

Received 13 August 2024, accepted 9 September 2024, date of publication 12 September 2024, date of current version 30 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3459411



RESEARCH ARTICLE

Automated Identification of Malaria-Infected Cells and Classification of Human Malaria Parasites Using a Two-Stage Deep Learning Technique

DHEVISHA SUKUMARRAN¹, EE SAM LOH¹, ANIS SALWA MOHD KHAIRUDDIN^{2,3} (Member, IEEE), ROMANO NGUI⁴, WAN YUSOFF WAN SULAIMAN⁵, INDRA VYTHILINGAM⁵, PAUL CLIFF SIMON DIVIS⁴, AND KHAIRUNNISA HASIKIN^{1,2}

¹Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia

²Centre of Intelligent Systems for Emerging Technology (CISET), Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia

³Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur 50603, Malaysia

⁴Malaria Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak 94300, Malaysia

⁵Department of Parasitology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur 50603, Malaysia

Corresponding author: Khairunnisa Hasikin (khairunnisa@um.edu.my)

ABSTRACT The gold standard for diagnosing malaria remains microscopic examination; however, its application is frequently impeded by the lack of a standardized framework that guarantees uniformity and quality, particularly in scenarios with limited resources and high volume. This study suggests a novel and highly effective automated diagnostic approach that employs deep-learning object detectors to improve the accuracy and efficiency of malaria-infected cell detection and *Plasmodium* species classification to overcome these challenges. *Plasmodium* parasites were detected within thin blood stain images using the YOLOv4 and YOLOv5 models, which were optimized for this purpose. YOLOv5 obtains a slightly higher accuracy on the source dataset ($mAP@0.5=96\%$) than YOLOv4 ($mAP@0.5=89\%$), but YOLOv4 exhibits superior robustness and generalization across diverse datasets, as demonstrated by its performance on an independent validation set ($mAP@0.5=90\%$). This robustness emphasizes the dependability of YOLOv4 for deployment in a variety of clinical settings. Furthermore, an automated process was implemented to produce bound single-cell images from YOLOv4's localization outputs, thereby eradicating the necessity for conventional and time-consuming segmentation methods. The DenseNet-121 model, which was optimized for species identification, obtained an impressive overall accuracy of 95.5% in the subsequent classification stage, indicating excellent generalization across all malaria species. Accurate classification of *Plasmodium* species on microscopically thin blood films is essential for guiding appropriate therapy and preventing unnecessary anti-malarial treatments, which can lead to adverse effects and contribute to drug resistance. This research contributes to the field of automated malaria diagnosis by offering a comprehensive framework that substantially improves clinical decision-making, particularly in resource-limited environments.

INDEX TERMS Machine learning, deep learning, biosurveillance, AI-monitoring, detection.

I. INTRODUCTION

Malaria remains a public health concern, as the rate of improvement in cases and deaths has stagnated in several

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung .

nations with moderate or high transmission. The situation was exacerbated, particularly in sub-Saharan Africa, by the COVID-19 epidemic and other humanitarian crises [1]. It is one of the top 10 causes of mortality in developing countries and one of the world's leading health concerns [1], [3]. Even though 100 countries have eradicated malaria and are

now malaria-free, the World Health Organization (WHO) report 2022 estimates that 249 million malaria cases occurred in 85 malaria-endemic countries [4]. One year after the COVID-19 pandemic and service disruptions, the estimated number of malaria cases rose to 241 million, an additional 14 million cases compared to 2019 [4].

Currently, malaria diagnostic tools are available in various forms [5]. Microscopy examination of stained thick and thin blood films to identify the *Plasmodium* parasites is the gold standard for malaria diagnosis [6]. The microscopy evaluations are being set as the gold standard due to their cost-effectiveness, less time consumption, and ease of preparation. However, millions of blood films are checked for malaria yearly, requiring a qualified microscopist to manually count parasites and infected red blood cells. Many blood smears are tested for malaria annually, with a skilled parasitologist physically counting the infected red blood cells and parasites. Identification of *Plasmodium* species is required for malaria diagnosis, evaluating drug resistance, determining therapeutic efficacy, and grading disease severity. The outcome from microscopy examination, on the other hand, is not standardized because it relies heavily on the expertise and skills of parasitologists [7].

In addition, it is common for parasitologists to work in isolation in low-resource settings without a standardized framework to ensure the consistency of their skills and, consequently, the diagnostic quality. This results in erroneous diagnoses and inappropriate therapy for the patient [7]. False negatives waste anti-malarial drugs, entail further visits and time away from work, and can lead to severe malaria in extreme instances. A misdiagnosis may result in unnecessary anti-malarial medication, which may have unintended adverse effects on the patient [8].

With these limitations, the current study has sparked initiatives to automate malaria diagnosis because automatic malaria detection methods are more reliable and standardized. Moreover, it can serve many patients with less workload and be cost-effective [8]. Different critical processing stages are usually required to quantify parasitemia automatically, with the first and most crucial step being obtaining digital images of blood film.

The traditional automation of the malaria detection procedure is composed of complex techniques for image processing with manual-engineered features such as contour, color, intensity, area, and texture [9], [10], [11], as well as computer vision [12], [13], [14] and machine learning [8]. However, solutions for classifying the infected cells, which are both precise and efficient in terms of computational time, have not been researched until the maximum level.

Recent improvements in machine learning and deep learning algorithms for the microscopic identification of blood smear images have yielded encouraging results. This is due to the capability of the algorithms to automatically assess the input and extracted features through its model's hidden

layers. Machine learning has been explored to classify diseases in various healthcare applications [15], [16].

Classification methods ranging from decision trees and virtual artificial neural networks or support vector machines to random tree classifiers have been popular in classifying malaria parasites into infected and noninfected red blood cells. Most studies reviewed used SVM, RNN, and KNN architectures in the classification process [17], [18], [19]. Some studies have performed the classification of parasites conforming to the *Plasmodium* species with the use of K-nearest neighbors (KNN), nearest mean (N.M.), support vector machines (SVM), 1-NN, and Fisher based on features such as gradient, granulometry, colour histogram, and flat texture [20]. Moreover, state-of-the-art nonlinear classifiers such as multilayer perceptron neural networks (MLP) and support vector machines (SVM) were used to classify the infected erythrocytes and identify the infection stages [21]. These machine-learning models have been extensively utilized and embedded in many intelligent sensors in various applications, such as lab-on-chip hybrid sensors [22], [23] or rapid real-time detection [24].

Deep learning techniques have been the latest trend in automated malaria identification, which usually employs the Convolutional Neural Network (CNN) with different architectures. Liang, et al. [25] are the first to employ deep learning to diagnose malaria, using a convolutional neural network to distinguish between infected and uninfected cells in thin blood smears. Many studies have developed customized CNN to classify red blood cells into infected or noninfected cells accurately.

Most studies have used transfer learning networks instead of developing a customized CNN that is not too deep or too shallow in layers and matches the classification goal of malaria-infected blood cells. Transfer learning is a deep learning technique in which a model pre-trained on an extensive dataset is reused and adjusted as the basis for a model for a different job. Examples include VGGNet, ResNet, AlexNet, DenseNet, and other transfer learning models [26], [27], [28].

In addition to customized CNN and transfer learning, the You Look Only Once Model (YOLO) has been used to detect malaria parasites from blood film images. For example, cascaded YOLOv2 has been used to detect the infected cells from thin blood smear images and achieve a mAP of about 79.22% [19]. Moreover, transfer learning has been used to classify erythrocytic life-cycle stages of malaria [26]. However, the classification of infected cells into species has not been performed widely in most of the research.

Due to its advantages over traditional machine learning, deep learning has become a popular research subject in artificial intelligence. Deep learning automatically enables computers to extract, evaluate, and comprehend relevant information from raw data. Unlike machine learning, it does not require manually extracted or constructed features. Deep learning collects features from raw data, processes them, and decides automatically [29]. Compared to conventional

machine learning techniques, it can also bypass the segmentation process in some cases [8]. Table 1 summarizes the machine learning and deep learning approaches in 17 studies on automated malaria diagnostic systems. Most studies have explored binary classification between infected and noninfected cells of one species (*i.e.*, *P. falciparum*). Only a few studies reported on multi-species classifications where Malihai et al. [20] classified the *Plasmodium* species into four classes using the conventional machine learning technique of KNN. Another recent multi-specified classification study was reported in [28], where transfer learning was employed to identify *P. falciparum* and *P. vivax*-infected cells. In the meantime, Among the studies that employed deep learning for malaria species classification, Krishnadas, et al. [30] devised an enhanced solution for classifying malaria's type and progression stage using YOLOv5. For classifications of malaria species, the proposed YOLOv5 model attained an accuracy of 78.5%. However, for stage progression classification of gametocytes, schizonts, trophozoites, and rings, the proposed model only attained a mean average precision (mAP) of 0.399. Although Krishnadas, et al. [30]'s proposed solution demonstrated a marked improvement over previously published works, the imbalanced class of progression stage is difficult to resolve. The improper handling of the overlapping red blood cells reduces the precision and practicability of the proposed work. In more recent studies, Garba et al. [55] employed customized CNN for species classification however they reported an extreme bias of the model with an accuracy of 33.73% and 39% for *P. vivax* and *P. ovale* respectively, in contrast, 99.97% for *P. falciparum* classification. Similarly, Guemas et al. [56] employed the real-time detection transformer and recorded biases of 19.9% and 15% for the *P. ovale* and *P. malariae* classification respectively with an overall mAP of 63.8%.

