Development of deep learning-based precision tools for Plasmodium and Cryptosporidium parasites analysis from microscopic images

Ziheng Yang

To cite this version:

Ziheng Yang. Development of deep learning-based precision tools for Plasmodium and Cryptosporidium parasites analysis from microscopic images. Micro and nanotechnologies/Microelectronics. Université de Lille, 2022. English. NNT : 2022ULILN049 . tel-04221290

HAL Id: tel-04221290
https://theses.hal.science/tel-04221290

Submitted on 28 Sep 2023

Ecole doctorale Science de l'ingénierie et des systèmes (ENGSYS-632)
IEMN - Institut d'électronique de microélectronique et de nanotechnologie

Thèse présentée par

ZIHENG YANG

En vue de l'obtention du grade de

Docteur de l'Université de Lille

Spécialité: Electronique, microélectronique, nanoélectronique et micro-ondes

Développement d'outils de précision basés sur l'apprentissage profond pour l'analyse des parasites Plasmodium et Cryptosporidium à partir d'images microscopiques

Présentée et soutenue publiquement le 5 Décembre 2022

Membres du jury:

Directeur de thèse	M. Dominique COLLARD	DR CNRS, Université de Lille
Rapporteurs	M. Abdeldjalil Ouahabi	Professeur, Ecole Polytechnique de l'Université de Tours
	M. Hedi Tabia	Professeur, Université Paris-Saclay
Examinateurs	Mme Samia Ainouz	Professeur, INSA de Rouen (président du jury)
	M. Maxime Devanne	MCF, Université de Haute-Alsace
	M. Abdelmalik Taleb-Ahmed	Professeur, Université Polytechnique Hauts-de-France
Encadrants	M. Halim BENHABILES	Ph.D, Enseignant-chercheur, Junia ISEN Lille
	Mme Feryal WINDAL	Ph.D, Enseignante-chercheuse, Junia ISEN Lille

Doctoral school engineering and system sciences (ENGSYS-632)
IEMN - Institute of Electronics, Microelectronics and Nanotechnology

Thesis submitted by

ZIHENG YANG

In order to obtain the grade of

Doctor from the University of Lille

Specialty: Electronics, microelectronics, nanoelectronics and microwaves

Development of deep learning-based precision tools for Plasmodium and Cryptosporidium parasites analysis from microscopic images

Presented and publicly supported on December 5, 2022

Jury members:

Thesis director	Mr. Dominique COLLARD	DR CNRS, University of Lille
Reporters	Mr. Abdeldjalil Ouahabi	Professeur, Ecole Polytechnique de l'Université de Tours
	Mr. Hedi Tabia	Professeur, Université Paris-Saclay
Examiners	Mme Samia Ainouz	Professeur, INSA de Rouen (jury president)
	Mr. Maxime Devanne	MCF, Université de Haute-Alsace
	Mr. Abdelmalik Taleb-Ahmed	Professeur, Université Polytechnique Hauts-de-France
Thesis supervisors	Mr. Halim BENHABILES	Ph.D, Enseignant-chercheur, Junia ISEN Lille
	Mme Feryal WINDAL	Ph.D, Enseignante-chercheuse, Junia ISEN Lille

Résumé

Dans cette thèse, nous avons proposé deux contributions principales liées à l'analyse par apprentissage profond des parasites *Plasmodium* et *Cryptosporidium* à partir d'images microscopiques.

Plus spécifiquement, dans la première contribution, nous avons proposé un framework pour diagnostiquer une infection par le parasite *Plasmodium* chez les humains en utilisant des images microscopiques de frottis sanguins fins. Comparé aux méthodes de l'état de l'art, notre framework est basé sur des approches de segmentation et de classification permettant l'analyse directe du parasite au lieu de la cellule qui le contient. Dans ce sens, le framework permet de segmenter directement le parasite *Plasmodium* et de distinguer son espèce parmi quatre classes principales : P. Falciparum, P. Ovale, P. Malaria et P. Vivax. Nous démontrons l'efficacité de notre framework et notamment son potentiel de généralisation sur des données interclasses en exploitant plusieurs jeux de données publiques. De plus, nous montrons que la technique d'augmentation des données que nous proposons, appelée Local Parasite Texture Scanning (LPTS), améliore davantage la précision de notre modèle de classification.

Dans la deuxième contribution, nous avons proposé un framework pour diagnostiquer une infection par le parasite *Cryptosporidium* chez les vaches laitières en utilisant des images microscopiques de fluorescence. À cette fin, nous avons proposé une méthodologie originale de segmentation des parasites basée sur une approche grossière à fine, qui atteint une grande précision sur notre jeu de données généré de *Cryptosporidium* et qui permet de dépasser en termes de performance les méthodes de segmentation de l'état de l'art. Nous avons également proposé un modèle de classification à haut pouvoir discriminant permettant d'identifier efficacement le stade de vie des parasites parmi 4 stades asexués : oocyste, trophozoïte, méronte, et forme libre. Nous montrons à travers une étude expérimentale que notre modèle atteint une grande précision en analysant uniquement le parasite lui-même et sans avoir besoin d'informations supplémentaires liées à la taille et au nombre de noyaux qui sont nécessaires au biologiste afin de réaliser sa classification.

Mots clés : Parasite *Plasmodium*, Parasite *Cryptosporidium*, Apprentissage profond, Analyse d'images microscopiques 2D, BIO-MEMS.

Abstract

In this thesis, we have proposed two main contributions related to the *Plasmodium* parasite and *Cryptosporidium* parasite analysis from microscopic images using deep learning techniques.

More specifically, in the first contribution, we have proposed a framework for diagnosing *Plasmodium* infection in humans using microscopic images of thin blood smears. Compared to the state-of-the-art studies, it is rather based on a straightforward segmentation and classification approaches, permitting the analysis of the parasite itself instead of the cell containing it. In this sense, the framework permits to directly segment the *Plasmodium* parasite and to distinguish its species among four major classes: P. Falciparum, P. Ovale, P. Malaria and P. Vivax. We demonstrate the efficiency of our framework and notably its potential of generalization over interclass data by exploiting several public datasets. Moreover, we show that our proposed data augmentation technique named Local Parasite Texture Scanning (LPTS) further improves the accuracy of our classification model.

In the second contribution, we have proposed a framework for diagnosing *Cryptosporidium* infection in dairy cows using fluorescence microscopic images. To this end, we have proposed an original parasite segmentation methodology based on a coarse-to-fine approach which achieves high accuracy on our generated dataset of *Cryptosporidium* and permits to outperform segmentation methods from the state-of-the-art. We have also proposed a classifier with a high discriminatory power that is used to efficiently distinguish the life stages of the parasites among 4 asexual stages: oocyst, trophozoite, meront, and free form. We show through an experimental study that our classifier achieves high accuracy by analyzing only the parasite itself and without the need of additional information related to the size and the number of nuclei which are required by the biologist to establish the classification.

Key words: *Plasmodium* parasite, *Cryptosporidium* parasite, Deep learning, 2D microscopic image analysis, BIO-MEMS.

Acknowledgements

The research presented in my thesis has been completed under the support of numerous people: supervisors, partners, colleagues, family and friends, to whom I would like to express my gratitude. Their support and encouragement were crucial to my success in completing the thesis.

First and foremost, I would like to express my special thanks of gratitude to my supervisors (Dr. Halim BENHABILES, Dr. Feryal WINDAL) as well as my thesis director (Dr. Dominique COLLARD) who gave me the opportunity to work with them in the departmemt DSLS of Junia. My supervisors helped me a lot during my thesis and I became more autonomous in my research work under their guidance. They were interested in the progress of my work and they always listened patiently to my questions and actively offered their help.

Besides, I would like to thank european Interreg program for funding my thesis research at the University of Lille. Especially the partners of the H4DC project have helped me a lot with my research. I would like to thank Dr. Jérôme, who gave me a lot of knowledge on *Cryptosporidium* parasites as the H4DC project leader. As well as Ms. Anna-Charlotte, who provided me with the raw data needed for the experiments and spent a lot of time labeling them for me.

I would like to show my gratitude to my great colleagues, Mr. Ruiwen HE, Mr. Tarek MAYLAA and Mr. Bilel GUETARNI, who have helped me a lot during my PhD research.

Last but not least, I am also grateful to my family and friends. My success would not have been possible without their constant love and encouragement.

Author's publications

International journals

- **Ziheng YANG**, Halim BENHABILES, Karim HAMMOUDI, Feryal WINDAL, Ruiwen HE, Dominique COLLARD, *"A generalized deep learning-based framework for assistance to the human malaria diagnosis from microscopic images"*, **Neural Computing and Applications**, Springer DOI https://doi.org/10.1007/s00521-021-06604-4, published online, 2021. IF: 5.102, Scopus highest percentile: Q1

International conferences

- **Ziheng YANG**, Halim BENHABILES, Feryal WINDAL, Jérôme FOLLET, Anne-Charlotte LENIERE, Dominique COLLARD, *"A coarse-to-fine segmentation methodology based on deep networks for automated analysis of* Cryptosporidium *parasite from fluorescence microscopic images"*, **1st International Workshop on Medical Optical Imaging and Virtual Microscopy Image Analysis (MOVI) MICCAI 2022**, DOI https://doi.org/10.1007/978-3-031-16961-8_16.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Parasitic diseases have a profound impact on global public health. The WHO (World Health Organization) reported in its global Health Estimates of 2020 [30] that parasitic and infectious diseases are a major cause of death in Africa. In this thesis, we have been interested in the development of precision analysis tools for two parasites, namely the *Plasmodium* parasite and the *Cryptosporidium* parasite. In this context, we will present briefly these two parasites:

- *Plasmodium* parasite causes the *Malaria* disease which is an acute febrile human disease. The *Malaria* is one of the epidemics that possess the potential to threaten human life [18, 102]. It is widespread in tropical and subtropical areas, particularly in Asia, Sub-Saharan Africa and Latin America [10]. The same trend has been observed until now (as shown in Fig 1.1). According to the WHO, in Southeast Asia and Africa regions, the number of mortality due to *Malaria* has declined from approximately 530,000 (2010) to 380,000 (2018). Nevertheless, since 2016, the decrease in mortality rate has decelerated [104]. Therefore, there is a need to develop an automated *Malaria* diagnostic system that is not only reliable, but can be faster and more economical, especially in resource-limited regions [111].

- *Cryptosporidium* parasite causes Cryptosporidiosis which is a diarrheal disease. The parasite may live in the intestine of several mammals. It is important to note that *Cryptosporidium* is the 5th causative agent of diarrhoeal disease in young children under 5 years of age

[49]. On farms, parasites are excreted in the external surroundings through the faeces of infected cows as a form called oocyst, which directly contaminates water or food and causes further infections (illustrated in Fig 1.2) [20, 34].

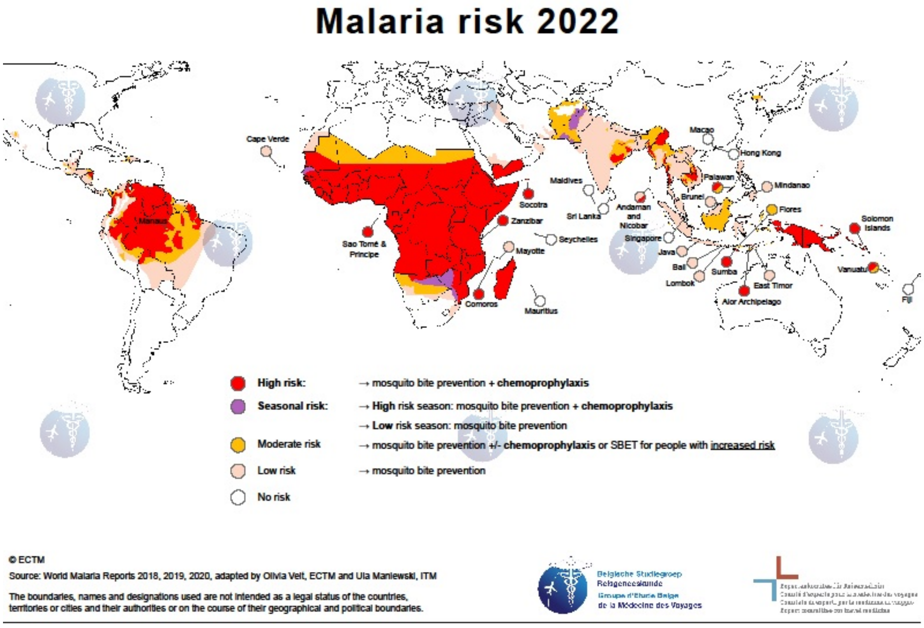


FIGURE 1.1: Malaria world map 2022 (extracted from WANDA).

Both parasites are generally diagnosed by experts through a visual analysis of microscopic images. These images represent for the *Malaria* case, blood samples and for the *Cryptosporidium* case, infection of HCT-8 cell lines with Cryptosporidium parvum. Nevertheless, the manual analysis method does not allow for a transition into large-scale tests. To face this issue, adopting deep learning methods is a natural choice in reason of their high performance demonstrated in medical image analysis [47, 56, 36]. However, direct exploitation of the existing methods is not sufficient to develop a reliable tool due to the particular visual aspect of these parasites.

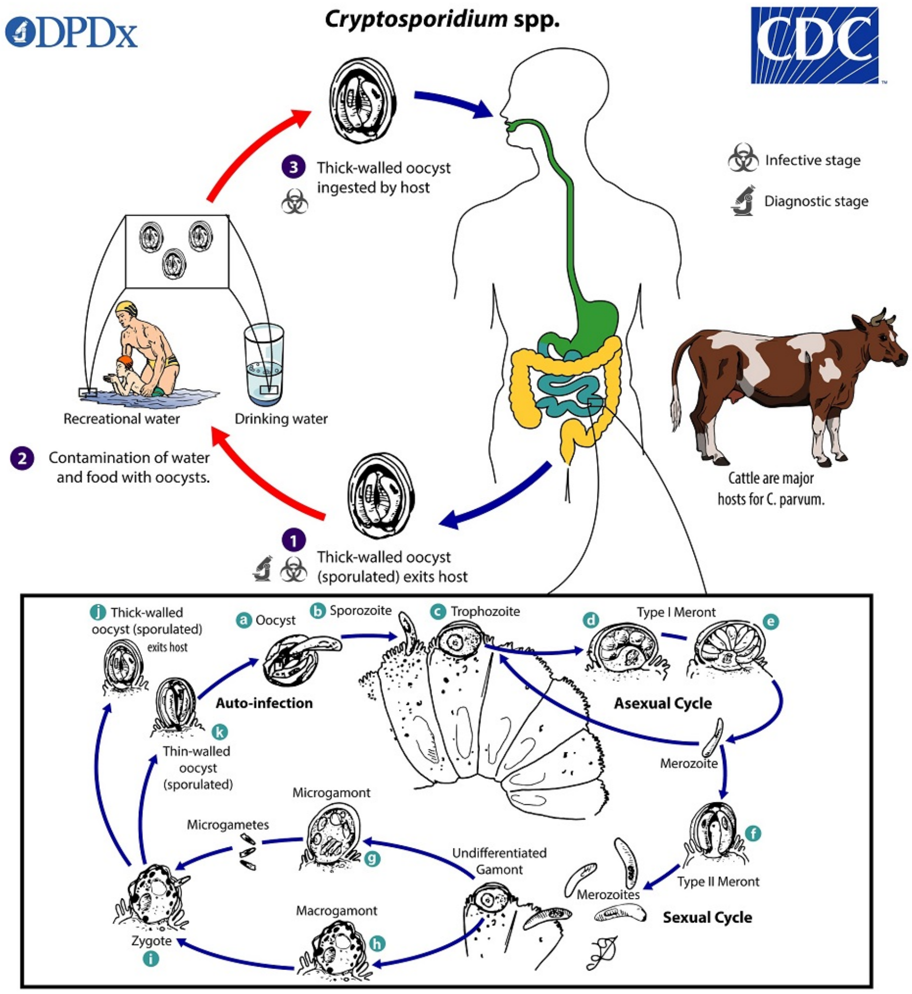


FIGURE 1.2: *Cryptosporidium parvum* life cycle (extracted from CDC).

This thesis is under the frame of the H4DC (Health for Dairy Cows) project [1] that has been funded by the european Interreg program and which includes mainly 4 academic partners (Junia[2], University of Kent[3], CNRS[4],

---

[1] `https://h4dc-interreg2seas.eu/`

[2] `https://www.junia.com/fr/`

[3] `https://www.kent.ac.uk/`

[4] `https://www.cnrs.fr/fr/page-daccueil`

University of East Anglia[5]) and 3 industrials (Inagro[6], ZLTO[7], CVE[8]). The project consists in working on the cryptosporidiosis disease within the cows (caused by *Cryptosporidium* parasite infection) by reducing its impact in farms. Indeed, the project aims to increase farm productivity in a number of different countries, making these businesses more efficient and successful, with a lower impact on human health. For this purpose, the project aims to provide pilot farms, detection tools and technological devices that can accelerate the discovery of new drugs at the lowest possible cost.

Nevertheless, as it was raised in the beginning of this introduction, we have been also interested in the development of analysis tools of the *Malaria* parasite. Indeed, we have started our thesis in the midst of the COVID-19 crisis which lasted for almost 2 years and which has involved several lockdowns, leading in our case to a major issue in the process of data acquisition from the partners of the project. Besides, no public *Cryptosporidium* data were available at that time. For these reasons, we choose to first address the *Malaria* parasite which belongs to the same group as the *Cryptosporidium* parasite namely Apicomplexa group and for which several public datasets were already available.

1.1 Contributions

In this thesis, we offer two different contributions: (1) the first one is related to the study of the human *Malaria* diagnosis from microscopic images, and (2) the second one is related to the study of automated *Cryptosporidium* parasite analysis from fluorescence microscopic images.

Deep learning-based precision tool for *Plasmodium* parasite analysis – We have proposed a framework for diagnosing *Malaria* infection in humans using microscopic images of thin blood smears. Compared to the

---

[5] `https://www.uea.ac.uk/`

[6] `https://inagro.be/`

[7] `https://www.zlto.nl/`

[8] `https://www.cvegroup.com/`

state-of-the-art studies, it is rather based on a straightforward segmentation and classification approaches, permitting the analysis of the parasite itself instead of the cell containing it. In this sense, the framework permits to directly segment the *Malaria* parasite and to distinguish its species among four major classes: P. Falciparum, P. Ovale, P. Malaria and P. Vivax. We demonstrate the efficiency of our framework and notably its potential of generalization over interclass data by exploiting several public datasets. Moreover, we show that our proposed data augmentation technique named Local Parasite Texture Scanning (LPTS) further improves the accuracy of our classification model.

Deep learning-based precision tool for *Cryptosporidium* parasite analysis – We have proposed a framework for diagnosing *Cryptosporidium* infection in dairy cows using fluorescence microscopic images. To this end, we have proposed an original parasite segmentation methodology based on a coarse-to-fine approach which achieves high accuracy on our generated dataset of *Cryptosporidium* and permits to outperform segmentation methods from the state-of-the-art. We have also proposed a classifier with a high discriminatory power that is used to efficiently distinguish the life stages of the parasites among 4 asexual stages: oocyst, trophozoite, meront, and free form. We show through an experimental study that our classifier achieves high accuracy by analyzing only the parasite itself and without the need of additional information related to the size and the number of nuclei which are required by the biologist to establish the classification.

1.2 Outline

The manuscript is organized as follows:

- **Chapter 2** introduces the state of the art of parasite analysis from microscopic images in a general manner. It is worth mentioning that the *Plasmodium* parasite studies take up the major part of this chapter. Indeed, there are few vision-based works for the analysis of *Cryptosporidium* parasite, especially exploiting deep learning techniques.

- **Chapter 3** presents a framework based on deep learning for assisting in the diagnosis of human *Malaria* infection. The chapter shows through a comprehensive experimental study the high performance of our framework in term of generalization potential on interclass data. Specifically, the chapter highlights the fact that our global system including segmentation and classification modules reaches 100% of diagnosis accuracy on the test set of 17 infected patients.

- **Chapter 4** presents a framework based on deep learning to diagnose the infection of *Cryptosporidium* parasite in dairy cows. The chapter first points out the limitation of state of the art methods to achieve the goal of accurate parasite segmentation, then it explains how we addressed it. To this end, the chapter discribes our coarse-to-fine segmentation approach. It presents also a classification model to distinguish the life stages of the parasite. Finally, this chapter concludes with a set of experiments to demonstrate the efficiency of our framework.

- **Chapter 5** summarizes the major results obtained in the frame of this thesis and discusses some research lines to investigate which are mainly on the *Cryptosporidium* analysis using bright-field images (raw images not including the fluorescence).

Chapter 2

State of the art of parasite image analysis

2.1 Introduction

This chapter introduces the state of the art of parasite analysis from microscopic images in a general manner, including *Plasmodium*, *Cryptosporidium*, intestinal parasite, *Toxoplasma*, *Babesia* and *Leishmania*.

Moreover, we also present briefly the life cycle of *Cryptosporidium*. Our study concerns four morphologies of the asexual phase: oocyst, trophozoite, meront, and free form.

2.2 *Plasmodium* parasite image analysis

In the previous decade, there has been a growing interest regarding the development of microscopic image analysis for use as an aid in *Malaria* diagnosis [77, 81, 58, 45]. Fig 2.1 shows the life cycles of the *Plasmodium* parasite.

Various analysis systems, as proposed in different surveys [77, 81, 58, 45], include thick blood smears (a drop of blood over glass slide) and thin blood smears (a drop of blood scattered throughout the slide). In comparison with the two, the system of thick blood smears is mainly used to detection of the parasite presence, while thin blood smears enable the determination of the parasite species. In fact, thin blood smear systems

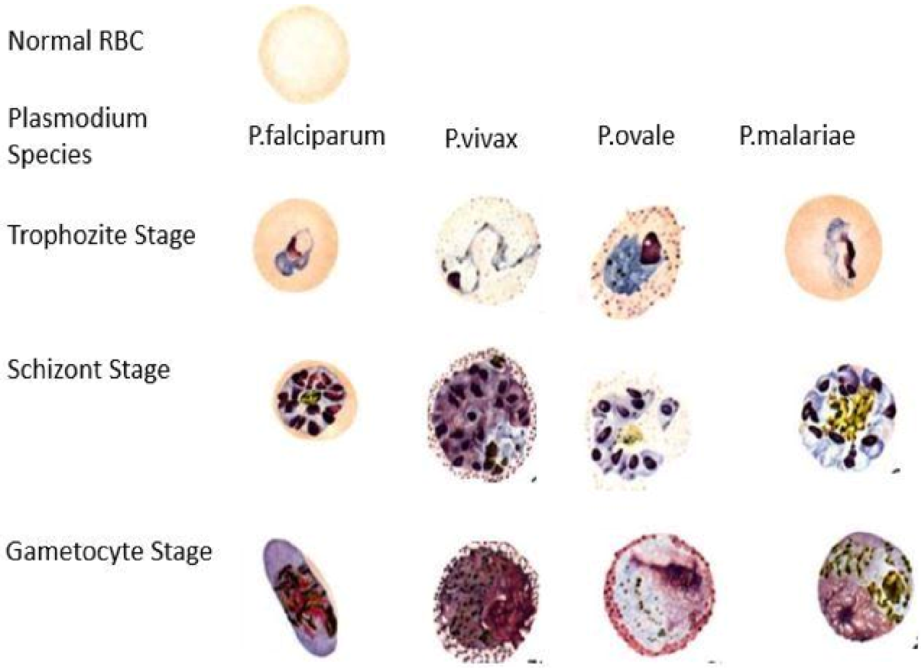


FIGURE 2.1: The life cycles of the Plasmodium parasite (extracted from [72]).

constitute the majority and consist of three main stages: pre-processing of the microscopic image, segmentation of the RBCs (red blood cells) and classification of the segmented cells in the image.

The purpose of the pre-processing is to enhance the image quality by decreasing unwanted noise. In fact, the capture system and its setup influence the rendition of the images directly. With the same blood sample, the illumination, texture and contrast and illumination of the image will vary depending on the used capture protocol, which may impact the discriminative features for segmenting blood cells and detecting the possible infection of parasites in the image. To address this issue, the proposed system leverages different techniques of image filtering for example morphological operations [83], contrast stretching [72] and median filtering [110].

The preprocessed image consists of a population of RBCs, each of which may be infected with the parasite. Hence, current analysis systems have presented segmentation of isolated RBCs as an intermediate stage. In order to achieve this goal, the presented systems have employed a large number of segmentation methods, primarily including conventional methods such as: Watershed segmentation [19], Gaussian mixture model [1] and Otsu binarization [38]. These methods initially outline the RBC in the image and crop out a set of images where each image represents an isolated cell.

The final stage is performed through the classification of each RBC (cropped image) in order to determine its infection status. At this status, the presented system includes the analysis from a simple binary classification to more complex classifications. To be precise, in the case of binary classification, the system distinguish if the blood cells are infected with parasites [82, 38, 1]. Nevertheless, with classification for multiple class, the system is able to recognize the parasite species [78] and, the life stage of it under certain circumstances [24]. Moreover, recent references [77, 81, 58, 45] show that nearly all classification methods proposed use machine learning techniques to characterize parasites using hand-designed characteristics.

Consecutive research efforts on *Malaria* diagnostic systems based on thin blood smear image analysis [70, 22, 91, 65, 3] suggest that it is still a challenging task and more research is needed to strengthen the robustness of the system.

