un globule rouge parasité par *Plasmodium. falciparum* d'un globule rouge non infecté ou d'un autre composant sanguin ?

QR2 Grâce à un modèle d'apprentissage profond avec une sensibilité élevée, est-il possible d'estimer de manière automatique, fiable et rapide une parasitémie sur des images de frottis sanguins minces au x500 ?

1 Synthèse des contributions

QR1 Dans le chapitre 3, notre principal objectif était de construire un jeu de données représentatif de la réalité diagnostique pour le paludisme à l'aide du système MALARIS. Nous avons développé un ensemble de données CNR composé d'images de globules rouges infectés par *P. falciparum* et d'images de composants sanguins non infectés provenant d'un échantillon de patients diversifié. Nous avons ensuite évalué les performances de plusieurs modèles, dont le modèle ResNet-50 entraîné avec le jeu de données NIH. Bien que ce dernier ait montré de faibles performances, les cinq autres modèles ont affiché de bons scores de précision, atteignant ou dépassant 99 % pour les modèles VGG-19-CNR, ResNet-50-CNR et EfficientNet-B7-CNR, et 96,9 % pour VGG-19-NIH et 93,9 % pour EfficientNet-B7-NIH. Nous avons également comparé nos modèles CNR avec les modèles NIH existants. Les modèles entraînés sur les données CNR ont montré une meilleure précision et un plus faible taux de faux positifs par rapport aux modèles NIH. De

plus, nous avons souligné les limites des modèles NIH en matière de classification infecté/non infecté par patient, avec un grand nombre de faux positifs et une présence de formes particulières ou de taches dans les globules rouges qui induisent ces erreurs. Nous avons également souligné l'importance de la validation des modèles au niveau du patient, en évaluant nos modèles CNR sur un ensemble de patients indépendant. Nos modèles ont montré de bonnes performances dans des conditions réelles, avec une spécificité plus élevée que les modèles NIH.

QR2 Le chapitre 4 apporte plusieurs contributions significatives dans le domaine de l'estimation de la parasitémie pour les infections à *P. falciparum*. La détermination du nombre optimal de champs requis pour estimer une parasitémie sur des images microscopiques de frottis minces au x500, en utilisant le système MALARIS. La mise en évidence des limites des limitations des techniques manuelles actuelles utilisées dans l'estimation de la parasitémie. En comparant différentes méthodes, on souligne l'importance de l'automatisation pour réduire la variabilité et les erreurs de comptage. Enfin, une évaluation comparative de trois techniques : L'étude évalue et compare trois techniques pour l'estimation de la parasitémie, en l'occurrence le système proposé MALARIS, la cytométrie en flux semi-automatique et la cellule de Miller. Cette évaluation comparative permet de mieux comprendre les performances et les avantages de chaque méthode. Ce chapitre contribue à l'amélioration de l'estimation de la parasitémie pour les infections à *P. falciparum* en introduisant un système basé sur la vision et l'apprentissage profond. Il permet d'avoir une estimation précise de la parasitémie dans un délai compatible avec la prise en charge clinique du patient. Cependant, MALARIS peut présenter des limites pour les faibles densités parasitaires (<1%) car 5 000 images peuvent ne pas être suffisantes. Pour surmonter cette limitation, la cytométrie en flux, plus efficace et sensible pour le décompte de la parasitémie faible entre 0,1 % et 1 % [Malleret et al., 2011], pourrait être utilisée en combinaison avec MALARIS. Cependant, le coût élevé, la nécessité de disposer de microscopistes formés à son utilisation et la faible disponibilité de la cytométrie en flux, en particulier dans les zones endémiques avec des ressources limitées, limitent son utilisation dans cette indication.

2 Perspectives de recherche à développer

2.1 Amélioration de la détection des cellules

Nos réponses aux questions **QR1** et **QR2** sont incomplètes, et nous sommes encore loin d'avoir résolu le problème du diagnostic microscopique automatique du paludisme. Dans le chapitre 3, nous avons vu comment différentes architectures de CNN pouvaient répondre au problème d'identification des globules rouges parasités par *P. falciparum* et des composants sanguins non infectés. Toutefois, la détection de ces cellules repose sur une étape de segmentation préalable permettant d'extraire les cellules à partir de leur contour. Il convient de souligner que la méthode de segmentation par watershed [Beucher, 1992], bien que largement utilisée dans le domaine de l'analyse d'images, présente certaines limites. L'une de ces limites concerne les cas où les contrastes entre les cellules et leur environnement sont faibles, ou lorsque les contours des cellules sont mal définis. Dans de telles situations, la segmentation par watershed peut rencontrer des difficultés et produire des résultats imprécis.

La méthode de watershed se base sur la détection de lignes de partage des eaux, qui sont déterminées par des gradients d'intensité dans l'image. Si les contrastes entre les cellules et leur fond sont faibles, les gradients d'intensité peuvent être peu prononcés, ce qui rend la détection des lignes de partage des eaux moins fiable. De même, lorsque les contours des cellules sont flous ou

mal définis, la méthode peut avoir du mal à segmenter correctement les cellules, entraînant ainsi des erreurs de détection et de quantification de la parasitémie (Figure 6.1).

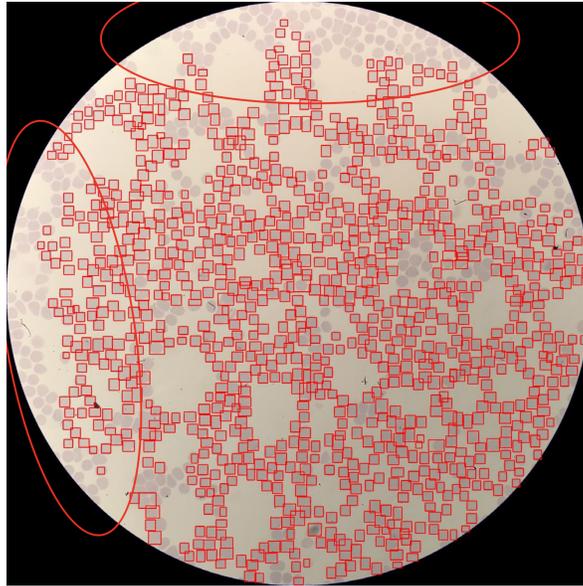


FIGURE 6.1 – Résultat de segmentation des cellules par watershed sur une image de frottis sanguin mince au x500.

Il est donc essentiel de prendre en compte cette étape cruciale et ces limitations lors du développement de la segmentation pour un tel système automatique du diagnostic. Des approches alternatives, telles que l'utilisation de modèles d'apprentissage profond spécifiquement conçus pour la détection de cellules sanguines, pourraient être explorées pour améliorer la précision de la segmentation et, par conséquent, la fiabilité du système proposé.

Afin d'améliorer la précision de la segmentation et de maximiser la détection des cellules parasitées et non parasitées, nous avons exploré l'utilisation du modèle YOLOv8 [Dumitriu et al., 2023] en combinaison avec la méthode de segmentation par watershed. YOLO (You Only Look Once) est un modèle d'apprentissage profond largement utilisé pour la détection de parasites dans les images [Koirala et al., 2022, Shal and Gupta, 2022]. Contrairement aux méthodes traditionnelles de détection d'objets, YOLO propose une approche en temps réel en prédisant directement les boîtes englobantes et les classes des objets dans une seule passe de réseau neuronal. YOLO fonctionne en divisant l'image en une grille de cellules et en prédisant simultanément les boîtes englobantes et les classes des objets présents dans chaque cellule. Il utilise une architecture en flux, ce qui signifie qu'il traite l'ensemble de l'image en une seule passe, ce qui le rend plus rapide que les approches basées sur les régions d'intérêt.

Pour entraîner notre modèle de détection des cellules sanguines, nous avons utilisé le jeu de données BCCD [Shenggan, 2017]. Ce jeu de données est un ensemble d'images spécifiquement conçu pour la détection des cellules sanguines. Il comprend différents types de cellules sanguines, tels que les globules rouges, les globules blancs et les plaquettes. Le BCCD est largement utilisé dans le domaine de la détection des cellules sanguines en raison de sa diversité et de sa représentativité des différentes caractéristiques cellulaires. Il contient des images de bonne qualité capturées à l'aide de microscopes optiques, ce qui en fait une ressource intéressante pour l'entraînement d'un premier modèle de détection d'objets. Cependant, il convient de noter que les données du

BCCD ne sont pas acquises à un grossissement de x500 et ne sont pas capturées dans les mêmes conditions que nos images de diagnostic du paludisme.

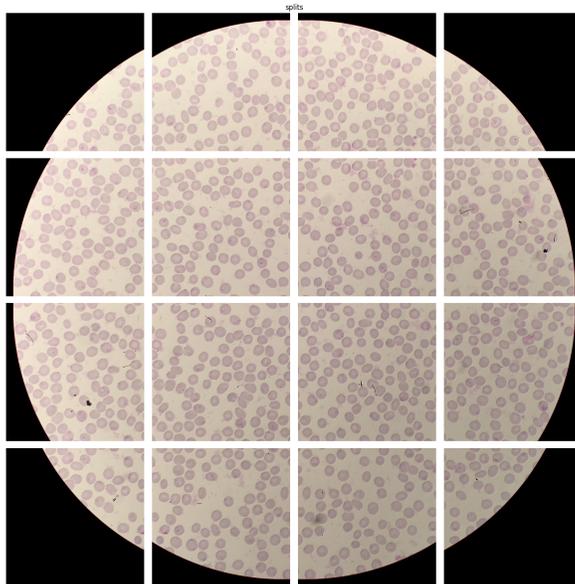


FIGURE 6.2 – Division d’une image de frottis sanguin mince au x500 en 16 imagettes.

Dans notre étude, nous avons adapté les images de nos champs au grossissement de x500, afin de les rendre compatibles avec ce modèle de détection. Il est également important de souligner que les modèles de détection tels que YOLOv8 ne sont pas conçus pour détecter simultanément des objets sur un nombre aussi important (environ 1 000 cellules dans notre cas). Par conséquent, nous avons divisé nos images de champs en sous-ensembles plus petits d’imagettes (Figure 6.2), permettant ainsi au modèle YOLO de détecter les cellules sanguines sur chaque imagette (Figure 6.3).