Given the substantial amount of literature available and the research gaps highlighted in Table 1, our study emphasizes the urgent requirement for a two-stage model that can successfully address the significant obstacles associated with the identification of infected cells and the classification of malaria species, such as *P. falciparum*, *P. ovale*, *P. vivax*, and *P. malariae*. The motivation for doing this research is based on significant problems within malaria treatment and clinical management. One of the prevailing concerns pertains to the possible progression of moderate infections into severe illnesses, hence presenting a substantial quandary regarding the appropriate timing for admitting patients to intensive care units (ICUs). To effectively address these problems, it is imperative to establish standards, scoring systems, or prediction models that can assist healthcare practitioners in accurately predicting severe disease and complications. This will facilitate prompt intervention and timely management. To our knowledge, prevailing automated classification methods for malaria detection mainly prioritize binary classification, distinguishing between classes Class 1 representing malaria-infected cells and Class 0 representing normal cells.

Nevertheless, a notable deficiency in the existing body of research is the absence of a thorough methodology for species classification. This inadequacy is especially noteworthy while handling slender blood films since proficient microscopists face significant difficulties. The key issue pertains to the observation that infected cells in the early stages of malaria infection in all species share a uniform appearance, rendering it impracticable to distinguish them merely based on ring morphology. However, the potential of species classification becomes apparent when examining different trophozoites and other stages of parasites.

Hence, the analysis of single-cell infections proves to be unfeasible inside clinical environments. This study aims to address substantial limitations in existing approaches related to the identification and classification of malaria. The proposed model was established by implementing an end-to-end automated system for detecting cells infected with malaria. To enhance the accuracy of malaria detection, a two-stage methodology was proposed. In the initial stage, the proposed model operates by detecting the infected cells and effectively overcoming the difficulties associated with the coexistence of overlapping red blood cells and multiple infections inside a single sample. Subsequently, after identifying the infected cells, the specific regions containing these cells were extracted by extracting the coordinates of each bounding box (left_x, top_y, width, and height) generated by the model on an image. These individual predicted cells were then stored separately. Subsequently, the cropped infected cells are fed into the next detection stage, wherein they are classified according to their respective species, namely *P. falciparum*, *P. ovale*, *P. vivax*, and *P. malariae*. This innovative approach presents a potentially viable resolution to enduring obstacles within the discipline.

II. PROPOSED METHOD

The proposed methodology involved two stages of infected cell identification and malaria parasite species classification. In this study, we identified the two major issues in malaria diagnosis: i) rapid identification of infected cells from malaria thin blood smear microscopic images, and ii) human malaria species classification for four species, *P. falciparum*, *P. ovale*, *P. vivax*, and *P. malariae*.

A. DATASET PREPARATION

The dataset used in this study was acquired from the public dataset, namely the Malaria Parasite Image Database (MP-IDB) [31]. Additionally, as described in previous studies, archived thin blood films of malaria infections were obtained from the Malaria Research Centre, Universiti Malaysia Sarawak (MRC-UNIMAS), Sarawak [32]. Both datasets contain thin blood smear images of four main human malaria species (*P. falciparum*, *P. ovale*, *P. malariae*, *P. vivax*) in four different erythrocytic stages (ring-form or early trophozoite, mature trophozoite, schizont, gametocyte). This research focuses on localizing malaria-infected red blood cells and classifying them into four *Plasmodium* genera.

Every image in MP-IDB is saved in JPG format, with a resolution of 2592×1944 pixels and a color depth of 24 bits, with a total file size of about 717 MB. This data was collected solely from thin blood smears stained with Giemsa, as discussed in [31]. The dataset was first released in 2019, updated in 2021, and further verified by our parasitologist experts. Meanwhile, the thin blood films from MRC-UNIMAS, were fixed using absolute methanol (BDH, England) for 10 seconds and subsequently air-dried at room temperature. Later, the films were subjected to staining using a 10% Giemsa solution (BDH, England) in Gurr® buffered water with a pH of 7.2 (BDH, England) for 30 minutes, as recommended by the protocol outlined by the World Health Organization (WHO, 2016). All thin blood films were examined using a microscope (Olympus Model BX53) at a total magnification of $\times 1,000$ with immersion oil. Images were captured using a digital camera (Olympus DP25) mounted to the microscope and analyzed using the Digital Imaging Solution Cell B software (Olympus).

The datasets in this study were randomly separated into three categories: training, testing, and cross-test datasets. The training dataset was used to construct a deep-learning model. In contrast, the testing dataset validated which epochs in the training process produced the best deep-learning model, while the cross-test dataset was used to validate the overall model's performance. The cross-test dataset was isolated from the training and testing dataset until the learning model was completed.

To validate the datasets using a limited number of images, we employed a random distribution method for stage 1, as outlined in Table 2. A comprehensive set of 446 images was obtained from the publicly available MP-IDB dataset and the MRC-UNIMAS collection in Sarawak. Of the entire dataset, 80% of the MP-IDB data were subjected to random assignment for training purposes, while the remaining 20% were utilized as testing data. To enhance the learning capabilities of the developed model and assess the capabilities of the models in classifying unseen data, 100% of newly acquired data from MRC-UNIMAS, Sarawak, were used as testing data, as summarized in Table 2.

As shown in Table 2, in the first stage, the MP-IDB images were partitioned into 80:20 for training and testing the models respectively. Only the training images were augmented with different combinations of geometric transformations such as zoom-in, zoom-out, flip left-right, flip top-bottom, rotation 90° , and 180° at different probabilities. The training images were augmented to enhance the diversity of the training images and improve the object detectors' ability to learn the features of the infected cells across the different malaria species. This procedure introduces variations in the images, helping the models generalize better and recognize infected cells more effectively.

During the first stage, the proposed model was developed to classify the red blood cells as infected cells. The center

TABLE 1. Summary of deep learning approaches in automated malaria diagnostic system.

Author	Technique	Species	Findings
Díaz, et al.[21]	MLP, SVM	<i>P.falciparum</i>	Classification into infected and non-infected cells: SVM Specificity - 99.7% Sensitivity - 94% Classification of infected cells into stages: SVM Specificity - 91.2% Sensitivity - 78.8%
Linder, et al. [33]	SVM	<i>P.falciparum</i>	Sensitivity - 84.9% Specificity - 99.9% NPV - 99.9% PPV - 74.2% AUC ROC - 0.997
Telang and Sonawane [34]	SVM, RNN, KNN	<i>P.falciparum</i>	RNN best performance. Accuracy - 96.1% Precision - 97.8% Recall - 94.4% f1 score - 96.1%
Malihi, et al. [20]	SVM, NM, KNN, I-NN, Fisher	<i>P.falciparum</i> <i>P. vivax</i> <i>P. ovale</i> <i>P. malariae</i>	KNN best performance. Accuracy - 91%
Kudisthalert, et al [35]	WELM, MLP, SVM, LDA	<i>P.falciparum</i>	WELM best performance. Accuracy - 96.18%
Liang, et al. [25]	Customized CNN	<i>P.falciparum</i>	Accuracy - 97.37% Sensitivity - 96.99% Specificity - 97.75% Precision - 97.73% F1 score - 97.36% MCC - 94.75%
Fuhad, et al. [36]	Autoencoder	<i>P.falciparum</i>	Accuracy - 99.23%
Rajaraman, et al. [37]	Customized CNN compared to AlexNet, VGG-16, Xception, ResNet-50, DenseNet-121	<i>P.falciparum</i>	ResNet-50 best performance. Accuracy - 0.957 ± 0.007 AUC - 0.990 ± 0.004 Sensitivity - 0.945 ± 0.020 Specificity - 0.969 ± 0.009 F1 score - 0.957 ± 0.008 MCC - 0.912 ± 0.014
Umer, et al. [38]	Stacked CNN	<i>P.falciparum</i>	Accuracy - 99.96% Precision - 100% Recall - 99.92% F1 score - 99.96%
Arshad, et al. [36]	VGG16, VGG19, ResNet50v2, DenseNet169, DenseNet201	<i>P. vivax</i>	Classification directly into stages: DenseNet201 - Accuracy 74.56%
			Classification into infected noninfected them into stages: ResNet50v2 - Accuracy 79.61%
Maqsood, et al. [39]	Proposed CNN, VGG-16, VGG-19, Xception, DenseNet-121, DenseNet-169, DenseNet-201, Inception_v3, Inc.Resnet_v2, ResNet-50, ResNet-101, ResNet-152, SqueezeNet	<i>P.falciparum</i>	Proposed CNN best performance. Accuracy - 0.9682 Specificity - 0.9778 Sensitivity - 0.9633 Precision - 0.9682 F1-score - 0.9682 MCC - 0.9364

TABLE 1. (Continued.) Summary of deep learning approaches in automated malaria diagnostic system.