Molina et al. [70] proposed a system following the processing pipeline presented previously. Precisely, the system considers only the green component of the microscopic image and applies a set of filters to improve its quality: i) an adaptive histogram equalization to enhance the contrast between RBCs and the background and ii) median and Gaussian filters to reduce the background noise. The resulting image is then binarised using Otsu thresholding and RBCs are segmented and cropped using watershed algorithm. A cascade of 3 classifiers are then used to classify each cropped image into 6 classes: normal RBC, RBC with parasite inclusion and 4 classes covering RBC with other erythrocyte inclusions (non parasite). To train and test the classifiers, authors built their own *Malaria* image dataset from 87 patients and split it into 4633 RBC images for training and 11027 of RBC images for testing. To feed the classifiers, each image has been characterized by a set of 2852 handcrafted features including shape, texture and color information. Authors reported a cell segmentation performance of 97.4% without mentioning the used metric and an overall parasite infection classification accuracy of 97.7% using SVM classifiers.

The system proposed by Delgado-Ortet et al. [22] permits to classify each segmented RBC into two classes namely *Malaria* infected or not. To this end, authors designed their own CNN model to learn an RGB pixel based cell segmentation. The segmentation model is composed of 3 sequential blocks of convolution layers: i) a down-scaling block for extracting the RBCs pixel features from the RGB input image (population of RBCs), ii) an up-scaling block to align the size and the shape of the extracted information with those of the input image, iii) a classification block to predict the belonging to a RBC or not for each pixel of the input image. The cell segmentation model has been trained on 169 images from 4 patients and validated on 17 images of 1 patient, the overall representing a total of 5911 RBCs. Authors reported a Jaccard coefficient ranged between 0.80 (lowest) and 0.94 (highest) for the 17 tested images. For the classification of each cropped RBC with respect to the parasite infection, authors proposed also their own CNN model which has been trained from scratch and tested using another public *Malaria* dataset composed of 27 558 cropped RBCs images [82]. The system reached on the test set (20% of this dataset) a classification accuracy of 95%. However, the classification reached 75.39% of accuracy when tested on the dataset used for training the segmentation model. Authors explained that this drop in performance is due to the difference between the acquisition parameters of the two datasets.

A similar deep learning based binary classifier (infected or not) as in [22] has been proposed by Sriporn et al. [91]. However, authors considered only the classification task and didn't address the segmentation one. To train the classifier, they used the public dataset composed of 27 558 cropped RBCs images [82] and have conducted a comparative experimental study between 6 popular deep architectures namely AlexNet, VGG-16, NasNetMobile, ResNet-50, Inception-V3 and Xception. To this end they explored different transfer learning strategies together with several optimizers. According to the authors, the best performance in term of classification accuracy has been obtained using Xception architecture. Precisely, they reached an accuracy of 99.28% using Nadam optimizer and Mish activation function. However, authors didn't precise the size of the cropped RBCs test set.

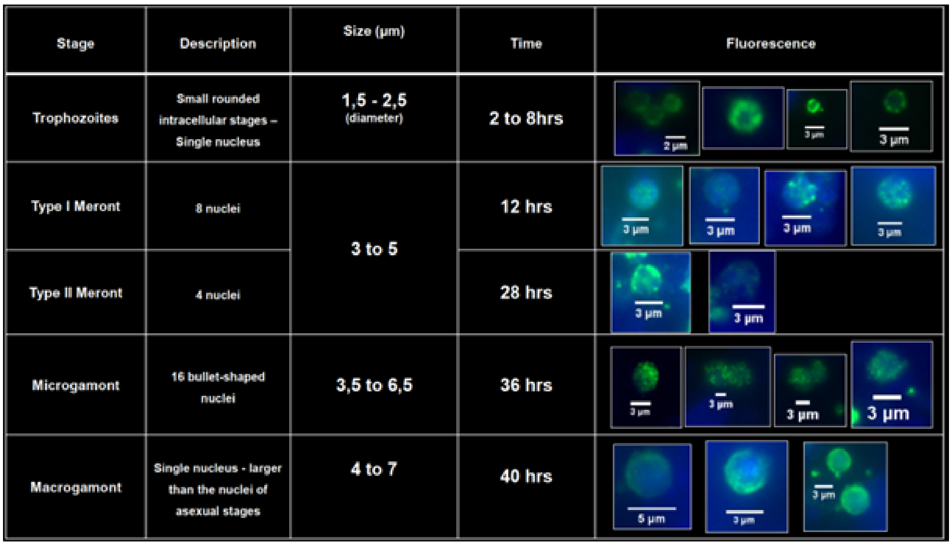
The system proposed by Maity et al. [65] allows to classify the RBCs into 6 classes corresponding to Normal RBCs, Falciparum species with two life stages and Vivax species with three life stages. To segment the RBCs from the input image, authors proposed to first correct the image color using an automatic equalization color technique then to train a pixel based MLP binary classifier (a Multi Layer Perceptron artificial network). The network is fed with pixels associated feature vectors, each one initially composed of 230 handcrafted descriptors covering geometrical and textural information. Authors exploited the training loss function versus one feature exclusion from the 230 in order to keep the most significant features (top 20). The final trained classifier outputs a pixel probability map of the input image which is post-processed using: i) a graph-cut algorithm to refine the borders of each segmented RBC and ii) a watershed algorithm to separate overlapped RBCs. Authors used for training the cell segmentation model their own dataset composed of 210 images and representing a total of 2720 RBCs. They used as a benchmark another dataset of 38 images (524 cells) and reported a Jaccard coefficient of 0.98 on the resulting segmentation but they didn't indicate how they exploited the data in order to reach this result. To classify the cropped images of RBCs, the authors proposed to train a classifier based on an adapted CapsNet deep architecture. To this end, they gathered the two datasets used for segmentation into one set (3244 RBCs) and augmented it to reach 141600 cropped images of RBCs. According to the authors, the classifier achieved an overall accuracy of 98.7% on the test set which corresponds to 15% of the augmented set.

Abbas and Dijkstra [3] proposed a system that permits to directly segment the parasite (only the falciparum species) instead of the RBC and classify its life stage. To segment the parasite, the authors designed a pixel based segmentation model using a Random forest binary classifier. To this end, a set of 9 handcrafted features is extracted for each pixel of the input image covering mainly Gaussian blur and average intensities while considering the image channels separately. The training of the segmentation model permits to output a binary mask which is refined by filling the holes of detected parasites. Authors trained and tested the segmentation model on their own publicly shared dataset composed of 837 images of RBCs collected from 16 patients infected with Falciparum. They reported an average Dice coefficient of 0.82 obtained on the whole dataset using a cross validation evaluation method. For the life stage classification of segmented falciparum parasites, authors considered 112 x 3 channels global handcrafted features extracted on the parasite cropped images including intensity and area related features. They trained a random forest classifier to learn 8 classes: 6 parasite life stages, white blood cell and debris. Authors reported an average classification accuracy of 58.8% when classifying all the life stages and 82.7% when simplifying the life stages to the three main ones namely ring, trophozoite and schizont.

Some research works have proposed parasite-related analysis approaches from the dataset MP-IDB [26, 25, 27, 65]. Di Ruberto et al. [26] proposed a method called Edge Boxes for detecting and quantifying red cells from a microscopic blood image. Di Ruberto et al. [25] proposed an efficient system for detecting white blood cells and classifying them as being affected by leukemia or healthy. Another article proposed by Di Ruberto et al. [27] leveraging the dataset MP-IDB to detect and quantitatively analyze red and white blood cells from blood images. Besides, Maity et al. [65] have also evaluated the feasibility of their algorithm on MP-IDB dataset. After Red Blood Cells (RBCs) segmentation, a trained model Capsule Network is applied in order to classify segmented blood cells to identify species and stages of the *Malaria* parasites. In particular, only two *Malaria* species (Falciparum and Vivax) and three stages (Ring, Gametocyte and Schzonat) are considered.

2.3 *Cryptosporidium* parasite image analysis

A brief biological introduction – *Cryptosporidium* is an intracellular parasite that occupies intracellular but extracytoplasmic vesicles in the host [6, 79, 73]. Its life cycle begins with the asexual stage, followed by the sexual stage [8]. The detailed information of different life stages is illustrated in Fig 2.2 and Fig 2.3.



Stage	Description	Size (µm)	Time	Fluorescence
Trophozoites	Small rounded intracellular stages – Single nucleus	1,5 - 2,5 (diameter)	2 to 8hrs	2 µm 3 µm 3 µm
Type I Meront	8 nuclei	3 to 5	12 hrs	3 µm 3 µm 3 µm 3 µm
Type II Meront	4 nuclei		28 hrs	3 µm 3 µm
Microgamont	16 bullet-shaped nuclei	3,5 to 6,5	36 hrs	3 µm 3 µm 3 µm 3 µm
Macrogamont	Single nucleus - larger than the nuclei of asexual stages	4 to 7	40 hrs	5 µm 3 µm 3 µm

FIGURE 2.2: Detailed information of the *Cryptosporidium parvum* life stages.

Our study concerns four morphologies of the asexual phase: oocyst, trophozoite, meront, and free form. As shown in Fig 2.4, ingestion of an oocyst leads to a process called excystation, releasing four parasites named free forms [85, 84]. The free form will enter the cell and be converted into a trophozoite [14]. Subsequently, the trophozoite evolves into type I meront. Afterwards, certain type I meronts develop into type II meronts and release free forms.

Existing work – Up to now, only one research has been conducted based on the analysis of microscopic images of *Cryptosporidium parvum*. In

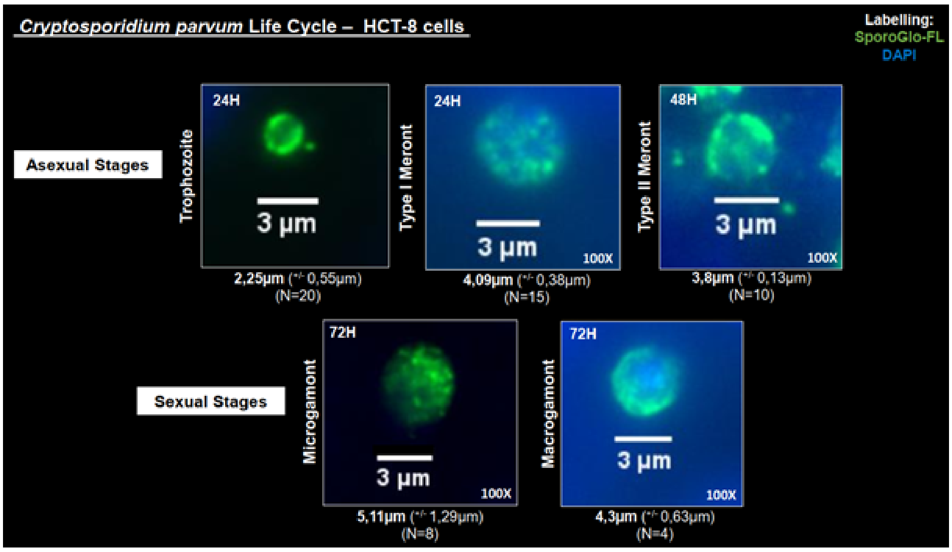


FIGURE 2.3: *Cryptosporidium parvum* life cycle examples.

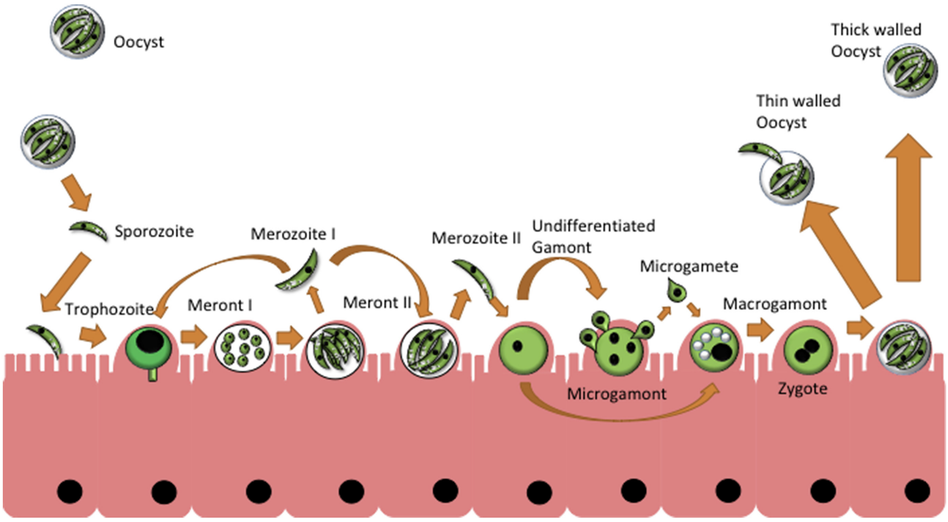


FIGURE 2.4: Outline of the *Cryptosporidium parvum* life stages (extracted from [8]).

[106], artificial neural networks (ANNs) were employed to identify the *Cryptosporidium parvum* oocysts. A total of 525 images were cropped to 36 x 36 pixels, and the cropped images were divided into two categories:

oocyst and non-oocyst images. To convert the digital images into the format used by the neural network, the authors chose a relatively simple algorithm: a histogram measuring the pixel intensity of grayscale images. Afterwards, this network was optimised by applying different quantities of training images and hidden neurons. The results indicated that the artificial neural network obtained satisfactory performance in test images. In several tests, the correct recognition rates ranged from 81% to 97% for oocyst images and 78% to 82% for non-oocyst images.

2.4 Other parasite image analysis

In addition to the existing works on the *Malaria* and *Cryptosporidium*parasite image analysis, we also found a number of relevant works related to other parasites [86, 63, 76, 113, 39]. Three studies regarding the classification of intestinal parasite images used different methods [86, 63, 76]. Fig 2.5 shows some samples of parasite species addressed in [63].

FIGURE 2.5: A sample of each parasite species addressed in [63].

Roder et al. [86] were the first to experiment with Deep Boltzmann Machine and Deep Belief Network for parasite classification, and the robustness of the model was confirmed by experiments using 3 different datasets including fecal impurity surrounds. Machaca et al. [63] presented a method for generating synthetic images of parasites using Deep Convolutional Generative Adversarial Network (DCGAN), followed by experiments using the pre-trained deep learning models Resnet34, Resnet50, VGG16, Densenet121, and Inceptionv3. The experimental results demonstrated the accuracy improvement with the use of synthetic datasets. While a hybrid method was proposed by Osaku et al. [76], which combines the views from two decision systems (support vector machines and deep neural networks VGG16) with complementary properties so as to enhance the overall efficiency.

In addition, in a study on microorganisms [113], some novel deep learning methods for parasite microscopic image recognition integrated the geometric morphology of parasites into deep learning framework to reduce the dependence on labeled data [54, 53, 46]. For example in [53], microscopic parasites (Toxoplasma, Plasmodium, Babesia) were linked to their morphologically similar macroscopic objects (banana, ring, and pear) by deep cycle transfer learning (illustrated in Fig 2.6). The results showed the high accuracy of the proposed method, with an average accuracy of more than 95% for all parasite types.

What's more, Górriz et al. [39] trained a U-Net model that can successfully segment leishmania parasites (displayed in Fig 2.7) and classify them into 3 classes: promastigotes, amastigotes, and adhered parasites. The authors thought that the improvement could be applied by using larger databases, keeping the balance of the classes, and providing a precise annotation.

2.5 Conclusion

Among the state of the art of parasite analysis from microscopic images presented in this chapter, it is worth mentioning that the *Malaria* studies take up the major part. Although the performances of the existing systems are promising, comparisons among different systems are quite challenging. The generality of the systems is also difficult to guarantee since most of the datasets used for assessing performance are not publicly available.

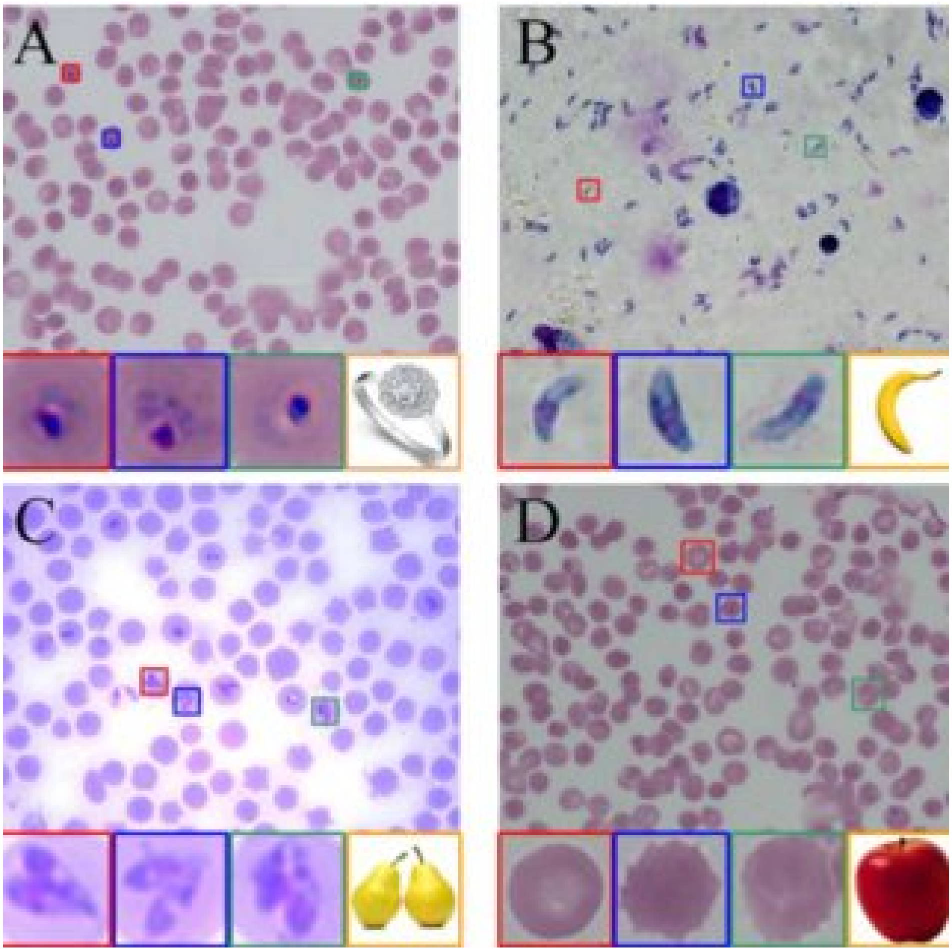


FIGURE 2.6: Microscopic parasites (Toxoplasma, Plasmodium, Babesia) were linked to their morphologically similar macroscopic objects (banana, ring, and pear) (extracted from [53]).

While for *Cryptosporidium*, there has been few vision-based works for the analysis of this parasite. Additionally, we have also presented the life cycle of *Cryptosporidium* and the evolution of each life stage in the asexual phase.

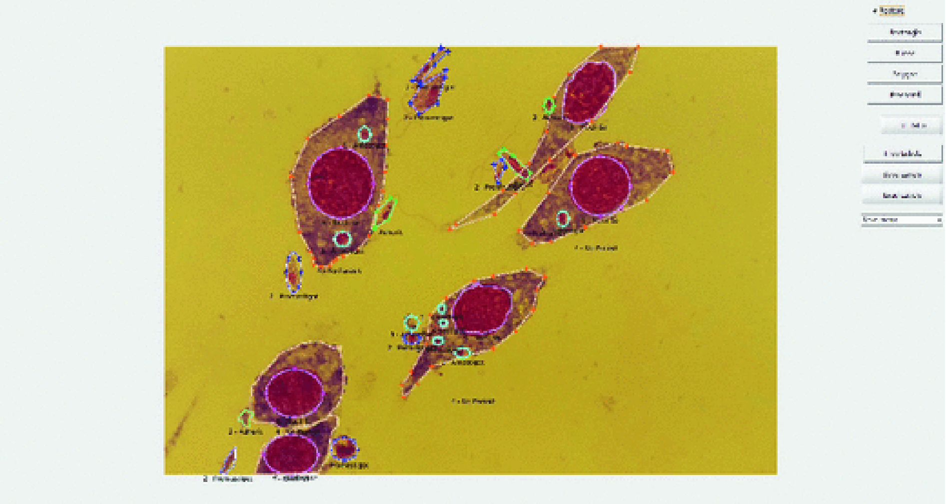


FIGURE 2.7: Annotation of leishmaniasis parasite (extracted from [39]).

Chapter 3

Generalized deep learning for *Plasmodium* parasite image analysis

3.1 Introduction

Malaria is one of the epidemics that possess the potential to threaten human life [18, 102]. In fact, during the year of 2018 there were an estimate of 228 million cases of *Malaria*, resulting in over 400,000 fatalities around the world. Of this group, children younger than 5 years of age represent to be the most fragile, with about one death under 2 minutes, constituting over 65% for all deaths due to *Malaria* [100, 105]. *Malaria* is caused by Plasmodium parasite and is transmitted through the bite of infected female Anopheles mosquitoes [99]. *Malaria* is widespread in tropical and subtropical areas, particularly in Asia, Sub-Saharan Africa and Latin America [10]. Globally, nearly 85% of *Malaria* deaths occurred in Africa and India during 2018, with 24% in Nigeria [105]. In WHO's Southeast Asia and Africa region, the number of mortality has declined from approximately 530,000 (2010) to 380,000 (2018). Nevertheless, since 2016, the decrease in mortality rate has decelerated [104].

There are a couple of ways for diagnosing *Malaria* [67], of which microscopic examination of thick or thin blood smears is the gold standard [67, 66]. To be more specific, blood samples are colored according to the

Giemsa method [80] and then a well-trained expert analyzes them with a microscope to detect parasites in the blood cells as well as to determine their species. In fact, human *Malaria* infections are due to Plasmodium parasites consisting of no less than six species, namely P. falciparum, P. malariae, P. ovale (two variants), P. vivax and P. knowlesi [90, 52]. These species may contribute to the death of patients through infection with red blood cells [61]. Therefore, regardless of the species of *Malaria* parasite, an early diagnosis is necessary in order to i) enable the patient to receive timely treatment [69], ii) limit the transmission of the infection via mosquitoes within the community [13]. However, specialist microscopic examination is still extremely time-consuming and challenging since parasites are typically tiny and potentially able to be mistaken with non-parasitic objects [95, 68, 43]. Furthermore, the examination is suggested to be completed by two specialists [5] to achieve a more solid diagnostic result. There is indeed a significant risk to the patient that inaccurate diagnosis may lead to inappropriate treatment [29].

The World Health Organization (WHO) has encouraged research on appropriate approaches for a fast and economical diagnosis to treat *Malaria* [75]. Hence, there is a need to develop an automated *Malaria* diagnostic system that is not only reliable, but can be faster and more economical, especially in resource-limited regions [111].

Artificial intelligence (AI) currently provides the possibility to achieve this goal; whereas in scientific community the employment of deep learning techniques is now the obvious alternative for medical image analysis [56]. All these techniques offer strong aids for specialists to execute high-level medical tasks, for instance, diabetic retinopathy detection [40] or skin cancer classification [32]. In this study, we propose a framework for diagnosing *Malaria* infection in humans using microscopic images of thin blood smears. To be specific, the framework aims at segmenting the Plasmodium parasites and distinguishing the species into four major classes: P. Falciparum, P. Ovale, P. Malaria and P. Vivax.

The main contributions of our study are as follows:

- A straightforward segmentation and classification method which relies on the parasite itself. Noteworthy is the fact that our method is complements to those proposed methods [65, 22, 70]. Although these methods are designed to quantify the amount of infected RBCs, our method enables to outline the shape of the parasite within these cells, thereby providing an accurate analytical tool.

- Our framework demonstrates the efficiency of generalization to interclass data on 2 *Malaria* datasets [59, 2]. These datasets consist of images that have extremely different features owing to variations in data collection conditions (e.g., microscope characteristics, illumination conditions, and magnification).

- Our experimental study takes into account six data sources [59, 2, 21, 82, 96, 57], offering an overview of *Malaria* feature detection, segmentation and classification.

- We named our data augmentation technology called Local Parasite Texture Scanning (LPTS), which allows for greatly increasing the number of images of cell images. LPTS technology generates local variations of images from authentic datasets with simple but effective cropping and decentering techniques. LPTS greatly improved the precision of classification in experiments with deep architectures.

3.2 Exploited CNN-based architectures

In this section, we describe in detail the related architectures for image segmentation as well as image classification that have been employed in our work.

Object segmentation – Convolutional neural networks (CNN) can achieve image classification, but how to identify objects in specific parts of images through deep learning has always been a problem. Until 2015, Long et al. [60] proposed the Fully Convolutional Networks, leading the field of image semantic segmentation. FCNs can classify images at the pixel level, thereby solving the problem of semantic segmentation. They have received increasing attention and have been employed successfully with good results in biomedical images, for example, liver and lesion segmentation in CT [17] and cardiac segmentation in MRI [98].

Unlike convolutional neural networks, there is no fully connected layer in FCNs. FCNs adopt deconvolution layers to upsample the feature map of the last convolutional layer in order to the same size as the input image. In this way, a prediction is generated for each pixel, while retaining the spatial information in the original input image. Finally, the pixel-wise classification is implemented based on the feature map obtained after the upsampling operation.

Fig 3.1 (Row A) illustrates the general network of a FCN architecture. The network consists of the following basic layers, including convolution (conv), pooling (pool), activation and deconvolution (deConv) [50]. Furthermore, the authors proposed two other FCN architectures that vary in upsampling scale: FCN-16s (Row B) and FCN-8s (Row C). They fuse the pooling information from different layers and can provide better semantic segmentation results than that of the original. The network architectures of FCN-16s and FCN-8s are also shown in (Row B is FCN-16s and Row C is FCN-8s). For instance, in FCN-8s, the coarse output of the FCN model is initially upsampled by a factor of 4, while the pool4 image is upsampled by a factor of 2. Subsequently, these upsampled images are fused with the pool3 layer image, which is finally upsampled by a factor of 8 to yield the predicted image of the same size as the input.