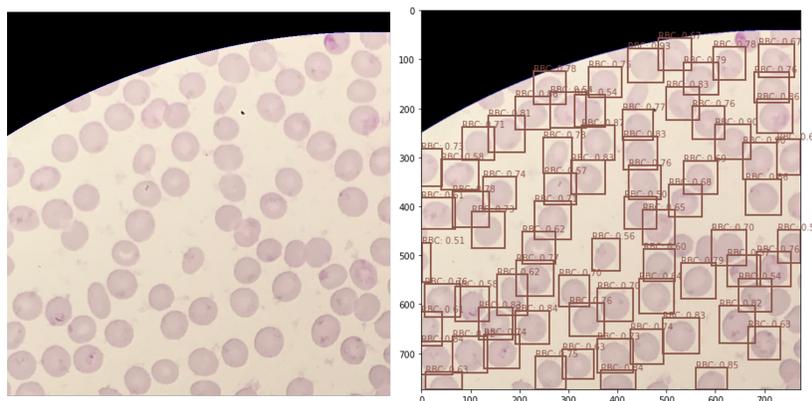


FIGURE 6.3 – Résultat de détection des cellules par YOLOv8 pour chaque imagette.

En combinant YOLOv8 avec la segmentation par watershed, nous avons pu bénéficier des avantages des deux approches. La méthode de segmentation par watershed nous permet d’extraire les cellules à partir de leur contour, en prenant en compte les informations de gradient d’intensité dans l’image. Cela nous permet d’obtenir des contours précis et de segmenter les cellules indivi-

duelles de manière efficace notamment au centre de l'image. Les cellules présentes sur les bords de l'image ont un contraste plus faible (Figure 6.4). D'autre part, YOLOv8 se révèle être un atout précieux en termes de détection des cellules basée sur leur faible contraste et les motifs caractéristiques des cellules (Figure 6.5).

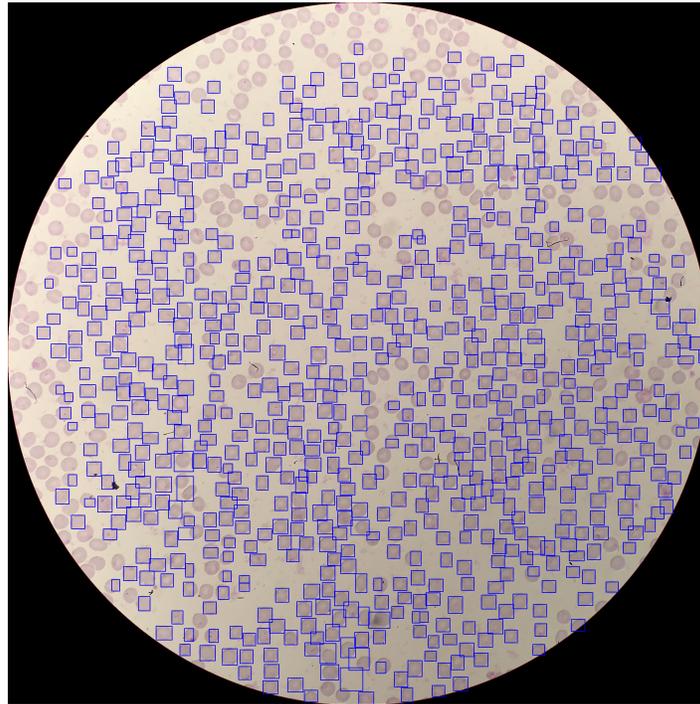


FIGURE 6.4 – Résultat de segmentation des cellules par watershed.

En combinant ces deux approches, nous sommes en mesure d'améliorer la précision de la détection des cellules sanguines, ce qui contribue à une meilleure fiabilité de la détection des cellules d'intérêt mais aussi de l'estimation de la parasitémie. Par ailleurs, pour garantir des résultats précis et éviter les détections multiples, nous avons utilisé une technique de suppression non maximale (NMS). La NMS nous permet de supprimer les détections redondantes en conservant uniquement les boîtes englobantes les mieux ajustées pour chaque objet détecté avec un seuil de recouvrement inférieur à 46 %. Cela nous permet d'obtenir des résultats plus précis et cohérents dans la localisation des cellules sanguines (Figure 6.6).

Il est important de noter que la combinaison de ces deux méthodes peut augmenter le temps de traitement nécessaire pour détecter et segmenter les cellules sanguines par champ. Sur une évaluation de 18 images issues de 18 patients, en moyenne, l'exécution de la méthode de segmentation par watershed prend environ 16 secondes, tandis que l'utilisation du modèle YOLOv8 nécessite environ 10 secondes. L'obtention d'un résultat combinant les deux approches prend en moyenne 28 à 30 secondes (voir le tableau détaillé en annexe 11).

Nous avons évalué la qualité de la segmentation en utilisant l'indice d'intersection sur l'union (IoU) sur un échantillon de 31 images sélectionnées de manière aléatoire provenant de 31 patients différents. L'IoU mesure le degré de chevauchement entre la segmentation automatique et la segmentation manuelle réalisée par des annotateurs. Après avoir comparé les résultats, nous avons obtenu une IoU moyenne de 95 % (voir le tableau détaillé en annexe 12), ce qui indique un meilleur accord entre la segmentation automatique et la vérité terrain manuelle.

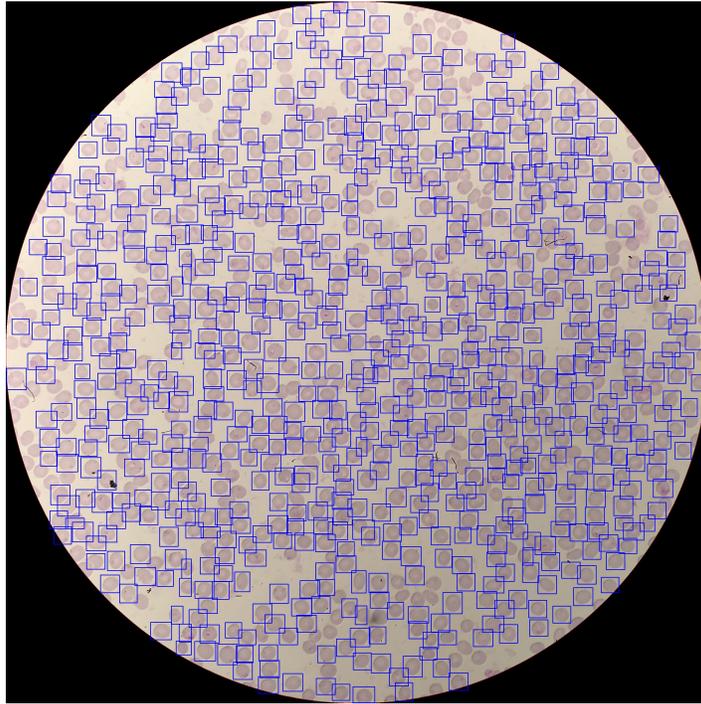


FIGURE 6.5 – Résultat de segmentation des cellules par YOLOv8.

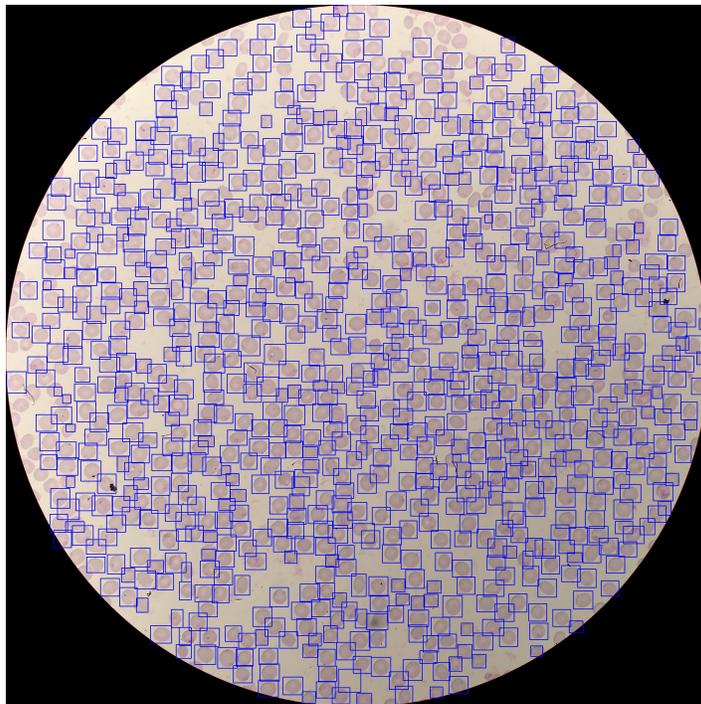


FIGURE 6.6 – Résultat de segmentation des cellules combinant les deux méthodes watershed et YOLOv8.

Bien que notre approche combinée de segmentation par watershed et de détection par YOLOv8 ait donné des résultats prometteurs, il existe toujours des perspectives d'amélioration pour optimiser davantage cette étape de segmentation avant de l'intégrer au sein du système proposé. Voici quelques pistes à explorer :

1. Amélioration de la précision de la détection : bien que YOLOv8 soit un modèle puissant pour la détection d'objets, il peut y avoir des cas où des cellules sont mal détectées en faible quantité, ou des faux positifs sont générés. En augmentant la taille du jeu de données d'entraînement initialement utilisé BCCD, nous pourrions améliorer la précision de la détection et généraliser la segmentation sur plus de contrastes.
2. Exploration de modèles de détection pour la segmentation et la classification directe des cellules : Bien que YOLOv8 ait été utilisé pour la segmentation et non la classification des cellules, il serait intéressant d'envisager d'appliquer les modèles de détection directement pour la classification des images en plusieurs classes. Cela suggère un effort d'annotation manuelle important ou l'utilisation d'un modèle de détection d'objet comme premier annotateur automatique. Cependant, il est important de noter que les globules rouges non infectés sont bien plus représentés que les globules rouges parasités, donnant lieu à un déséquilibre des classes.
3. Validation clinique approfondie : Pour garantir l'applicabilité clinique de notre méthode, une validation approfondie sur un ensemble de données multi-centrique et diversifié est nécessaire. Cela permettra de vérifier la robustesse de cette approche sur différentes images avec des colorations différentes et des sources microscopiques variées tout en évaluant la fiabilité de la mesure de la parasitémie.

En explorant ces perspectives d'amélioration, nous pourrions continuer à améliorer la performance de notre méthode de segmentation et de détection, ce qui aurait un impact significatif sur la précision et la fiabilité du système MALARIS.

2.2 Amélioration du jeu de données CNR : qualité et diversité

Pour améliorer le jeu de données CNR utilisé dans notre travail, il est important de prendre en compte la qualité et la diversité des échantillons. Actuellement, notre jeu de données se concentre principalement sur la détection et la quantification des cellules sanguines infectées par *Plasmodium falciparum*, qui est l'espèce de paludisme la plus mortelle. Cependant, il est essentiel de considérer d'autres co-infections qui peuvent survenir, telles que celles causées par d'autres espèces de *Plasmodium* ou d'autres pathogènes.