Rajaraman, et al. [37]	customized CNN (VGG-19 al SqueezeNet)	<i>P.falciparum</i>	Accuracy - 99.51% ± 0.1 AUC - 99.92 ± 0.1 MSE - 0.63 ± 0.1 Precision - 99.84 ± 0.1 F1 score - 0.05 ± 0.1 MCC - 99.0 ± 0.2
Pattanaik, et al. [40]	Proposed CAD - FLANN al SSAE	<i>P.falciparum</i>	Accuracy - 89.10% Sensitivity - 93.90% Specificity - 83.10%
Zhao, et al. [28]	VGG16, VGG19, ResNet50, Inception v3, Xception, InceptionResNet v2, DenseNet121, MobileNet v2	<i>P.falciparum</i> <i>P. vivax</i>	VGG16 best performed. Accuracy - 0.9653 ± 0.0043 Sensitivity - 0.9500 ± 0.0067 Specificity - 0.9807 ± 0.0025 AUC - 0.9940 ± 0.0010 F1 score - 0.9648 ± 0.0043 MCC - 0.9330 ± 0.0082
Irmak [41]	Novel CNN	<i>P.falciparum</i>	Accuracy - 95.28%
Krishnadas, et al. [30]	YOLOv5 and Scaled YOLOv4	<i>P.falciparum</i> <i>P. vivax</i> <i>P. ovale</i> <i>P. malariae</i>	Parasite classification based on object detection model: Precision -74.1% Recall - 79.7% mAP - 78% Stage classification: Precision – 44.7% Recall - 55.6% mAP - 39.9%
Garba et al. [55]	Customized CNN	<i>P.falciparum</i> <i>P. vivax</i> <i>P. ovale</i> <i>P. malariae</i>	Overall accuracy:99.9 <i>P.falciparum</i> : 99.97 <i>P.ovale</i> : 39% <i>P.malariae</i> : 64.6% <i>P.vivax</i> : 33.73%
Guenas et al. [56]	Real-time detection transformer	<i>P.falciparum</i> <i>P. vivax</i> <i>P. ovale</i> <i>P. malariae</i>	Overall accuracy: 63.8% <i>P.falciparum</i> :85.8 <i>P.ovale</i> : 19.9% <i>P.malariae</i> :64.4% <i>P.vivax</i> :15%

TABLE 2. Data distribution from MP-IDB and MRC-UNIMAS, Sarawak, for Stage 1.

Dataset	Total images before augmentation	Training data before augmentation	Training images after augmentation	Test images
MP-IDB	210	168 (80%)	1000	42 (20%)
MRC-UNIMAS	236	-	-	236 (100%)

coordinates of the infected cells were identified, and the infected cells' localized areas were cropped for the second stage classification of four species of *Plasmodium* genus (*P. falciparum*, *P. ovale*, *P. vivax*, and *P. malariae*). Then, the second stage was conducted using the same dataset sources to analyze the infected cells. As for the second stage, the detected infected cells were divided into training, validation, and testing, as listed in Table 3.

TABLE 3. Data distribution from MP-IDB and MRC-UNIMAS, Sarawak, for Stage 2.

Original dataset (Combination of MP-IDB and MRC-UNIMAS, Sarawak): Before augmentation						
Plasmodium species	MP-IDB dataset	MRC-UNIMAS, Sarawak	Total	Training data (70%)	Validation data (10%)	Testing data (20%)
<i>P. falciparum</i>	779	-	779	545	78	156
<i>P. malariae</i>	66	75	141	99	13	29
<i>P. ovale</i>	30	88	118	83	11	24
<i>P. vivax</i>	98	85	183	129	17	37

Before Augmentation			After Augmentation			
Plasmodium species	Training data	Validation data	Testing data	Training data	Validation data	
<i>P. falciparum</i>	545	78	156	7630*	78	156
<i>P. malariae</i>	99	13	29	1386*	91*	203*
<i>P. ovale</i>	83	11	24	1162*	77*	168*
<i>P. vivax</i>	129	17	37	1834*	119*	259*

*The images were augmented 14 times, which include 90° clockwise direction, 90° anti-clockwise direction, 180° clockwise direction, 270° clockwise, brightness levels of 0.8 & 1.2, Contrast level of 0.5, Brightness level 1.5, flips horizontal & vertical, saturation levels of 0.5 & 1.5, sharpness levels 0.5 & 1.5 (Figure 1).

The images were augmented 7 times, which include contrast 1.5, sharpness 5.0, blur 1.5, 90° clockwise direction, 180° clockwise direction, and 270° clockwise direction (Figure 2).

To cater to the imbalance between classes, the training samples were augmented 14 times, as shown in Fig. 1. In comparison, the validation and testing samples were augmented 7 times, as shown in the example in Fig. 2. Due to the limited samples of *P. malariae*, *P. ovale*, and *P. vivax*, we augmented the validation and testing samples by 7 times, which include the original image, contrast 1.5, sharpness 5.0, blur 1.5, 90° clockwise direction, 180° clockwise direction, 270° clockwise direction (Fig. 2). Only these augmentation processes were used for the validation and testing samples to expose the developed model to a wider range of variations that could be encountered in real-world scenarios. By testing with augmented data, we can evaluate how the developed model is generalized to different conditions and its prediction robustness to these variations. Testing using heightened contrast aims to replicate images that exhibit intense lighting conditions, hence assessing the capability of the proposed model to handle such scenarios effectively. Applying sharpness augmentation and blur techniques can replicate the fluctuations in image quality that may arise inside real medical settings. The assessment of the performance of the proposed model under these specified parameters ensures its capability to handle variations in image clarity effectively. Rotation augmentations are employed to ensure that the proposed model can accurately classify objects or patterns, irrespective of their orientation. This attribute is crucial in the context of identifying the infected malaria species.

B. TWO-STAGE AUTOMATED IDENTIFICATION AND CLASSIFICATION OF JUMAN MALARIA PARASITES

1) STAGE 1: IDENTIFICATION OF INFECTED AND NONINFECTED CELLS (BINARY CLASSIFICATION)

After data preparation, the initial phase of binary classification was conducted. This phase started with annotating

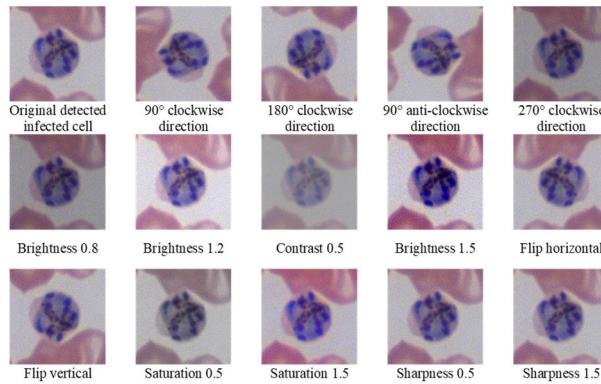


FIGURE 1. Example of data augmentation on the detected infected cell conducted on the training dataset.

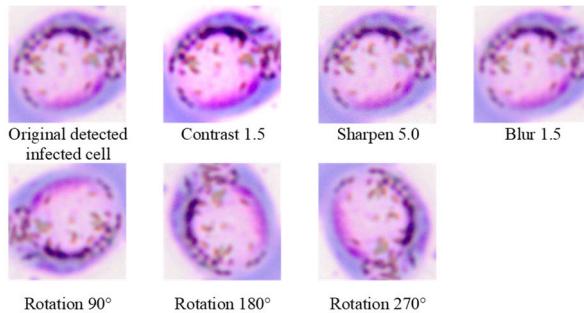


FIGURE 2. Example of data augmentation on the detected infected cell conducted on the validation and testing datasets.

the infected cells within the whole thin blood smear images using LabelImg. Data annotation involves drawing bounding boxes around the infected cells, which is crucial for training object detectors. The bounding boxes help the models to focus on relevant features and ignore irrelevant information beyond the bounding boxes. Each bounding boxes is labeled as “Infected Cell”, which guides the models in learning to predict and classify these cells’ features accurately.