After series of successful applications of FCNs in semantic segmentation, numerous new algorithms based on FCNs techniques have been proposed. They are extensively investigated in the fields of image segmentation, classification, etc [103, 64]. As a variant of FCNs, U-Net architecture [87] has been widely used in the field of cellular segmentation and biomedical image processing [88, 92].

The architecture U-Net is shown in Fig 3.2. It includes repeatedly applying two 3x3 convolution operations, ReLU activations, and a downsampling operation through 2x2 max-pooling with stride 2. Between some

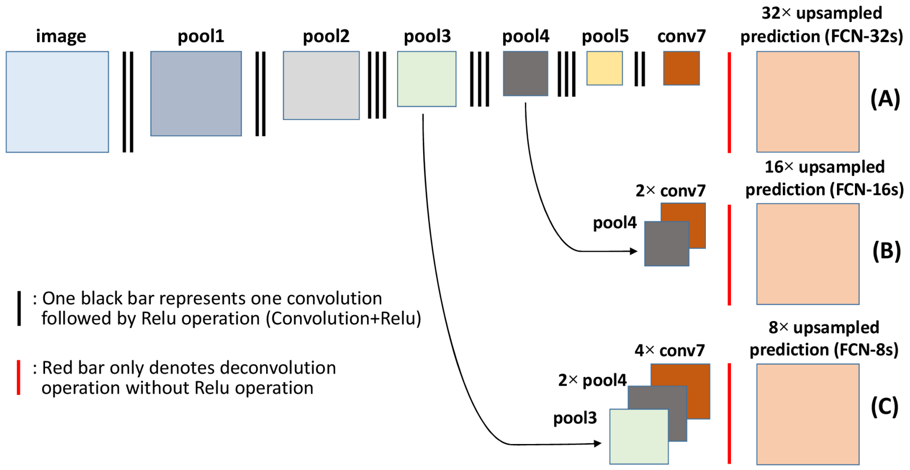


FIGURE 3.1: Fully convolutional neural networks (FCNs) [60].

convolutional layers, dropout is adopted. The deconvolution process comprises an upsampling of the feature maps, subsequent a 2x2 convolution, concatenation of the corresponding feature maps from the previous convolutional layers, and two 3x3 convolutions with ReLU activations. In contrast to the original U-Net, a 3x3 convolution (including two filters) layer is applied before the output. The last layer is employed to map feature vectors to the required number of classes to generate the result of pixel-wise segmentation.

Object classification – In this section, I will compare several architectures that are frequently used in recent years.

AlexNet

AlexNet [50] is the first large-scale convolutional neural network architecture performing well on ImageNet [23] classification. AlexNet participated in the competition and outperformed previous non-deep learning based models with remarkable margins.

AlexNet was trained on ImageNet and the input size was 227 x 227 x 3. The AlexNet architecture consists of convolutional layers, pooling layers and fully connected layers. As is displayed in Fig 3.3), there are five convolutional layers, starting with a convolutional layer with a kernel

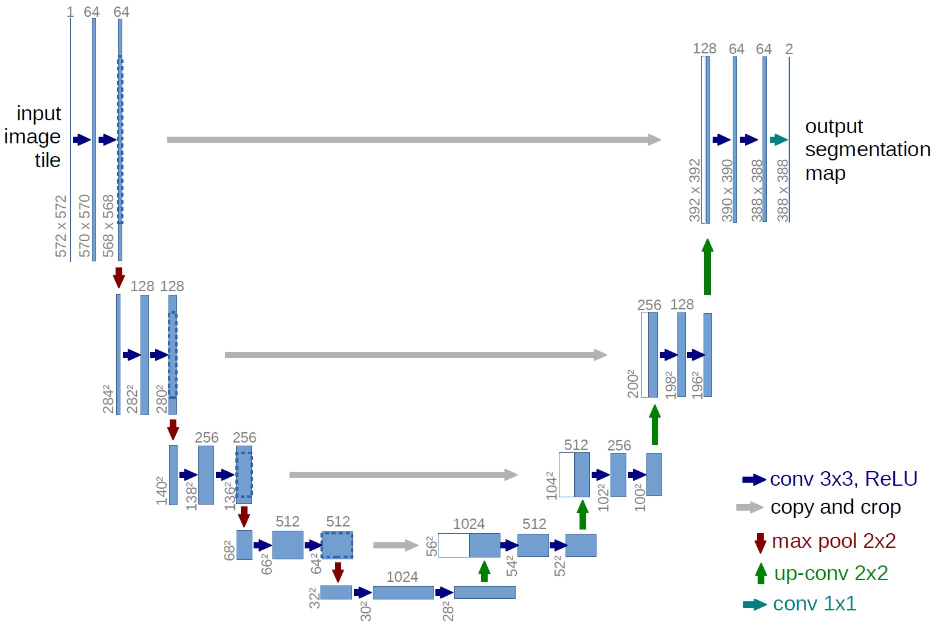


FIGURE 3.2: U-Net architecture [87].

size of 11 x 11, followed by 5 x 5 and some 3 x 3. The last layer of the fully connected layer applies the softmax activation function, which maps to 1,000 ImageNet classes.

VGGNet

In LSVRC-2014, VGG16Net [89] won the competition on ImageNet with 92.7% accuracy, reducing the error of the AlexNet by a factor of 2. There are two versions of VGGNet, VGG16 and VGG19. The VGG19 architecture (also known as VGGNet-19) is very similar to the VGG16, but has more convolutional layers in it (see in Fig 3.4).

Inception

Inception v3 [93] is a combination of many ideas developed by several researchers for many years and has been proven to yield an accuracy of over 78.1% on the imageNet.

Inception consists of symmetric and asymmetric blocks, including convolutional layers, max pooling, average pooling, dropout and fully connected layers. Its main idea is to merge convolutional kernels in parallel, as shown in Fig 3.5. Each inception block contains convolutional layers of different kernel sizes to extract features in different scales.

FIGURE 3.3: AlexNet architecture [50].

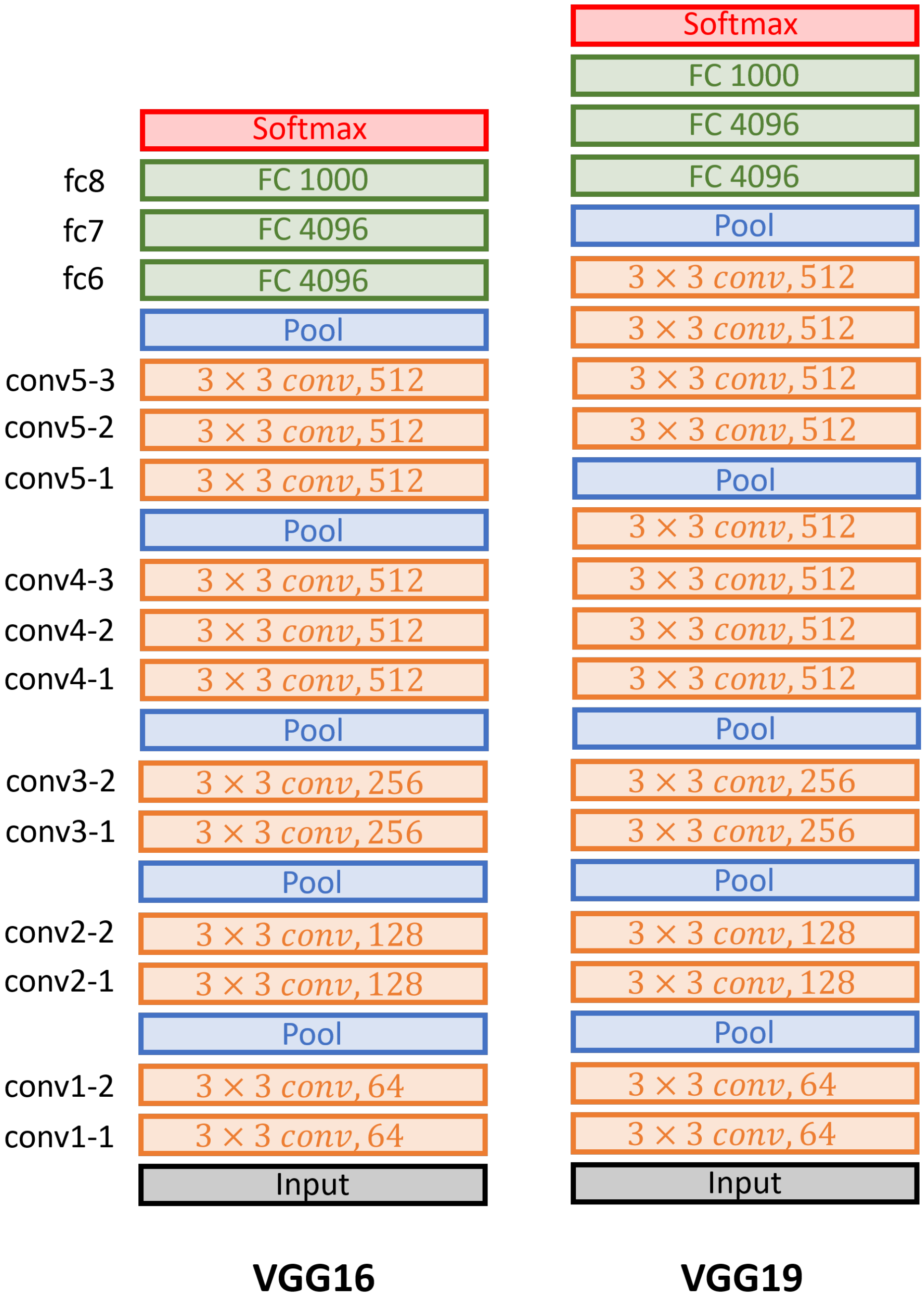


FIGURE 3.4: Two versions of VGGNet, VGG16 and VGG19 [89].

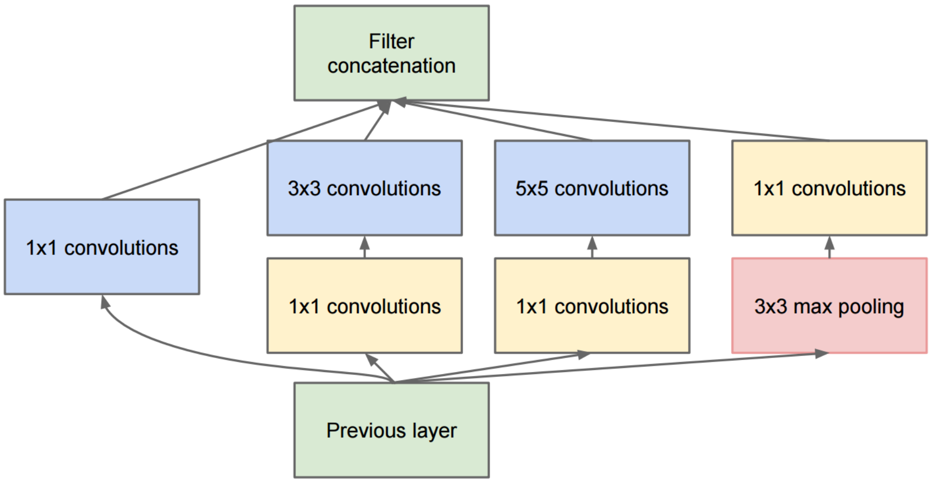


FIGURE 3.5: Inception module [93].

ResNet

ResNet [44] won first place in the ILSVRC and COCO 2015 competition with an error rate of 3.6%. It even performed better than humans.

The main fundamental element of ResNet is the residual block (illustrated in Fig 3.6). As the layers get deeper, the calculations become more complex. The layers are stacked on top of each other, each trying to learn some underlying mapping. ResNet-18 is shown in Fig 3.7 as an example.

3.3 Methodology

Our *Malaria* diagnostic system consists of a number of steps, as shown in Fig 3.8. To be specific, taking as input an image stack of the patient sample obtained by microscope, the linear interpolation-based downscaling operation (Step 1) is applied to these images in order to achieve 224 by 224. In fact, through the analysis of the *Malaria* public datasets, we observed that the captured images have a high resolution (ranging from 640 x 480 to 2592 x 1944), which is unsuitable for direct processing with deep architectures. In order to achieve this goal, we have adopted a small square resolution, which has been widely used in the field of deep learning.

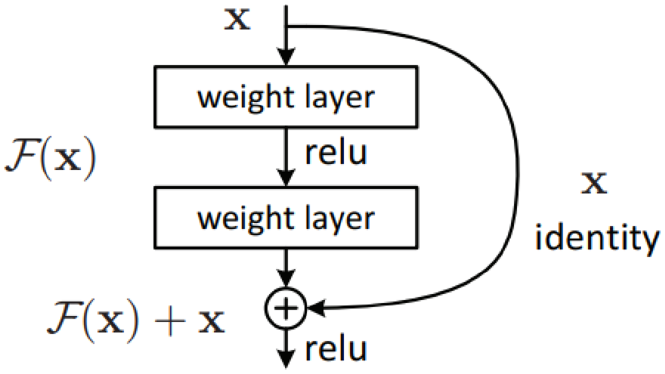


FIGURE 3.6: Residual-Block used for ResNet architecture [44].

Next, we propose a new approach to the segmentation of the possible parasites in each image, which allows for the generation of relevant binary masks (Step 2). Then, the resulting binary mask is enlarged (with the identical linear interpolation technique used in Step 1 to achieve the original resolution) and applied to mark the detected parasites (Step 3). Both the detected parasites and their bounding boxes in the binary mask are mapped onto original image. (Step 4).

Based on this, the experts can take advantage of two possibilities: (1) the classification by the analysis to distinguish (Step 5.1 - 5.3), (2) to supply indicators, like the number of parasites detected and fine delineation of parasite shape (Step 5 ’).

For the prediction of parasite species, the parasite with the largest area is chose out of each image and extracted through a square window of 224 x 224 pixels centered on the parasite before cropping. Then, we classify the parasite crops into 4 classes: P. Falciparum, P. Malariae, P. Ovale, and P. Vivax. Finally, a major voting method is used to determine the infected parasite species based on the previous classification results. If there is no parasite found in the segmentation step, then the classication step will output no *Malaria* infection.

It should be noted that the selection of 224 x 224 for image crops is appropriate in this situation, as it has been found that in the majority of

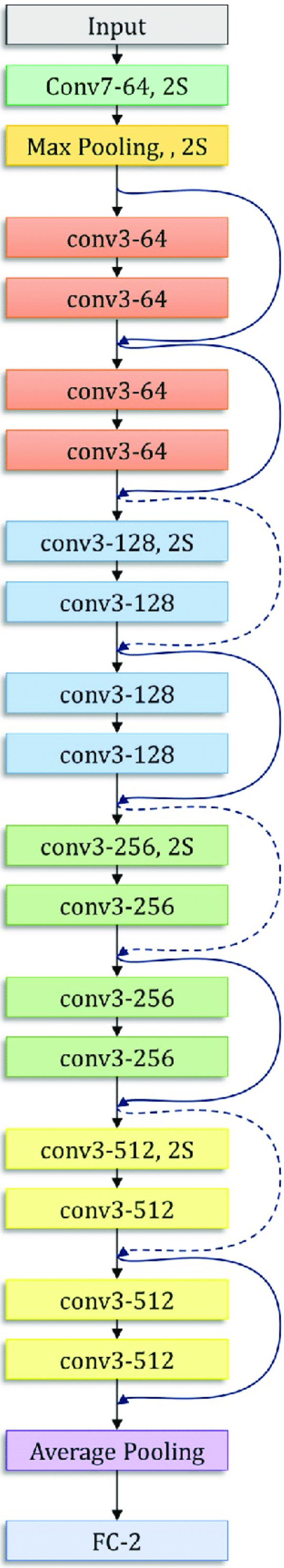


FIGURE 3.7: ResNet-18 architecture [51].

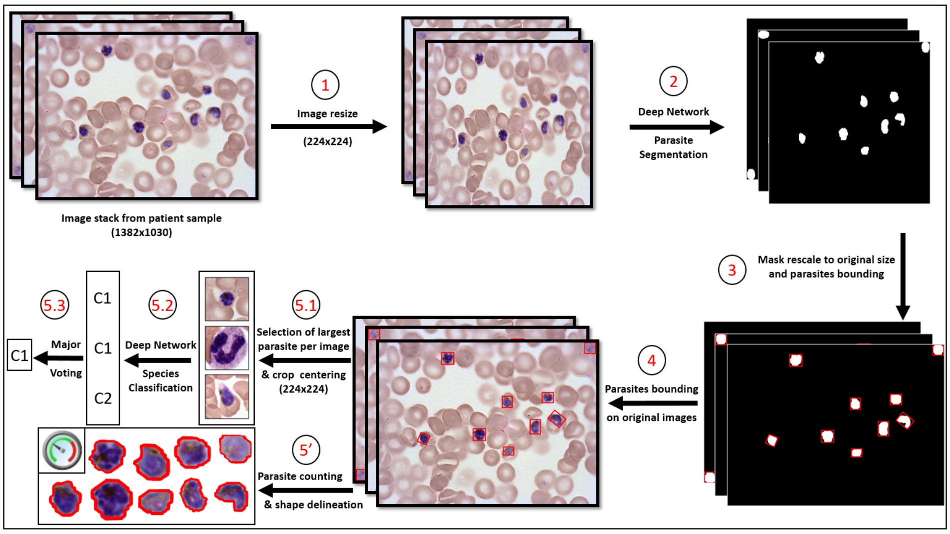


FIGURE 3.8: Overview of our proposed *Malaria* diagnosis assistance system.

cases this resolution is enough to include a single parasite (99% of the parasite in all data sets is less than 224 × 224). We also downscale the larger parasite crops to 224 x 244 (their width or height is no more than 350 pixels). In addition, we consider only the largest parasite in each image, so as to improve the probability of detection of a true positive.

3.3.1 Parasite segmentation

In order to extract potential parasites from the input image, we used a pixel based method to segment them. Specifically, a label is allocated to every pixel in an input image, i.e., a foreground and a background. The foreground category includes the pixels of the subject of interest, which are parasites, whereas the background class includes other pixels in the picture, such as uninfected cells, extracellular fluid. Therefore, the classification of the individual pixels results in the creation of a binary mask that allows to delimit the the contour of the parasite from the image.

We use the U-Net architecture as the kernel of our image segmentation algorithm. This architecture consists of an encoder and a symmetric decoder. The encoder allows the input image to be scaled down by passing it into a series of convolution-layer blocks (64, 128, 256, 512), with an intermediate maxpooling operation followed by a stack of 1024 small contextual feature maps as output. The symmetric decoder allows to upgrade the generated feature maps from the encoder by passing them into an inverted sequence of convolution layers (512, 256, 128, 64). Each of these blocks is preceded by an intermediate up-convolution operation and produces a stack of feature maps with size 64 as output. In order to improve the positioning accuracy of the target, the authors proposed that the convolution layer block of the encoder be connected to it, which is the equivalent (in terms of the number of layers). Then there is a second set of 64 convolution layers, which is followed by the binary mask.

In order to improve the performance of the U-Net, we customized a variant by integrating a deep architecture VGG19 [89] that is tailor-made for image classification. In particular, we used VGG19 instead of the original U-Net encoder. By this way, we followed the same principles as U-net to construct the symmetric decoder.

Based on the Jaccard Index of Eelbode et al. [31] we kept this measure and modified it to solve the problem of segmentation. In particular, the variation has been trained to minimize the Jaccard loss (Jl) between an image's predicted binary mask and its associated ground truth mask. To define this loss, we first formulate our parasite segmentation problem as follow:

Let bm_a be the binary mask obtained by our approach for an image $Img(w,h)$, where (w,h) represent the image's width and height. We denote the label of a pixel $p_{i,j}$ positioned at indexes (i,j) of Img by $l_{bm_a}^{i,j}$ in the binary mask bm_a, where:

$$l_{bm_a}^{i,j} = \begin{cases} 1 & if\ p_{i,j} \in parasite \\ 0 & else \end{cases} \tag{3.1}$$

Similarly we define bm_{gt} as the ground-truth binary mask that corresponds to a manual segmentation of Img performed by a specialist.

The Jl between two masks is defined as:

$$Jl(bm_a, bm_{gt}) = 1 - Ji(bm_a, bm_{gt}) \tag{3.2}$$

Ji represents Jaccard index and is defined as:

$$Ji(bm_a, bm_{gt}) = \frac{bm_a \cap bm_{gt}}{bm_a \cup bm_{gt}} with \tag{3.3}$$

$$\begin{cases} (bm_a \cap bm_{gt}) = \sum_{i,j}^{h,w} I(l_{bm_a}^{i,j} = l_{bm_{gt}}^{i,j} = 1), \\ I : \text{binary indicator} \\ (bm_a \cup bm_{gt}) = \sum_{x=a,gt} \sum_{i,j}^{h,w} I(l_{bm_x}^{i,j} = 1) - (bm_a \cap bm_{gt}) \end{cases}$$

The Jl permits the measurement of the difference between the two segmentations and outputs in the range $[0,1]$ where 0 indicates that the two segmentations are the same.

To minimize Jl, the architecture's weights are optimized according to the adaptive moment optimization technique (Adam) [15], which is a stochastic gradient descent technique based on an adaptive estimation of the first and second order moments. We denote w_{ij_t} a weight to update between two neurons x_i and x_j of our architecture for step $t+1$ of the training:

$$w_{ij_{t+1}} = w_{ij_t} - \frac{\eta \cdot \widehat{m}_{t+1}}{\sqrt{\widehat{v}_{t+1}} + \varepsilon} with \tag{3.4}$$

$$\begin{cases} \widehat{m}_{t+1} = m_{t+1}/(1-\beta_1^{t+1}) & \\ \widehat{v}_{t+1} = v_{t+1}/(1-\beta_2^{t+1}) & \\ m_{t+1} = \beta_1 m_t + (1-\beta_1) \cdot \delta_{j_t}, & \beta_1 = 0.9, m_0 = 0 \\ v_{t+1} = \beta_2 v_t + (1-\beta_2) \cdot \delta_{j_t}^2, & \beta_2 = 0.999, v_0 = 0 \\ \eta = 10^{-3}, & \text{learning rate} \\ \varepsilon = 10^{-8}, & \text{calculation stability} \end{cases}$$

Here the (β_1, β_2) are the exponential decay rates for the first and second moment respectively and δ_{j_t} is the gradient error back-propagated from the loss Jl over the neuron x_j at step t of the backward pass.

3.3.2 Parasite classification

To implement our classification module, we developed our own deep architecture called Light-Net, that is based on a standard CNN. As shown in Figure 3.9, the architecture consists of 3 blocks of convolutional layers, each followed by a $Relu : max(x,0)$ activation function and a max pooling operation. The convolutional sequence is followed by a global average pooling, then by a dense layer of 64 units, with a $Relu$ activation and an output dense layer of 4 units that corresponds to the four species to be predicted. For this purpose, the output layer is related to a $softmax$ function to calculate the species' probability. The architecture has been trained to minimize the categorical cross entropy loss L for the crop classification defined below:

for the crop classification of an image defined as follows.

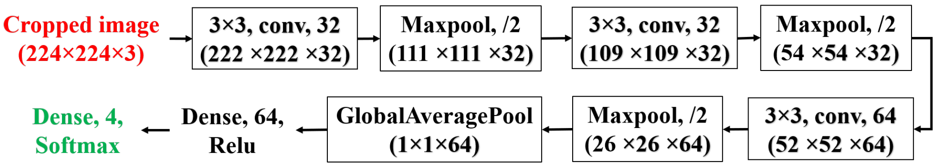


FIGURE 3.9: Our proposed CNN classification architecture named Light-Net.

$$L = - \sum_{c=1,4} y_c log(p_c) \tag{3.5}$$

where y_c is a binary indicator whether the class c is the correct one for the given crop and p_c is the predicted probability by the model on the same crop for being of class c. The weights have been updated based on a Root Mean Square propagation technique (RMSprop) [15]. We denote w_{ij_t} a weight to update between two neurons x_i and x_j of our architecture for step $t+1$ of the training:

$$\begin{cases} w_{ij_{t+1}} = w_{ij_t} - \left(\frac{\eta}{\sqrt{v_t+\varepsilon}} * \delta_{j_t}\right) & \\ v_t = \beta v_{t-1} + (1-\beta)\delta_{j_t}^2, & \beta = 0.9, v_0 = 0 \\ \eta = 10^{-3}, & \text{learning rate} \\ \varepsilon = 10^{-8}, & \text{calculation stability} \end{cases} \tag{3.6}$$

where δ_{j_t} is the gradient error of the neuron x_j back-propagated from the loss L on the predicted class for the input image at step t of the backward pass.

