Convolutional Neural Networks Deep Learning Based for Malaria Detection and Diagnosis

Josue Ouedraogo[1], Ferdinand T. Guinko[2], Kiswendsida K. Kabore[3]
{jos.weder@gmail.com[1], fguinko@ujkz.bf[2], kisito@ujkz.bf[3] }

Université Joseph KI-ZERBO, Ouagadougou, Burkina Faso[1,2,3].

Abstract. Malaria is a disease that occurs worldwide, especially in tropical regions where a high prevalence is observed. Difficulties are encountered especially in developing countries where resources in terms of equipment and trained personnel are limited. Until today, microscopic analysis is the standard method for diagnosing *Plasmodium falciparum*, which is the causative agent of malaria. In this paper, we proposed a malaria detection and diagnosis system using a deep learning technique which is a convolutional neural network called YOLOv5. Model learning was performed using a combination of two given image sources, Delgado Dataset B and Dijkstra Dataset, as a dataset containing thin smear images. We then evaluated the performance of the model by comparing it with other state-of-the-art results on deep learning. We obtained for the detection, a mean Average Precision of 96.71%.

Keywords: Malaria, Deep Learning, Object detection, YOLOv5, *Plasmodium falciparum*.

1 Introduction

Malaria is a parasite disease that causes fever and digestive disorders. It is transmitted to humans by the bite of an infected mosquito (female anopheles), and more rarely by a blood transfusion or transmission from mother to child during pregnancy. At the global level, the number of malaria cases is estimated at 241 million in 2020 in 85 malaria-endemic countries. The World Health Organization African Region alone accounts for about 95% (228 million) of the estimated cases in 2020 [1]. Six of these countries alone accounted for nearly 55% of cases: Nigeria (27%), the Democratic Republic of Congo (12%), Uganda (5%), Mozambique (4%), Angola (3.4%), and Burkina Faso (3.4%) [2]. In 2020, according to figures from the "Programme National de Lutte Contre le Paludisme au Burkina Faso", more than 11 million cases of malaria were recorded in the country's health facilities, resulting in nearly 4,000 deaths [3]. The peak of malaria cases is usually observed between July and August. The number of cases remains high until October, and even beyond that in some localities where the epidemic curve begins to decline in February or March [4]. Several methods exist for the diagnosis of malaria, namely the Rapid Diagnostic Test (RDT), Polymerase Chain Reaction (PCR)

JRI 2023, December 18-20, Ouagadougou, Burkina Faso

DOI 10.4108/eai.18-12-2023.2348130

techniques, Moop-mediated isothermal Amplification (LAMP) techniques, and Quantitative Buffy Coat (QBC). The detection of malaria parasites by light microscopy on Giemsa-stained blood films remains the reference method for malaria diagnosis [5], due to its accessibility even in developing countries. This method involves the observation of *Plasmodium* parasites on a blood smear stained with a chemical called Giemsa. The main limitations of this technique are that it is time-consuming, the results obtained are difficult to reproduce and the need for qualified personnel. In the rest of the article, we will first see the related work. Then we unveil the proposed method which is subdivided into 3 parts namely the datasets used, the algorithm used, and the implementation. The results and the discussion are presented by a third party and we end with the conclusion.

2 Related Works

Several works have been done on the detection of *Plasmodium* in blood. We chose to present some of them which took place between 2015 and June 2022, thus representing the state of the art. Previous work on the computational diagnosis of malaria has been reviewed [6], which the authors recommend to further improve the accuracy of detection. However, most of these attempts are based on manual engineering of feature extraction techniques, which requires skill and expertise. Table **??** presents a summary of the algorithms used. Devi et al [7], in their work, proposed a hybrid classification system for *Plasmodium falciparum* detection. These individual classifiers are SVM, kNN, and Naive Bayes. Their techniques thus use classical machine learning algorithms that have limitations on processing large data.

Table 1: Summary table of work in the domain of *Plasmodium* detection.