During this phase, significant features will be extracted within the bounding boxes drawn in the thin blood smear images and trained to detect the infected cells. A preliminary study was conducted to evaluate the effectiveness of various object detection model algorithms, revealing the exceptional proficiency of the You Only Look Once (YOLO) in identifying malaria-infected cells [42]. Therefore, we extend the analysis by developing the improved YOLO models with the optimized hyperparameter as tabulated in Table 4.

Both models were optimized to adeptly apply to the critical task of malaria-infected cell detection. The proposed model initiates by receiving an input image, generally a thin blood film, and subjecting it to a deep convolutional neural network (CNN) as its backbone (BottleNeck CSP) integrated with a Feature Pyramid Network (FPN) or the

TABLE 4. The optimized hyperparameters for model training in classifying infected blood cells.

Parameters	Optimized YOLOv4	Optimized YOLOv5
Batch size	64	32
Learning rate	0.01	0.01
Backbone	CSP-Darknet53	CSP-Darknet53
Neck	SPP & PANet	SPP & PANet
Number of epochs	6000	350

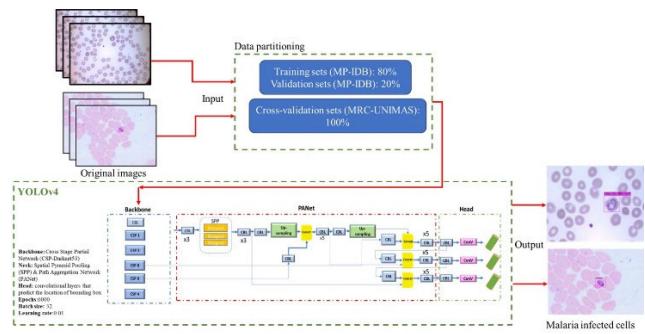


FIGURE 3. Stage 1 identification and the localization of malaria-infected cells.

neck to extract complex image characteristics (Fig. 3). These object detection heads of YOLO predict bounding boxes, class probabilities, and object scores for predicted objects present in the image, with the latter specifically referring to malaria-infected cells in this context. Anchor boxes play a crucial role in enabling accurate predictions of item position and size, while the object score serves as a filter to exclude detections with low confidence. Furthermore, the model assigns class probabilities to each bounding box, enabling it to differentiate between infected and normal cells. Non-maximum suppression (NMS) is employed to eliminate superfluous boxes, resulting in a conclusive output consisting of precisely localized malaria-infected cells, their corresponding confidence ratings, and class labels. This methodology enables the automatic and proficient identification of malaria within blood smear pictures, optimizing the diagnostic process and enhancing the efficacy of therapeutic treatment [42].

After identifying the infected cells, the localized bounding box region will be cropped to isolate individual infected cells. Upon the comparison of the models’ performances, the optimized YOLOv4 is chosen as the best-performing model. The model gives predictions in the form of bounding boxes with the coordinates of the bounding boxes in terms of (xmin, ymin, xmax, ymax). A separate algorithm is integrated upon predictions to automatically crop and save the cells within the bounding boxes only. These cells will then be fed into the second stage of the malaria species classification.

2) STAGE 2: MULTICLASS CLASSIFICATION OF 4 MALARIA SPECIES

The isolated individual infected cells extracted from Stage 1 were then classified into four classes: Class 0 indicates *P. falciparum*, Class 1 indicates *P. malariae*, Class 2 indicates *P. ovale*, and Class 3 indicates *P. vivax*. Four transfer learning models were utilized and optimized, namely Residual Neural Network (ResNet-50), Visual Geometry Group (VGG-16), Dense Convolutional Network (DenseNet-121) and AlexNet Convolutional Neural Network (Alexnet).

The AlexNet contained eight layers with learnable parameters: five convolution layers (C1, C2, C3, C4, C5) with a combination of max pooling layers followed by three fully connected layers (F6, F7, and an output layer). Each of the layers used Relu activation except the output layer. Meanwhile, VGGNet is a classic CNN architecture that uses only 3×3 convolutional layers to stack at each other for increasing depth. Max pooling layers were used to reduce the volume size of the model. VGG-16 and VGG-19 are named based on the number of weight layers in the network.

Meanwhile, the ResNet architecture introduces a skip connection, which acts as a “shortcut” to allow the input to travel from the previous layer to the next layer without having the trade-off of accuracy. Therefore, ResNet can have a deeper network due to this skip connection. The dataset was imported with the resolution fixed to 224×224 pixels to fit the requirement of the transfer learning model under PyTorch torch-vision. Besides, the PyTorch library comprises all the main functions to train, validate, and test the deep learning model.

Input images were imported and normalized to $([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])$, which is the standard normalization parameters according to ImageNet (as the transfer learning model used is pre-trained with ImageNet database). Normalization is required to normalize the input cell patches stained with different colours. The overview of each deep learning architecture is summarized in Table 5. These networks were tested, and hyperparameters were optimized as outlined in Table 6.

After the 25 training and validation epochs were run, the model weightage with the best training and validation accuracy was chosen and saved. The saved model was then loaded again to classify the test dataset, each model’s performance was determined, and the confusion matrix was plotted for further analysis. These models were then compared regarding their accuracy, precision, recall, specificity, and F1-score performances. The confusion matrices were plotted, and time complexity was analyzed to select the best model.

III. RESULTS

The performances of the proposed models in Stages 1 and 2 are evaluated in terms of a confusion matrix. This confusion matrix is a foundation for computing the predictability performance of classifiers. It is in the form of a “contingency table” that visualizes how the findings are disseminated over

TABLE 5. Overview of deep learning models used in this study.

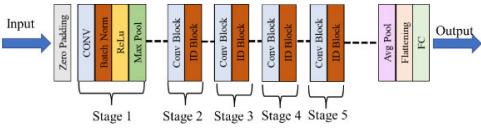
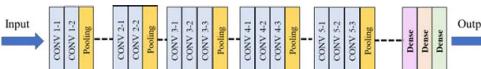
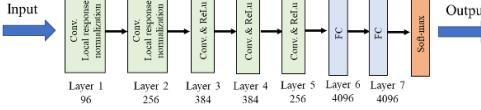
Model	Description
ResNet-50	ResNet-50 is a variant of the ResNet model, which has 48 convolution layers along with 1 MaxPool and 1 Average Pool layer. ResNet-50 is an artificial neural network (ANN) that forms networks by stacking residual blocks. Identity connections take the input directly to the end of each residual block and learn the features to be classified as the desired output (in this case, the malaria-infected cells).
	
VGG-16	VGG-16 is a convolutional neural network architecture with convolutional layers of 3×3 filters with stride 1 and constant padding and maxpool layers of 2×2 filters with stride 2 used. This convolutional and max pool layer arrangement is followed uniformly throughout the entire architecture. The end has 2 F.C. (fully connected layers) and a softmax for output. The number 16 in VGG16 indicates that it has 16 layers with weights.
	
DenseNet-121	Each layer in DenseNet receives additional input from all preceding layers and passes on its feature maps to all subsequent layers. Concatenation is employed. Each layer gets “collective knowledge” from the layers above it. Because each layer receives feature maps from all previous layers, the network can be made thinner and more compact (fewer channels can exist).
	
Alexnet	AlexNet comprises five convolutional layers, three max-pooling layers, two normalization layers, two fully connected layers, and one softmax layer. Convolutional filters and a nonlinear activation function ReLU are used in each convolutional layer. The pooling layers are used for maximum pooling. The input size is fixed because of the presence of fully connected layers.
	
MobileNet	MobileNet is a family of neural network architectures specifically crafted for mobile and embedded devices with constrained computational resources. It achieves high efficiency through depth-wise separable convolutions, which independently process spatial and channel dimensions, thereby significantly reducing parameters and computational load (Howard. G et al., 2017).
	

TABLE 5. (Continued.) Overview of deep learning models used in this study.

EfficientNet	EfficientNet is a convolutional neural network that introduces a systematic approach to scaling network dimensions. Instead of arbitrarily adjusting the depth, width, and resolution of the network, EfficientNet uses a compound method. This approach uniformly scales all three dimensions, depth, and resolution.