3.3.3 Proposed parasite-oriented data augmentation technique

Data augmentation by applying our local parasite texture scanning technique (LPTS) to generate a set of derived parasite crops from an image I containing one centered parasite P, according to the illustration displayed in Figure 3.10, can be noted by the following:

$$LPTS(I_{1\times P}) = \bigcup_{i=0}^{8} I_{p_i} \tag{3.7}$$

where p_0 corresponds to the initial position of the frame f (parasite centered) and $p_{i_{(i\neq 0)}}$ corresponds to all possible horizontal and/or vertical rigid displacements of f from p_0; namely p_1: $a=0$ and $d=0$, p_2: $a=0$, p_3: $a=0$ and $b=0$, p_4: $b=0$, p_5: $b=0$ and $c=0$, p_6: $c=0$, p_7: $c=0$ and $d=0$, p_8: $d=0$ with a, b, c, d representing respective distances in between the parasite bounding box and the boundaries of f; in this case, $card(LPTS(I_{1\times P})) = 9$ parasite crops. In situations where the parasites are located at the image borders, possible displacements of f are limited, thus reducing the amount of derived crops. Overall, considering N parasites contained in an image, the maximum number of parasite crops derived from LPTS is:

$$card(LPTS(I_{N\times P})) = 9 \times N \tag{3.8}$$

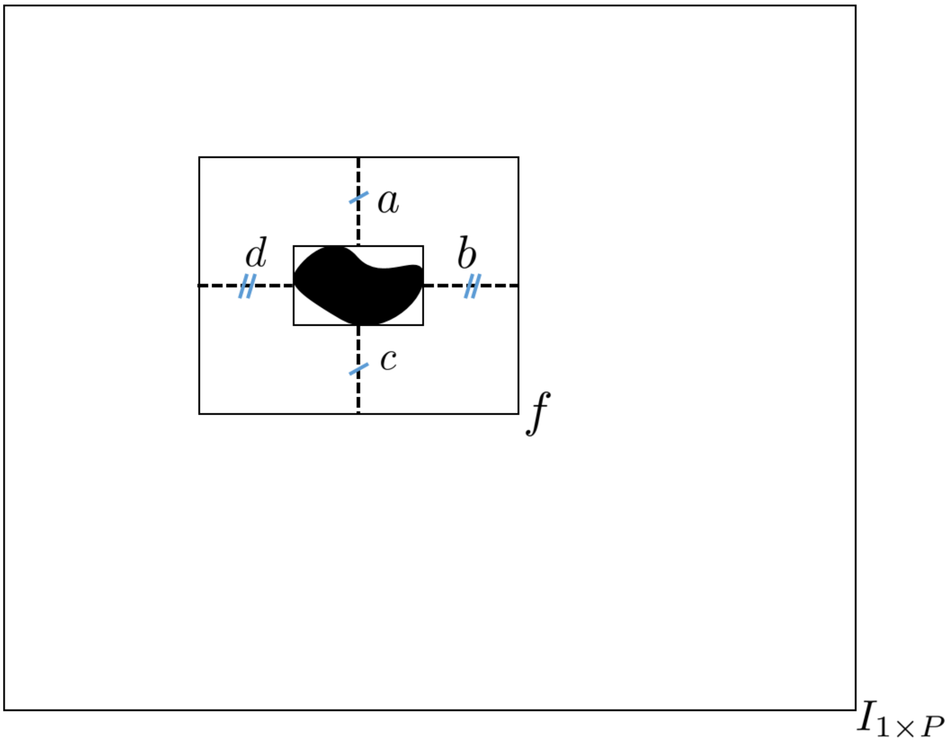
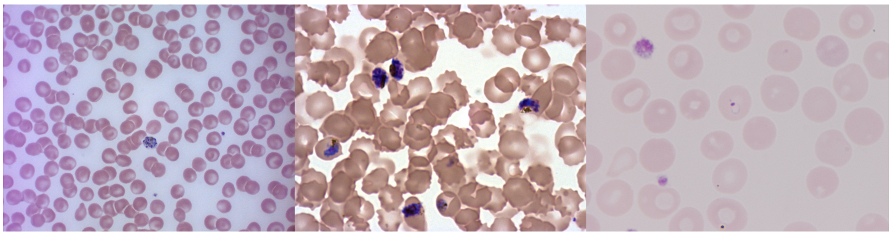


FIGURE 3.10: Principle of the *LPTS* technique by horizontal and/or vertical displacements of the rigid frame f around the parasite $P \in I_{1\times P}$ varying a, b, c, d parameters from an initial position p_0.

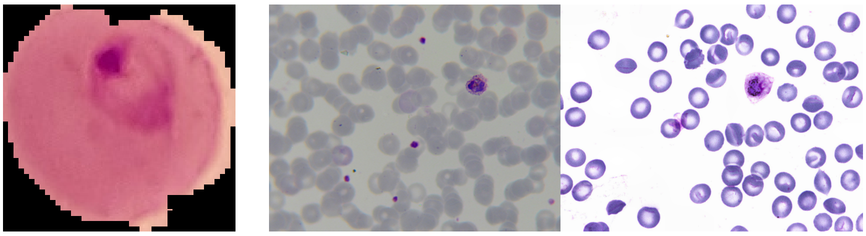
3.4 Experimental study

3.4.1 Datasets analysis

To explore the possibility of generalization, we have explored 6 public *Malaria* datasets. It should be noted that, the *Malaria* diagnosis in all samples was performed according to the May-Grünwald-Giemsa (MGG) method [80]. However, the images were obtained with a variety of microscopes and different parameters, in particular with respect to magnification and illumination. Even within the dataset, these parameters may differ. As a result, the images display large differences in resolution, colour, contrast, brightness and texture (shown in Figure 3.11).



(A) $Dataset_1$[59] (B) $Dataset_2$[2] (C) $Dataset_3$[21]



(D) $Dataset_4$[82] (E) $Dataset_5$[96] (F) $Dataset_6$[57]

FIGURE 3.11: Image examples from the experimented public datasets.

Figure 3.12 visualizes the six datasets by data dimensionality reduction. In particular, PCA (Principal Component Analysis) and t-SNE (t-distributed stochastic neighbor embedding) [62] were applied to each dataset to generate a point distribution from the 200 image samples. This figure allows us to observe the intra- and inter-class relationships of the images in the different datasets. It is worth noting that the visual exploration shows a high intra-class homogeneity for the $Datasets_{3,4}$; while the others are more moderate. The $Dataset_4$ is isolated; it confirms the specificity of its images, as they involve single cells, while the other datasets consist of multicellular images. The $Datasets_{2,3,5,6}$ are connected by overlapping regions. In these datasets, the regions of $Datasets_{2,6}$ are different from the other datasets because both are homogeneously stretched and relatively overlapping. Despite $Dataset_1$ represents multicellular images, it demonstrates a large inter-class gap compared to the other datasets.

These 6 datasets are summarized in Table 3.1.

The table illustrates that only $Dataset_1$ and $Dataset_2$ propose the ground-truth for parasite segmentation. Furthermore, only $Dataset_1$ includes 4

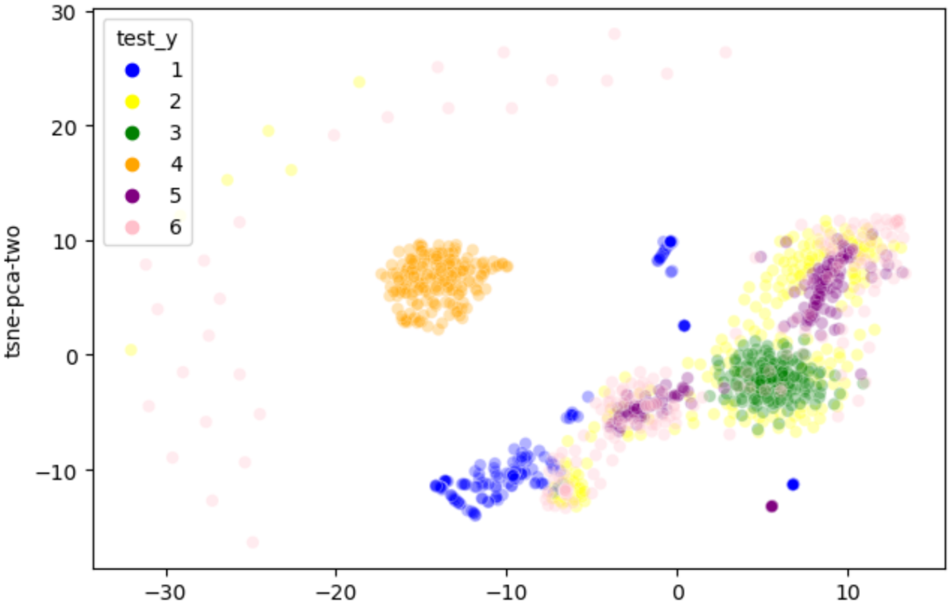


FIGURE 3.12: The global visualization of six datasets by the distribution of points from 200 image samples of each dataset.

TABLE 3.1: Characteristics of the public malaria datasets used in our experiments.

Dataset reference	Infected/non-infected patients	Type of image*	Total of images	Image resolution (W x H)	Total of parasites	Parasite species**	Parasite segmentation truth
$Dataset_1$ [59]	17/0	PC	210	2592 x 1944	1437	F,M,O,V	Yes
$Dataset_2$ [2]	17/0	PC	883	1382 x 1030	3586	F	Yes
$Dataset_3$ [21]	5/0	PC	331	2400 x 1800	-	NA	No
$Dataset_4$ [82]	151/50	IC	27 558	Heterogeneous	13 779	F	No
$Dataset_5$ [96]	-	PC	654	640 x 480	-	F	No
$Dataset_6$ [57]	-	PC	1328	1944 x 1328	-	V	No

*PC: Population of Cells, IC: Individual Cell. **F: Falciparum, M: Malariae, O: Ovale, V: Vivax, NA: Not Available.

species of parasites. For these reasons, we used these two datasets as basis for the training / validation / test of our generalized segmentation and classification modules and we utilized $Datasets_{3,6}$ for additional tests.

3.4.2 Results and evaluation of the generalized segmentation module

3.4.2.1 Data preparation

Figure 3.13 illustrates how we split the training / validation / test sets to experiment with $Dataset_1$ and $Dataset_2$. The images in each folder are chosen randomly. Nevertheless, each set in $Dataset_1$ included the images for each parasite species, which is an essential condition for learning its shape segmentation. As displayed in Figure 3.13, two patients (with Falciparum species) were excluded from the validation set of $Dataset_1$ because they both had only three images. Therefore, we selected one image for each of them as the test set and added the remaining two images to the training set.

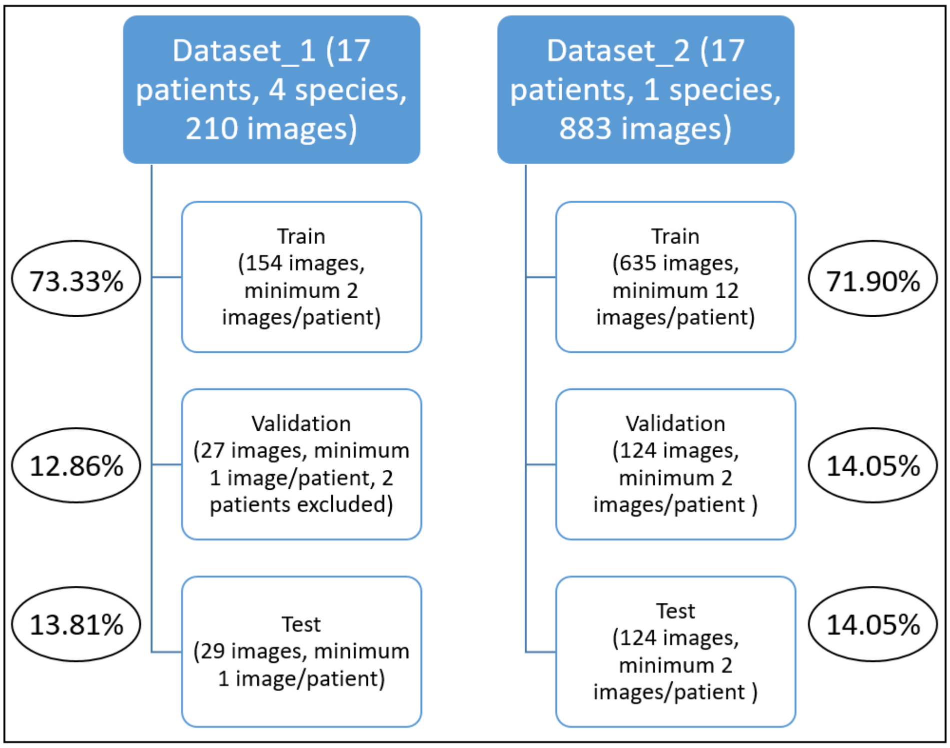


FIGURE 3.13: Data split for segmentation module.

We also created a global training set (hybrid set) by combining the two datasets named $Dataset_{12}$, including 1559 images. We artificially increased the training set of $Dataset_1$ from 154 to 924 images so as to relatively balance the two datasets (924 and 635 respectively) in a relative manner. For this purpose, we applied five transformations on $Dataset_1$, as follows: (1) two spatial transformations, involving horizontal and vertical flips; (2) three pixel-based transformations, covering integrated contrast and brightness, CLAHE and Gaussian noise.

3.4.2.2 Segmentation performance analysis

We assessed the quality of our segmentation module using the Average Precision metric (AP), which is often employed for the assessment of biomedical image segmentation (e.g [88, 92]). In our case, the metric allows to calculate the ratio $AP = TP/(TP+FP+FN)$ between the number of correctly detected parasites True Positive (TP) and incorrectly detected parasites False Positive (FP) as well as missed ones False Negative (FN). For this purpose, TP, FP and FN are determined compared to the ground-truth of parasites based on a matching threshold which corresponds to the Jaccard index (see equation 3.3) in the range of $[0.5, 1]$. The value 0.5 is a common threshold for assessing segmentation quality. In our case, if the parasite detected has at least 50% of the area matches the area of the parasite in the ground truth, then it is detected as a TP.

Our segmentation architecture called U-Net_VGG19 has been trained and validated on:

- The training / validation sets of the two datasets separately which allow for the production of two specialized segmentation models U-Net_VGG19_$Dataset_1$ and U-Net_VGG19_$Dataset_2$
- The training / validation sets of the two datasets together ($Dataset_{12}$) which allow for the production of generalized segmentation model U-Net_VGG19_$Dataset_{12}$.

We emphasize that among the five widely used optimizers (Adam, SGD, RMSprop, Adagrad, Adadelta), we observed that the Adam performed the best.

For ease of reading, we denote U-Net_VGG19_$Dataset_1$, U-Net_VGG19_$Dataset_2$ and U-Net_VGG19_$Dataset_{12}$ parasite segmentation models as model 1, model 2 (specialized models) and generalized model respectively.

Figure 3.14 illustrates the AP curves of specialized and generalized models generated on the test sets of $Dataset_1$ on the left (179 parasites) and $Dataset_2$ on the right (485 parasites). The curves indicate that while the specialized models are incapable of detecting parasites on the external data ($AP_{Ji@0.5} = 0.03$ for model 1 on $Dataset_2$ and model 2 with an $AP_{Ji@0.5} = 0.04$ on $Dataset_1$), the generalized model successfully segmented parasites on both datasets with promising precision ($AP_{Ji@0.5} = 0.73$ for both datasets). Additionally, the curves reveal that the generalized model outperforms the specialized model 1 on the test set of $Dataset_1$ and performs similarly compared to the specialized model 2 on the test set of $Dataset_2$ (with a difference of about 0.01 at each step of the Ji threshold). It can be also observed that the precision of the models drop when the Ji threshold increases, suggesting that it is difficult to align the ground-truth parasite contours with the automatic ones, especially for $Dataset_1$. Nevertheless, as demonstrated in Figure 3.15, the segmentation results obtained by the generalized model on both datasets (bottom rows of subfigures 3.15a and 3.15b) seem to be very close to the ground-truth in shape and number, regardless of the parasite species (top rows of subfigures 3.15a and 3.15b). The early drop in model precision on $Dataset_1$ might be partly explained by the small size of the parasites areas. Indeed, in contrast to $Dataset_2$, most parasites in $Dataset_1$ are characterized by a small surface which may make it challenging to match the area between automatic and ground-truth parasites. Furthermore, due to the small area of the parasites, we believe that it is difficult to obtain a high level of area matching even between two human experts' segmentations.

Figure 3.16 displays the results of the generalized model for parasite detection and segmentation on both test sets of $Dataset_1$ and $Dataset_2$ with respect to TP, FP and TN. Indeed, the figure shows the high capacity of

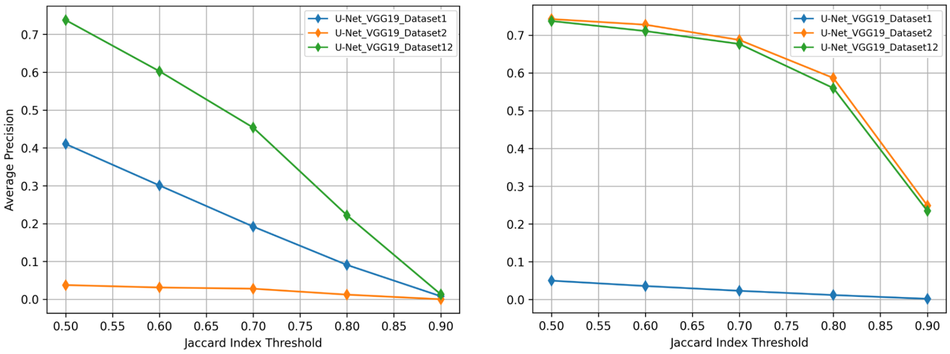
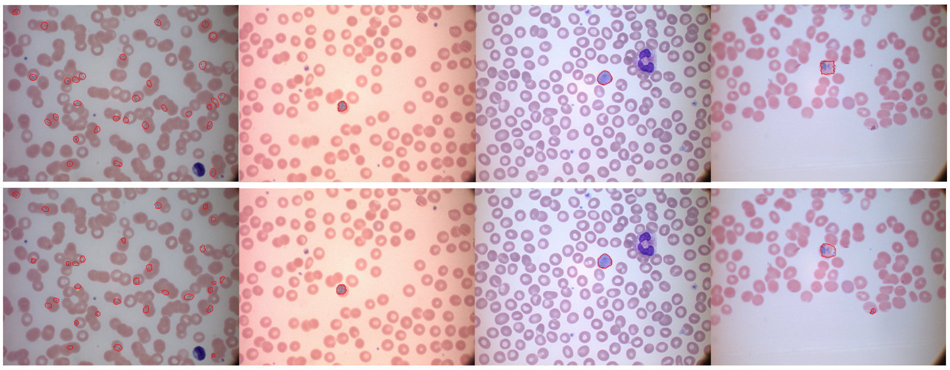


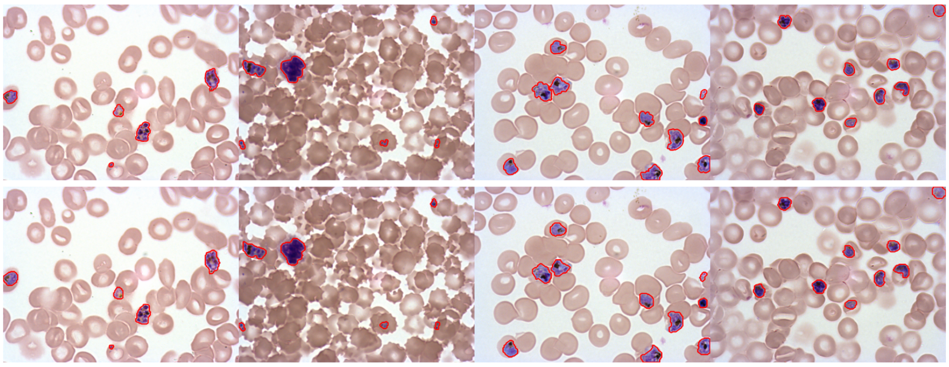
FIGURE 3.14: Average Precision curves of specialized and generalized models obtained over the test set from $Dataset_1$ (left side) and the test set from $Dataset_2$ (right side).

the generalized model to segment parasites on images having varied interclass and intra-class features with equivalent performance. To be specific, the model detected 166/179 parasites and 426/485 parasites on $Dataset_1$ and $Dataset_2$ respectively (TP parasites). It can be observed that the generalized model detects the parasites located within the same cell and close to each other as one parasite, which may lead to an increase in FN. Figure 3.17 shows an example for each dataset (see delineated parasites in yellow on the groud-truth on the left and the corresponding parasites detected by the model on the right). The model might be improved if there were more examples of such cases. We can also observe on Figure 3.16 that the amount of false parasites (FP) is relatively low. After visually analyzing the results for both datasets, we noticed that FP parasites are not necessarily false. In fact, as indicated in Figure 3.17 for both datasets some parasites (delineated in orange) have been missed in the ground-truth. We note that these parasites are usually small and located at the boundaries of the images.

Except for our proposed U-Net_VGG19 generalized model, we also trained two other state-of-the-art architectures, the original U-Net and the U-Net_ResNet34 following the same training / validation rules. Figure 3.18 illustrates the AP curves of the three architectures computed on a test set corresponding to the merging of the test sets from $Dataset_1$ and



(A) $Dataset_1$ test set: from left to right image examples representing Falciparum, Malariae, Ovale and Vivax species.



(B) $Dataset_2$ test set: image examples of Falciparum species.

FIGURE 3.15: Parasite segmentation results of *Generalized model*. For each dataset, ground-truth on top row, automatic segmentation on bottom row.

$Dataset_2$. The curves show that U-Net_VGG19 has approximately 10% at Ji=0.5 than the other two architectures.

We also compared the parasite segmentation results of our generalized model with the results of Abbas and Dijkstra's segmentation method [3] (owners of $Dataset_2$ [2]). Since the authors trained their segmentation model on $Dataset_2$ using a cross-validation method it would be impossible to make a fair comparison on this dataset. For this reason, we compared the two segmentation models on a sample of 13 images from the CDC (Control Disease Center), and the authors have shared the associated automated segmentation results as figures in the supplementary material document (referenced in [3]). To be specific, we used the screenshots of these images

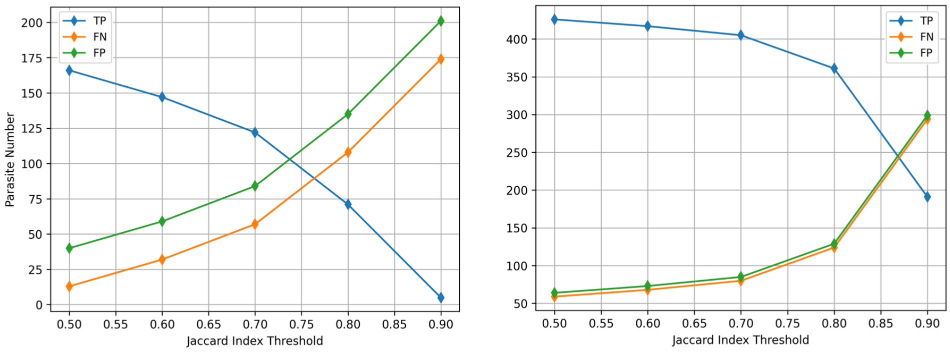
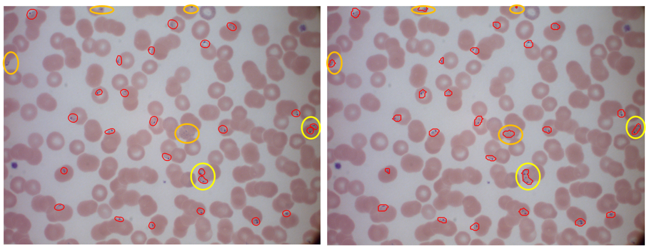
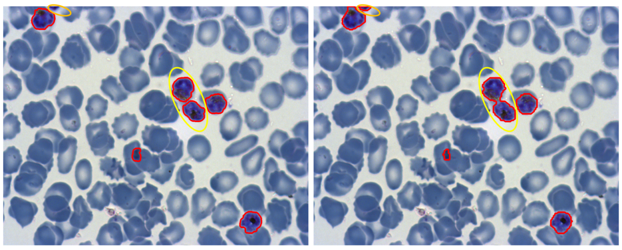


FIGURE 3.16: From left to right, parasite detection metrics of the *Generalized segmentation model* on test sets from *Dataset*$_1$ on the left and *Dataset*$_2$ on the right.



(A) *Dataset*$_1$



(B) *Dataset*$_2$

FIGURE 3.17: For each dataset, ground-truth segmentation on the left vs automatic segmentation on the right: illustration of some missed parasites shown in orange and merged ones shown in yellow.

to test our model and generate associated segmentation results. Figure 3.19 displays comparative examples of the respective results for the first three images shared in their document (more comparative results can be found in the supplementary file). In addition to the fact that we tested our model on

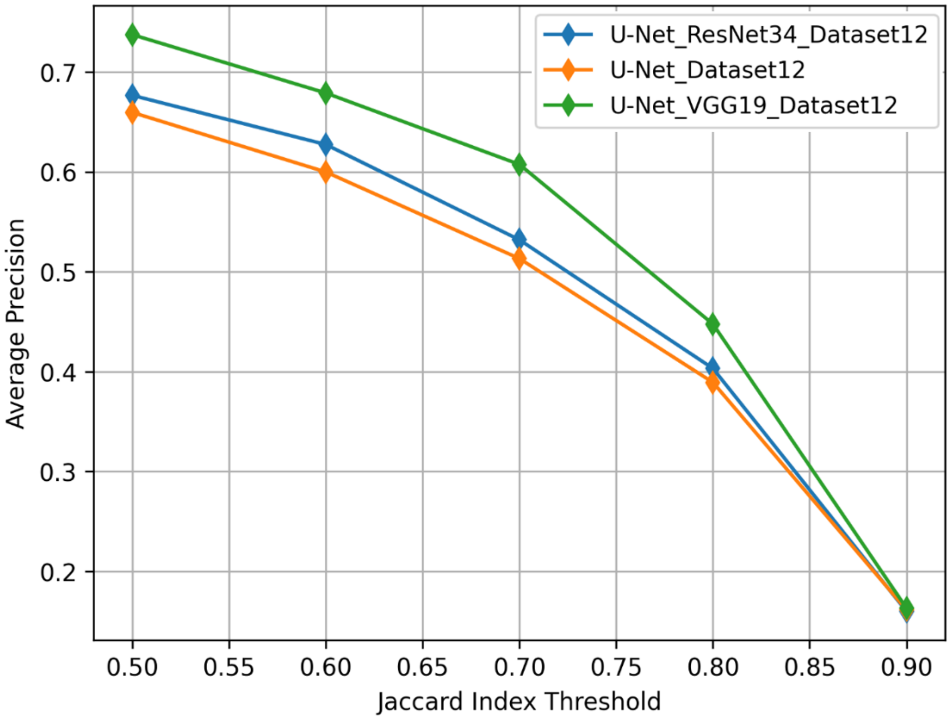


FIGURE 3.18: Performance comparison between generalized parasite segmentation models based on 3 deep architectures on a merged test set from $Dataset_1$ and $Dataset_2$.

screen-shot images, it is worth mentioning that we did not apply any cell diameter-based image rescaling or surface threshold-based object suppression, as was done in [3]. One may note that our method has finer contour segmentation. However, some parasites are missed, probably partly due to the degraded quality of the screen-shot images.