Une des perspectives d'amélioration serait d'étendre le jeu de données CNR en incluant des échantillons représentatifs de différentes co-infections possibles. Par exemple, il serait intéressant d'inclure des échantillons avec des globules rouges parasités par *Plasmodium vivax*, qui peuvent poser des défis supplémentaires pour la détection de globules rouges infectés. De plus, les globules rouges non infectés par des *Plasmodium* contenant des corps de Howell-Jolly, des inclusions anormales souvent observées dans certaines conditions médicales, pourraient également être inclus comme des classes de sortie des modèles d'apprentissage pour enrichir la diversité des échantillons.

Actuellement, notre jeu de données utilise une résolution de x500, ce qui était déjà un défi important pour l'identification des globules parasités. Cependant, il serait bénéfique d'améliorer la qualité des données en augmentant la résolution. Cela pourrait être réalisé en utilisant des moyens d'acquisition tels que des microscopes avancés ou d'autres techniques de numérisation de lames.

Une meilleure résolution permettrait une visualisation plus détaillée des cellules sanguines, ce qui faciliterait leur identification et leur segmentation précise. Nous reconnaissons que l'utilisation de techniques d'imagerie avancées, pourrait potentiellement améliorer la qualité des données en offrant une visualisation plus détaillée des cellules sanguines tout en ayant un effectif plus important.

Cependant, nous sommes également conscients que ces avancées technologiques peuvent entraîner des coûts élevés et poser des problèmes d'accessibilité pour les pays aux ressources limitées. Dans ce contexte, il est important de trouver un équilibre entre l'amélioration de la qualité des données et la disponibilité des ressources. Par exemple, plutôt que de se concentrer uniquement sur l'augmentation de la résolution, il serait possible d'explorer d'autres techniques d'amélioration d'image qui peuvent être plus abordables et accessibles, tout en maintenant une qualité acceptable des frottis.

En résumé, pour améliorer le jeu de données CNR, il est essentiel d'élargir la diversité des échantillons en incluant différentes co-infections possibles et en collaborant avec des partenaires et des institutions afin de collecter des échantillons représentatifs de différentes populations et conditions médicales. Cela permettrait d'obtenir une plus grande variété de cellules sanguines.

2.3 Optimisation des performances et déploiement du système

Nous pourrions également aborder la question de la vitesse de traitement et d'analyse des champs sanguins. Une perspective d'amélioration consisterait à explorer des approches de parallélisation du traitement, en utilisant des techniques telles que le traitement distribué ou le calcul sur GPU, afin d'accélérer le processus global. Cela permettrait de réduire le temps nécessaire pour analyser un grand nombre d'images, ce qui est particulièrement crucial dans des contextes où le temps de traitement est un facteur critique pour le diagnostic du paludisme. Parallèlement, il serait intéressant d'étudier l'impact de la réduction de la taille des images sur les performances des modèles. En diminuant la résolution des images, nous pourrions potentiellement accélérer le traitement sans compromettre de manière significative la précision de la détection et de la segmentation. Une analyse comparative entre différentes résolutions d'images pourrait fournir des informations précieuses sur le rapport entre la qualité des résultats et les ressources nécessaires en termes de temps de calcul et de stockage.

Une perspective d'amélioration consiste à développer une plateforme en ligne qui permettrait aux utilisateurs d'accéder facilement au système MALARIS via un navigateur web. Cette plateforme pourrait offrir une interface conviviale où les utilisateurs pourraient télécharger leurs images de frottis sanguins, les soumettre au système MALARIS et recevoir rapidement les résultats de la détection et de la quantification de la parasitémie. L'utilisation d'une plateforme en ligne faciliterait l'accès au système, en permettant aux utilisateurs de différentes régions et institutions d'utiliser le système sans avoir à installer de logiciel supplémentaire. Cependant, il est important de souligner que le déploiement d'un tel système en ligne demande des ressources importantes en matière de sécurité et de stockage cloud. Étant donné que le système est destiné à être utilisé dans un contexte médical, il est essentiel de garantir la confidentialité et la protection des données des utilisateurs. Cela nécessite la mise en place de mesures de sécurité avancées, telles que le chiffrement des données, les contrôles d'accès appropriés et la gestion rigoureuse de la conformité réglementaire. De plus, le déploiement en ligne d'un dispositif médical tel que le système d'aide au diagnostic requiert également une conformité réglementaire stricte. Il est nécessaire de respecter les normes et les réglementations en vigueur dans le domaine de la santé pour assurer la sécurité des utilisateurs et la fiabilité du système. Cela implique une validation et une certification approfondies,

ainsi qu'une surveillance continue pour garantir la conformité aux exigences réglementaires.

En conclusion, notre étude met en évidence plusieurs perspectives d'amélioration pour le système proposé. Nous envisageons l'intégration de modèles de détection pour la classification directe des cellules, l'enrichissement du jeu de données avec une diversité biologique incluant différentes espèces de *Plasmodium* ainsi que des échantillons négatifs plus variés. L'amélioration de la qualité des frottis et des échantillons, l'optimisation des performances en termes de vitesse et de robustesse des modèles, le déploiement en ligne du système, et la prise en compte des aspects de sécurité et de conformité réglementaire sont également des axes importants d'amélioration. En poursuivant ces pistes, nous pourrions renforcer les performances et l'utilité du système dans la détection et l'identification de globules rouges infectés par le paludisme, contribuant ainsi à une meilleure lutte contre cette maladie.

- A. Abdul Nasir, M. Mashor, and Z. Mohamed. Segmentation based approach for detection of malaria parasites using moving k-means clustering. In *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, pages 653–658, Langkawi, Malaysia, Dec. 2012. IEEE. ISBN 978-1-4673-1666-8 978-1-4673-1664-4. doi: 10.1109/IECBES.2012.6498073. URL <http://ieeexplore.ieee.org/document/6498073/>.
- A. Acherar, I. Tantaoui, M. Thellier, A. Lampros, R. Piarroux, and X. Tannier. Real-life evaluation of deep learning models trained on two datasets for plasmodium falciparum detection with thin blood smear images at 500x magnification. *Informatics in Medicine Unlocked*, 35 :101132, 2022.
- A. Alassaf and M. Yacin Sikkandar. Intelligent Deep Transfer Learning Based Malaria Parasite Detection and Classification Model Using Biomedical Image. *Computers, Materials & Continua*, 72(3) :5273–5285, 2022. ISSN 1546-2226. doi: 10.32604/cmc.2022.025577. URL <https://www.techscience.com/cmc/v72n3/47470>.
- A. Albertini, E. Lee, S. O. Coulibaly, M. Sleshi, B. Faye, M. L. Mationg, K. Ouedraogo, A. G. Tsadik, S. M. Felleke, I. Diallo, et al. Malaria rapid diagnostic test transport and storage conditions in burkina faso, senegal, ethiopia and the philippines. *Malaria Journal*, 11 :1–9, 2012.
- V. A. Alegana, E. A. Okiro, and R. W. Snow. Routine data for malaria morbidity estimation in africa : challenges and prospects. *BMC medicine*, 18 :1–13, 2020.
- L. F. W. Anthony, B. Kanding, and R. Selvan. Carbontracker : Tracking and Predicting the Carbon Footprint of Training Deep Learning Models. *arXiv :2007.03051 [cs, eess, stat]*, July 2020. URL <http://arxiv.org/abs/2007.03051>.
- O. Bailo, D. Ham, and Y. Min Shin. Red blood cell image generation for data augmentation using conditional generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 0–0, 2019.
- M. K. Bashar. Automated Classification of Malaria Parasite Stages Using Convolutional Neural Network-Classification of Life-cycle Stages of Malaria Parasites. In *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, pages 1–5, Vancouver BC Canada, Aug. 2019. ACM. ISBN 978-1-4503-7625-9. doi: 10.1145/3387168.3387185. URL <https://dl.acm.org/doi/10.1145/3387168.3387185>.
- H.-P. Beck. Digital microscopy and artificial intelligence could profoundly contribute to malaria diagnosis in elimination settings. *Frontiers in artificial intelligence*, 5 :510483, 2022.
- M. Beckmann, N. F. Ebecken, and B. S. Pires de Lima. A knn undersampling approach for data balancing. *Journal of Intelligent Learning Systems and Applications*, 7(04) :104–116, 2015.
- S. Bempah, A. Curtis, G. Awandare, and J. Ajayakumar. Appreciating the complexity of localized malaria risk in ghana : Spatial data challenges and solutions. *Health & Place*, 64 :102382, 2020.
- S. Beucher. The watershed transformation applied to image segmentation. *Scanning Microscopy*, 1992(6) : 28, 1992.