TABLE 6. Hyperparameters for the pre-trained deep learning model training in classifying infected blood cells according to 4 species (4 classes).

Parameters	ResNet-101	VGGNe t-19	DenseN et-121	AlexNet	Mobile Net	EfficientN et
Batch size	8	8	8	8	8	8
Num of workers	2	2	2	2	2	2
Input features	2048	4096	1024	4096	1208	1208
Output features	4	4	4	4	4	4
Learning rate	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
Step size	7	7	7	7	7	7
Gamma	0.1	0.1	0.1	0.1	0.1	0.1
Number of epochs	25	25	25	25	25	25

the actual class (represented in rows) and predicted class (represented in columns). The matrix comprises four instances: “True Positive” (TP) and “False Positive” (FP) are observations of correct and incorrect predictions per actual classes, accordingly, while “True Negative” (TN) and “False Negative” (FN) are instances of right and wrong rejections per actual classes, respectively. Based on this matrix, the following evaluation metrics for classification tasks were used to assess the prediction performance of all models proposed in this study: (i) precision, (ii) recall, (iii) specificity, and (iv) F1 score. Table 6 summarizes the model evaluation metrics.

1) STAGE 1: IDENTIFICATION OF INFECTED AND NON-INFECTED CELLS

The evaluation of the object detection models’ performances is presented in Table 8, reflecting their efficacy in detecting malaria-infected cells using the testing dataset. To validate the robustness of the developed models, a comprehensive analysis was conducted on previously unseen data sourced from MRC-UNIMAS. These data were specifically utilized as cross-testing images, allowing us to assess the adaptability and generalizability of the models. While the optimized YOLOv5 model demonstrated improved performance across all dimensions when evaluated on MP-IDB images, different trends were identified when analyzing cross-testing images. When tested on the new and unseen images, the optimized YOLOv4 was more robust and attained mAP (IoU=0.5) of 90% as compared to 59% attained by the optimized YOLOv5.

The performance of the models was further evaluated across various IoU thresholds, as detailed in Table 8. Both

TABLE 7. Model evaluation metrics.

Metrics	Formula	Description
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$	A fraction of correctly classified cases over the total samples.
Precision	$(TP) / (TP + FP)$	A proportion of positive classes were correctly classified.
Recall	$(TP) / (TP + FN)$	A ratio of all positive samples correctly predicted as positive.
Specificity	$(TN) / (TN + FP)$	A proportion of true negatives that are correctly identified by the model
F1-score	$2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$	A combination of precision and recall as their harmonic mean.

TABLE 8. Performance comparison of the developed object detection.

Models	IoU	Datasets	mAP (%)	Precision (%)	Recall (%)
Optimized YOLOv4	0.3	MP-IDB	94.33	89	92
		MRC_images	90.51	66	91
Optimized YOLOv5	0.5	MP-IDB	97.2	94.9	94.6
		MRC_images	60.7	57.2	61.5
Optimized YOLOv4	0.5	MP-IDB	89	85	88
		MRC_images	90	66	90
Optimized YOLOv5	0.5	MP-IDB	96	94	93
		MRC_images	59	59	58
Optimized YOLOv4	0.7	MP-IDB	43.01	54	56
		MRC_images	65.07	52	72
Optimized YOLOv5	0.7	MP-IDB	60.7	68.3	66.2
		MRC_images	50.1	53.3	51.2

YOLOv4 and YOLOv5 models achieved higher mAP with an IoU threshold of 0.3, compared to a threshold of 0.5. Conversely, model accuracy decreased significantly at an IoU threshold of 0.7. Notably, even with a more lenient IoU threshold of 0.3, YOLOv4 demonstrated superior accuracy across cross-dataset evaluations. This observation highlights the robustness of YOLOv4 and suggests that YOLOv5 has reduced generalization capabilities, given its moderate accuracy at a lower threshold (Fig. 4). An IoU of 0.3 is generally advantageous for mAP in object detection, as it allows for a lower overlap between the ground truth and predicted bounding boxes, thereby capturing more potential detections. Conversely, an IoU threshold of 0.7 is less effective, leading to a substantial decrease in model accuracy and an increase in false positive predictions.

The contrasting performance observed in the cross-testing images from MRC-UNIMAS reflects the dataset variability.

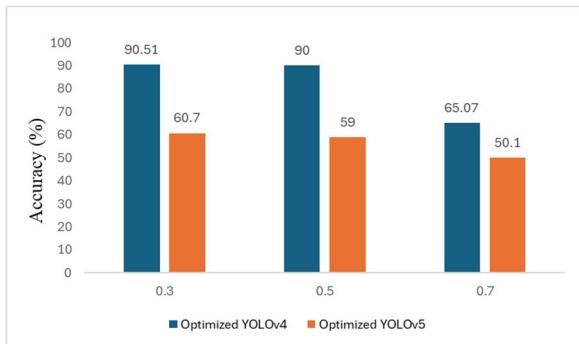


FIGURE 4. Comparisons of YOLOv4 and YOLOv5 mAP on the MRC-UNIMAS Images at various thresholds.

The new dataset introduces variations in imaging conditions that were different from the training images. The images from the new dataset were brighter and had high illumination as compared to the training images (Fig. 1 and 2). The disparity results in YOLOv4 surpassing YOLOv5, suggesting that the architecture of YOLOv4 is more adept at accommodating the distinct attributes that are inherent in the cross-testing dataset.

The visual representation of the identification of infected and non-infected cells obtained from the initial step of the proposed model is depicted in Fig. 5. The cells infected with malaria were isolated and confined within the bounding box, as shown in Fig. 5. The proposed YOLOv4 model demonstrates a notable ability to accurately detect infected cells (indicated by the green arrow) within an image. In contrast, YOLOv5 exhibits limitations in detecting multiple infected cells simultaneously. Remarkably, despite the presence of debris in the image, the optimized YOLOv4 model adeptly identifies and accurately detects the infected cells, as illustrated in Fig. 5.

When considering the detection of malaria-infected cells, the optimized YOLOv4 exhibits several benefits in comparison to the optimized YOLOv5 model. The robustness of the method against variations in cell appearances and staining techniques, which are frequently encountered in images of malaria-infected cells, enables the detection of distinct cell variations in a more consistent manner. Significantly, the proposed optimized YOLOv4 demonstrates an exceptional capability to manage image debris or noise, facilitating precise detection of infected cells amidst congested backgrounds. This capability is particularly critical in real-world microscopy images that are susceptible to artefacts. Furthermore, in situations involving datasets that closely resemble the learned representations, the optimized YOLOv4 exhibits superior performance compared to YOLOv5, indicating its specialized capability to accurately identify particular cell attributes. The model's robust capacity for generalization empowers it to efficiently adjust to novel and unobserved variations that may arise in various datasets. This enhances its versatility in managing a wide range of cell types, stains, and imaging conditions commonly encountered in malaria cell detection tasks. Designed with feature extraction and

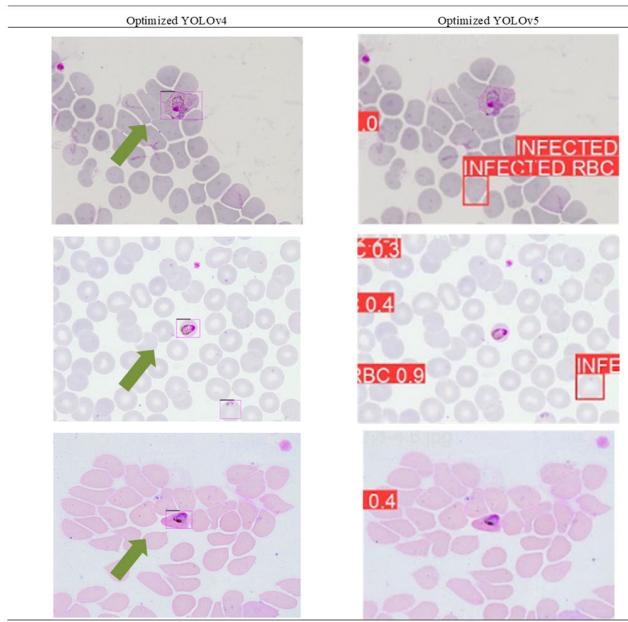


FIGURE 5. The identification and localization of the malaria-infected cells attained from Stage 1 classification.

object scale management in mind, the optimized YOLOv4's architecture significantly contributes to its ability to identify malaria-infected cells of diverse sizes and shapes. As a result, the optimized YOLOv4 was chosen for malaria cell detection tasks on account of its collective strengths, which emphasize adaptability, robustness, and consistent performance across a wide range of datasets.