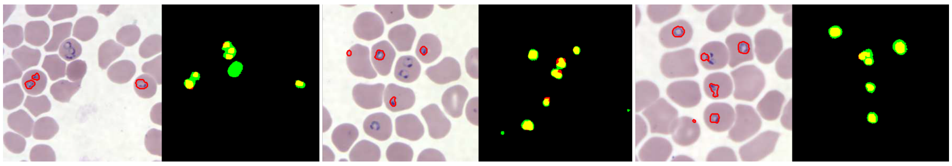


FIGURE 3.19: Parasite segmentation results from our generalized model (colored images) vs Abbas and Dijkstra method [3] (quadtone masks from CDC blind data; yellow pixels: TP, green pixels: FP, red pixels: FN, black: background).

3.4.3 Results and evaluation of the generalized classification module

3.4.3.1 Data preparation

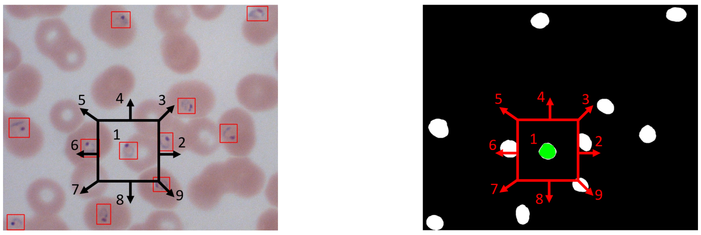
TABLE 3.2: Quantity of parasite images using a cascade of augmentations with species balancing.

Parasite species	Global image $dataset_1$			Crops LPTS augmentation*			Crops LPTS + standard augmentation	
	Train	Validation	Test	Train	Validation	Test	Train	Validation
Falciparum	74	14	16	7363	1333	1524	11289	2047
Malariae	27	5	5	297	45	45	10692	2075
Ovale	21	4	4	261	32	36	11000	2052
Vivax	32	4	4	441	54	81	11178	2044

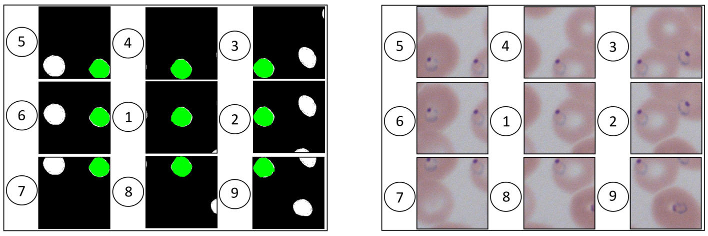
*11512 generated crops using Eq. 3.7 over the 1437 parasites of $Dataset_1$ (close to maximal value 12933 provided by Eq. 3.8).

We mainly used $Dataset_1$ because it is the only dataset that contains 4 species. To be specific, we used the same training / validation / test sets as the segmentation part (see $Dataset_1$ in Figure 3.13). The first three columns of Table 3.2 show the data split details of the species. It can be observed that the size of the training set is too small to train our classification architecture. For this reason, we proposed a first step for parasite data augmentation using our LPTS (Local Parasite Texture Scanning) approach.

Local Parasite Texture Scanning (LPTS) – Figure 3.20 shows an example of the output when our method is applied to a parasite. To be more specific, the parasite is first bounded on the original image (red box in subfigure 3.20a) according to the associated segmentation in the binary mask (subfigure 3.20b). Then the bounding box is enlarged to a 224 x 224 square window in the center of the parasite and shifted in 8 directions (black square in subfigure 3.20a and 3.20b) in order to generate 9 binary masks crops (subfigure 3.20c) and associated original image crops (subfigure 3.20d) according to the method given in section 3.3.3. Though the method is based on image shifting, it allows to increase effectively the number of images and to diversify the configuration of parasites by locally changing their position / shape / number. The last three columns of Table 3.2 indicate the amount of parasites crops per species obtained after applying our approach to the original data of $Dataset_1$.



(A) Original image. (B) Ground-truth.



(C) Generated mask crops. (D) Original crops.

FIGURE 3.20: An output example of our LPTS parasite augmentation technique.

LPTS data species standard augmentation and balancing – The last two columns of the Table 3.2 indicate the quantity of parasite crops for each species in the training / validation sets after a last augmentation and balancing step. To be exact, we increased the first species (Falciparum) by adding parasite image crops from the training set of $Dataset_2$ used in the segmentation part (see $Dataset_2$ in Figure 3.13). This can improve the generalization potential of the classifier. For the other species, we performed several spatial and pixel-based transformations on the crops while maintaining species sizes balance. To be specific, we combined horizontal and vertical flips with Gaussian noise, histogram equalization, contrast, brightness, median blurring, CLAHE, gray level, random rain, fog, shadow and snow transformations.

3.4.3.2 Classification performance analysis

To assess the performance of the parasite species classifier, we employed standard scoring metrics including accuracy, precision, recall and confusion matrix [101]. Our classification architecture has been trained and validated on the "Crops LPTS + standard augmentation" training / validation sets (see last 2 columns of Table 3.2) and tested on the test set (see 6^{th} column of Table 3.2). We highlight that among the five commonly tested optimizers (Adam, RMSprop, SGD, Adadelta, Adagrad), the best performance was observed for RMSprop. To determine whether our LPTS technique helps to improve the performance of the classifier, we trained our architecture on a reduced training set, keeping only the window crops in the center of the parasite (see position 1 in subfigures 3.20c and 3.20d) and removing the crops in the remaining 8 positions. We also compared the performance of our classification architecture with the state-of-the-art by experimenting Xception [16], ResNet50 [44] and VGG16/19 [89] architectures. For this purpose, we employed a transfer learning strategy to fine-tune the associated models pre-trained on the ImageNet dataset [23]. The architectures have been tuned by keeping the default sequence of convolution layer blocks, adding a global average pooling and an output layer of four units (corresponding to the four classes of the parasite). The fine-tuning step has been performed during the training by freezing the ImageNet weights for the sequence of convolutional layer blocks and optimizing the weights of the newly added dense layer. These adapted versions have been trained, validated on the same "Crops LPTS + standard augmentation" training / validation sets as our architecture and tested on the same test set. Among the four tested architectures, only VGG16 achieves competitive performance compared to our architecture.

Table 3.3 displays the performance obtained by our architecture on two training strategies (with and without LPTS crops) and VGG16 on the validation set. It can be noticed that our architecture achieved the best performance for the four species. Furthermore, the table shows that our LPTS technique enabled our architecture and VGG16 to outperform the degraded version (Light-Net*), improving considerably their discrimination ability

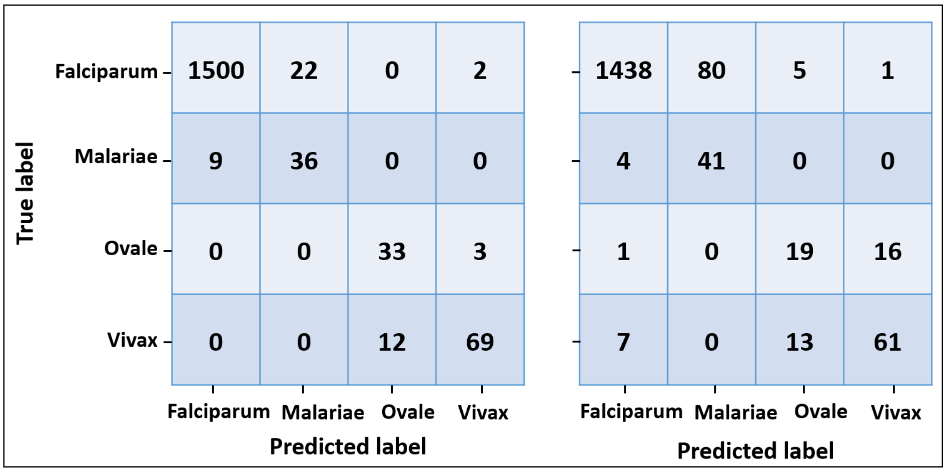
TABLE 3.3: Classifiers performance comparison on the Crops LPTS + standard augmentation validation set using standard metrics

Parasite	Accuracy			Precision			Recall		
species	VGG16	Light-Net*	Light-Net	VGG16	Light-Net*	Light-Net	VGG16	Light-Net*	Light-Net
Falciparum	0.96	0.65	**0.98**	0.95	0.89	**0.98**	0.96	0.65	**0.98**
Malariae	**0.99**	0.94	**0.99**	0.96	0.85	**0.98**	**0.99**	0.94	**0.99**
Ovale	0.72	0.47	**0.83**	0.77	0.53	**0.84**	0.72	0.47	**0.83**
Vivax	0.77	0.72	**0.84**	0.76	0.57	**0.83**	0.77	0.72	**0.84**

VGG16 applied from [89].*Trained on reduced training set (without 8 positions of LPTS crops).

for different species, especially for Falciparum and Ovale species.

Figure 3.21 displays the confusion matrices obtained on the test set by our architecture and the VGG16 [89] when trained on the “Crops LPTS + standard augmentation” training set. The matrices reveal that our architecture outperforms the VGG16 on three species (Falciparum, Ovale and Vivax), while the VGG16 is slightly better at identifying the Malariae species. The average accuracy on the four species is 88% for our architecture vs 78% for VGG16.



True label	Falciparum	Malariae	Ovale	Vivax
Falciparum	1500	22	0	2
Malariae	9	36	0	0
Ovale	0	0	33	3
Vivax	0	0	12	69

True label	Falciparum	Malariae	Ovale	Vivax
Falciparum	1438	80	5	1
Malariae	4	41	0	0
Ovale	1	0	19	16
Vivax	7	0	13	61

FIGURE 3.21: From left to right, confusion matrices obtained on the parasite species test set from $Dataset_1$ by Light-Net and VGG16 [89] respectively.

A final test of the two architectures has been done on a test set from $Dataset_2$ consisting of 5889 Falciparum parasite crops. These crops have

TABLE 3.4: Classification results obtained by the two architectures on a test set of 5889 Falciparum crops from $Dataset_2$.

Classifier	True Positive	False Negative		
	Falciparum	Malariae	Ovale	Vivax
VGG16 [89]	5798	32	19	40
Light-Net	**5864**	5	1	19

been generated from the validation / test sets used in the segmentation part (see $Dataset_2$ in Figure 3.13). Table 3.4 displays the classification results. We observe that the discrimination ability of our architecture is confirmed as it detects 99.5% of crops as Falciparum species and outperforms the VGG16. Additionally, the results show that our architecture has a high generalization potential on $Dataset_1$ and $Dataset_2$.

3.4.4 Global system results and evaluation

3.4.4.1 Patient level evaluation

TABLE 3.5: Results of the global diagnosis system at the patient level.

Dataset		Total of images	Total of parasites		Species*	
			GT**	Predicted	GT**	Predicted
$Dataset_1$ [59]	Patient 1	2	55	56	F	F
	Patient 2	2	33	42	F	F
	Patient 3	2	2	2	M	M
	Patient 4	2	2	2	O	O
	Patient 5	2	3	5	V	V
$Dataset_2$ [2]	Patient 1	8	24	29	F	F
	Patient 2	6	28	27	F	F
	Patient 3	22	48	44	F	F
	Patient 4	8	29	33	F	F
	Patient 5	6	19	24	F	F

*F: Falciparum, M: Malariae, O: Ovale, V: Vivax. **GT: Ground-Truth.

We remind that we used the same test sets from $Dataset_1$ and $Dataset_2$ in the segmentation and classification modules, so we can use these test sets to evaluate the global system (combined segmentation and classification) at the patient level (see Figure 3.8). Table 3.5 displays examples of the results obtained by our global system on 10 patients from $Dataset_1$ and $Dataset_2$. One can note that all the patients have been diagnosed with the correct species. In fact, this result is confirmed for all patients in both datasets. It can also be observed that the predicted number of parasites is close to the ground-truth. All these results clearly indicate that merging inter-class data represents a promising direction of investigation to develop powerful generalized systems based on deep learning.

3.4.4.2 Performance analysis on other datasets with partial ground-truth

TABLE 3.6: Global system performance in term of parasite detection and species classification.

Test dataset	Infected patients	Total of images	Parasite detection rate	Predicted patients	Ground-truth species	Species classification
$Dataset_3$ [21]	5	331	0.99	5/5	-	-
$Dataset_4$ [82]	151	13779	-	-	Falciparum	0.63
$Dataset_5$ [96]	-	654	0.99	-	Falciparum	0.84
$Dataset_6$ [57]	-	1328	0.85	-	Vivax	0.31

This section presents our global system results on the 4 test datasets ($Dataset_{3,4,5,6}$) in Table 3.1 by its segmentation and classification modules. Since these datasets do not provide a segmentation ground-truth of the parasite, we assessed our system by calculating the parasite detection rate. It corresponds to the ratio of the number of images in which the segmentation module detected at least one parasite (i.e. image detected as infected) divided by the total number of images ($TP/(TP+FN)$). We precise that all images of these 4 datasets are labeled as infected, which implies that all of them should contain at least one parasite. The results are summarized in the Table 3.6. Notably, the module is not suitable for segmenting $Dataset_4$, as it contains individual cell images. The table demonstrates that our segmentation module has successfully detected at least one parasite in each image, with high detection rates (over 99%) in $Dataset_3$, $Dataset_5$ and an acceptable one (85%) in $Dataset_6$. Figure 3.22 shows some qualitative results of the segmentation obtained by our module on the datasets.

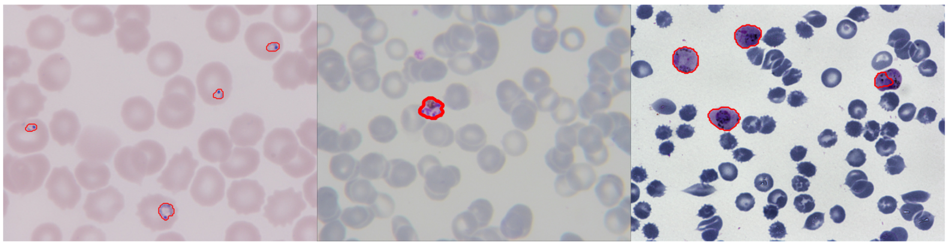


FIGURE 3.22: Parasite segmentation results of our system over samples of $Dataset_3$, $Dataset_5$ and $Dataset_6$ (from left to right).

For the classification part, our system selected the largest parasite detected by the segmentation module from each image and classified its species, then following a major voting on the images of the same patient and the system output the final predicted species. The Major voting is performed only in $Dataset_4$ because the images are organized by patient, while in $Dataset_5$ and $Dataset_6$, where no information is provided about the patient, the system outputs the predicted species of the largest parasite detected in each image. In $Dataset_3$ it is impossible to assess our classification module because no information on species is provided. The results are summarized in the Table 3.6. The table reveals that for $Dataset_4$, the classification module has a reduced discrimination capacity for P. Falciparum. This is due to the particular form of images considered (individual cell images rather than LPTS crops). Nevertheless, the module succeeded in identifying the Falciparum species in $Dataset_5$, reaching 84%. We believe in this data set, the accuracy of our system for species identification at the patient level may be higher, as classification decisions will be based on a major voting. Finally, we note that there is a decrease of performance in $Dataset_6$ which involves Vivax species. This result was expected because our generalized classification module has been trained on mixed data for Falciparum species (data from $Dataset_1$ and $Dataset_2$) but not for the other species due to lack of data. We believe that mixing data for Malariae, Ovale and Vivax species may further improve the generalization potential of our classification module.

3.5 Conclusion

We have proposed a framework for diagnosing *Malaria* infection in humans using microscopic images of thin blood smears. To be specific, this framework aims to segment the Plasmodium parasite and to differentiate between its species in four dominant classes: P. Falciparum, P. Ovale, P. Malaria and P. Vivax. For this purpose, two deep learning architectures based on CNN, called U-Net_VGG19 and Light-Net, have been performed for the segmentation and classification steps of parasites, respectively. Both architectures have been trained with the generalization of the resulting models in mind, especially on inter-class data. By combining two public datasets [59, 2] and a parasite data augmentation technique called LPTS, our training strategy demonstrated its efficiency in improving the performance of the global system. Indeed, the resulting generalized segmentation model performed well for detecting parasites (accuracy between 85% and 99%) in a couple of public datasets [21, 96, 57]. However, more interclass data for parasite species could further optimize the classification module of the global system.

Chapter 4

Granularity guided deep learning for precise *Cryptosporidium* parasite image analysis

4.1 Introduction

Cryptosporidiosis is a diarrheal disease caused by the *Cryptosporidium* parasite, which may live in the intestine of several mammals. Currently 40 species were discribed in this gender. Among these last, C. parvum has been a primary issue of concern because it appears in the majority of human and livestock infections. Moreover, this species presents zoonotic properties and can therefore be transmitted from humans to animals and vice versa [74, 33]. In humans, infected people might present with some symptoms, particularly diarrhea, abdominal pain and acute gastroenteritis [42, 37]. It is important to note that *Cryptosporidium* is the 5th causative agent of diarrhoeal disease in young children under 5 years of age. Furthermore, its presence has been implicated in increased mortality among young children [49]. Within livestock, cattle (especially calves) are considered to be the main reservoir of this parasite. In fact, over one million oocysts per gram of faeces can be excreted by an infected animal [7]. The VIDA (Veterinary Investigation Diagnosis Analysis) 2014 report indicated that *Cryptosporidium* was one of the four main causes of diarrhoea in calves during the period 2007 to 2011. This parasite also contributes to the death of small animals and it has a significant economic impact by reducing the rentability of farms [4]. Apart from labour expenses, the treatment of a sick calf costs a minimum of £34 [97].

It is therefore clear that controlling parasite colonisation would offer genuine benefits to the economy and would contribute to limiting the risk to health of humans. Nevertheless, achieving this goal is a challenge due to: i) the high level of resistance of the parasite to environmental condition, ii) the lack of effective medicines for infected animals as well as humans [11]. In such cases, developing a completely effective drug that targets this parasite would allow to prevent its reproduction on farms in the presence of infection. This lack of drugs is mainly attributed to the lack of automated and easily accessible tools to screen drug libraries. There are currently no highthroutput system to simplify the process of characterising new anti-parasitic compounds.

The researchers initiate their experiments in vitro through the employment of cell layer infections when they wish to evaluate the efficacy of drugs against pathogens. After a period of incubation (varying from 24 hours to 72 hours within case of *Cryptosporidium* studies) the infected cells are fixed. The parasites are then specifically identified on the microscope slide using fluorescent-labelled antibodies [55] (as shown in Fig 4.1). Microscopic observation will result in an assessment of the reduction of parasite infections and differentiate their life stages (from asexual to sexual life stages). Nevertheless, it is a error-prone task for biologists to count the amount of parasites. For example, in a microscopic image captured with a 63 magnification (an image representing a view of relatively a small region of the infected sample), the number of parasites can reach on average a hundred after 24 hours of infection. Consequently, such manual method of analysis does not allow for a transition into large-scale trials in which thousands of molecules are used to accelerate drug discovery.

Additionally, direct exploitation of the existing methods is not sufficient to develop a reliable tool for the analysis of *Cryptosporidium* images due to the particular visual aspect of this parasite. Indeed, as illustrated

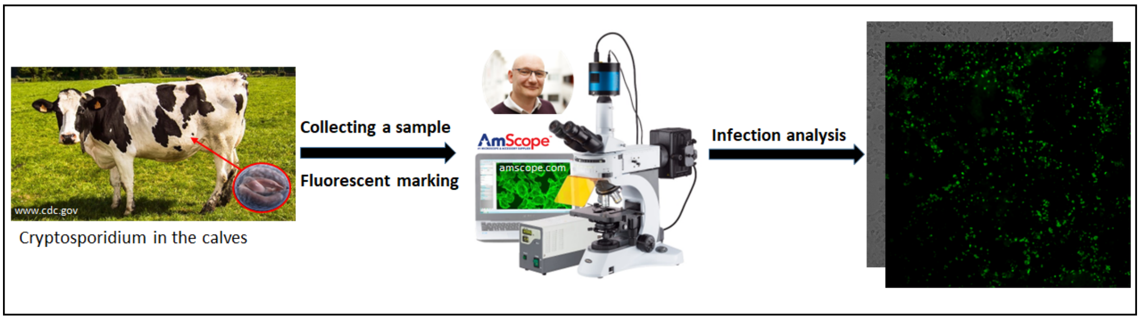


FIGURE 4.1: *Cryptosporidium* analysis by biologist expert.

in Fig 4.2, the parasites have irregular shapes with confused contours notably for stuck parasites and their surfaces as well as their fluorescence are highly varied. All these observations make the characterization of the parasite by existing deep architectures a real challenge.

To this end, we propose a coarse-to-fine segmentation approach to precisely segment a population of *Cryptosporidium parvum* parasites from a microscopic image by the first step. Then, a subsequent classifier with high discriminatory power is used to distinguish the life stages of the parasites among 4 asexual stages: oocyst, trophozoite, meront, and free form. To the best of our knowledge, this work is the first to address *Cryptosporidium parvum* analysis from microscopic images and offers a precision tool for assisting the biologists.

The main contributions of our work are summarized as follows:

- An original coarse-to-fine segmentation approach that permits to delineate precisely the contours of parasites whatever their shapes, their sizes and their situations in the image (grouped or isolated). To this end, the approach goes through three successive processing steps namely 1) a coarse segmentation step that outputs either individual parasites or grouped ones (i.e. they share common contours), 2) shape classification that permits to identify grouped parasites 3) contour refinement that permits to isolate each parasite from the

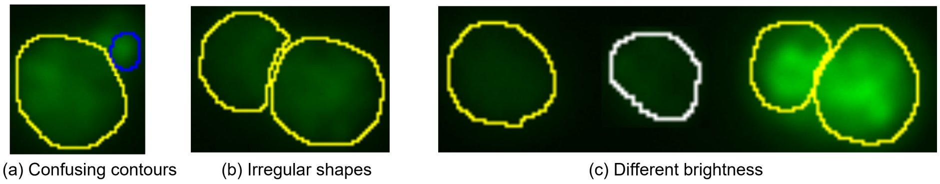

(a) Confusing contours
(b) Irregular shapes
(c) Different brightness

FIGURE 4.2: Huge challenges present in detecting parasite contours, observed from the labelings of experts. Different situation as follows: (a) Confusing contours, (b) Irregular shapes and (c) Different brightness

grouped ones. The approach allowed to significantly improve the performance in comparison with the state of the art ones.

- A classifier with high discriminative power for distinguishing the life stages of the parasites based only on the analysis of their visual appearance (reflected in the green channel of the microscopic image) and without the need of information related to the size and the number of nuclei which are required by the biologist. Indeed, the high performance obtained by the classifier on the experimented dataset shows that our deep learning architecture succeeded to characterize the parasite life stage from its visual appearance.
- A dataset of *Cryptosporidium parvum* parasite publicly available. The dataset is composed of 58 microscopic images acquired in our laboratory from infected HCT-8 cell lines (Declinai HCT). The images contain over 3,000 parasites which have been segmented manually and classified among four life stages by a biologist expert.

4.2 Exploited Transformer-based architectures

In this section, we mainly present the concept of Vision Transformer (ViT) and the related architectures applied to our research in this chapter. For other architectures of image classification and segmentation, we have already presented in the previous chapter.

In the last decade, deep learning-based networks have been broadly employed for image segmentation in the medical field and have demonstrated high performance. However, they suffer from a lack of understanding of long-range dependencies in images owing to the induction bias inherent among convolutional architectures. To address this limitation, we have discovered that a growing number of recent studies have been proposed to encode long-range dependencies based on Transformer architectures exploiting self-attention mechanism over the state-of-the-art [112, 114, 108, 107].

Dosovitskiy et al.[28] first proposed the concept of Vision Transformer (ViT), that is a pure transformer for direct application to image classification task. ViT follows the original transformer as much as possible in its design, as shown in Fig 4.3. The authors split the image into patches with fixed-size, added position embeddings and then fed these vector sequences to the Transformer encoder.