Authors	Algorithms	Data	Result
S. S. Devi et al. [7]	SVM, kNN and Naive Bayes	200 thin smear images	98.38% of precision
Mehedi Masud et al. [8]	CNN	27.558 red cells	97% of precision
M. Umer et al. [9]	CNN	27.558 red cells	98.5% of precision
Abdurahman et al. [10]	YOLOV3 and YOLOV4	948 thick smear images	6.14% YOLOV3-MOD-2, 96.32% YOLOV4-MOD of Precision
Mohamed Sawadogo et al. [11]	YOLOV3	236 thick smear images	92,22% mAP
Feng Yang et al. [12]	Cascading YOLOv2	2.567 thin smear images	79.22% mAP
Oliver S. Zhao et al. [13]	SSD	1.364 thin smear images	90.4% mAP

Their classification accuracy is 98.38% for SVM, 97.35% for kNN, and 97.23% for Naive Bayes. The tasks are therefore limited to the classification into two classes of erythrocytes already isolated from blood smear images. The work of Mehedi Masud et al [8] and M. Umer et al [9] was based on the classification of infected or uninfected erythrocytes. These studies used preprocessed

images. To further reduce data preprocessing, object detection algorithms were later used. In the work of Abdurahman et al [10], Mohamed Sawadogo et al [11], Feng Yang et al [12], and Oliver S. Zhao et al [13], object detection algorithms were tested and their results give an average accuracy of 96.32%, 92.22%, 79.22%, and 90.4% respectively. These algorithms have a low accuracy compared to simple classification algorithms but offer the advantage of detecting several objects in an image. We have therefore chosen to work with object detection algorithms, in particular with Yolo in its version 5.

Table 2: Comparison of various malaria datasets.

Malaria Datasets	*Plasmodium* species	Type of diagnosis	Type of images	Number of images
NIH Dataset [14]	*Plasmodium falciparum*	Thin smear	Binary	27 558
Nigeria dataset [15]	*Plasmodium falciparum*	Thick smear	Multi-class	2 986
MaMic Image Dataset [16]	Plasmodium falciparum	Thin smear	Binary	16 991
Malaria-655 [17]	*Plasmodium falciparum vivax, malariae, ovale*	Thin smear	Multi-class	4 363
MP-IDB [18]	*Plasmodium falciparum*, *vivax, malariae, ovale*	Thin smear	Multi-class	840
IML-Malaria Dataset [19]	*Plasmodium vivax*	Thin smear	Multi-class	38 449
UGANDA 2703 Dataset [20]	*Plasmodium falciparum*	Thick smear	Multi-class	2 703
Cachar distric dataset [7]	*Plasmodium falciparum* and *vivax*	Thin smear	Multi-class	120
MORU Bangkok Dataset [12]	*Plasmodium vivax*	Thin smear	Multi-class	2 567
Delgado Dataset B [21]	*Plasmodium falciparum*	Thin smear	Multi-class	331
Abbas and Dijkstra Dataset [22]	*Plasmodium falciparum*	Thin smear	Multi-class	883
Broad Institute Dataset [23]	*Plasmodium vivax*	Thin smear	Multi-class	1 300

In addition, the data used for training and testing the selected algorithms are included. Table 2 shows a summary of these data. These data can be categorized in several ways. For example, we have categorized them according to the species of *Plasmodium* detected in the thin smear image. Thus NIH Dataset, Nigeria dataset, MaMic Image Dataset, UGANDA 2703 Dataset, Delgado Dataset B, and Abbas and Dijkstra Dataset are composed only of images containing *Plasmodium falciparum* while MORU Bangkok Dataset and Broad Institute dataset are composed only of *Plasmodium vivax*. The other datasets are composed of a mixture of plasmodial species namely Malaria-655, MP- IDB, and the Cachar district dataset. Our study is based on *Plasmodium falciparum* and for a broad consideration of datasets, we consider the combination of at least two datasets namely Delgado Dataset B [21] and Abbas and Dijkstra Dataset [22].