- O. Bouchaud, F. Bruneel, E. Caumes, S. Houzé, P. Imbert, B. Pradines, C. Rapp, and C. Strady. Management and prevention of imported malaria. 2018 update of the 2007 french clinical guidelines. *Medecine et maladies infectieuses*, 50(2) :161–193, 2020.
- R. N. Bronzan, M. L. McMorrow, and S. Patrick Kachur. Diagnosis of malaria : challenges for clinicians in endemic and non-endemic regions. *Molecular diagnosis & therapy*, 12(5) :299–306, 2008.
- F. Chollet. Xception : Deep Learning with Depthwise Separable Convolutions. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1800–1807, Honolulu, HI, July 2017. IEEE. ISBN 978-1-5386-0457-1. doi: 10.1109/CVPR.2017.195. URL <http://ieeexplore.ieee.org/document/8099678/>.
- D. K. Das, A. K. Maiti, and C. Chakraborty. Automated system for characterization and classification of malaria-infected stages using light microscopic images of thin blood smears. *Journal of Microscopy*, 257(3) :238–252, Mar. 2015. ISSN 1365-2818. doi: 10.1111/jmi.12206.
- J. De Fauw, J. R. Ledsam, B. Romera-Paredes, S. Nikolov, N. Tomasev, S. Blackwell, H. Askham, X. Glorot, B. O’Donoghue, D. Visentin, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, 24(9) :1342–1350, 2018.
- C. Dong, C. C. Loy, and X. Tang. Accelerating the Super-Resolution Convolutional Neural Network. *arXiv :1608.00367 [cs]*, Aug. 2016. URL <http://arxiv.org/abs/1608.00367>.
- A. Dumitriu, F. Tatui, F. Miron, R. T. Ionescu, and R. Timofte. Rip current segmentation : A novel benchmark and yolov8 baseline results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1261–1271, 2023.
- I. Ersoy, F. Bunyak, J. Higgins, and K. Palaniappan. Coupled edge profile active contours for red blood cell flow. 2012.
- G. A. Farcas, K. J. Zhong, T. Mazzulli, and K. C. Kain. Evaluation of the realart malaria lc real-time pcr assay for malaria diagnosis. *Journal of Clinical Microbiology*, 42(2) :636–638, 2004.
- G. P. Gopakumar, M. Swetha, S. S. Gorthi, and G. Sai Subrahmanyam. CNN based Malaria Diagnosis from Focus-stack of Blood Smear Images Acquired using Custom-built Slide Scanner. *Journal of Biophotonics*, 11 :e201700003, Aug. 2017. doi: 10.1002/jbio.201700003.
- K. S. Griffith, L. S. Lewis, S. Mali, and M. E. Parise. Treatment of malaria in the united states : a systematic review. *Jama*, 297(20) :2264–2277, 2007.
- B. T. Grimberg. Methodology and application of flow cytometry for investigation of human malaria parasites. *Journal of immunological methods*, 367(1-2) :1–16, 2011.
- J.-O. Guintran, C. Delacollette, P. I. Trigg, W. H. Organization, et al. Systems for the early detection of malaria epidemics in africa : an analysis of current practices and future priorities. *World Health Organization*, 2006.
- K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Dec. 2015. URL <http://arxiv.org/abs/1512.03385>.
- M. A. Hearst. Trends & Controversies : Support Vector Machines. *IEEE Intell. Syst.*, 1998. doi: 10.1109/5254.708428.
- J. Heide, K. C. Vaughan, A. Sette, T. Jacobs, and J. Schulze zur Wiesch. Comprehensive review of human plasmodium falciparum-specific cd8+ t cell epitopes. *Frontiers in Immunology*, 10 :397, 2019.
- A.-K. Heuschen, G. Lu, O. Razum, A. Abdul-Mumin, O. Sankoh, L. Von Seidlein, U. D’Alessandro, and O. Müller. Public health-relevant consequences of the covid-19 pandemic on malaria in sub-saharan africa : a scoping review. *Malaria journal*, 20 :1–16, 2021.
- H. Hopkins, I. J. González, S. D. Polley, P. Angutoko, J. Ategeka, C. Asimwe, B. Agaba, D. J. Kyabayinze, C. J. Sutherland, M. D. Perkins, et al. Highly sensitive detection of malaria parasitemia in a malaria-endemic setting : performance of a new loop-mediated isothermal amplification kit in a remote clinic in uganda. *The Journal of infectious diseases*, 208(4) :645–652, 2013.
- G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

- J. Hung and A. Carpenter. Applying Faster R-CNN for Object Detection on Malaria Images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 808–813, Honolulu, HI, USA, July 2017. IEEE. ISBN 978-1-5386-0733-6. doi: 10.1109/CVPRW.2017.112. URL <https://ieeexplore.ieee.org/document/8014846/>.
- M. R. Islam, M. Nahiduzzaman, M. O. F. Goni, A. Sayeed, M. S. Anower, M. Ahsan, and J. Haider. Explainable transformer-based deep learning model for the detection of malaria parasites from blood cell images. *Sensors*, 22(12):4358, 2022.
- C. J. Janse and P. H. Van Vianen. Flow cytometry in malaria detection. In *Methods in cell biology*, volume 42, pages 295–318. Elsevier, 1994.
- N. A. Khan, H. Pervaz, A. K. Latif, A. Musharraf, et al. Unsupervised identification of malaria parasites using computer vision. In *2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pages 263–267. IEEE, 2014.
- T. A. Khartabil, Y. B. de Rijke, R. Koelewijn, J. J. van Hellemond, and H. Russcher. Fast detection and quantification of plasmodium species infected erythrocytes in a non-endemic region by using the sysmex xn-31 analyzer. *Malaria Journal*, 21(1):119, 2022.
- A. Kiszewski, B. Johns, A. Schapira, C. Delacollette, V. Crowell, T. Tan-Torres, B. Ameneshewa, A. Teklehaimanot, and F. Nafo-Traoré. Estimated global resources needed to attain international malaria control goals. *Bulletin of the World Health Organization*, 85(8):623–630, 2007.
- A. Koirala, M. Jha, S. Bodapati, A. Mishra, G. Chetty, P. K. Sahu, S. Mohanty, T. K. Padhan, J. Mattoo, and A. Hukkoo. Deep learning for real-time malaria parasite detection and counting using yolo-mp. *IEEE Access*, 10:102157–102172, 2022.
- A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, May 2017. ISSN 0001-0782, 1557-7317. doi: 10.1145/3065386. URL <https://dl.acm.org/doi/10.1145/3065386>.
- P.-C. Kuo, H.-Y. Cheng, P.-F. Chen, Y.-L. Liu, M. Kang, M.-C. Kuo, S.-F. Hsu, H.-J. Lu, S. Hong, C.-H. Su, D.-P. Liu, Y.-C. Tu, and J.-H. Chuang. Assessment of Expert-Level Automated Detection of Plasmodium falciparum in Digitized Thin Blood Smear Images. *JAMA network open*, 3(2):e200206, Feb. 2020. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2020.0206.
- M.-T. Le, T. R. Bretschneider, C. Kuss, and P. R. Preiser. A novel semi-automatic image processing approach to determine plasmodium falciparum parasitemia in giemsa-stained thin blood smears. *BMC cell biology*, 9:1–12, 2008.
- T. A. Leski, C. R. Taitt, A. G. Swaray, U. Bangura, N. D. Reynolds, A. Holtz, C. Yasuda, J. Lahai, J. M. Lamin, V. Baio, et al. Use of real-time multiplex pcr, malaria rapid diagnostic test and microscopy to investigate the prevalence of plasmodium species among febrile hospital patients in sierra leone. *Malaria journal*, 19(1):1–8, 2020.
- Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. A. Hossain, A. Sameer, R. J. Maude, J. X. Huang, S. Jaeger, and G. Thoma. CNN-based image analysis for malaria diagnosis. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 493–496, Dec. 2016. doi: 10.1109/BIBM.2016.7822567.
- N. Linder, R. Turkki, M. Walliander, A. Mårtensson, V. Diwan, E. Rahtu, M. Pietikäinen, M. Lundin, and J. Lundin. A malaria diagnostic tool based on computer vision screening and visualization of plasmodium falciparum candidate areas in digitized blood smears. *PLoS One*, 9(8):e104855, 2014a.
- N. Linder, R. Turkki, M. Walliander, A. Mårtensson, V. Diwan, E. Rahtu, M. Pietikäinen, M. Lundin, and J. Lundin. A malaria diagnostic tool based on computer vision screening and visualization of Plasmodium falciparum candidate areas in digitized blood smears. *PloS One*, 9(8):e104855, 2014b. ISSN 1932-6203. doi: 10.1371/journal.pone.0104855.
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD : Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, and M. Welling, editors, *Computer Vision – ECCV 2016*, volume 9905, pages 21–37. Springer International Publishing, Cham, 2016a. ISBN 978-3-319-46447-3 978-3-319-46448-0.

- doi:10.1007/978-3-319-46448-0_2. URL http://link.springer.com/10.1007/978-3-319-46448-0_2.
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. SSD : Single Shot MultiBox Detector. *arXiv :1512.02325 [cs]*, 9905 :21–37, 2016b. doi: 10.1007/978-3-319-46448-0_2. URL <http://arxiv.org/abs/1512.02325>.
- J. Long, E. Shelhamer, and T. Darrell. Fully Convolutional Networks for Semantic Segmentation. *arXiv :1411.4038 [cs]*, Mar. 2015. URL <http://arxiv.org/abs/1411.4038>.
- G. Madhu. Computer vision and machine learning approach for malaria diagnosis in thin blood smears from microscopic blood images. *Machine learning for intelligent decision science*, pages 191–209, 2020.
- B. Malleret, C. Claser, A. S. M. Ong, R. Suwanarusk, K. Sriprawat, S. W. Howland, B. Russell, F. Nosten, and L. Rénia. A rapid and robust tri-color flow cytometry assay for monitoring malaria parasite development. *Scientific reports*, 1(1) :118, 2011.
- P. Manescu, M. J. Shaw, M. Elmi, L. Neary-Zajiczek, R. Claveau, V. Pawar, I. Kokkinos, G. Oyinloye, C. Bendkowski, O. A. Oladejo, et al. Expert-level automated malaria diagnosis on routine blood films with deep neural networks. *American Journal of Hematology*, 95(8) :883–891, 2020.
- B. A. Mathison and B. S. Pritt. Update on malaria diagnostics and test utilization. *Journal of clinical microbiology*, 55(7) :2009–2017, 2017.
- C. Mehanian, M. Jaiswal, C. Delahunt, C. Thompson, M. Horning, L. Hu, S. McGuire, T. Ostbye, M. Mehanian, B. Wilson, C. Champlin, E. Long, S. Proux, D. Gamboa, P. Chiodini, J. Carter, M. Dhorda, D. Isaboke, B. Ogutu, W. Oyibo, E. Villasis, K. M. Tun, C. Bachman, and D. Bell. Computer-Automated Malaria Diagnosis and Quantitation Using Convolutional Neural Networks. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 116–125, Venice, Oct. 2017a. IEEE. ISBN 978-1-5386-1034-3. doi: 10.1109/ICCVW.2017.22. URL <http://ieeexplore.ieee.org/document/8265232/>.
- C. Mehanian, M. Jaiswal, C. Delahunt, C. Thompson, M. Horning, L. Hu, S. McGuire, T. Ostbye, M. Mehanian, B. Wilson, C. Champlin, E. Long, S. Proux, D. Gamboa, P. Chiodini, J. Carter, M. Dhorda, D. Isaboke, B. Ogutu, W. Oyibo, E. Villasis, K. M. Tun, C. Bachman, and D. Bell. Computer-Automated Malaria Diagnosis and Quantitation Using Convolutional Neural Networks. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 116–125, Venice, Oct. 2017b. IEEE. ISBN 978-1-5386-1034-3. doi: 10.1109/ICCVW.2017.22. URL <http://ieeexplore.ieee.org/document/8265232/>.
- Y. Michalakis and F. Renaud. Malaria : Evolution in vector control. *Nature*, 462 :298–300, 2009.
- A. Moreno-Sabater, D. Sterlin, L. Imamovic, F. Bon, A.-C. Normand, C. Gonnin, M. Gazzano, M. Bensalah, K. Dorgham, E. B. Salah, et al. Intestinal candida albicans overgrowth in iga deficiency. *Journal of Allergy and Clinical Immunology*, 2023.
- H. Mwenesi, C. Mbogo, N. Casamitjana, M. C. Castro, M. A. Itoe, F. Okonofua, and M. Tanner. Rethinking human resources and capacity building needs for malaria control and elimination in africa. *PLOS Global Public Health*, 2(5) :e0000210, 2022.
- A. A. Nasir, M. Mashor, and Z. Mohamed. Segmentation based approach for detection of malaria parasites using moving k-means clustering. In *2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences*, pages 653–658. IEEE, 2012.
- K. M. Nazi. The miller disk : An improvement in the performance of manual reticulocyte counts. *Laboratory Medicine*, 17 :742–744, Dec 1986.
- A.-C. Normand, A. Chaline, N. Mohammad, A. Godmer, A. Acherar, A. Huguenin, S. Ranque, X. Tannier, and R. Piarroux. Identification of a clonal population of aspergillus flavus by maldi-tof mass spectrometry using deep learning. *Scientific Reports*, 12(1) :1575, 2022.
- O. Ogunfowokan, A. I. Nwajei, and B. A. Ogunfowokan. Sensitivity and specificity of malaria rapid diagnostic test (mrtd carestattm) compared with microscopy amongst under five children attending a primary care clinic in southern nigeria. *African Journal of Primary Health Care and Family Medicine*, 12(1) :1–8, 2020.