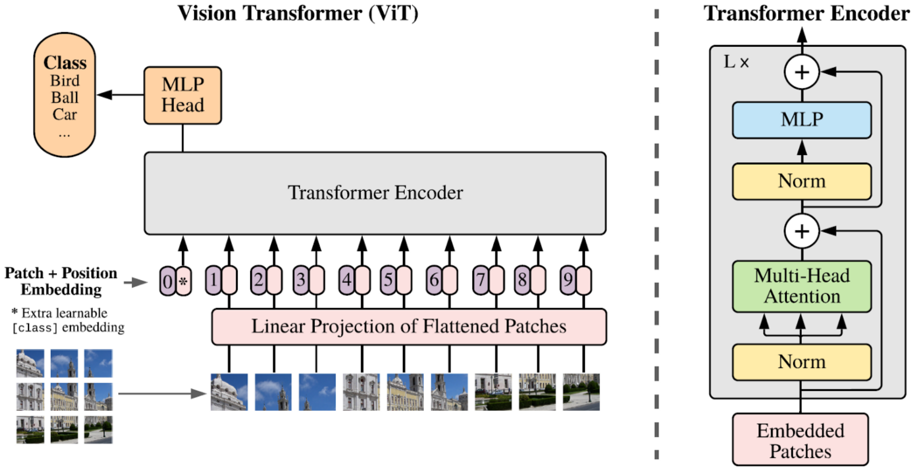


FIGURE 4.3: The Vision Transformer (ViT) framework [28].

Chen et al. [12] proposed the first medical image segmentation architecture named TransUNet. Unlike [28], the authors did not split the image into patches, but used a series of convolutional operations for feature extraction. As displayed in Fig 4.4, TransUNet adopts a hybrid CNN-Transformer architecture, in which feature maps are first generated by

CNN and then patch embedding is performed based on these feature maps. The authors believe that Transformer can be leveraged as a powerful encoder for image segmentation, combining with U-Net for the enhancement of finer details by recovering local spatial information.

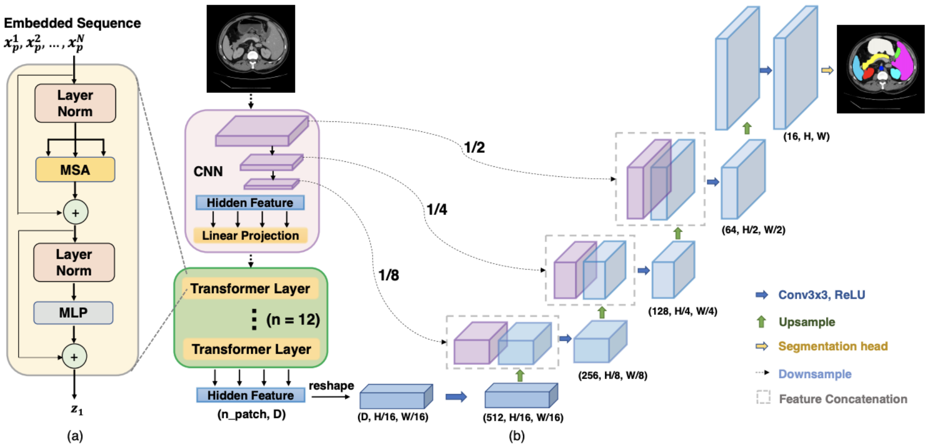


FIGURE 4.4: Architecture TransUNet[12].

Cao et al. [9] proposed a pure Unet-like Transformer architecture for image segmentation, named Swin-Unet (as can be seen in Fig 4.5). Tokenized patches of the images are sent to the architecture via a skip connection for local and global feature learning. To be specific, hierarchical Swin Transformer with shifted windows is used for extracting image features as encoder. Also, the decoder part employs a Swin Transformer-based design, which includes a patch expanding layer to conduct upsampling operation to recover the spatial information of the feature maps.

4.3 Methodology

Fig 4.6 illustrates the overview of our framework, which consists of a coarse-to-fine segmentation module, a loop control module and a module of parasite life stage classifier. Our coarse-to-fine segmentation method applies deep learning models by three consecutive steps of analysis: 1) a coarse segmentation model which inputs a microscopic image and outputs a mask with delimited contours that delineate individual or grouped

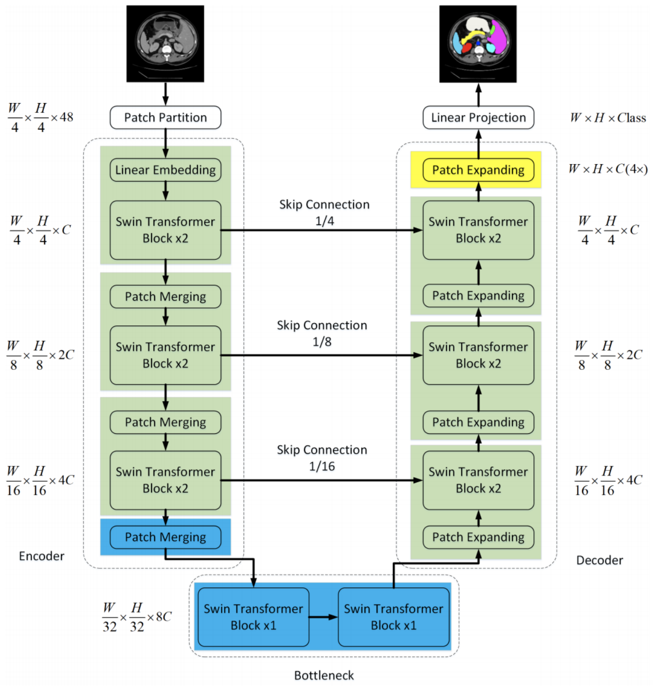


FIGURE 4.5: Architecture Swin-Unet [9].

parasites. 2) a shape classification model which allows to input cropped images of parasites that have been detected by previous step and outputs their classes (grouped or individual). 3) a fine segmentation model which inputs only the parasite crop identified as grouped in previous step and isolates them by delimiting each of their contours. Subsequently, a loop control module helps to detect outlier parasites. In addition, a classifier with high discriminatory power is used to distinguish the life stages of the parasites.

4.3.1 Coarse-to-fine segmentation module

Coarse-to-fine segmentation model includes the following steps: coarse segmentation and parasite cropping, parasite shape classification after cropping and fine segmentation.

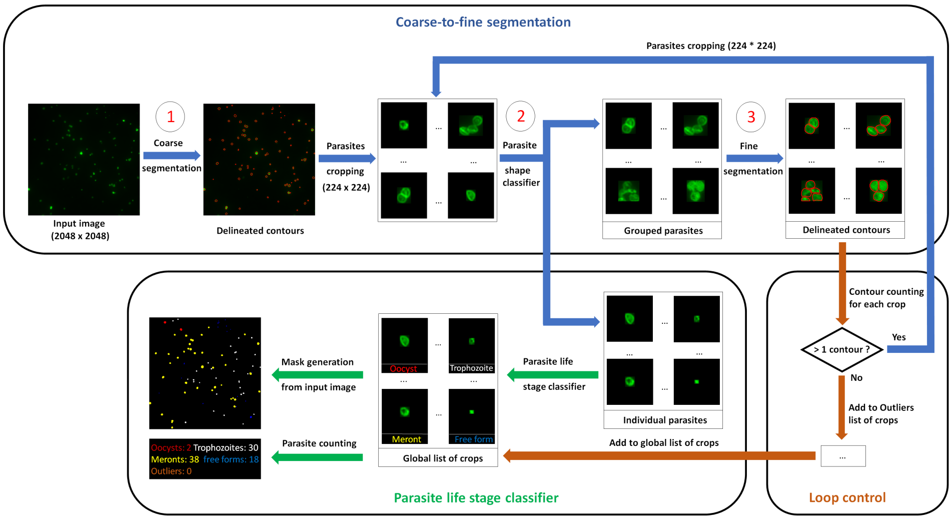


FIGURE 4.6: An overview of our method, which consists of a coarse-to-fine segmentation module, a loop control module and a module of parasite life stage classifier.

4.3.1.1 Coarse segmentation

The segmentation architecture is designed to be identical for both coarse and fine segmentation step. The major difference lies in the image training set employed by the different models. To be more specific, the coarse segmentation model learns from a collection of images comprising a group of segmented parasites (the whole microscopic image), while the fine segmentation model learns from a collection of image crops consisting of a small group of segmented parasites with common outlines.

Fig 4.7 represents our implemented architecture TransUNet [12]. It is made up of two different processing modules, namely the encoder and the decoder. The encoder conducts feature extraction through a series of convolutional layers (16, 32, 64, 128, 256, 512), each of which is followed by a max pooling operation. Afterwards, the linear projection operation converts the feature map from the convolution operation into matrices for passage over a series of transformer blocks. After rescaling the output of the transformer block, the hidden features are decoded by cascading several upsampling blocks to come up with the final segmentation mask,

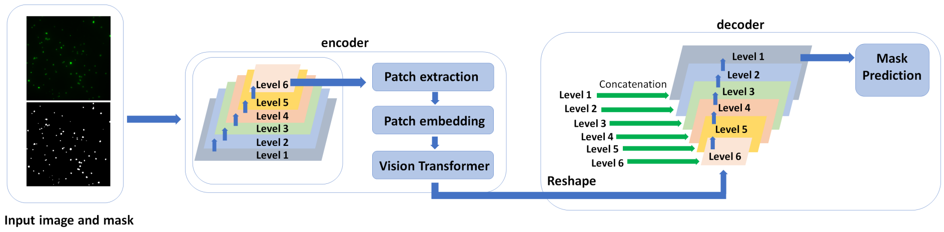


FIGURE 4.7: Our implemented segmentation architecture TransUNet.

each containing a 2× upsampling operator, a 3×3 convolutional layer, as well as a ReLU layer.

We propose to employ a hybrid loss function to train the network, consisting of a Jaccard loss and a Cross-entropy loss, so that both local and global features are optimized:

$$Loss_{Jac} = 1 - \frac{\sum y_{ij}t_{ij}}{\sum y_{ij} + \sum t_{ij} - \sum y_{ij}t_{ij}} \tag{4.1}$$

where $y_{ij} \in [0,1]$ and $t_{ij} \in [0,1]$ denote the output and target of each pixel at position (i, j). We adopt the Cross-entropy loss as an auxiliary loss in order to allow optimizing the local region of the image. It is formulated as:

$$Loss_{CE} = -\sum_{i}^{C} t_i log(s_i) \tag{4.2}$$

where t_i and s_i are the ground truth and the CNN score for each class i in C. Based on the two above losses, our loss function is as follows:

$$Loss_{our} = \alpha Loss_{Jac} + (1-\alpha) Loss_{CE} \tag{4.3}$$

where $\alpha \in$ [0, 1]. In order to minimize the loss function, the Adam [48] optimizer is employed.

4.3.1.2 Parasite shape classifier after cropping

After coarse segmentation, we crop the detected parasites and regenerate them as small images with a resolution of 224 x 224.

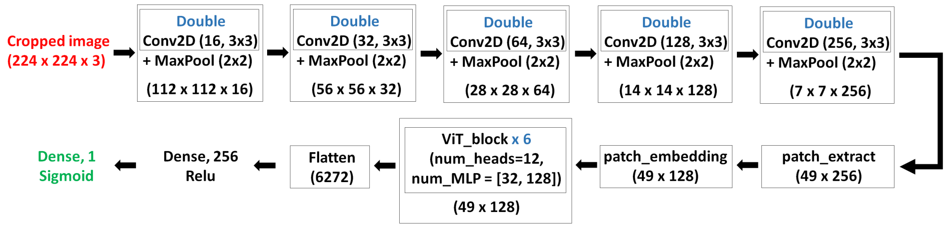


FIGURE 4.8: Our proposed classification architecture TransCNN.

By analyzing the cropped parasite images, we find that most of the parasites exist in isolation, however there are some parasites in the form of parasite groups, as some parasites are interconnected. Even certain parasite groups contain parasites of different life stages (as displayed in Fig 4.2(a)). Therefore, it is necessary to separate each parasite in the parasite group by a more refined segmentation before classifying the parasite life stages.

Compared to isolated parasites, we discover that parasite groups are usually more irregular in shape rather than close to circular. Therefore, we designed a custom architecture that incorporates CNN and transformer blocks for the construction of a parasite shape classifier. The inspiration for the architecture comes from the encoder of the segmentation architecture introduced in previous section.

As illustrated in Fig 4.8, the architecture first consists of 10 convolutional layer blocks (5 successive double blocks), each of which is followed by a ReLU as activation function and a max pooling operation. After transforming the feature maps into matrices, the network is then connected to the six ViT (Vision Transformer) blocks. A flatten operation follows ViT blocks, which is followed by a dense layer of 256 units with a ReLU activation function. Eventually, an output dense layer of 1 unit is applied to gain the prediction probability for the shape of the parasite using a sigmoid activation function.

4.3.1.3 Fine segmentation

Considering the relatively regular shape of the parasites, we consider contour refinement of the parasite groups so as to isolate each parasite from

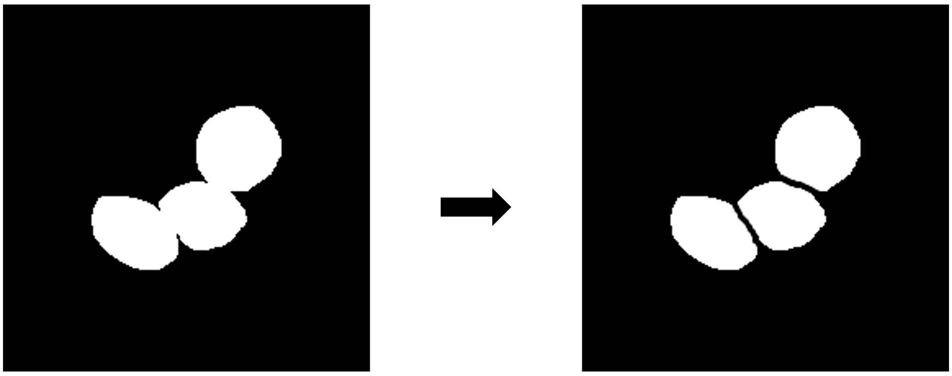


FIGURE 4.9: Parasite fine segmentation for grouped parasites.

the grouped parasites. Given that the shape of the parasites is known to be approximately circular, we select grouped parasite masks and perform parasite segmentation by manually filling black pixels for the contour-connected parasites (as displayed in Fig 4.9). Afterwards, we apply the architecture as in coarse segmentation to train a deep learning network for fine segmentation using the parasite masks. The only difference is that the input size of the network is changed to 224 x 224.

Subsequently, according to the fine segmentation of the mask, we obtain the result of the fine segmentation on the corresponding pixels of the original image.

4.3.2 Loop control module

After fine segmentation, we add a loop control module for detecting the number of parasite contours in each image. In the normal case, the connected parasites in the parasite group are separated and the module should detect more than 1 parasite at this point. However, if only 1 parasite contour is detected, it means that the previous coarse-to-fine segmentation module cannot handle the current parasite image. In this circumstance, we consider the parasite image as an outlier, and all subsequent outliers will be detected by the expert.

4.3.3 Parasite life stage classification module

To classify all detected individual parasites in the framework, we train the deep learning network using approximately the same architecture as the parasite shape classifier. By comparison, the classification of parasite life stages requires the output to be changed to 4 instead of 1, and the corresponding activation function to be changed to softmax. In addition, the loss function is changed to categorical cross entropy instead of binary cross entropy.

4.4 Experimental study

4.4.1 Data preparation

Our dataset for the analysis *Cryptosporidium parvum* comprises 58 microscopic images. The images have been obtained by infection of HCT-8 cell lines with *Cryptosporidium parvum* in the laboratory (24 hours of infection). During the infection, the cells were fixed with methanol. Then the parasites are killed and stained with specific fluorescent markers. The parasite was detected with FITC-labeled Sporoglow antibodies marketed by Waterborne Inc. The stained parasites can be visible in the green channel of microscope. In the life cycle of *Cryptosporidium parvum*, we studied the asexual stage, containing four species: oocyst, trophozoite, meront, and free form[35, 94]. All images have been manually annotated by a parasitologist to delimit the outline of each parasite (as shown in Fig 4.10).

Table 4.1 indicates the size of the training/validations sets and image resolutions exploited during each step of our coarse-to-fine segmentation training process. A few data augmentation techniques have been implemented for the coarse segmentation only. The 3 scenarios are as follows.

Scenario1: Spatial transformation based data augmentation. In this case, we performed rotation, vertical flip and horizontal flip on both the original images and the associated masks.

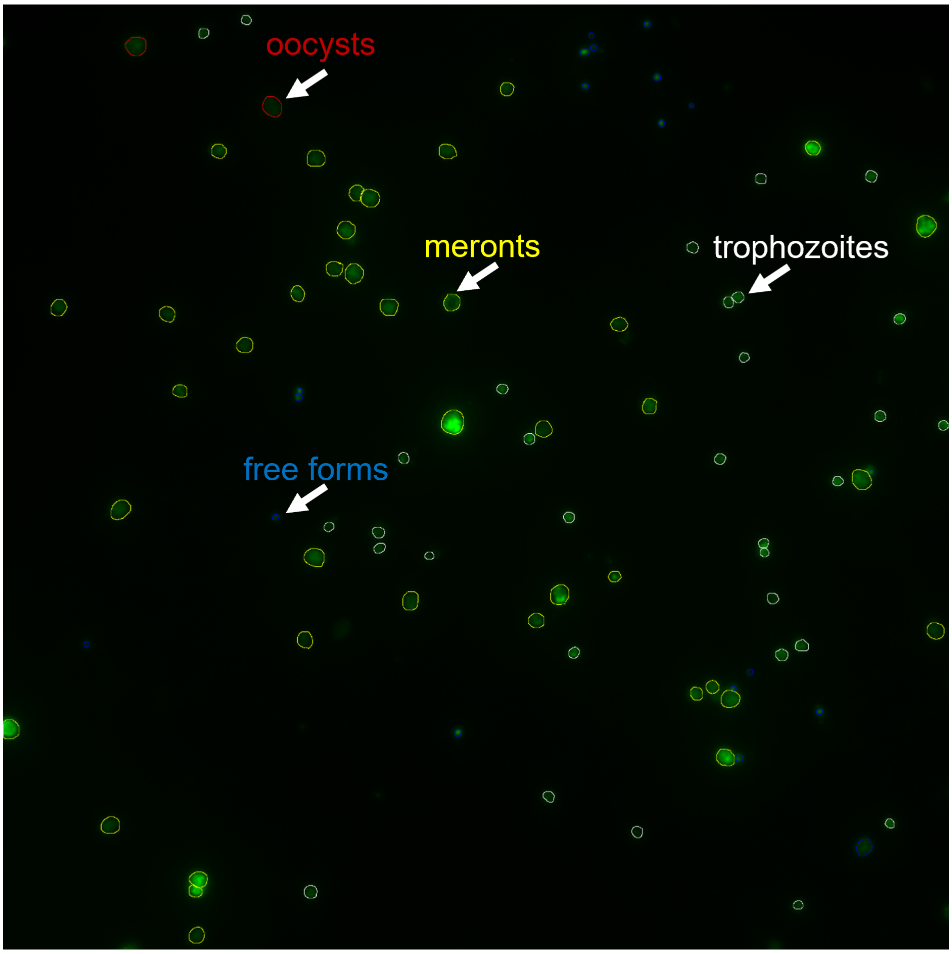


FIGURE 4.10: Example of parasite image annotated by biologist using 4 different colors.

Scenario2: Texture transformation based data augmentation. We performed gaussian blur, contrast enhancement and brightness enhancement for the original images only and maintain the associated mask constant.

Scenario3: Spatial and texture transformation based data augmentation.

TABLE 4.1: Detailed information of dataset used in different steps

Steps	Training set	Validation set	Image resolution
Coarse segmentation	51 images (2510 parasites)	7 images (589 parasites)	1024 x 1024
Shape classification	244 crops (122 per class)	60 crops (30 per class)	224 x 224
Fine segmentation	122 crops (263 parasites)	30 crops (65 parasites)	224 x 224
Species classification	293 parasites per class	111 parasites per class	224 x 224

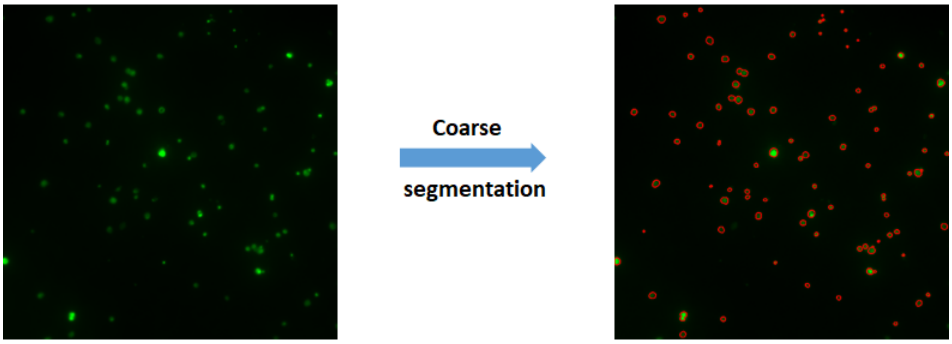


FIGURE 4.11: The coarse segmentation step.

4.4.2 Experimental results

4.4.2.1 Coarse segmentation

As illustrated in Fig 4.11, the coarse segmentation step that outputs the contours of each parasite, whether it exists in individual or grouped forms.

We compared our approach with five other approaches, including three deep learning approaches (U-Net_VGG19, U-Net_ResNet34 and Swin-Unet[9]) and two conventional approaches (watershed algorithm[71] and HSV color filter). For deep learning approaches, the same scenarios as our approach were applied. First, all models were trained based on original data (without our proposed scenarios 1, 2, and 3). Fig 4.12 illustrates the AP (Average Precision) curves for six architectures computed on the validation set of the coarse segmentation (data augmentation under scenario1 in the training, with the best performance for all methods relative to the other augmentation scenarios and the base dataset). AP is defined as AP=TP / (TP + FP + FN), where TP is true positive, FP is false positive, FN is false negative in term of parasite object compared to the ground truth.

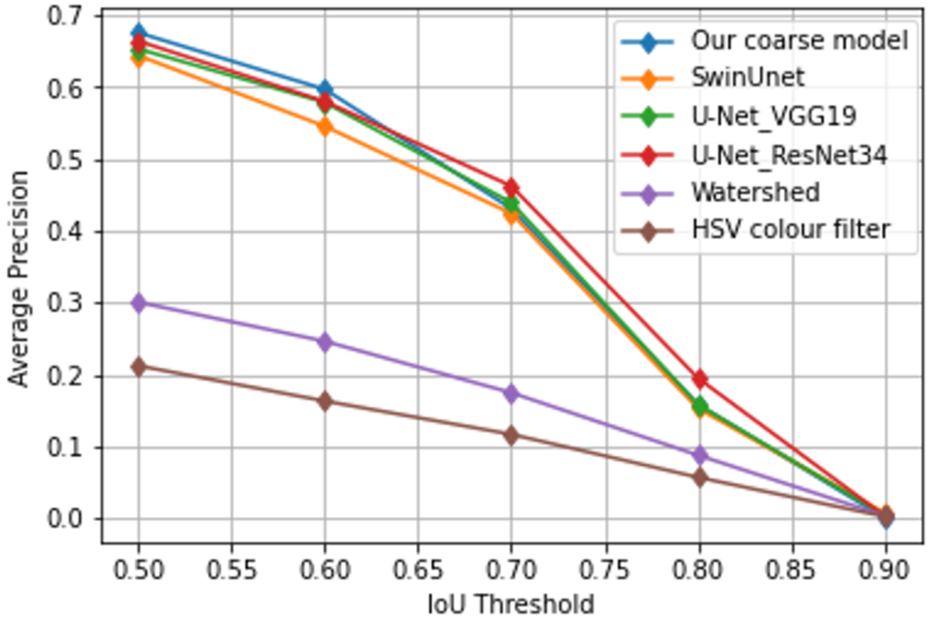


FIGURE 4.12: AP curves on the validation set of the coarse segmentation.

They are computed with the IoU (Intersection over Union) metric ranging from a threshold of 0.5 to 1. The curves demonstrate that our model is outperforming the others. In fact, our coarse model achieved 67% AP with an IoU threshold of 0.5. It can also be noted that the performance of conventional segmentation approaches (watershed and color filter) are very low compared to CNN, which clearly shows that use of simple features, such as color, is not sufficient to solve the issue of parasite segmentation.

Fig 4.13 shows the results of the our model in term of TP, FP and TN for parasite coarse segmentation on the validation set. It could be observed that the number of the FP and FN is relatively close to each other. When the IoU threshold is more than 0.7, both FP and FN surpass the number of the TP.

4.4.2.2 Parasite shape classification

The shape classification step that permits to identify grouped parasites (as displayed in Fig 4.14).

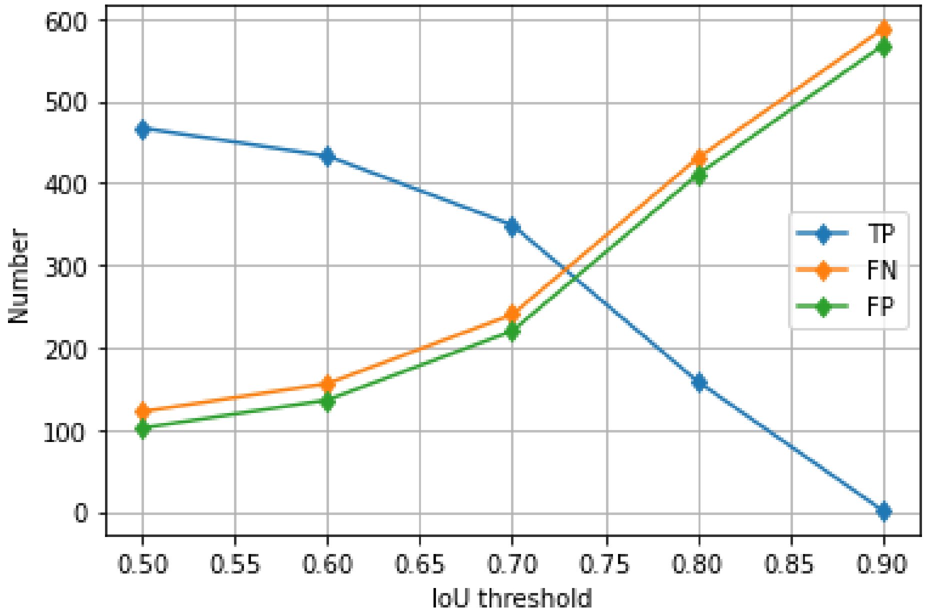


FIGURE 4.13: Parasite detection metrics of the our parasite coarse segmentation model on the validation set.