2) STAGE 2: MULTICLASS CLASSIFICATION OF 4 MALARIA SPECIES

As for the multiclass classification of 4 *Plasmodium* species, Class 0 indicates *P. falciparum*, Class 1 indicates *P. malariae*, Class 2 indicates *P. ovale*, and Class 3 indicates *P. vivax*. As it is a multiclass classification system, thus the precision, recall, specificity, and F1-score of the proposed model are evaluated based on each specific class. The confusion matrices in classifying infected cells into four classes (four species) are shown in Fig. 5. The performance metrics of the trained and optimized CNN models are tabulated in Table 9.

As seen in Fig. 6 and listed in Table 9, all deep learning models obtained more significant than 90% accuracy. However, the suggested model's performance attributes, such as precision, recall, specificity, and F1-score, vary for each class. Since precision reflects a model's accuracy in identifying a sample as positive, all presented models have great precision in predicting Class 0. (*P. falciparum*) and Class 1 (*P. malariae*), with DenseNet-121 attaining 92.68% and 100% precision, respectively. As for Class 2 (*P. ovale*) AlexNet attained the highest precision of 97.59, whereas the MobileNet attained the highest accuracy of 97.21 for Class 3 (*P. vivax*).

All presented models have demonstrated exceptional recall rates for all class predictions, indicating a high level of sensitivity. Nevertheless, the VGG-16 model achieved a recall rate of just 85.33% for class 3, specifically for *P. vivax*. All proposed models demonstrated excellent performance in terms of specificity for each class, with values over 90%. Notably, DenseNet-121 earned the highest specificity of 100% for Classes 1 and 2, respectively. The F1-score is a metric that quantifies the harmonic mean of precision and recall, serving as an alternative to accuracy for assessing the performance of a model. The F1-score is a suitable metric for achieving a trade-off between precision and recall, particularly in imbalanced class distribution. Based on the obtained data, it can be observed that the F1-score of all the proposed models for Classes 0 to 3 exhibit a consistently high level of performance. In addition, the DenseNet-121 model achieved the highest score among all the models for Classes 0 and 1. In the case of Classes 2 and 3, AlexNet demonstrated the highest F1-score of 97.0% and MobileNetv2 with a score of 95.87% respectively. A high F1 score for a particular class prediction signifies the model's strong performance, characterized by high recall and precision.

Based on the results, the MobileNetv2 model achieved the highest overall accuracy of 95.67%, which is 0.12% higher than that of the DenseNet-121 model. However, it is crucial to evaluate the performance of these models across all classes to ensure that the overall accuracy is not disproportionately influenced by performance in only a few classes. The DenseNet-121 outperforms the MobileNetv2 with the highest F1-score in Class 0 and Class 1. Whereas, MobileNetv2 has a higher F1-score in Class 2 and Class 3. The DenseNet-121 model demonstrated higher precision than MobileNetv2 in classifying Class 0, Class 1, and Class 2. High precision across all classes is essential for accurately classifying cells according to their respective categories. Analysis of the confusion matrices reveals that the DenseNet-121 model correctly classified 12 more cells than MobileNetv2 in Class 1 and Class 2 (Fig. 6), whereas MobileNetv2 has shown better performance only in Class 2 and that has mainly contributed to its high accuracy.

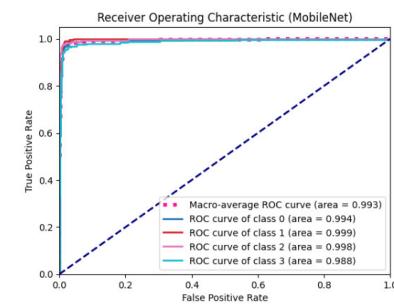
In a multi-class classification, evaluating the performance of the models across different classes is crucial to ensure that overall accuracy is not skewed by classes with more instances. Macro-averaged metrics such as precision, recall, and F1-score are calculated by averaging the scores of individual classes. This approach treats all classes equally, regardless of their frequency, and provides a balanced view of the model performance across all classes. When comparing the macro-averaged accuracies for DenseNet-121 and MobileNetv2, both models show similar individual accuracies for Class 1, Class 2, and Class 3 (Fig. 7). Notably, DenseNet-121 achieves a perfect accuracy of 100% for Class 1, while MobileNetv2 achieves 99.9%. This difference is mainly due to DenseNet-121's superior precision in classifying Class 2, despite having a lower recall rate.

	ResNet-50				VGG-16					
	0	1	2	3	0	1	2	3		
True class	0	155	0	0	1	0	155	0	0	1
	1	7	193	1	2	1	8	194	1	0
	2	0	0	156	12	2	3	2	157	6
	3	14	3	4	238	3	27	2	9	221

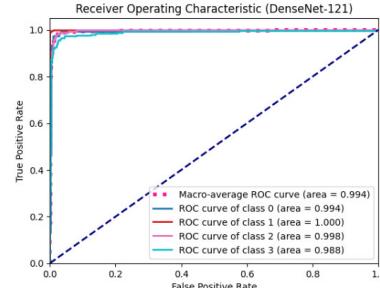
	MobileNet-v2				EfficientNet					
	0	1	2	3	0	1	2	3		
True class	0	154	1	0	1	0	153	2	0	1
	1	11	191	1	0	1	13	189	1	0
	2	0	0	163	5	2	0	1	132	35
	3	9	1	5	244	3	16	1	3	229

	DenseNet-121				AlexNet					
	0	1	2	3	0	1	2	3		
True class	0	152	0	2	2	0	155	0	0	1
	1	4	199	0	0	1	14	185	3	1
	2	0	0	152	16	2	1	1	162	5
	3	8	0	3	248	3	13	1	1	244

FIGURE 6. 4 × 4 Confusion matrix of ResNet-50, VGG-16, DenseNet-121 and AlexNet, MobileNet-v2 and EfficientNet.



(a)



(b)

FIGURE 7. Macro-average accuracies of MoblineNetv2 (a) and DenseNet-121 (b) across Classes.

As tabulated in Table 9, all models scored highly in most performance parameters. EfficientNet has the lowest accuracy, precision, recall, specificity, and F1-score, while DenseNet-121 and MobileNetv2 have the highest performance. Although the EfficientNet is shallower than the other CNN models besides MobileNetv2, it did not outperform the other deeper models. This indicates that a shallower model does not guarantee its excellence in a classification problem. Such as the architecture of the EfficientNet focusses on optimizing the trade-off between accuracy and efficiency without necessarily increasing its depth. As a result, its performance is balanced across these dimensions without necessarily achieving the highest accuracy compared

TABLE 9. Summary of performance results of proposed two-stage on classifying infected blood cells into four species.

	Performance metrics (%)	Proposed model				
		Mobi leNet	Efficient Net	ResNet-50	VGG-16	DenseNet-121
Class 0	Accuracy	95.67	90.71	94.40	92.49	95.55
	Precision	88.51	84.07	88.0	80.31	92.68
	Recall / Sensitivity	98.72	98.08	99.36	99.36	97.44
	Specificity	96.82	95.39	96.67	93.97	98.09
Class 1	F1-score	93.33	90.53	93.37	88.83	95.00
	Precision	98.96	97.93	98.47	97.98	100.00
	Recall / Sensitivity	94.09	93.10	95.07	95.57	98.08
	Specificity	99.56	99.31	99.48	99.31	100.00
Class 2	F1-score	96.46	95.45	96.74	96.76	99.00
	Precision	96.45	97.06	96.89	94.01	96.82
	Recall / Sensitivity	97.02	78.57	92.86	93.45	90.48
	Specificity	99.02	99.35	99.19	98.38	99.19
Class 3	F1-score	96.74	86.84	94.83	93.73	93.54
	Precision	97.60	86.91	94.07	96.93	93.23
	Recall / Sensitivity	94.21	92.28	91.89	85.33	95.75
	Specificity	98.86	93.16	97.15	98.67	96.58
F1-score		95.87	89.51	92.97	90.76	94.48
						95.69

to deeper models. In contrast, the DenseNet-121's greater depth and dense connectivity can significantly contribute to a multi-class classification problem.

Given the results of the CNN models, the DenseNet-121 model indeed outperforms the other models in terms of its overall accuracy and generalization across the malaria species. This study focuses on achieving the highest possible accuracy; therefore, the DenseNet-121 model is chosen as the best-performing model given its consistent performance across classes. If the primary goal of the study is efficiency and speed, the MobileNetv2 can be a potential selection of model as it is computationally efficient for deployment on mobile or embedded devices. However, although the DenseNet-121 is deeper than the MobileNetv2, the DenseNet-121 model still has a small size of 27.2Mb and G-FLOPs of 2.89 only that is relatively lower than other shallower models such as ResNet-50 that has a size of 87.8 and G-FLOPs of 4.13. This indicates that although DenseNet-121 is deeper it is indeed computationally efficient and has an efficient design that leverages the use of computational resources.