- W. P. O'Meara, M. Barcus, C. Wongsrichanalai, S. Muth, J. D. Maguire, R. G. Jordan, W. R. Prescott, and F. E. McKenzie. Reader technique as a source of variability in determining malaria parasite density by microscopy. *Malaria Journal*, 5(1) :118, Dec. 2006. ISSN 1475-2875. doi: 10.1186/1475-2875-5-118. URL <https://doi.org/10.1186/1475-2875-5-118>.
- A. Oviedo, A. Knipes, C. Worrell, L. M. Fox, L. Desir, C. Fayette, A. Javel, F. Monestime, K. Mace, M. A. Chang, et al. Combination of serological, antigen detection, and dna data for plasmodium falciparum provides robust geospatial estimates for malaria transmission in haiti. *Scientific Reports*, 10(1) :8443, 2020.
- C. V. Plowe. Malaria chemoprevention and drug resistance : a review of the literature and policy implications. *Malaria Journal*, 21(1) :104, 2022.
- M. Poostchi, I. Ersoy, K. McMenamin, E. Gordon, N. Palaniappan, S. Pierce, R. J. Maude, A. Bansal, P. Srinivasan, L. Miller, et al. Malaria parasite detection and cell counting for human and mouse using thin blood smear microscopy. *Journal of Medical Imaging*, 5(4) :044506–044506, 2018a.
- M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. Thoma. Image analysis and machine learning for detecting malaria. *Translational Research*, 194 :36–55, 2018b. ISSN 1931-5244. doi: <https://doi.org/10.1016/j.trsl.2017.12.004>. URL <https://www.sciencedirect.com/science/article/pii/S193152441730333X>. In-Depth Review : Diagnostic Medical Imaging.
- B. Pöschl, J. Waneesorn, O. Thekisoe, S. Chutipongvivate, and K. Panagiotis. Comparative diagnosis of malaria infections by microscopy, nested pcr, and lamp in northern thailand. *The American journal of tropical medicine and hygiene*, 83(1) :56, 2010.
- A. Rahman, H. Zunair, M. S. Rahman, J. Q. Yuki, S. Biswas, M. A. Alam, N. B. Alam, and M. Mahdy. Improving malaria parasite detection from red blood cell using deep convolutional neural networks. *arXiv preprint arXiv:1907.10418*, 2019.
- S. Rajaraman, S. K. Antani, M. Poostchi, K. Silamut, M. A. Hossain, R. J. Maude, S. Jaeger, and G. R. Thoma. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. *PeerJ*, 6 :e4568, 2018. ISSN 2167-8359. doi: 10.7717/peerj.4568.
- S. Rajaraman, S. Jaeger, and S. K. Antani. Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PeerJ*, 7, 2019.
- M. Ren. Greater political commitment needed to eliminate malaria. *Infectious diseases of poverty*, 8(02) : 1–4, 2019.
- S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn : Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28, 2015.
- C. Rypien, B. Chow, W. W. Chan, D. L. Church, and D. R. Pillai. Detection of plasmodium infection by the illumi gene malaria assay compared to reference microscopy and real-time pcr. *Journal of clinical microbiology*, 55(10) :3037–3045, 2017.
- S. Sathpathi, A. K. Mohanty, P. Satpathi, S. K. Mishra, P. K. Behera, G. Patel, and A. M. Dondorp. Comparing leishman and giemsa staining for the assessment of peripheral blood smear preparations in a malaria-endemic region in india. *Malaria journal*, 13(1) :1–5, 2014.
- S. Savkare and S. Narote. Automated system for malaria parasite identification. In *2015 international conference on communication, information & computing technology (ICCICT)*, pages 1–4. IEEE, 2015a.
- S. Savkare and S. Narote. Automated system for malaria parasite identification. In *2015 International Conference on Communication, Information & Computing Technology (ICCICT)*, pages 1–4, Mumbai, India, Jan. 2015b. IEEE. ISBN 978-1-4799-5522-0 978-1-4799-5521-3. doi: 10.1109/ICCICT.2015.7045660. URL <http://ieeexplore.ieee.org/document/7045660/>.
- N. Schwalbe and B. Wahl. Artificial intelligence and the future of global health. *The Lancet*, 395(10236) : 1579–1586, 2020.
- D. Selvarajah, C. Naing, N. H. Htet, and J. W. Mak. Loop-mediated isothermal amplification (lamp) test for diagnosis of uncomplicated malaria in endemic areas : a meta-analysis of diagnostic test accuracy. *Malaria journal*, 19 :1–10, 2020.

- N. Sengar, R. Burget, and M. K. Dutta. A vision transformer based approach for analysis of plasmodium vivax life cycle for malaria prediction using thin blood smear microscopic images. *Computer Methods and Programs in Biomedicine*, 224 :106996, 2022.
- A. Shal and R. Gupta. A comparative study on malaria cell detection using computer vision. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pages 548–552. IEEE, 2022.
- Y. Sharma, V. Arya, S. Jain, M. Kumar, L. Deka, and A. Mathur. Dengue and typhoid co-infection—study from a government hospital in north delhi. *Journal of Clinical and Diagnostic Research : JCDR*, 8(12) : DC09, 2014.
- Shenggan. Bccd dataset, 2017. URL https://github.com/Shenggan/BCCD_Dataset.
- D. O. Simba, D. Kakoko, T. Nyamhanga, Z. Mrango, and P. Mujinja. Improving prompt access to malaria diagnostics and treatment in rural remote areas using financial benefit for community health workers in kilosa district, tanzania. *Research and Reports in Tropical Medicine*, pages 137–146, 2018.
- K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv :1409.1556 [cs]*, Apr. 2015a. URL <http://arxiv.org/abs/1409.1556>.
- K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv :1409.1556 [cs]*, Apr. 2015b. URL <http://arxiv.org/abs/1409.1556>.
- J. Talapko, I. Škrlec, T. Alebić, M. Jukić, and A. Včev. Malaria : the past and the present. *microorganisms* 7 (6) : 179, 2019.
- M. Tan and Q. V. Le. EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks, Sept. 2020. URL <http://arxiv.org/abs/1905.11946>.
- K. Torres, C. M. Bachman, C. B. Delahunt, J. Alarcon Baldeon, F. Alava, D. Gamboa Vilela, S. Proux, C. Mehanian, S. K. McGuire, C. M. Thompson, et al. Automated microscopy for routine malaria diagnosis : a field comparison on giemsa-stained blood films in peru. *Malaria journal*, 17 :1–11, 2018.
- R. Varo, N. Balanza, A. Mayor, and Q. Bassat. Diagnosis of clinical malaria in endemic settings. *Expert Review of Anti-infective Therapy*, 19(1) :79–92, 2021.
- T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. *arXiv :1711.11585 [cs]*, Aug. 2018. URL <http://arxiv.org/abs/1711.11585>.
- WHO. Giemsa staining of malaria blood films — who.int. <https://www.who.int/publications/i/item/HTM-GMP-MM-SOP-07a>, 2016a.
- WHO. Malaria parasite counting — who.int. <https://www.who.int/publications/i/item/HTM-GMP-MM-SOP-09>, 2016b.
- WHO. Microscopy examination of thick and thin blood films for identification of malaria parasites — who.int. <https://www.who.int/publications/i/item/HTM-GMP-MM-SOP-08>, 2016c.
- WHO. *World malaria report 2021*. Organisation Mondiale de la Santé, 2021.
- K. J. Wicht, S. Mok, and D. A. Fidock. Molecular mechanisms of drug resistance in plasmodium falciparum malaria. *Annual review of microbiology*, 74 :431–454, 2020.
- C. Wongsrichanalai, M. J. Barcus, S. Muth, A. Sutamihardja, and W. 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.
- Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le. Self-training with Noisy Student improves ImageNet classification, June 2020. URL <http://arxiv.org/abs/1911.04252>.
- F. Yang, H. Yu, K. Silamut, R. J. Maude, S. Jaeger, and S. Antani. Smartphone-Supported Malaria Diagnosis Based on Deep Learning. In H.-I. Suk, M. Liu, P. Yan, and C. Lian, editors, *Machine Learning in Medical Imaging*, volume 11861, pages 73–80. Springer International Publishing, Cham, 2019. ISBN 978-3-030-32691-3 978-3-030-32692-0. doi: 10.1007/978-3-030-32692-0_9. URL http://link.springer.com/10.1007/978-3-030-32692-0_9.

J. Yoon, J. A. Kwon, S. Y. Yoon, W. S. Jang, D. J. Yang, J. Nam, and C. S. Lim. Diagnostic performance of Cel-laVision DM96 for Plasmodium vivax and Plasmodium falciparum screening in peripheral blood smears. *Acta Tropica*, 193 :7–11, May 2019. ISSN 1873-6254. doi: 10.1016/j.actatropica.2019.02.009.

H. Yu, F. Yang, S. Rajaraman, I. Ersoy, G. Moallem, M. Poostchi, K. Palaniappan, S. Antani, R. J. Maude, and S. Jaeger. Malaria Screener : a smartphone application for automated malaria screening. *BMC Infectious Diseases*, 20(1) :825, Dec. 2020. ISSN 1471-2334. doi: 10.1186/s12879-020-05453-1. URL <https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05453-1>.

Z. Zhang, P. Chen, M. McGough, F. Xing, C. Wang, M. Bui, Y. Xie, M. Sapkota, L. Cui, J. Dhillon, et al. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *Nature Machine Intelligence*, 1(5) :236–245, 2019.