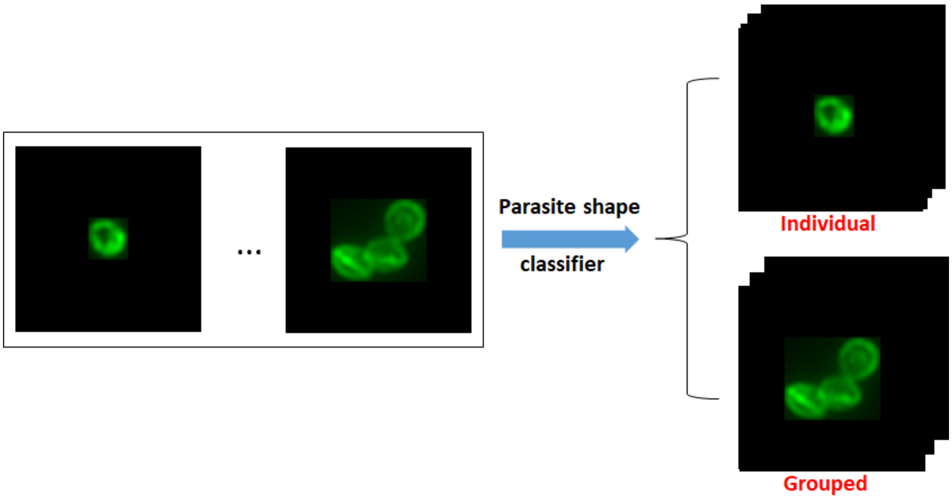


FIGURE 4.14: The shape classification step.

TABLE 4.2: Comparison of the performance and the number of parameters between our classification architecture and some CNN architectures

Model	Accuracy	Precision	Recall	Parameters number
VGG16	1	1	1	134M
ResNet50	0.98	0.98	0.98	23M
Simple CNN (6 convolution layers)	0.96	0.96	0.96	25M
Our model	1	1	1	7M

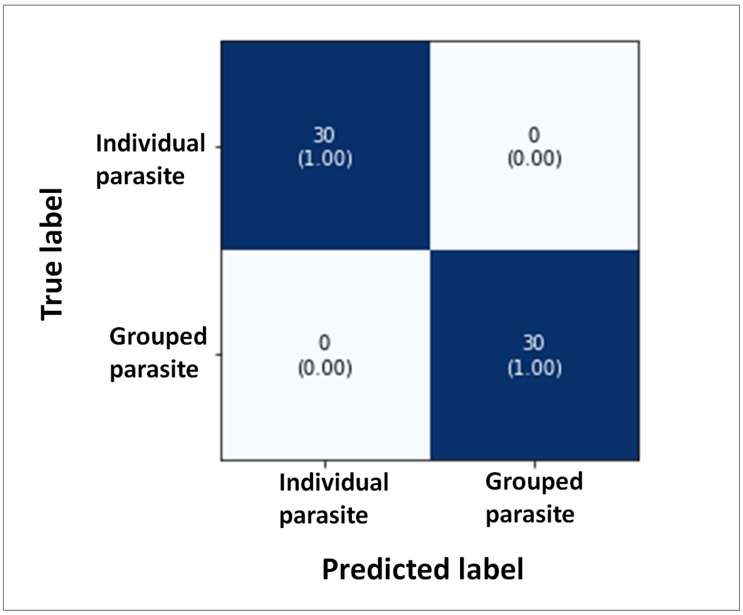


FIGURE 4.15: Confusion matrices obtained by our model on the parasite shape validation set.

We compared our model (CNN combined with ViT blocks) with several CNN classifiers from the state of the art. Table 4.2 displays the performance of the different models. Our model successfully classified all the validation data, which is identical to VGG16. Nevertheless, as indicated in the table, the model VGG16 is 19 times larger than our model (134M vs. 7M).

The confusion matrix on the validation set illustrates that our model has performed perfectly on the classification of the parasite shape. From the Fig 4.15, our model succeeds to classify all the validation data.

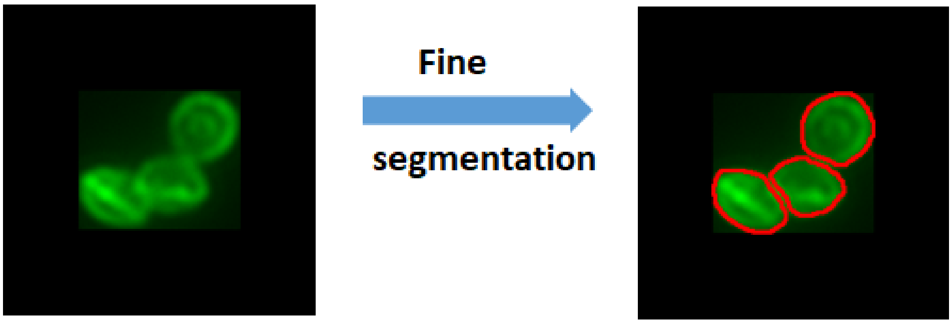


FIGURE 4.16: The fine segmentation step.

4.4.2.3 Fine segmentation

Fig 4.16 shows the fine segmentation step that permits to isolate each parasite from the grouped ones.

Similar to the coarse segmentation, we compared our model with three deep learning approaches. We followed also the same training protocol along with the evaluation metric (Average Precision). Fig 4.17 demonstrates the AP curves computed on the validation set for our fine segmentation model. The curves indicate that our model gives a better performance than the other models. It is notable that even with a high IoU threshold, the AP value of our model maintains the best AP value amongst the other models, which reveals a high accuracy in parasite separation.

It can be seen also in the metrics of our parasite fine segmentation model obtained over the validation set (illustrated in Fig 4.18). When the IoU Threshold is no more than 0.8, all the TP are successfully detected and there is no FN and FP. The number of FN and FP will both equal to 2 when the IoU threshold is 0.9.

4.4.2.4 Direct segmentation vs Coarse-to-fine segmentation

The purpose of this section is to illustrate the contribution of our coarse-to-fine segmentation method compared to the standard method that performs segmentation directly. Based on aboved steps, we reconstructed the final binary mask to evaluate our coarse-to-fine segmentation framework. First we performed post-processing on the mask obtained by coarse segmentation. Comparing with the ground truth, we found that the predicted mask

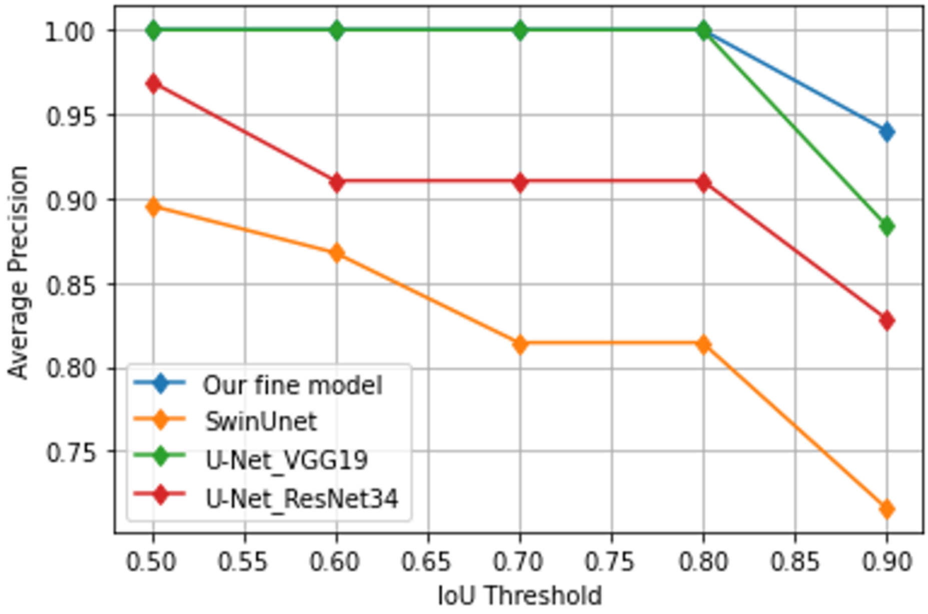


FIGURE 4.17: AP curves computed on the validation set of fine segmentation.

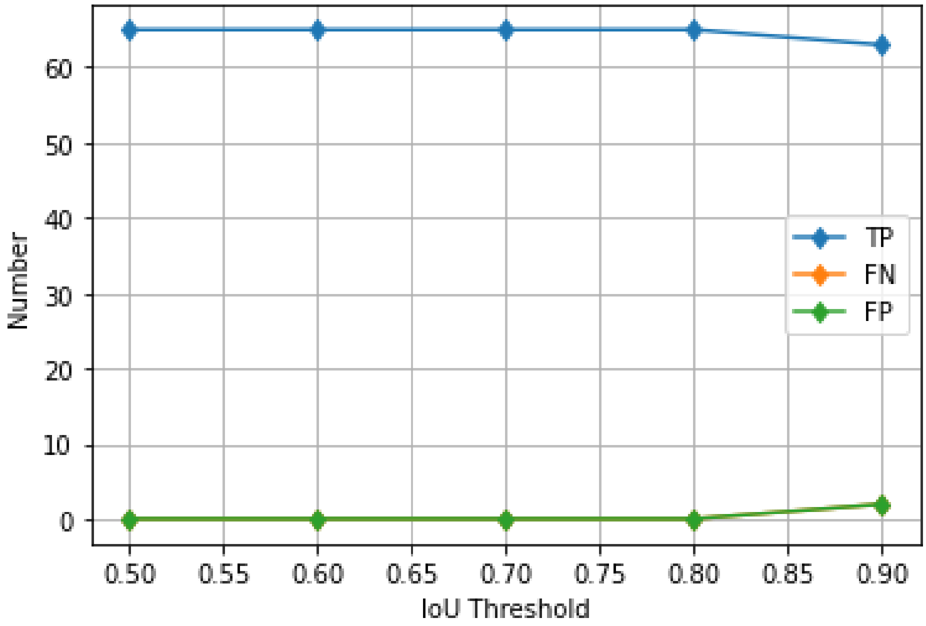


FIGURE 4.18: Parasite detection metrics of our parasite fine segmentation model obtained over the validation set.

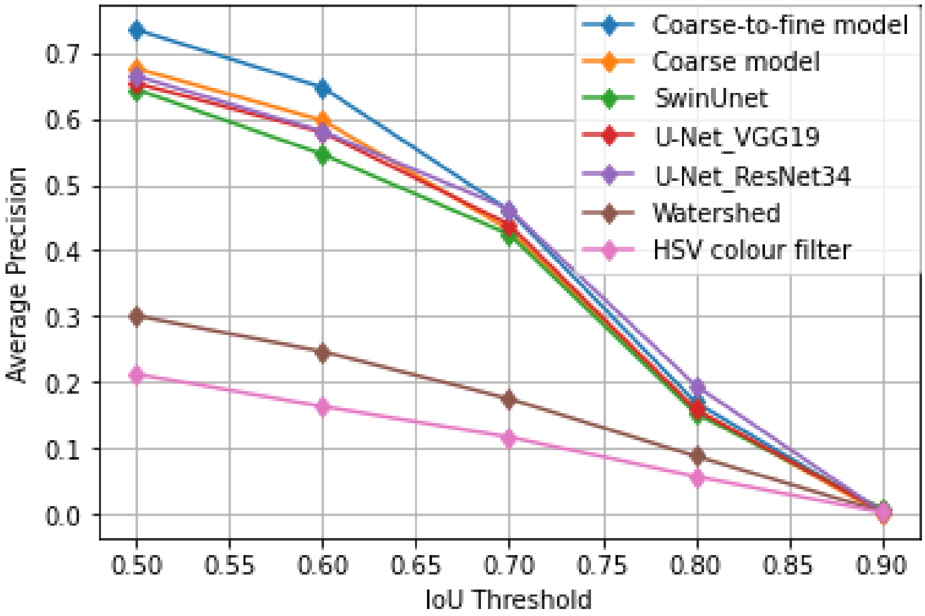


FIGURE 4.19: AP curves over the validation set of coarse-to-fine segmentation.

detected some parasites on the edges of the image, however, the ground truth has only a very small number of parasites. To reduce the number of false positives, we manually removed the parasites detected at the edges in the predicted mask. In addition, to prevent the detection of individual group of pixels with small areas that could be considered as parasites, we added a filter to detect them and then filled them with black. After all post-processing steps, we found 11 outliers from 7 validation images.

As can be seen from the Fig 4.19, with an IoU threshold of 0.5, our coarse-to-fine model achieved an AP of 73.45%, exceeding all comparative segmentation approaches. Furthermore, the model allows to obtain over 5% of AP in comparison with the coarse model. While in Fig 4.20, compared with the result in Fig 4.13, the number of both FN and FP has decreased to less than 100 (IoU threshold=0.5).

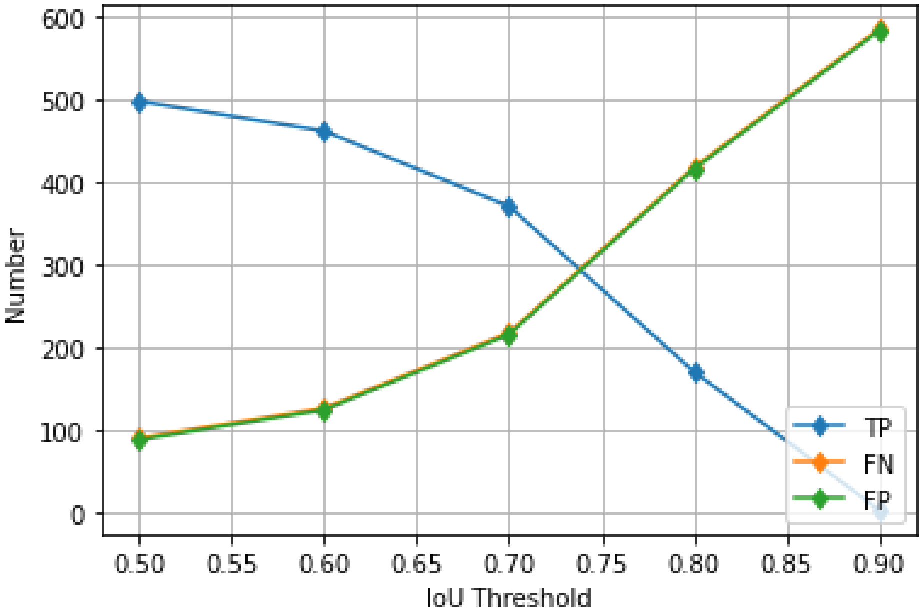


FIGURE 4.20: Parasite detection metrics of our parasite coarse-to-fine segmentation model on validation set.

4.4.2.5 Parasite life stage classification

Fig 4.21 shows the parasite life stage classification step that permits to distinguish the *Cryptosporidium parvum* parasite among 4 asexual stages: oocyst, trophozoite, meront, and free form.

Similar to the parasite shape classification experiments, we also compared our classification model with the 3 other CNN methods. As can be seen from the table 4.3, the accuracy of all models exceeds 92%, while our model outperforms other models, with 95.5% as accuracy.

The confusion matrix on the validation set illustrates that our model has well performed on all 4 classes (as displayed in Fig 4.22). For the oocyst and free form, the accuracy is much better than the other two parasite species, which is 97% and 99% respectively. While for the trophozoite and meront, both accuracy are 93%, which shows a little bit of a gap compared to the other two species. This is because a few images may be confused between them.

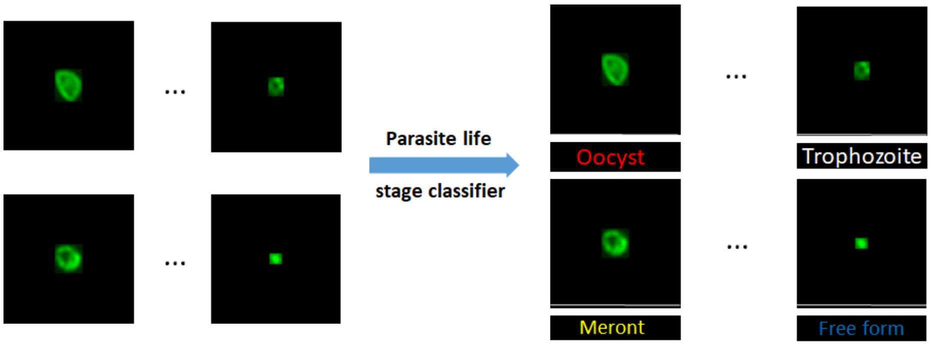


FIGURE 4.21: The parasite life stage classification step.

TABLE 4.3: Comparison of the performance between our classification model and 3 other CNN methods

Model	Accuracy	Precision	Recall
VGG16	0.946	0.947	0.946
ResNet50	0.923	0.923	0.923
Simple CNN (6 convolution layers)	0.953	0.953	0.953
Our model	0.955	0.955	0.955

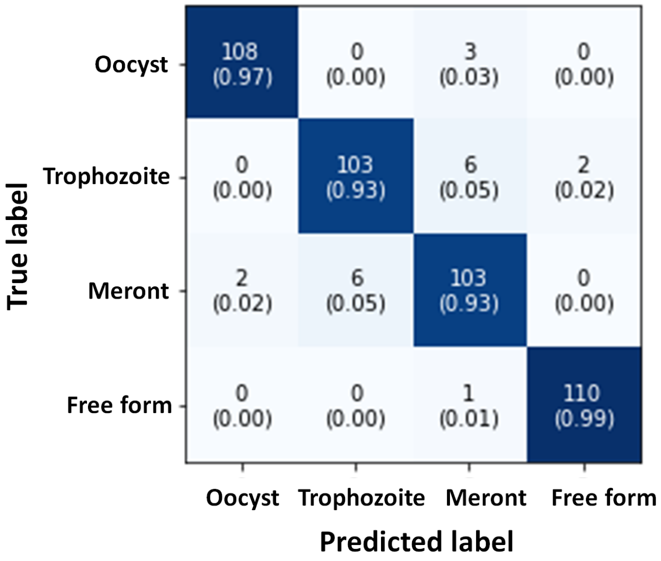


FIGURE 4.22: Confusion matrices obtained by our model on the parasite life stage validation set.

4.5 Conclusion

This chapter has presented a deep learning-based framework to automatically analyze and diagnose the *Cryptosporidium parvum* infection. The framework allows to segment precisely a group of parasites from microscopic images by employing a coarse-to-fine segmentation approach, followed by a subsequent classifier with high discriminatory power to distinguish the parasite life stages among 4 asexual stages: oocyst, trophozoite, meront, and free form. Our coarse-to-fine segmentation module has achieved an average precision of 73.45% on our validation dataset and permits to obtain over 5% of average precision in comparison with the direct segmentation. Additionally, for the parasite life stage classification module, our model also outperformed better than other state-of-the-art models, with 95.5% of accuracy.

Chapter 5

Conclusion and future work

5.1 Summary

The thesis focused on the development of new deep learning based methods for the analysis of two parasites namely *Plasmodium* and *Cryptosporidium* from microscopic images. In this sense, we proposed two major contributions.

For the *Plasmodium* parasite analysis, we have proposed a generic framework that permits to efficiently segment the parasite and to differentiate between four of its species: P. Falciparum, P. Ovale, P. Malaria and P. Vivax. Indeed, the global system reached 100% of diagnosis accuracy on a test set of images representing 17 infected patients. We believe that our segmentation module can be exploited as pre-segmentation tool by the scientific community to speed up the labeling of more data sources.

For the *Cryptosporidium* parasite analysis, we have proposed a framework that permits to precisely segment a population of parasites and classify their life stages into 4 asexual stages: oocyst, trophozoite, meront and free form. More specifically, our coarse-to-fine segmentation approach has permitted to achieve an AP (Average Precision) of 73.45% on our validation dataset, leading to an improvement of 5% compared to the top 1 state-of-the-art method. In addition, for the parasite life stage classification task, our model also outperformed better than other state-of-the-art models, with 95.5% of accuracy.

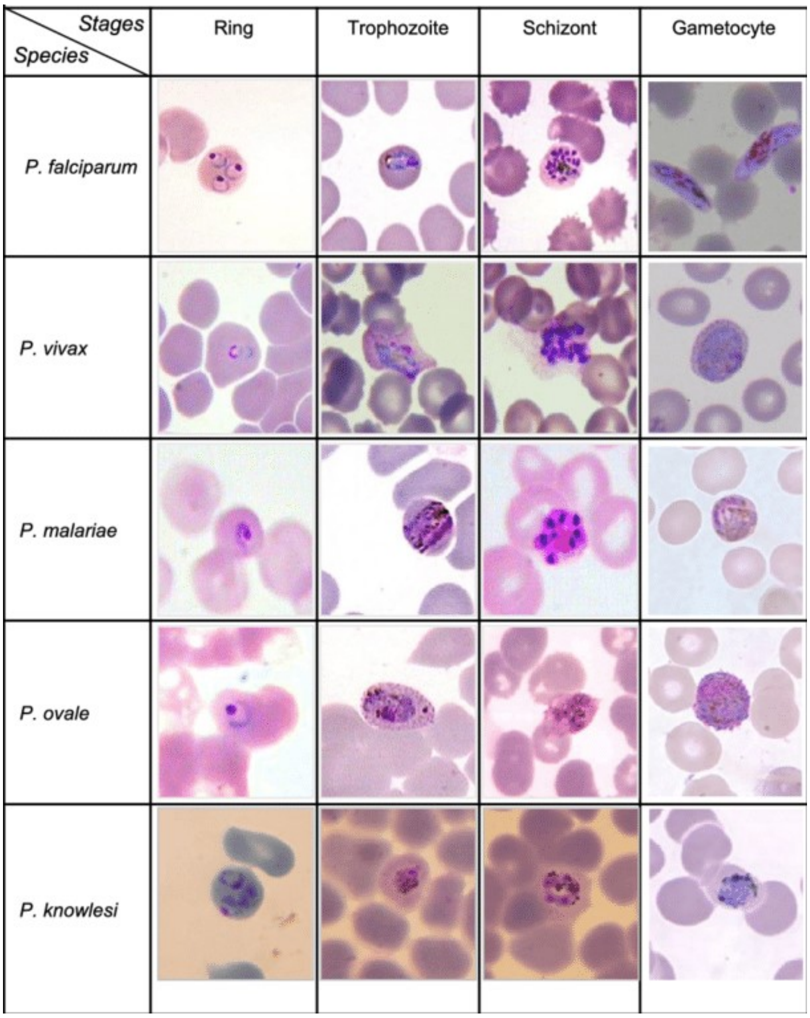


FIGURE 5.1: *Malaria* parasite species and its life stage (extracted from [81]).

5.2 Perspective for improvement of parasite analysis

5.2.1 Data enrichment and model optimization

For the *Plasmodium* parasite analysis, our study involved four morphologies of its species: P. Falciparum, P. Ovale, P. Malaria and P. Vivax. As illustrated in Fig 5.1, there exists a fifth species named P. Knowlesi which also cause infection within human and may threaten their lives.

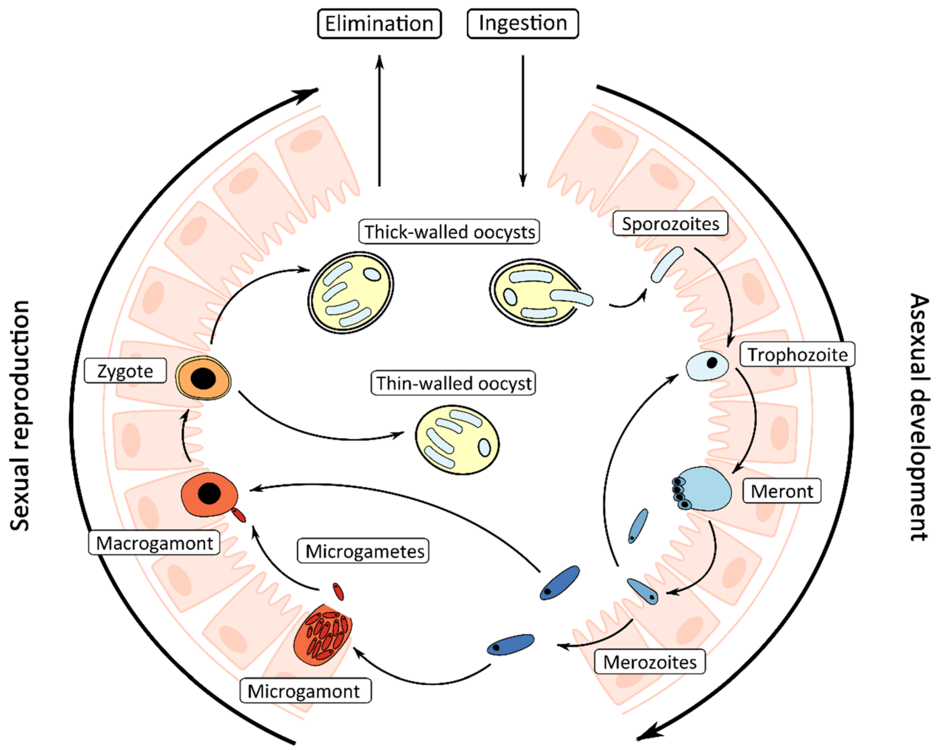


FIGURE 5.2: *Cryptosporidium* parasite life stages (extracted from [41]).