According to Sarkar, et al. [44], DenseNet-121 uses dense connectivity, unlike ResNet-50, which uses sparse connectivity between layers [39], [45]. Each layer of DenseNet-121 receives the preceding feature maps from former layers as the input, and all its output feature maps are passed into their subsequent layers. DenseNet-121 has various advantages: it strengthens feature propagation, encourages feature reuse, leads to more diversified features with richer patterns, and

reduces vanishing problems. Due to the architecture and those advantages, DenseNet-121 tends to provide smoother decision boundaries [46], [47]. The DenseNet-121 has achieved consistent results, although not the highest in all performance parameters (precision, recall, specificity, and F1-score).

While most deep learning models have demonstrated comparable predictive analytical capabilities, the feasibility of a model is contingent upon its ability to accurately classify unlabeled and randomly occurring species of malaria parasites visible in images of thin blood smears. The end-to-end deep learning model was presented and evaluated on publicly available and previously unobserved datasets from MRC, UNIMAS. Including images in the dataset ensures that all classes of malaria are evenly represented, hence allowing for the assessment of the proposed models' ability to generalize. The DenseNet-121 model was chosen as the most suitable representation for the Stage 2 classification task due to its consistent precision and good performance across most classes.

In addition, we have performed a time complexity analysis to quantify the processing time of a set of algorithms relative to the amount of input. Fig. 8 compares deep learning models' training and testing times in malaria species classification. As the number of parameters and weight increases, the training time of the deep learning models increases. The time complexity is evaluated based on the training time of the deep learning models used in this study for the number of epochs of 25.

According to Fig. 8, the overall training time for all deep learning models in classifying the four malaria species requires more than one hour, whereas AlexNet trained the shortest. For the training time of classification systems, VGG-16 has the highest time consumption compared to other models. VGG-16 takes around 112 minutes for the malaria species classification. According to Krizhevsky et al. [48], VGG-16 has reduced filter size and increased the network depths, producing more accurate results with about 138 million trainable parameters. Besides, this study has a weight file size of about 524MB, considered an extensive network, which explains why the VGG-16 has the longest training time compared to other models.

In contrast, AlexNet requires 67 minutes of training time, while the longest testing time per image is 125 milliseconds. According to Krizhevsky, et al. [48], AlexNet is a shallow CNN comprising eight layers (5 convolutional layers and three fully connected layers) with about 60 million trainable parameters. However, with complex image patterns of malaria species, AlexNet requires more time to be computed for the testing phase.

Moreover, the DenseNet-121 has a slightly higher training time than ResNet-50 in this study. However, DenseNet has been shown to outperform ResNet in feature use efficiency, and DenseNet has fewer parameters than ResNet [44]. Besides, DenseNet-121 has a weight file size of about 27 M.B., much smaller than the weight file size of ResNet-50 of 97 M.B. in this study. However, due to the dense



FIGURE 8. Training and testing time comparisons of proposed deep learning models in malaria species identification.

concatenation features of DenseNet, it requires higher GPU memory than ResNet [49]. Moreover, DenseNet uses many small convolutions in the network, causing it to run slower on GPU, leading to more training time. Thus, DenseNet-121 will have a longer training time than ResNet-50, although DenseNet-121 has fewer trainable parameters and smaller weight files.

Regarding the duration of testing for the proposed deep learning models in classifying infected cells into four distinct classes based on species, the time expended is relatively consistent. This timeframe spans almost over 100 milliseconds for all models except for the VGG-16 model, which exhibits the shortest testing time of 66 milliseconds. However, in terms of accuracy, this model demonstrates the lowest level of precision at 92.49%. Based on the findings of this study, there is no statistically significant variation in the testing duration among the recommended deep learning models for each categorization system. The models successfully achieved detection and classification tasks within less than one minute, exhibiting remarkable potential for their clinical relevance.

IV. DISCUSSION

In this study, we demonstrated the capability of end-to-end machine learning for the automated identification of *Plasmodium* parasites in thin blood smear images. The developed system is divided into two main stages where the first stage focuses on the localization of the malaria-infected red blood cells. Then, in the second detection stage, we classified the *Plasmodium* parasites according to the four species with different proposed deep-learning approaches. The performance estimates of all models can correctly discriminate the species. We summarized the contributions of this study are summarized as follows:

A. AUTOMATED IDENTIFICATION OF MALARIA-INFECTED CELLS AND CLASSIFICATION OF HUMAN MALARIA PARASITES USING A TWO-STAGE DEEP LEARNING TECHNIQUE

This study introduces an innovative framework, “Automated Identification of Malaria-infected Cells and Classification of

Human Malaria Parasites Using a Two-stage Deep Learning Technique,” to revolutionize malaria diagnosis and species identification. Initially, the framework precisely localizes malaria-infected red blood cells within complex thin blood smear images, even when cells overlap using YOLOv4 architecture. This localization lays the groundwork for the second stage, in which sophisticated deep-learning techniques were tested to classify *Plasmodium* parasites into different species, such as *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*. The ResNet-50, VGG-16, DenseNet-121, AlexNet, MobileNetV2, and EfficientNet deep learning models were developed, and their performance is analyzed in terms of their accuracy, recall, precision, and specificity.

Findings indicate that the YOLOv4 model can accurately detect the infected cells within the whole blood smear images with a mAP of 89%. The YOLOv4 model has shown supremacy against the YOLOv5 model on the independent dataset indicating its robustness and generalization capability. On the other hand, the DenseNet-121 is identified as the best model in classifying the malaria-infected cells into four aforementioned species with an accuracy of 95.5%, other performance measurements as tabulated in Table 9. No extensive preprocessing is required to increase the image quality using the suggested deep-learning approach.

Prior works have employed deep-learning object detectors to detect the infected cells within the whole thin blood smear images (Table 10). In comparison, a study by Yang et al. [19] utilized the YOLOv2 model to localize the *P. vivax*-infected cells. However, this model achieved only an accuracy of 71.34%, which raises concerns about the capability of the model to generalize across different types of malaria infections. Recent advancements in YOLO models have led to investigations into newer models such as YOLOv5 and YOLOv8. For instance, Zedda et al. [50] achieved an accuracy of 84.6% in detecting the *Pfalciparum*-infected cells from the MP-IDB images using the YOLOv5. Similarly, Liu et al. [51] employed YOLOv5 for infected cell detection and achieved an accuracy of 90.8%. Zedda et al. [52] employed YOLOv8 to detect the *P. vivax*-infected cells and attained an accuracy of 85.9%. However, upon testing on an independent dataset of the same parasitic infection, the model's accuracy significantly reduced to 56.2% indicating limited generalizability and robustness of the model. Ozbilge et al. [53] employed YOLOv8 to achieve a 90.31% accuracy in detecting cells specifically at the ring stage of infection. This accuracy is limited, as the model is only detecting cells from one stage of infection and there are fewer variations for the models to detect. In a more recent study, Xiong and Wu [54] employed the MAS-Net model, an adaptation of YOLOv5, which achieved an accuracy of 75.9% for detecting *P. vivax*-infected cells.

Newer models like YOLOv5 [50], [51], [54] and YOLOv8 [52], [53] have demonstrated lower robustness and generalization on malaria datasets compared to the YOLOv4 utilized in this study. Additionally, many of these studies focus on detecting cells infected by only one or two types of malaria

TABLE 10. Performance comparison of the object detectors on prior published works for infected cell prediction from thin blood smear images.

References	Datasets	Species	Model	mAP(%)
Hung et al. [13]	Images containing 100,000 cells	<i>P. vivax</i>	Faster R-CNN	59
Yang, Quizon, et al. [19]	2567 images	<i>P. vivax</i>	YOLOv2	71.34
Zedda et al. [50]	MP-IDB	<i>P. falciparum</i>	YOLOv5	84.6
Zhao et al. [28]	1,364 images from Broad Institute	<i>P. vivax</i>	SSD	90.4
R. C. Liu et al. [51]	150 images from SmartMalariaNet	Not mentioned	YOLOv5	90.8
Zedda et al. [52]	MP-IDB	<i>P. falciparum</i>	YOLOv8	MP-IDB images: 85.9 IML dataset: 56.3
Ozbilge et al. [53]	Images from 25 patients, 21 diagnosed with <i>P. falciparum</i> , 3 with <i>P. vivax</i> and 1 with <i>P. ovale</i>	<i>P. falciparum</i> <i>P. vivax</i> <i>P. ovale</i>	YOLOv8	90.31
Xiong et al. [54]	National Institutes of Health (NLM)	<i>P. vivax</i>	MAS-Net	75.9
This study		<i>P. falciparum</i> <i>P. vivax</i> <i>P. ovale</i> <i>P. malaria</i>	YOLOv4	MP-IDB images: 89 MRC-UNIMAS images: 90

parasites (Table 10). The moderate accuracies achieved in detecting cells infected by a single malaria parasite, which typically exhibits similar morphologies, raise concerns about the models' ability to accurately identify cells infected by different malaria parasites with distinct morphologies.