3 Proposed Method

3.1 YOLOv5

The "You Only Look Once", or YOLO is a family of models is a one-step detection algorithm and is a series of end-to-end deep learning models designed for fast object detection, developed by Joseph Redmon et al. Published by Glenn Jocher who also is the inventor of the Mosaic data augmentation, explained in YOLOv4 [24]. Our model [25] is composed of nine (9) layers of standard convolutions interspersed by eight (08) layers C3 or CSP, a layer of SPPF, two (02) layers of UpSample that allow the oversampling of data, four (04) layers Concat that allow concatenation of a list of tensors and a final layer Detect for detection. The C3 or CSP layers each group three (03) standard convolution layers. The SPPF layer consists of two convolution layers and a Maxpooling layer.

3.2 Data

The quality of a model training is related to the quantity and quality of the images used. We worked with two data sources. First, a dataset containing 331 digital images [21] of Giemsa-stained peripheral blood smears (MGG). These were compiled during the daily work in the central laboratory of the Hospital Clinic of Barcelona from five patients diagnosed with malaria infection. In addition, a dataset containing 883 Giemsa-stained images [22] from 17 malaria-infected patients was collected. We then proceeded to the labeling of the images with the open-source software CVAT [1]. We were assisted during the labeling process by a medical expert from the "Centre National de Recherche et de Formation sur le Paludisme". **Figure 1** shows an example of the labeling process with the CVAT tool where we can see the infected red cells framed in yellow and the uninfected red cells in red. In summary, our labeled data set is composed of 564 smear images, 1,984 objects representing infected erythrocytes, and 24,431 objects representing healthy erythrocytes. The separation is done as follows: 85% of images in the training set, 10% in the validation set, and 5% in the test set.

4 Implementation

The dataset, after the preparation of the training data, is used as input for our model. The model is then tested after the training phase, using images from the test dataset and we analyze the results based on the evaluation metrics. The hyperparameters must be configured before any training. The number of classes is set to 03 (three) and placed the NMS threshold at 0.7. Model training was performed on 1200 epochs using the stochastic gradient downward method (SGD) as the optimization algorithm. The training took about 3h 44m 58s hours. We used the Kaggle platform which provides us with 19.6GB of free storage and 16GB of RAM and a Tesla P100-PCIE-16GB GPU. For every 100th iteration of learning, a checkpoint for reuse in case of a sudden process stop and patience of 600 epochs are used to control the learning improvement. At the end of the training, we obtained a pre-trained model. We developed a web platform with the Django framework that

[1] `https://cvat.org/`

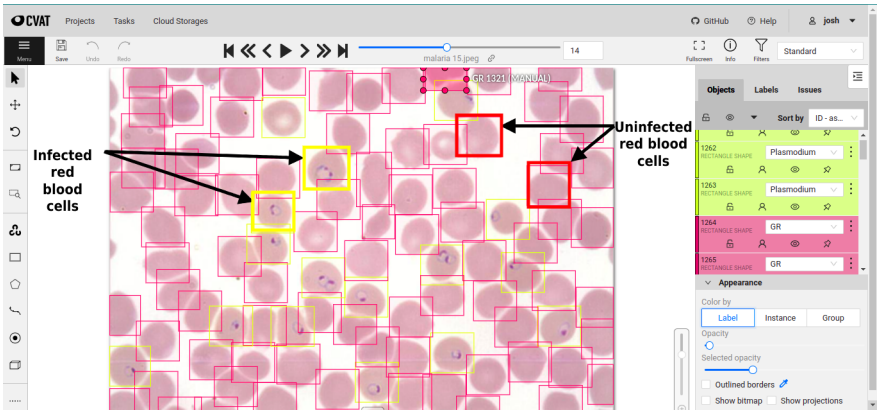

Fig. 1. Labeling with the CVAT tool.

allows us to predict images and that can be used by medical staff in medical centers. **Figure 2** shows a screenshot of the web platform.