Future work could incorporate this species in the parasite analysis process. Moreover, life stages of the five species should also be integrated to have a finer and more precise analysis tool. Similarly, as for the *Malaria* framework, our *Cryptosporidium* framework can be optimized to offer a complete analysis tool by integrating the sexual life stages (microgamont and macrogamont as shown in Fig 5.2) in addition to the asexual ones (oocyst, trophozoite, meront, and free form). Nevertheless, it is clear that to develop such tools, more data are required with particularly finer annotations. In this sense, our proposed frameworks can be employed as pre-segmentation tools for the new data to accelerate the annotation process.

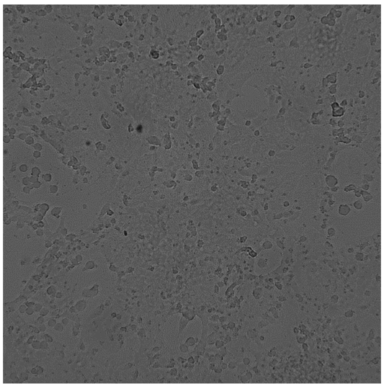


FIGURE 5.3: An example of BF microscopic image of *Cryptosporidium* parasite.

5.2.2 Exploitation of a new form of data for *Cryptosporidium*: bright-field microscopic images

Context and motivations – During this thesis, we mainly exploited the fluorescence microscopic image for the analysis of the *Cryptosporidium* parasite. Nevertheless, there exists a raw data named Bright-Field (BF) image (without fluorescent markers) that is often used in the study of *Cryptosporidium* parasites. Indeed, the raw format permits to study the entire life cycle (asexual and sexual) of the parasites without killing the experimented cell cultures following the usage of the fluorescent marker. However, as illustrated in Fig 5.3, it is difficult for biologists to identify parasites using this raw format. Therefore, artificial intelligence (AI) technology could be exploited to face issue [109].

Future work for analyzing this type of image can be carried out in two successive steps: The first one is to identify whether the BF image presents an infection or not. In the positive case, the second step is to provide a coarse quantification of that infection. For this second step and as shown in Fig 5.4, the deep learning model proceeds by subdividing into small square crops the input BF image and then detect possible presence of the parasite into each crop. By this way, the AI is able to calculate the rate of infection by the parasite.

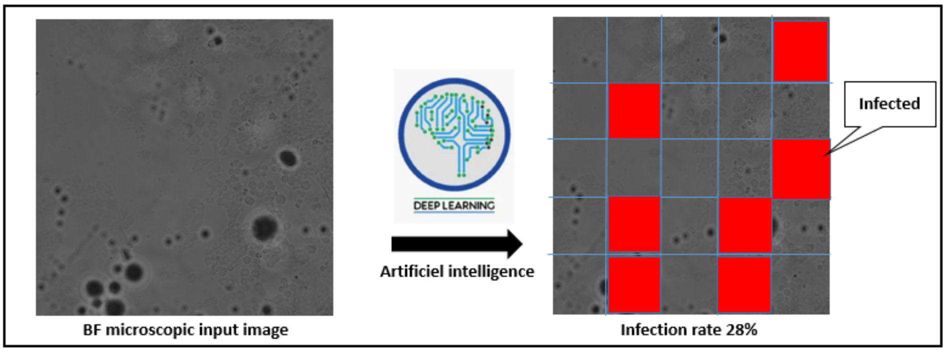


FIGURE 5.4: Deep learning method for parasite detection and quantification.

A potential AI framework for BF image analysis – Fig 5.5 illustrates a design of an AI framework for parasite detection and quantification. First, we use as input BF images associated with their ground truths as input examples to train Model 1 for a global infection detection. The output of Model 1 permits to decide whether a finer analysis is required or not. Then Model 2 gives as output a binary information to indicate if an individual crop contains or not some parasites. An infection rate could be finally calculated based on the previous output of Model 2 over all the crops of an image. The main difference between the two models is the data used to feed them during the training process (global BF image vs crop BF image). It is worth mentioning that the choice of quantifying the positive patches instead of segmenting individually the parasites permits to overcome the challenge of segmenting manually these parasites by the biologists which is required for preparing the training data.

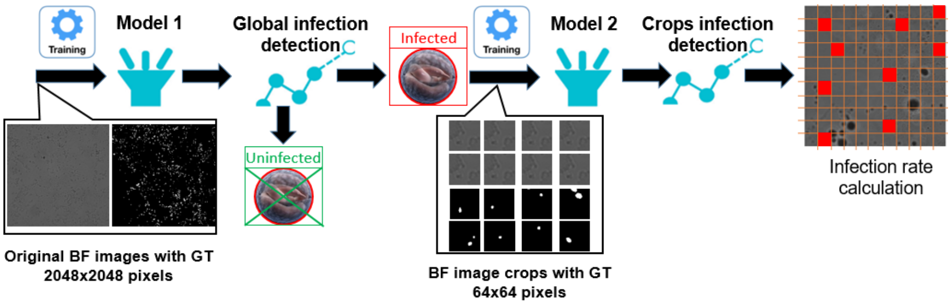


FIGURE 5.5: AI framework for BF image analysis.

Bibliography

[1] Naveed Abbas et al. "Machine aided malaria parasitemia detection in Giemsa-stained thin blood smears". In: *Neural Computing and Applications* 29.3 (2018), pp. 803–818.

[2] Syed Saiden Abbas and Tjeerd Dijkstra. *Malaria-Detection-2019*. Mendeley Data, V1, doi: 10.17632/5bf2kmwvfn.1, `https://data.mendeley.com/datasets/5bf2kmwvfn/1`. 2019.

[3] Syed Saiden Abbas and Tjeerd MH Dijkstra. "Detection and stage classification of Plasmodium falciparum from images of Giemsa stained thin blood films using random forest classifiers". In: *Diagnostic pathology* 15.1 (2020), pp. 1–11.

[4] APHA. "Veterinary Investigation Diagnosis Analysis (VIDA) report, 2014". In: (2014).

[5] J Wendi Bailey et al. "Guideline: the laboratory diagnosis of malaria". In: *British journal of haematology* 163.5 (2013), pp. 573–580.

[6] Tamara V Beyer et al. "Cryptosporidium parvum (Coccidia, Apicomplexa): some new ultrastructural observations on its endogenous development". In: *European Journal of Protistology* 36.2 (2000), pp. 151–159.

[7] Brian G Blackburn et al. "Cryptosporidiosis associated with ozonated apple cider". In: *Emerging Infectious Diseases* 12.4 (2006), p. 684.

[8] Alexander J Bones et al. "Past and future trends of Cryptosporidium in vitro research". In: *Experimental parasitology* 196 (2019), pp. 28–37.

[9] Hu Cao et al. "Swin-unet: Unet-like pure transformer for medical image segmentation". In: *arXiv preprint arXiv:2105.05537* (2021).

[10] Hector Caraballo and Kevin King. “Emergency department management of mosquito-borne illness: malaria, dengue, and West Nile virus.” In: *Emergency medicine practice* 16.5 (2014), pp. 1–23.

[11] Prinessa Chellan, Peter J Sadler, and Kirkwood M Land. “Recent developments in drug discovery against the protozoal parasites Cryptosporidium and Toxoplasma”. In: *Bioorganic & Medicinal Chemistry Letters* 27.7 (2017), pp. 1491–1501.

[12] Jieneng Chen et al. “Transunet: Transformers make strong encoders for medical image segmentation”. In: *arXiv preprint arXiv:2102.04306* (2021).

[13] Junying Chen et al. “A study of image segmentation algorithms combined with different image preprocessing methods for thyroid ultrasound images”. In: *International Conf. on Imaging Systems and Techniques*. IEEE. 2017, pp. 1–5.

[14] Xian-Ming Chen et al. “Cryptosporidium parvum is cytopathic for cultured human biliary epithelia via an apoptotic mechanism”. In: *Hepatology* 28.4 (1998), pp. 906–913.

[15] Dami Choi et al. “On empirical comparisons of optimizers for deep learning”. In: *arXiv preprint arXiv:1910.05446* (2019).

[16] François Chollet. “Xception: Deep learning with depthwise separable convolutions”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1251–1258.

[17] Patrick Ferdinand Christ et al. “Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields”. In: *International conference on medical image computing and computer-assisted intervention*. Springer. 2016, pp. 415–423.

[18] Alan F Cowman et al. “Malaria: biology and disease”. In: *Cell* 167.3 (2016), pp. 610–624.

[19] Dev Kumar Das et al. “Machine learning approach for automated screening of malaria parasite using light microscopic images”. In: *Micron* 45 (2013), pp. 97–106.

[20] Valeria F Del Coco, María A Córdoba, and Juan A Basualdo. “Cryptosporidium infection in calves from a rural area of Buenos Aires, Argentina”. In: *Veterinary parasitology* 158.1-2 (2008), pp. 31–35.

[21] Maria Delgado et al. *Dataset B: 331 digital images of MGG-stained blood smears from five malaria-infected patients*. Mendeley Data, V1, doi: 10.17632/2v6h4j48cx.1, `https://data.mendeley.com/datasets/2v6h4j48cx/1`. 2020.

[22] Maria Delgado-Ortet et al. “A Deep Learning Approach for Segmentation of Red Blood Cell Images and Malaria Detection”. In: *Entropy* 22.6 (2020), p. 657.

[23] Jia Deng et al. “Imagenet: A large-scale hierarchical image database”. In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee. 2009, pp. 248–255.

[24] Salam Shuleenda Devi, Rabul Hussain Laskar, and Shah Alam Sheikh. “Hybrid classifier based life cycle stages analysis for malaria-infected erythrocyte using thin blood smear images”. In: *Neural Computing and Applications* 29.8 (2018), pp. 217–235.

[25] Cecilia Di Ruberto, Andrea Loddo, and Giovanni Puglisi. “Blob Detection and Deep Learning for Leukemic Blood Image Analysis”. In: *Applied Sciences* 10.3 (2020), p. 1176.

[26] Cecilia Di Ruberto, Andrea Loddo, and Lorenzo Putzu. “A region proposal approach for cells detection and counting from microscopic blood images”. In: *International Conf. on Image Analysis and Processing*. Springer. 2019, pp. 47–58.

[27] Cecilia Di Ruberto, Andrea Loddo, and Lorenzo Putzu. “Detection of red and white blood cells from microscopic blood images using a region proposal approach”. In: *Computers in biology and medicine* 116 (2020), p. 103530.

[28] Alexey Dosovitskiy et al. “An image is worth 16x16 words: Transformers for image recognition at scale”. In: *arXiv preprint arXiv:2010.11929* (2020).

[29] DN Durrhelm et al. “Diagnostic disagreement–the lessons learnt from malaria diagnosis in Mpumalanga.” In: *South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde* 87.5 (1997), pp. 609–611.

[30] Owen Dyer. *African malaria deaths set to dwarf covid-19 fatalities as pandemic hits control efforts, WHO warns*. 2020.

[31] Tom Eelbode et al. “Optimization for medical image segmentation: theory and practice when evaluating with Dice score or Jaccard index”. In: *IEEE Transactions on Medical Imaging* 39.11 (2020), pp. 3679–3690.

[32] Andre Esteva et al. “Dermatologist-level classification of skin cancer with deep neural networks”. In: *nature* 542.7639 (2017), pp. 115–118.

[33] Yaoyu Feng, Una M Ryan, and Lihua Xiao. “Genetic diversity and population structure of Cryptosporidium”. In: *Trends in parasitology* 34.11 (2018), pp. 997–1011.

[34] Yuanyuan Feng et al. “Prevalence and genotypic identification of Cryptosporidium spp., Giardia duodenalis and Enterocytozoon bieneusi in pre-weaned dairy calves in Guangdong, China”. In: *Parasites & vectors* 12.1 (2019), pp. 1–9.

[35] Lisa J Funkhouser-Jones, Soumya Ravindran, and L David Sibley. “Defining stage-specific activity of potent new inhibitors of Cryptosporidium parvum growth in vitro”. In: *Mbio* 11.2 (2020), e00052–20.

[36] Jun Gao et al. “Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview”. In: *Mathematical Biosciences and Engineering* 16.6 (2019), pp. 6536–6561.

[37] Elisabetta Gerace, Vincenzo Di Marco Lo Presti, and Carmelo Biondo. "Cryptosporidium infection: epidemiology, pathogenesis, and differential diagnosis". In: *European Journal of Microbiology and Immunology* 9.4 (2019), pp. 119–123.

[38] Gopalakrishna Pillai Gopakumar et al. "Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner". In: *Journal of biophotonics* 11.3 (2018), e201700003.

[39] Marc Górriz et al. "Leishmaniasis parasite segmentation and classification using deep learning". In: *International Conf. on Articulated Motion and Deformable Objects*. Springer. 2018, pp. 53–62.

[40] Varun Gulshan et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs". In: *Jama* 316.22 (2016), pp. 2402–2410.

[41] Samantha Gunasekera et al. "Organoids and bioengineered intestinal models: potential solutions to the Cryptosporidium culturing dilemma". In: *Microorganisms* 8.5 (2020), p. 715.

[42] Kareem Hatam-Nahavandi et al. "Cryptosporidium infections in terrestrial ungulates with focus on livestock: a systematic review and meta-analysis". In: *Parasites & vectors* 12.1 (2019), pp. 1–23.

[43] Riddhi Hathiwala et al. "LED fluorescence microscopy: Novel method for malaria diagnosis compared with routine methods". In: *Journal of Infection and Public Health* 10.6 (2017), pp. 824–828.

[44] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[45] Zahoor Jan et al. "A review on automated diagnosis of malaria parasite in microscopic blood smears images". In: *Multimedia Tools and Applications* 77.8 (2018), pp. 9801–9826.

[46] Hao Jiang et al. "Geometry-aware cell detection with deep learning". In: *Msystems* 5.1 (2020), e00840–19.

[47] Justin Ker et al. “Deep learning applications in medical image analysis”. In: *Ieee Access* 6 (2017), pp. 9375–9389.

[48] Diederik P Kingma and Jimmy Ba. *Adam: A method for stochastic optimization. CoRR abs/1412.6980*. 2014.

[49] Karen L Kotloff et al. “Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study”. In: *The Lancet* 382.9888 (2013), pp. 209–222.

[50] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. “Imagenet classification with deep convolutional neural networks”. In: *Advances in neural information processing systems* 25 (2012).

[51] Rohit Kundu et al. “Pneumonia detection in chest X-ray images using an ensemble of deep learning models”. In: *Plos one* 16.9 (2021), e0256630.

[52] Albert Lalremruata et al. “Species and genotype diversity of Plasmodium in malaria patients from Gabon analysed by next generation sequencing”. In: *Malaria journal* 16.1 (2017), p. 398.

[53] Sen Li et al. “Parasitologist-level classification of apicomplexan parasites and host cell with deep cycle transfer learning (DCTL)”. In: *Bioinformatics* 36.16 (2020), pp. 4498–4505.

[54] Sen Li et al. “Transfer learning for toxoplasma gondii recognition”. In: *Msystems* 5.1 (2020), e00445–19.

[55] Jeff W Lichtman and José-Angel Conchello. “Fluorescence microscopy”. In: *Nature methods* 2.12 (2005), pp. 910–919.

[56] Geert Litjens et al. “A survey on deep learning in medical image analysis”. In: *Medical image analysis* 42 (2017), pp. 60–88.

[57] Vebjorn Ljosa, Katherine L Sokolnicki, and Anne E Carpenter. “Annotated high-throughput microscopy image sets for validation.” In: *Nature methods* 9.7 (2012), pp. 637–637.

[58] Andrea Loddo, Cecilia Di Ruberto, and Michel Kocher. “Recent advances of malaria parasites detection systems based on mathematical morphology”. In: *Sensors* 18.2 (2018), p. 513.

[59] Andrea Loddo et al. “MP-IDB: the malaria parasite image database for image processing and analysis”. In: *Sipaim–Miccai Biomedical Workshop*. Springer. 2018, pp. 57–65.

[60] Jonathan Long, Evan Shelhamer, and Trevor Darrell. “Fully convolutional networks for semantic segmentation”. In: *IEEE conf. on computer vision and pattern recognition*. 2015, pp. 3431–3440.

[61] Andrew A Lover et al. “Malaria elimination: time to target all species”. In: *The American journal of tropical medicine and hygiene* 99.1 (2018), pp. 17–23.

[62] Laurens van der Maaten and Geoffrey Hinton. “Visualizing Data using t-SNE”. In: *Journal of Machine Learning Research* 9.86 (2008), pp. 2579–2605. URL: `http://jmlr.org/papers/v9/vandermaaten08a.html`.

[63] Mila Yoselyn Pacompia Machaca et al. “Data Augmentation using Generative Adversarial Network for Gastrointestinal Parasite Microscopy Image Classification”. In: ().

[64] Emmanuel Maggiori et al. “Fully convolutional neural networks for remote sensing image classification”. In: *2016 IEEE international geoscience and remote sensing symposium (IGARSS)*. IEEE. 2016, pp. 5071–5074.

[65] Maitreya Maity et al. “Quantification of malaria parasitaemia using trainable semantic segmentation and capsnet”. In: *Pattern Recognition Letters* 138 (2020), pp. 88–94.

[66] Karan S Makhija, Samuel Maloney, and Robert Norton. “The utility of serial blood film testing for the diagnosis of malaria”. In: *Pathology* 47.1 (2015), pp. 68–70.

[67] Afoma Mbanefo and Nirbhay Kumar. “Evaluation of malaria diagnostic methods as a key for successful control and elimination programs”. In: *Tropical Medicine and Infectious Disease* 5.2 (2020), p. 102.

[68] Courosh Mehanian et al. “Computer-automated malaria diagnosis and quantitation using convolutional neural networks”. In: *IEEE International Conf. on Computer Vision Workshops*. 2017, pp. 116–125.

[69] Harshal Mehta et al. *Detection Of Malaria Parasite Using Deep Learning*. Tech. rep. EasyChair, 2020.

[70] Angel Molina et al. “Sequential classification system for recognition of malaria infection using peripheral blood cell images”. In: *Journal of Clinical Pathology* (2020).

[71] Laurent Najman and Michel Schmitt. “Watershed of a continuous function”. In: *Signal Processing* 38.1 (1994), pp. 99–112.

[72] Akshay Nanoti et al. “Detection of malaria parasite species and life cycle stages using microscopic images of thin blood smear”. In: *International Conf. on Inventive Computation Technologies*. Vol. 1. IEEE. 2016, pp. 1–6.

[73] Steven P O’Hara and Xian-Ming Chen. “The cell biology of Cryptosporidium infection”. In: *Microbes and infection* 13.8-9 (2011), pp. 721–730.

[74] Jennifer K O’Leary, Roy D Sleator, and Brigid Lucey. “Cryptosporidium spp. diagnosis and research in the 21st century”. In: *Food and Waterborne Parasitology* 24 (2021), e00131.

[75] World Health Organization. *Global technical strategy for malaria 2016-2030*. World Health Organization, 2015.

[76] Daniel Osaku et al. “Automated diagnosis of intestinal parasites: a new hybrid approach and its benefits”. In: *Computers in Biology and Medicine* 123 (2020), p. 103917.

[77] Priyadarshini Adyasha Pattanaik and Tripti Swarnkar. “Vision-based malaria parasite image analysis: a systematic review”. In: *International Journal of Bioinformatics Research and Applications* 15.1 (2019), pp. 1–32.

[78] Kristofer E Delas Peñas, Pilarita T Rivera, and Prospero C Naval. “Malaria parasite detection and species identification on thin blood smears using a convolutional neural network”. In: *International Conf. on Connected Health: Applications, Systems and Engineering Technologies*. IEEE. 2017, pp. 1–6.

[79] Franz Petry. “Structural analysis of Cryptosporidium parvum”. In: *Microscopy and Microanalysis* 10.5 (2004), pp. 586–601.

[80] Eric Piaton et al. “Technical recommendations and best practice guidelines for May-Grünwald-Giemsa staining: literature review and insights from the quality assurance”. In: *Annales de pathologie*. Vol. 35. 4. 2015, pp. 294–305.

[81] Mahdieh Poostchi et al. “Image analysis and machine learning for detecting malaria”. In: *Translational Research* 194 (2018), pp. 36–55.

[82] Sivaramakrishnan Rajaraman et al. “Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images”. In: *PeerJ* 6 (2018), e4568.

[83] Muhammad Imran Razzak. “Automatic detection and classification of malarial parasite”. In: *International Journal of Biometrics and Bioinformatics (IJBB)* 9.1 (2015), pp. 1–12.

[84] David W Reduker, CA Speer, and John A Blixt. “Ultrastructural changes in the oocyst wall during excystation of Cryptosporidium parvum (Apicomplexa; Eucoccidiorida)”. In: *Canadian Journal of Zoology* 63.8 (1985), pp. 1892–1896.

[85] DAVID W REDUKER, CA Speer, and JOHN A BLIXT. “Ultrastructure of Cryptosporidium parvum Oocysts and Excysting Sporozoites as Revealed by High Resolution Scanning Electron

Microscopy 1". In: *The Journal of protozoology* 32.4 (1985), pp. 708–711.

[86] Mateus Roder et al. "Intestinal parasites classification using deep belief networks". In: *International Conference on Artificial Intelligence and Soft Computing*. Springer. 2020, pp. 242–251.

[87] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2015, pp. 234–241.

[88] Uwe Schmidt et al. "Cell detection with star-convex polygons". In: *International Conf. on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2018, pp. 265–273.

[89] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2015. arXiv: `1409.1556 [cs.CV]`.

[90] Balbir Singh and Cyrus Daneshvar. "Human infections and detection of Plasmodium knowlesi". In: *Clinical microbiology reviews* 26.2 (2013), pp. 165–184.

[91] Krit Sriporn et al. "Analyzing Malaria Disease Using Effective Deep Learning Approach". In: *Diagnostics* 10.10 (2020), p. 744.

[92] Carsen Stringer et al. "Cellpose: a generalist algorithm for cellular segmentation". In: *Nature Methods* (2020), pp. 1–7.

[93] Christian Szegedy et al. "Rethinking the inception architecture for computer vision". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.

[94] Jayesh Tandel et al. "Life cycle progression and sexual development of the apicomplexan parasite Cryptosporidium parvum". In: *Nature microbiology* 4.12 (2019), pp. 2226–2236.

[95] Noppadon Tangpukdee et al. "Malaria diagnosis: a brief review". In: *The Korean journal of parasitology* 47.2 (2009), p. 93.

[96] F Boray Tek, Andrew G Dempster, and Izzet Kale. “Parasite detection and identification for automated thin blood film malaria diagnosis”. In: *Computer vision and image understanding* 114.1 (2010), pp. 21–32.

[97] Sarah Thomson et al. “Bovine cryptosporidiosis: impact, host-parasite interaction and control strategies”. In: *Veterinary Research* 48.1 (2017), pp. 1–16.

[98] Phi Vu Tran. “A fully convolutional neural network for cardiac segmentation in short-axis MRI”. In: *arXiv preprint arXiv:1604.00494* (2016).

[99] Renu Tuteja. “Malaria- an overview”. In: *The FEBS journal* 274.18 (2007), pp. 4670–4679.

[100] UNICEF. *World Malaria Day 2020 : Malaria - Status Update on Children*. Tech. rep. UN Children’s Fund, 2020. URL: `https://data.unicef.org/wp-content/uploads/2020/04/WMD-Covid_2020_CoverNote_Final.pdf`.

[101] Ryan J Urbanowicz and Jason H Moore. “ExSTraCS 2.0: description and evaluation of a scalable learning classifier system”. In: *Evolutionary intelligence* 8.2 (2015), pp. 89–116.

[102] Haidong Wang et al. “Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015”. In: *The lancet* 388.10053 (2016), pp. 1459–1544.

[103] Lijun Wang et al. “Visual tracking with fully convolutional networks”. In: *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 3119–3127.

[104] WHO. *World malaria report*. Tech. rep. WHO TEAM : Global Malaria Programme, 2019.

[105] WHO. *World malaria report*. Tech. rep. WHO TEAM : Global Malaria Programme, 2019. URL: `https://www.mmv.org/sites/default/files/uploads/docs/publications/World%20Malaria%20Report_0.pdf`.

[106] Kenneth W Widmer, Kevin H Oshima, and Suresh D Pillai. "Identification of Cryptosporidium parvum oocysts by an artificial neural network approach". In: *Applied and environmental microbiology* 68.3 (2002), pp. 1115–1121.

[107] Haiping Wu et al. "Cvt: Introducing convolutions to vision transformers". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, pp. 22–31.

[108] Tete Xiao et al. "Early convolutions help transformers see better". In: *Advances in Neural Information Processing Systems* 34 (2021).

[109] Fuyong Xing et al. "Deep learning in microscopy image analysis: A survey". In: *IEEE transactions on neural networks and learning systems* 29.10 (2017), pp. 4550–4568.

[110] Dahou Yang et al. "A portable image-based cytometer for rapid malaria detection and quantification". In: *PLoS One* 12.6 (2017), e0179161.

[111] Feng Yang et al. "Deep learning for smartphone-based malaria parasite detection in thick blood smears". In: *IEEE Journal of Biomedical and Health Informatics* 24.5 (2019), pp. 1427–1438.

[112] Kun Yuan et al. "Incorporating convolution designs into visual transformers". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021, pp. 579–588.

[113] Yang Zhang et al. "Deep learning for imaging and detection of microorganisms". In: *Trends in Microbiology* 29.7 (2021), pp. 569–572.

[114] Sixiao Zheng et al. “Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021, pp. 6881–6890.