O. S. Zhao, N. Kolluri, A. Anand, N. Chu, R. Bhavaraju, A. Ojha, S. Tiku, D. Nguyen, R. Chen, A. Morales, D. Valliappan, J. P. Patel, and K. Nguyen. Convolutional neural networks to automate the screening of malaria in low-resource countries. *PeerJ*, 8 :e9674, Aug. 2020. ISSN 2167-8359. doi: 10.7717/peerj.9674. URL <https://peerj.com/articles/9674>.

G. Zhou, N. Minakawa, A. K. Githeko, and G. Yan. Climate variability and malaria epidemics in the highlands of east africa. *Trends in parasitology*, 21(2) :54–56, 2005.

Annexes

A Annexe A

TABLE 1 – Carbon footprint per fold (5 folds of 30 epochs)

Model	Actual consumption	Predicted consumption
C-CNN-NRC	Actual consumption for 1 epoch(s) :	Predicted consumption for 30 epoch(s) :
	Time : 4 :40 :00	Time : 139 :59 :57
	Energy : 0.117341 kWh	Energy : 3.520237 kWh
	CO2eq : 34.522509 g	CO2eq : 1035.675277 g
	This is equivalent to : 0.286732 km travelled by car	This is equivalent to : 8.601954 km travelled by car
VGG-19-NRC	Actual consumption for 1 epoch(s) :	Predicted consumption for 30 epoch(s) :
	Time : 14 :33 :13	Time : 436 :36 :34
	Energy : 0.365499 kWh	Energy : 10.964984 kWh
	CO2eq : 107.532168 g	CO2eq : 3225.965038 g
	This is equivalent to : 0.893124 km travelled by car	This is equivalent to : 26.793730 km travelled by car
ResNet-50-NRC	Actual consumption for 1 epoch(s) :	Predicted consumption for 30 epoch(s) :
	Time : 50 :33 :20	Time : 1516 :39 :46
	Energy : 0.844492 kWh	Energy : 25.334750 kWh
	CO2eq : 248.454596 g	CO2eq : 7453.637889 g
	This is equivalent to : 2.063576 km travelled by car	This is equivalent to : 61.907291 km travelled by car
EfficientNet-B7-NRC	Actual consumption for 1 epoch(s) :	Predicted consumption for 30 epoch(s) :
	Time : 1 :18 :03	Time : 39 :01 :43
	Energy : 0.303549 kWh	Energy : 9.106474 kWh
	CO2eq : 89.306002 g	CO2eq : 2679.180051 g
	This is equivalent to : 0.741744 km travelled by car	This is equivalent to : 22.252326 km travelled by car

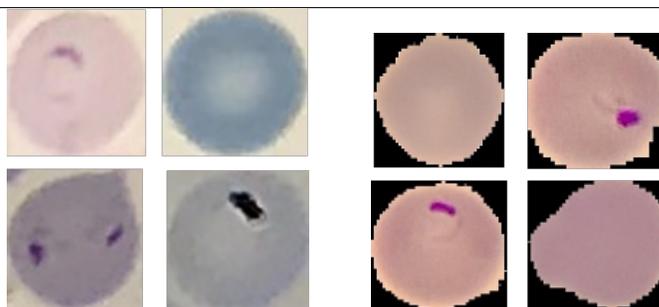
TABLE 2 – Les hyperparamètres des modèles expérimentaux et les métriques d'évaluation

Modèle	Nombre d'époques par pli	Optimiseur	Taux d'apprentissage
C-CNN (Master et Student)	30	RMSProp	1×10^{-5}
VGG-19	30	Adam	1×10^{-5}
ResNet-50	30	Adam	1×10^{-5}
EfficientNet-B7	30	Adam	1×10^{-5}

TABLE 3 – Caractéristiques des ensembles de données CNR et NIH

Caractéristiques	CNR	NIH
Espèce	<i>P. falciparum</i>	<i>P. falciparum</i>
Technique de coloration	Giemsa	Giemsa
Grossissement du microscope	x500	x1 000
Nombre de globules rouges par image de champ	≈1 000	≈250
Nombre total de patients	202	200
Nombre de patients infectés	101	150
Nombre de patients non infectés	101	50
Nombre de composants sanguins non infectés	54 705	13 779
Nombre de globules rouges infectés	13 453	13 779
Segmentation des globules rouges	Watershed [Beucher, 1992]	Level-sets [Ersoy et al., 2012]
Cellules superposées (overlapping)	Oui	Non
Soustraction de l'arrière-plan des cellules individuelles	Non	Oui
Taille des images des cellules sanguines segmentées	≈50 x 50 px	128 x 128 px

Échantillons d'images



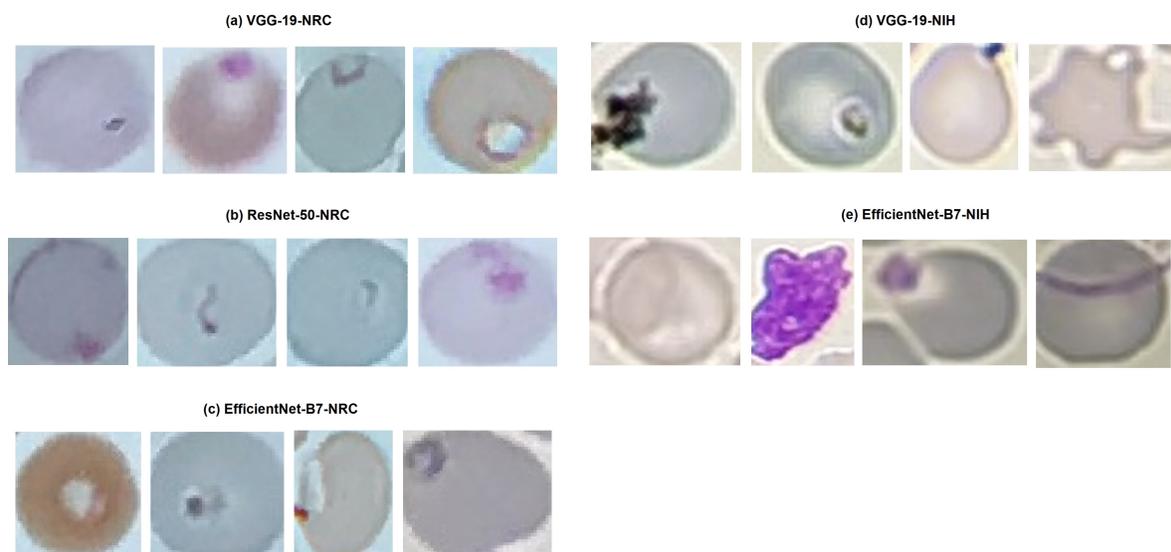


FIGURE A.1 – Échantillons de faux positifs mal classés par les trois modèles. Pour montrer le type d’erreurs de prédiction pour chaque modèle, des images de GR comprenant des artefacts, des chevauchements de plaquettes ou des précipitations de coloration (a), des images de GR comprenant des précipitations de coloration qui induisent une interprétation difficile (b), des images de GR comprenant des impuretés, des précipitations d’huile, une partie de globule blanc et un GR touchant une plaquette (c) et (d), des GR chevauchés, un globule blanc, un GR touchant une plaquette et un GR comprenant des précipitations de coloration (e).

TABLE 4 – Les mesures de parasitémie moyenne sur n champs (%) de chaque méthode pour les 14 patients

Patient	MALARIS PARs(%)	Miller PARs(%)	Std PARs(%)
P1	0.873	0.937	0.899
P2	0.749	1.028	0.886
P3	0.993	1.358	2.102
P4	1.690	1.613	1.357
P5	0.850	1.706	1.165
P6	1.904	2.353	2.650
P7	2.151	2.439	2.954
P8	3.504	3.627	3.441
P9	4.987	3.992	3.814
P10	5.3282	4.924	4.868
P11	3.1193	5.158	4.003
P12	3.135	8.752	9.190
P13	12.280	10.196	9.845
P14	13.4412	12.124	14

TABLE 5 – Les parasitémies (%) mesurées par le système MALARIS par champ pour un total de 10 champs

Patient	MALARIS1	MALARIS2	MALARIS3	MALARIS4	MALARIS5	MALARIS6	MALARIS7	MALARIS8	MALARIS9	MALARIS10
P1	0.000	1.087	2.041	0.000	0.000	0.000	0.000	0.935	1.099	0.000
P2	1.869	0.000	1.000	0.926	0.980	0.000	2.679	1.075	1.626	0.000
P3	0.000	0.870	2.521	0.870	0.000	0.870	0.800	0.000	0.870	0.000
P4	1.075	1.099	5.051	3.191	1.961	0.000	1.042	0.000	0.000	3.030
P5	0.917	0.000	1.818	0.901	0.000	0.000	0.000	0.909	0.781	0.746
P6	1.715	1.637	1.287	1.665	2.183	1.218	2.376	1.758	2.947	2.254
P7	2.655	1.961	4.762	2.000	4.211	0.000	1.869	0.870	1.818	2.679
P8	2.128	2.857	4.040	0.000	3.738	5.155	4.082	5.263	1.869	0.980
P9	4.587	7.071	5.455	3.093	2.655	8.772	1.961	4.762	2.000	6.422
P10	4.968	4.617	4.501	6.548	4.348	5.751	5.020	6.135	4.596	6.798
P11	2.617	2.789	1.730	3.177	3.364	3.623	3.453	3.596	3.980	2.864
P12	3.135	8.752	9.190	4.211	3.158	1.333	0.000	2.273	2.667	0.990
P13	10.811	11.250	9.783	9.091	9.756	10.000	20.548	13.187	12.903	9.195
P14	10.263	11.349	15.907	11.429	12.365	15.019	14.518	14.401	13.632	15.529

TABLE 6 – Les parasitémies (%) mesurées par 10 microscopistes sur 14 patients, la mesure est assistée par la cellule de Miller en microscopie optique