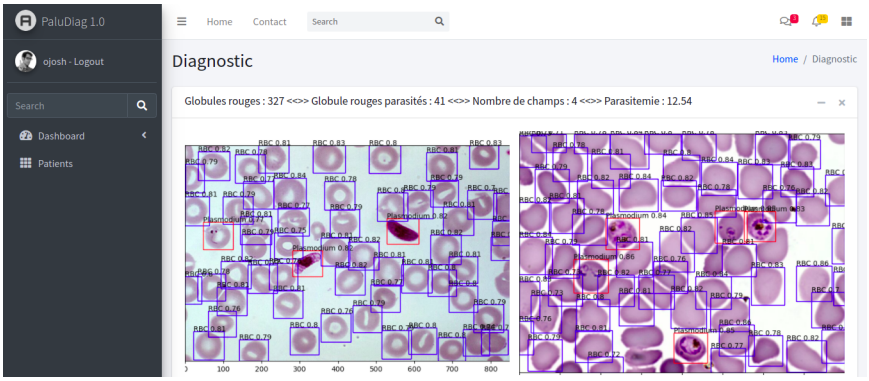

Fig. 2. Screenshot of our web platform.

5 Results and Discussion

5.1 Results

Deep learning techniques use various metrics to compare performance, accuracy, and others. To measure the quality of detection, we used the Confusion Matrix [26] shown in **Figure 3**, which is a summary of the prediction results of a model. The four (04) basic elements of the Confusion Matrix are true positives (TP) which denote infected erythrocytes; false negatives (FN) which denote

infected erythrocytes detected as uninfected erythrocytes; false positives (FP) which are uninfected erythrocytes detected as infected erythrocytes and true negatives (TN) which are uninfected erythrocytes.

		Actual Class	
		Positive (**P**)	Negative (**N**)
Predicted Class	Positive (**P**)	True Positive (**TP**)	False Positive (**FP**)
	Negative (**N**)	False Negative (**FN**)	True Negative (**TN**)

Fig. 3. Confusion Matrix. [26]

The most commonly used metrics for comparing object detection algorithms are :

- **the Precision (P)**: is the ratio of correctly predicted positive samples(VP) to predicted positive samples(VP+FP). It represents confidence in the detection of *Plasmodium*. It is calculated according to formula 1 :

$$P = \frac{VP}{VP+FP}[27] \tag{1}$$

- **the Recall (R)**: is the ratio of correctly predicted positive samples (VP) to the total number of samples that are actually positive (VP+FN) and is also called the detection rate. The detection rate reflects the ability of the model to recognize *Plasmodium*. It is calculated according to formula 2 :

$$R = \frac{VP}{VP+FN}[27] \tag{2}$$

- **the mean Average Precision (mAP)**: is the average of the average accuracies (AP) of each class. It compares the ground truth bounding box to the detected box and returns a score. The higher the score, the more accurate the model is in its detection. In computer vision, mAP is a popular evaluation metric used for object detection (i.e. localization and classification). It is calculated according to formula 3 :

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k[27] \tag{3}$$

where n is the Number of classes and AP_k, the average precision of class k.

The results obtained are measured according to the metrics that we have previously described, namely mAP, precision, and recall. This is useful to know the performance of our model and also to be able to compare it to other works.

We measured the performance of our model using mAP, precision, and recall. The Wandb tool, allowed us to observe the following values:

- the mAP: we obtain an mAP of 96.71% ;
- the precision: here we have obtained 94.83% as precision ;
- the recall: we get a recall of 91.26%.

Figure 4 shows the evolution of mAP during model training where the y-axis represents the percentage of accuracy and the x-axis the number of epochs.

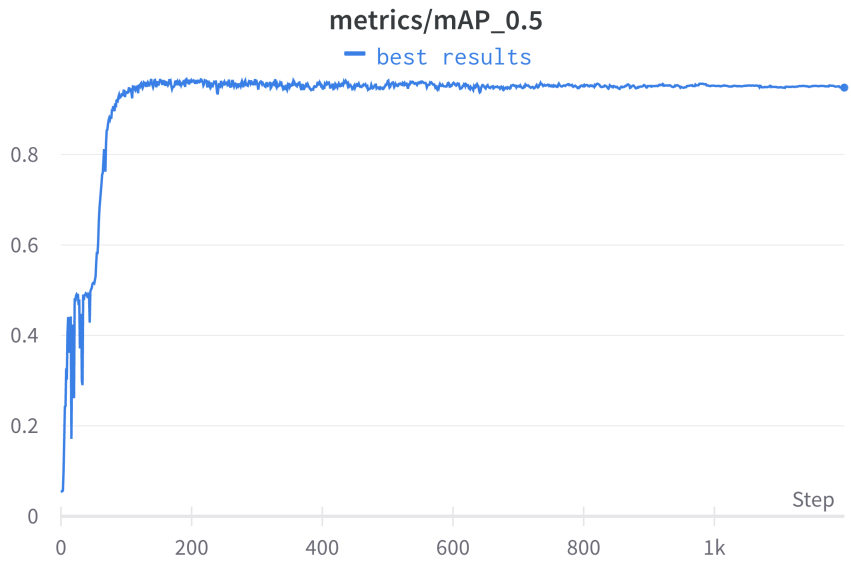

Fig. 4. mAP value evolution.

5.2 Discussion

Comparing works is difficult because the same methods are not used, nor are the same data used, and also the evaluation metrics. Nevertheless, a comparison can be made on the basis of works using similar machine learning techniques. **Table 3** compares the performance of our detection method with the methods examined in the state of the art. These works have a priori the same goal as ours, but the difference is in the approach techniques and the data used. Our model is more efficient compared to these works and this can be explained by the fact that we use a more recent object detection algorithm.

We proposed a detection model using YOLOv5, a deep learning technique that is a convolutional neural network. YOLOv5 has the advantage of meeting the speed of execution and satisfactory performance. We subsequently evaluated the performance of the model by comparing it with other state-of-the-art deep learning results. From this evaluation, we found our model to be as efficient in *Plasmodium* detection. Nevertheless, improvements are still possible through the parameterization of YOLOv5, and the quantity and quality of data used.

Table 3: Performance table of some models.

Works	Algorithms	Data	mAP
Zhao et al. [13]	SSD300	*Plasmodium vivax*	90.4%
Yang et al. [12]	Cascaded YOLO	*Plasmodium vivax*	79.22%
Our model	YOLOv5(x)	*Plasmodium falciparum*	96.71%

6 Conclusion

In this research, we worked on the detection and diagnosis of malaria using machine learning techniques. To do this, we first made a case for the need for such work, which benefits our health centers and the fight against malaria. The image processing pipeline of the system consists of three main steps: first the labeling of the images, then the parameterization of the object detection algorithm, and finally the model training for detecting and classifying infected and uninfected cells. We use a combination of two different sources of datasets namely Delgado Dataset B and Dijkstra Dataset for a total of 564 thin smear images. The object detection algorithm YOLO with its version 5 was used for training and detection. The results obtained in particular the mAP are up to 96.71%. Our system can be used in resource-limited areas without requiring specific expertise on malaria. Following the detection process, we have developed a web platform that allows the exploitation of the pre-trained model. In the future, we plan to improve our model by comparing the results of our system with the field results obtained by medical experts.

References

[1] The Effect of International Traffic on Malaria Case;. Available from: https://encyclopedia.pub/entry/18294.

[2] World malaria report 2020;. Available from: https://www.who.int/publications-detail-redirect/9789240015791.

[3] ANNUAIRE STATISTIQUE 2020;. Available from: https://www.sante.gov.bf/detail-documents?tx_news_pi1%5Baction%5D=detail&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Bnews%5D=581&cHash=ad2259d1454ee418fbca764a191f5392.

[4] Burkina Faso : pendant la saison des pluies, les populations déplacées luttent contre le paludisme et les maladies hydriques;. Available from: https://www.msf.fr/actualites/burkina-faso-pendant-la-saison-des-pluies-les-populations_deplacees-luttent-contre-le-paludisme-et-les-maladies-hydriques.

[5] Tangpukdee N, Duangdee C, Wilairatana P, Krudsood S. Malaria diagnosis: a brief review;47(2):93. Publisher: Korean Society for Parasitology.

[6] Rosado L, Correia da Costa JM, Elias D, S Cardoso J. A review of automatic malaria parasites detection and segmentation in microscopic images;14(1):11-22. Publisher: Bentham Science Publishers.

[7] Devi SS, Roy A, Singha J, Sheikh SA, Laskar RH. Malaria infected erythrocyte classification based on a hybrid classifier using microscopic images of thin blood smear;77(1):631-60. Publisher: Springer.