Patient	MILLER1	MILLER2	MILLER3	MILLER4	MILLER5	MILLER6	MILLER7	MILLER8	MILLER9	MILLER10
P1	0.76	0.87	1.04	0.83	0.73	1.3	0.89	1.43	0.73	0.79
P2	0.81	0.93	1.39	1.15	1.1	0.51	0.94	1.12	1.12	1.21
P3	1.51	1.12	1.1	1.04	1.45	1.71	0.89	1.32	2.02	1.42
P4	1.12	1.05	1.74	2.1	1.54	1.87	1.66	0.62	1.58	2.85
P5	0.95	0.9	1.07	0.88	0.97	0.65	8.17	1.18	1.68	0.61
P6	2.5	2.3	2.47	2.12	2.3	3.4	1.6	2.3	2.44	2.1
P7	2.52	2.49	2.63	2.53	1.99	2.58	2.46	2.46	2.34	2.39
P8	5.16	4.11	4.01	3.61	3.44	3.38	3.24	3.17	3.12	3.03
P9	4.86	4.27	4.19	3.99	3.94	3.85	3.84	3.73	3.66	3.59
P10	4.3	5.7	4.8	3.89	4.9	5.8	5.3	4.7	4.87	4.98
P11	5.5	3.4	6.0	5.0	5.8	4.8	4.4	4.9	7.12	4.66
P12	8.36	9.2	8.7	8.69	9.93	7.95	7.39	9.77	9.7	7.83
P13	10.14	9.35	11.42	11.14	10.7	11.47	10.27	8.8	9.66	9.01
P14	15.6	11.8	14.0	11.5	10.3	14.7	9.84	12.4	9.25	11.85

TABLE 7 – Les parasitémies (%) mesurées par 10 microscopistes sur 14 patients, la mesure est réalisée en suivant la procédure standard [WHO, 2016b]

Patient	STAND1	STAND2	STAND3	STAND4	STAND5	STAND6	STAND7	STAND8	STAND9	STAND10
P1	0.61	1.12	0.91	0.76	0.62	1.16	1.05	0.87	0.87	1.02
P2	1.03	0.75	0.52	1.08	0.66	0.56	1.02	1.08	0.96	1.20
P3	0.82	1.59	9.65	1.10	1.06	1.60	1.18	1.47	1.25	1.30
P4	0.98	1.43	1.81	1.67	1.06	1.24	1.12	1.56	1.64	1.06
P5	0.65	0.54	3.75	0.66	1.16	1.23	1.12	0.88	1.04	0.62
P6	1.8	2.7	1.0	2.7	2.5	5.3	2.8	2.75	2.32	2.68
P7	4.30	4.05	2.09	3.23	2.5	2.88	1.97	3.04	3.07	2.41
P8	1.3	4.17	4.58	4.9	2.5	3.75	3.08	4.10	2.41	3.62
P9	4.03	4.37	3.99	5.7	3.04	3.22	3.12	2.89	3.22	4.56
P10	7.4	4.8	3.0	4.08	5.9	5.5	5.5	4.1	4.0	4.4
P11	5.3	3.7	2.8	5.12	4.9	3.8	4.85	2.05	3.19	4.32
P12	6.11	11.46	8.60	10.48	14.03	11.84	9.21	6.63	7.95	5.59
P13	12.23	12.25	7.34	10.68	9.24	6.27	9.5	10.08	9.26	11.60
P14	15.9	23.3	9	11.9	11.8	19.1	12.1	10.6	16.0	11.9

TABLE 8: Les parasitémies (%) mesurées par les trois techniques sur 102 patients

Patient	MALARIS PARs(%)	Miller PARs(%)	Std PARs(%)
P1	0.0386	0.33	0.2
P2	0.1206	1.8	0.7
P3	0.173	1.5	1.4
P4	0.3242	0.5	0.4
P5	0.3878	0.3	0.28
P6	0.3888	2.7	3.3
P7	0.4179	0.2	0.5
P8	0.4412	0.4	0.15
P9	0.4965	1	1
P10	0.6277	0.9	1
P11	0.6905	0.53	0.3
P12	0.6938	2.7	3.16
P13	0.751	0.6	0.4
P14	0.8137	0.65	0.8
P15	0.9376	0.86	1.23
P16	1.004	1	1.3
P17	1.0044	1.1	1.27
P18	1.049	1.3	1.5
P19	1.0627	0.2	0.25
P20	1.2315	2.16	2.5
P21	1.6069	2.5	2
P22	1.623	1.4	1.5
P23	1.6441	2.6	1.75
P24	1.7655	1.17	1.11
P25	1.804	2.1	2.7
P26	2.089	1.9	2.4
P27	2.104	2.3	3.27
P28	2.2072	3.1	2.7
P29	2.2467	2.2	2.6
P30	2.2988	2.8	3.6
P31	2.3509	3.5	2.65
P32	2.395	3.5	4
P33	2.512	2	2.7
P34	2.563	1.7	2
P35	2.5949	1.7	2
P36	2.619	3.5	4
P37	2.62	2	2.8
P38	2.788	2.5	2.5
P39	2.814	3.4	3.9
P40	2.8334	3	2.87
P41	2.859	5	4.75
P42	2.875	3.3	3

Suite à la page suivante

Table 8 – Suite de la page précédente

Patient	MALARIS PARs(%)	Miller PARs(%)	Std PARs(%)
P43	2.937	5.7	3.38
P44	2.9505	2.7	3.02
P45	3.202	2	2.7
P46	3.2241	2.2	2.6
P47	3.385	6.4	7
P48	3.401	5.4	9.45
P49	3.58	3.5	4
P50	3.599	4.6	5.3
P51	3.61	4.96	4.5
P52	3.6598	1.95	4.5
P53	3.701	4.2	4.7
P54	3.705	3.03	4.25
P55	3.7875	3.1	2.7
P56	3.851	1.8	2.5
P57	3.916	4.9	5.2
P58	3.945	3.3	4.23
P59	4.013	6.8	6.8
P60	4.038	3.3	5.4
P61	4.1038	3.3	5.31
P62	4.281	3.8	4.3
P63	4.3075	4.6	5.7
P64	4.3883	3	3.4
P65	4.428	3.49	5.7
P66	4.5338	4.8	4.7
P67	5.2119	4.2	4.3
P68	5.239	5.5	6.2
P69	5.27	6.1	7.3
P70	5.2781	10.31	8.3
P71	5.422	4.6	4.8
P72	5.55	5.7	5.06
P73	5.6361	4.3	5.27
P74	5.6441	9.2	6.2
P75	6.2548	5.6	6
P76	6.507	3.4	3.3
P77	6.679	1.3	1.2
P78	6.982	9.8	10
P79	7.11	7.4	8.1
P80	7.191	8.8	8
P81	7.358	5	6.8
P82	7.4438	7.7	10.7
P83	7.6022	7.2	7.5
P84	8.0505	7.7	10
P85	8.5535	6.43	6.16
P86	8.6962	8	8.95

Suite à la page suivante

Table 8 – Suite de la page précédente

Patient	MALARIS PARs(%)	Miller PARs(%)	Std PARs(%)
P87	9.065	7.9	12.2
P88	9.2491	10	10.7
P89	9.336	7.3	6.54
P90	9.61	7.4	10.7
P91	10.157	14.1	10.3
P92	10.458	7.6	7.96
P93	11.0573	11	10.61
P94	11.57	8	10.6
P95	12.485	11	12
P96	13.976	11.7	12.2
P97	15.3881	12.8	14.2
P98	15.5166	17	16.75
P99	16.449	15.3	11.6
P100	19.62	20.2	20.2
P101	22.3341	15.7	18.8
P102	27.328	23	26.2

TABLE 9: Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients

Patient	MALARIS1	MALARIS2	MALARIS3	MALARIS4	MALARIS5	MALARIS6	MALARIS7	MALARIS8	MALARIS9	MALARIS10
Patient 1	0.085	0.000	0.000	0.097	0.204	0.000	0.000	0.000	0.000	0.000
Patient 2	0.000	0.238	0.432	0.222	0.000	0.000	0.107	0.103	0.104	0.000
Patient 3	0.078	0.000	0.230	0.345	0.302	0.077	0.146	0.308	0.244	0.000
Patient 4	0.893	0.396	0.284	0.287	0.468	0.182	0.189	0.177	0.178	0.188
Patient 5	0.686	0.104	0.373	0.812	0.166	0.431	0.675	0.094	0.264	0.274
Patient 6	0.749	0.678	0.279	0.368	0.275	0.139	0.282	0.421	0.500	0.198
Patient 7	0.477	0.679	0.516	0.297	0.506	0.360	0.629	0.148	0.351	0.216
Patient 8	0.000	0.224	0.352	0.293	0.624	0.477	0.205	0.443	0.811	0.983
Patient 9	0.327	0.476	0.082	0.560	0.343	0.363	1.470	0.556	0.207	0.582
Patient 10	0.470	0.743	0.469	0.316	0.534	0.549	1.024	0.684	0.935	0.553
Patient 11	1.228	0.273	1.215	0.432	0.812	0.679	0.945	0.403	0.404	0.514
Patient 12	0.417	0.900	0.504	1.154	0.521	0.552	1.134	0.516	0.416	0.824
Patient 13	0.918	0.478	1.130	0.988	0.460	0.575	0.642	0.615	0.831	0.873
Patient 14	0.546	1.493	0.705	0.684	0.548	0.795	1.199	1.294	0.540	0.333
Patient 15	1.877	1.104	1.096	0.714	0.636	0.779	0.941	0.829	0.743	0.657
Patient 16	0.833	1.910	0.836	0.676	1.089	1.001	0.719	0.958	1.330	0.688
Patient 17	0.336	0.713	1.735	0.761	1.254	2.055	0.602	0.995	1.151	0.442
Patient 18	0.635	1.717	1.336	0.722	1.155	0.877	0.814	1.195	0.992	0.788
Patient 19	0.980	2.194	2.849	0.626	1.250	0.941	0.498	0.758	0.429	0.102
Patient 20	1.662	0.962	0.933	1.084	0.766	0.958	1.592	1.589	1.574	1.195
Patient 21	1.023	1.225	1.054	2.281	1.464	1.658	1.866	2.046	2.433	1.019
Patient 22	2.170	2.244	1.918	1.247	1.292	2.224	1.281	1.552	1.260	1.045
Patient 23	1.402	1.909	2.404	1.426	1.230	1.027	0.991	2.139	2.169	1.744
Patient 24	1.678	1.439	1.547	2.491	3.177	1.022	1.355	2.431	1.410	1.105
Patient 25	2.128	1.414	2.420	1.977	1.919	1.591	1.687	1.468	2.386	1.054
Patient 26	1.558	2.832	1.746	2.330	1.049	2.652	2.464	1.952	2.330	1.981

Suite à la page suivante

TABLE 9: Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)

Patient	MALARIS1	MALARIS2	MALARIS3	MALARIS4	MALARIS5	MALARIS6	MALARIS7	MALARIS8	MALARIS9	MALARIS10
Patient 27	2.217	2.389	1.016	1.832	1.503	2.365	2.764	2.368	2.165	2.430
Patient 28	2.092	2.643	2.258	2.522	1.194	1.935	2.321	2.037	3.021	2.049
Patient 29	2.320	2.862	2.111	2.583	2.147	1.904	1.602	1.820	2.234	2.885
Patient 30	1.961	2.398	4.515	2.106	1.931	2.056	1.951	1.739	2.101	2.230
Patient 31	2.901	2.404	1.933	2.799	2.031	2.179	2.196	2.027	2.518	2.521
Patient 32	2.314	2.708	1.455	2.925	2.112	1.684	2.158	2.503	3.544	2.555
Patient 33	2.553	2.095	2.603	2.768	3.410	2.972	2.309	2.413	2.004	1.993
Patient 34	3.008	3.327	3.736	1.897	2.361	1.378	2.217	2.537	2.538	2.638
Patient 35	2.431	2.402	2.215	3.535	2.603	3.386	1.782	3.312	2.636	1.647
Patient 36	2.449	3.274	2.304	3.275	2.267	2.231	2.468	2.571	3.128	2.232
Patient 37	1.902	1.771	3.044	2.177	2.522	2.218	3.613	4.397	2.536	2.027
Patient 38	1.859	3.512	3.077	2.994	2.929	2.230	3.446	2.508	2.539	
Patient 39	2.221	2.391	2.473	3.443	2.976	3.000	2.363	3.170	3.689	2.414
Patient 40	2.842	2.657	2.182	2.610	3.131	2.582	2.819	3.907	2.775	2.829
Patient 41	3.062	3.046	3.748	2.933	3.223	2.470	2.385	2.775	2.008	2.947
Patient 42	2.612	3.272	2.500	3.128	2.741	3.195	2.458	2.788	3.612	2.445
Patient 43	3.569	3.195	2.977	2.291	2.522	1.323	3.799	3.618	2.477	3.606
Patient 44	2.953	3.336	3.050	3.371	2.059	2.232	3.438	3.114	2.524	3.428
Patient 45	4.528	2.541	3.211	2.911	2.929	2.705	3.359	3.539	3.623	2.674
Patient 46	2.244	3.749	3.654	2.493	2.541	2.962	3.337	3.321	3.559	4.381
Patient 47	2.738	2.970	4.452	2.421	4.572	3.188	3.279	3.020	3.944	3.267
Patient 48	2.526	3.704	3.249	3.790	3.866	3.861	3.068	3.317	3.585	3.044
Patient 49	3.581	3.839	3.220	3.333	3.485	3.333	3.500	3.480	4.312	3.718
Patient 50	3.177	4.340	3.148	3.486	3.680	4.225	3.141	4.711	2.867	3.221
Patient 51	3.828	3.583	3.803	2.580	4.802	4.299	3.923	3.033	2.703	3.550
Patient 52	3.214	4.047	4.234	3.162	4.048	3.806	2.526	4.378	4.163	3.02
Patient 53	2.941	3.453	4.939	4.886	3.735	3.098	3.870	3.178	3.805	3.108

Suite à la page suivante

TABLE 9: Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)

Patient	MALARIS1	MALARIS2	MALARIS3	MALARIS4	MALARIS5	MALARIS6	MALARIS7	MALARIS8	MALARIS9	MALARIS10
Patient 54	3.745	3.223	3.399	3.327	3.872	4.145	4.533	2.899	4.570	3.342
Patient 55	4.505	3.302	3.938	3.795	5.164	2.400	3.287	2.772	4.841	3.871
Patient 56	4.181	2.939	2.994	3.293	2.944	5.778	3.763	3.710	4.227	4.688
Patient 57	4.360	3.669	3.968	4.174	3.628	3.943	3.945	3.986	3.953	3.543
Patient 58	4.325	3.472	3.361	3.122	3.455	4.655	4.767	4.863	3.612	3.825
Patient 59	4.279	3.970	3.454	4.299	4.106	4.207	4.748	4.000	3.566	3.502
Patient 60	4.304	4.160	4.661	4.469	4.297	4.646	3.670	3.388	3.069	3.717
Patient 61	4.335	3.545	4.605	5.316	3.754	4.266	3.742	3.137	4.528	3.811
Patient 62	4.065	5.539	3.899	4.829	4.679	4.563	4.140	3.033	3.448	4.621
Patient 63	4.439	4.468	3.392	4.494	4.499	4.676	4.621	3.846	4.830	3.810
Patient 64	4.649	3.323	4.557	4.105	6.425	3.627	5.134	4.883	2.789	4.391
Patient 65	4.609	3.886	5.359	3.834	3.713	5.152	4.402	4.510	4.321	4.502
Patient 66	5.280	4.542	5.186	6.027	3.892	3.387	4.127	3.653	5.349	3.895
Patient 67	5.219	4.688	4.026	9.869	5.769	4.552	4.563	4.936	4.255	4.242
Patient 68	5.638	2.601	5.118	6.328	4.344	3.175	6.770	5.805	6.433	6.181
Patient 69	6.648	4.385	5.106	5.462	5.356	4.844	3.906	6.533	5.435	5.028
Patient 70	3.343	4.592	4.253	7.205	6.746	7.601	5.509	4.988	3.851	4.693
Patient 71	4.762	6.544	7.002	6.316	6.627	4.851	4.819	4.811	4.953	3.537
Patient 72	5.000	4.754	4.686	5.753	6.019	5.969	4.670	5.363	5.089	8.197
Patient 73	5.042	5.333	5.812	6.089	6.917	4.664	5.590	5.294	4.308	7.312
Patient 74	6.316	5.935	3.579	6.407	5.742	7.925	4.739	4.060	6.748	4.990
Patient 75	6.280	6.748	7.180	6.364	6.133	5.897	5.947	5.829	5.847	6.323
Patient 76	3.849	6.186	4.290	9.358	7.599	2.958	7.216	8.021	3.676	11.926
Patient 77	3.738	4.797	2.647	3.737	11.531	11.721	3.783	10.150	9.607	5.079
Patient 78	6.011	7.419	7.686	7.692	7.797	6.975	6.780	6.984	6.572	5.906
Patient 79	7.886	6.800	5.966	7.794	6.954	7.396	6.941	6.313	7.543	7.508
Patient 80	7.850	6.820	6.039	6.579	7.559	8.598	8.460	6.936	6.492	6.578

Suite à la page suivante

TABLE 9: Les parasitémies (%) mesurées sur 10 champs par le système MALARIS sur 102 patients (suite)

Patient	MALARIS1	MALARIS2	MALARIS3	MALARIS4	MALARIS5	MALARIS6	MALARIS7	MALARIS8	MALARIS9	MALARIS10
Patient 81	7.198	7.656	6.630	8.095	7.143	6.812	5.506	9.123	7.252	8.173
Patient 82	7.724	8.376	6.710	7.500	8.229	5.605	6.497	7.211	8.722	7.864
Patient 83	8.422	7.252	7.361	7.096	7.254	6.941	8.068	7.001	8.278	8.349
Patient 84	7.736	7.117	8.549	7.203	8.296	8.469	7.907	7.810	9.975	7.443
Patient 85	8.995	7.759	9.491	8.716	8.885	7.899	7.362	8.403	7.838	10.187
Patient 86	8.613	8.049	9.129	9.413	7.657	7.853	9.935	9.795	9.722	6.796
Patient 87	9.733	9.477	8.957	9.722	7.456	9.834	9.852	9.055	8.000	8.571
Patient 88	9.049	9.809	9.182	9.827	9.855	9.456	10.659	9.309	6.929	8.416
Patient 89	9.773	8.140	10.522	10.367	8.342	6.396	10.093	9.863	10.532	9.328
Patient 90	9.760	9.540	9.313	8.879	8.582	9.715	9.747	11.469	8.240	10.858
Patient 91	10.157	10.336	10.339	10.195	9.636	8.649	9.307	11.293	10.055	11.610
Patient 92	9.146	10.500	10.717	9.656	9.130	10.564	9.923	13.093	9.920	11.931
Patient 93	10.367	11.531	11.277	10.805	10.585	13.923	10.227	10.303	11.213	10.342
Patient 94	13.671	13.476	11.663	7.322	9.810	10.406	12.223	15.529	10.208	11.400
Patient 95	14.565	12.121	12.053	13.111	11.880	11.493	11.916	13.617	12.405	11.694
Patient 96	15.544	15.398	14.017	12.966	13.072	14.770	12.782	12.060	16.465	12.690
Patient 97	15.184	14.637	17.030	16.547	14.678	14.742	13.294	16.129	15.678	15.962
Patient 98	17.854	17.025	20.070	15.254	12.085	14.625	15.423	15.779	16.826	10.225
Patient 99	15.480	15.813	13.430	16.667	14.749	17.482	15.527	17.979	20.073	17.291
Patient 100	21.167	17.893	18.094	20.660	20.309	20.239	20.841	20.570	18.682	17.746
Patient 101	24.062	19.668	21.739	21.459	22.713	22.309	22.940	22.005	22.533	23.913
Patient 102	24.409	26.829	27.880	30.071	27.516	28.533	25.260	27.202	27.254	28.332

TABLE 10 – Évaluation de la sensibilité et de la spécificité à différents seuils de détection de la parasitémie, 4 % et 10 %.

Techniques	Sensibilité (%)		Spécificité (%)	
	Seuil de détection à 4 %	Seuil de détection à 10 %	Seuil de détection à 4 %	Seuil de détection à 10 %
MALARIS avec la cellule de Miller	84.1	83.3	86.2	97.8
MALARIS avec la cytométrie en flux	93.2	91.7	77.6	93.3
La cellule de Miller avec la cytométrie en flux	97.8	91.7	82.5	93.3

B Annexe B

Approche de détection	Watershed	YOLOv8	Combinaison
Nombre d'images analysées	31	31	31
Nombre de globules rouges segmentés	30 356	24 836	32 835
Temps d'exécution par image de champ (en secondes)	16-18	10-12	28-30

TABLE 11 – Comparaison entre les deux approches de segmentation watershed et YOLOv8.

Patient	IOU SCORE
IMG1	0.971
IMG2	0.98
IMG3	0.973
IMG4	0.968
IMG5	0.987
IMG6	0.90
IMG7	0.952
IMG8	0.97
IMG9	0.981
IMG10	0.986
IMG11	0.95
IMG12	0.97
IMG13	0.913
IMG14	0.972
IMG15	0.95
IMG16	0.961
IMG17	0.943
IMG18	0.917
IMG19	0.95
IMG20	0.98
IMG21	0.97
IMG22	0.971
IMG23	0.973
IMG24	0.968
IMG25	0.987
IMG26	0.90
IMG27	0.952
IMG28	0.97
IMG29	0.981
IMG30	0.986
IMG31	0.95
Moyenne	0.95

TABLE 12 – Scores IoU pour chaque patient.