[8] Leveraging Deep Learning Techniques for Malaria Parasite Detection Using Mobile Application;. Available from: https://www.hindawi.com/journals/wcmc/2020/8895429/.

[9] Umer M, Sadiq S, Ahmad M, Ullah S, Choi GS, Mehmood A. A novel stacked CNN for malarial parasite detection in thin blood smear images;8:93782-92. Publisher: IEEE.

[10] Abdurahman F, Fante KA, Aliy M. Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models;22(1):1-17. Publisher: BioMed Central.

[11] SAWADOGO SM, KABORET ML, BISSYANDE TFA, SAWADOGO S. Apprentissage Automatique à base de Réseaux de Neurones Profonds pour la Détection et le Diagnostic du Paludisme;.

[12] Yang F, Quizon N, Yu H, Silamut K, Maude RJ, Jaeger S, et al. Cascading YOLO: automated malaria parasite detection for Plasmodium vivax in thin blood smears. In: Medical Imaging 2020: Computer-Aided Diagnosis. vol. 11314. International Society for Optics and Photonics;. p. 113141Q.

[13] Zhao OS, Kolluri N, Anand A, Chu N, Bhavaraju R, Ojha A, et al. Convolutional neural networks to automate the screening of malaria in low-resource countries;8:e9674. Publisher: PeerJ Inc. Available from: https://peerj.com/articles/9674.

[14] LHNCBC Full Download List;. Available from: https://lhncbc.nlm.nih.gov/LHC-downloads/downloads.html#malaria-datasets.

[15] Expert-level automated malaria diagnosis on routine blood films with deep neural networks - Manescu - 2020 - American Journal of Hematology - Wiley Online Library;. Available from: https://onlinelibrary.wiley.com/doi/10.1002/ajh.25827.

[16] The MaMic Image Database;. Available from: http://fimm.webmicroscope.net/research/momic/mamic.

[17] Dossier - Google Disque;. Available from: https://drive.google.com/drive/folders/1EMJ7dg0TBs34sDWcj7Tj1wozXJC0wtbc.

[18] Loddo A, Di Ruberto C, Kocher M, Prod'Hom G. MP-IDB: The Malaria Parasite Image Database for Image Processing and Analysis;. p. 57-65.

[19] Arshad QA, Ali M, Hassan Su, Chen C, Imran A, Rasul G, et al. A Dataset and Benchmark for Malaria Life-Cycle Classification in Thin Blood Smear Images.

[20] Quinn J, Andama A, Munabi I, Kiwanuka F. Automated Blood Smear Analysis for Mobile Malaria Diagnosis. In: Mobile Point-of-Care Monitors and Diagnostic Device Design;. p. Pages 115-32. Journal Abbreviation: Mobile Point-of-Care Monitors and Diagnostic Device Design.

[21] Delgado M, Molina A, Alferez S, Rodellar J, Merino A. Dataset B: 331 digital images of MGG-stained blood smears from five malaria-infected patients;1. Publisher: Mendeley Data. Available from: https://data.mendeley.com/datasets/2v6h4j48cx/1.

[22] Abbas SS, Dijkstra T. Malaria-Detection-2019;1. Publisher: Mendeley Data. Available from: https://data.mendeley.com/datasets/5bf2kmwvfn/1.

[23] Broad Bioimage Benchmark Collection;. Available from: https://bbbc.broadinstitute.org/BBBC041.

[24] Bochkovskiy A, Wang CY, Liao HYM. Yolov4: Optimal speed and accuracy of object detection.

[25] YOLOv5 Structure. Ultralytics;. Original-date: 2020-05-18T03:45:11Z. Available from: https://github.com/ultralytics/yolov5/blob/d059d1da03aee9a3c0059895aa4c7c14b7f25a9e/models/yolov5x.yaml.

[26] Confusion Matrix, The Science of Machine Learning;. Available from: https://www.ml-science.com/confusion-matrix.

[27] Mean Average Precision (mAP) Explained: Everything You Need to Know;. Available from: https://www.v7labs.com/blog/mean-average-precision.