In contrast, in this study, the YOLOv4 model demonstrated impressive performances of 89% and 90% on the source and independent datasets respectively containing images of four types of malaria parasites from various stages of infections. YOLOv4 has set benchmarks in various classification tasks across multiple studies. The YOLOv4 model has undergone extensive training and testing by the developers, offering a high level of robustness. Therefore, despite the advancements in YOLO models, analyzing the performance of YOLOv4 in malaria diagnosis remains valuable.

With the two-stage deep learning architecture, capabilities in learning complex features, and performing more intensive computational tasks, manual feature extraction is not required in the proposed predictive model, resulting in varying-quality images. This framework presents a novel and highly effective strategy for addressing the complex challenges associated with malaria-infected cell detection and species classification, particularly in overlapping cells and multi-species infections. Combining precise localization

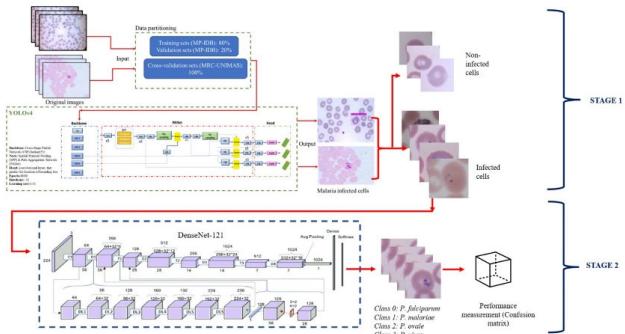


FIGURE 9. The proposed overall framework of two-stage end-to-end deep learning malaria classification.

with species-specific classification considerably improves the accuracy and dependability of malaria diagnosis, supplanting conventional methods. In addition, the framework's adaptability and fine-tuning capabilities enable researchers and healthcare professionals to tailor the model to specific datasets, making it applicable in various clinical and geographical contexts. Consequently, this innovative two-stage deep learning technique not only enhances our comprehension of automated malaria diagnosis but also promises to revolutionize the field by providing a potent tool for enhancing healthcare outcomes and contributing to the ongoing fight against this fatal disease. The complete proposed framework is shown in Fig. 9.

B. MULTI-SPECIES IDENTIFICATION AND PRESENCE OF OVERLAPPING CELLS

During the first stage of the proposed model, the main goal is to localize red blood cells infected with the malaria parasite. The localization process entails the determination of the exact borders of each infected cell in the picture. The algorithm is trained to accurately recognize and distinguish individual infected cells in the presence of overlapping cells. This process efficiently separates the overlapping cells (Fig. 10 (a)), even when they are entangled or partly concealed by one other. The model determines the image regions containing overlapping cells as part of the localization procedure. In subsequent phases, a more in-depth analysis is conducted on these regions. The ability of the model to precisely isolate these overlapping regions is crucial for managing multi-infection scenarios effectively, as shown in the example in Fig. 10 (b) and (c). After identifying infected cells, the system continues to categorize the *Plasmodium* parasites into distinct species during the subsequent phase. The second step of the process derives advantages from the accurate localization accomplished in the first stage. Despite overlapping cells, the model has successfully established distinct borders for each. This feature enables the model to categorize individual cells autonomously, considering their respective species, without encountering confusion or misclassification due to overlapping structures. By doing localization before

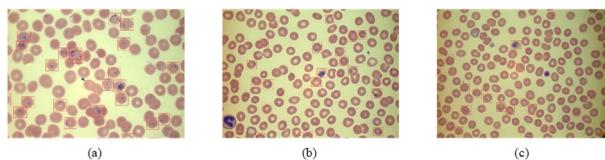


FIGURE 10. Example of identification of malaria-infected cells on (a) overlapping red blood cells (b) and (c) Multi-infection where *P. falciparum* and *P. malariae* are detected in one sample.

classification sequentially, the system enhances the accuracy and dependability of the species classification procedure. The issue of overlapping cells, which might have presented difficulties in conventional single-stage classification approaches, has now been effectively resolved with enhanced accuracy, resulting in a substantial decrease in the likelihood of misdiagnosis. Implementing this two-stage strategy is highly pragmatic in clinical situations characterized by the frequent occurrence of overlapping red blood cells. This technology automatically identifies and categorizes malaria parasites, even when cells are densely concentrated or intricately entwined. This advancement significantly enhances the overall efficacy and precision of malaria diagnosis.

C. LOCALIZATION OF MALARIA-INFECTED CELLS AND HUMAN MALARIA SPECIES CLASSIFICATION

To date, limited studies on species identifications have been reported. Predicting and identifying the species of human malaria parasites is essential for administering the correct treatment to a patient and preventing overtreatment and incorrect usage of anti-malarial drugs on patients. The proposed two-stage deep learning model will enable rapid and reliable diagnosis by localizing contaminated red blood cells in the first stage and then categorizing them into various species in the subsequent stage. This proposed solution proves to help facilitate accurate diagnosis of malaria. This capacity immediately results in treatment choices that are more customized and efficient. The proposed model's ability to differentiate between *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae* guarantees that individuals are administered the most suitable anti-malarial drugs, reducing the likelihood of drug resistance and severe consequences. Different malaria species necessitate individualized treatment strategies to minimize patient fatality risks. For example, infections caused by *P. falciparum* have the potential to advance swiftly and often need immediate intervention.

On the other hand, *P. vivax* can establish latent liver stages, necessitating particular treatment strategies. Precise species identification plays a crucial role in enabling healthcare practitioners to effectively administer suitable drugs, closely monitor patient progress, and proactively avoid the occurrence of severe sequelae. This, in turn, leads to enhanced efficacy in malaria treatment, decreased fatality rates, and better overall patient outcomes.

Based on the published works, there is a scarcity of research that has used the same dataset and employed a classification system for human malaria species. Only a few

published results [30], [54], [55], as summarized in Table 11, have categorized human malaria species into four distinct groups using DL techniques. Most of the published literature has been on detecting malaria-infected cells utilizing single-cell analysis, particularly on the *P. falciparum* species. In a previous study by Malihi, et al. [20], machine learning models were used to classify the four species of human malaria. This research necessitated the execution of pre-processing tasks before the classification process. Specifically, the Canny edge detection and Otsu thresholding techniques were implemented, which required human intervention to select the most optimal threshold for efficiently segmenting the infected cells.

The same MP-IDB public dataset was employed in the recent works of Krishnadas et al. [30]. The four human malaria species were identified using a YOLOv5 and a scaled YOLOv4 object detection instrument. Even though four human malaria species were used, model efficacy was only compared based on its ability to detect the object of interest (i.e., the four human malaria species). Thus, 78% of mAP was achieved in identifying the infected cells. The results also indicated that their proposed techniques are capable of classifying malaria stages. However, due to a paucity of sample size at each stage, the mAP for their work was only 39.9%.

Similar to Krishnadas et al. [30], Guemas et al. [56] compared the performances of YOLOv5 and YOLOv8 with the real-time detection transformer (RT-DETR) for malaria species classification. Their study revealed that the YOLOv5 and YOLOv8 models outperform the RT-DETR on the test images with mAPs of 67% and 72.7% respectively. Whereas the RT-DETR achieved a mAP of 63.8% on the same dataset. Analysis of individual class accuracies further highlighted the limitations of RT-DETR as it only achieved mAPs of 19.9% and 15% for *P. ovale* and *P. malariae* classification. This indicates that RT-DETR exhibits significant bias towards certain classes, as evidenced by its comparatively poor performance on the other two malaria classes. The accuracies obtained from object detectors for species classifications are relatively lower than the accuracies attained in this study. Consequently, the results exhibit that using the CNN models for the multi-class classification is relatively more effective than using the DL object detectors. The DL object detectors are more applicable and practical for disease diagnosis, whereas the CNN models for multi-class species classification as suggested in this study.

Besides that, notably, the study by Garba et al. [55] attained an accuracy of 99.9% in classifying the human-only *Plasmodium* species. However, upon analyzing the classification results by species, the customized CNN attained an accuracy of 64.6%, 39%, and 33.73% for the classes of *P. malariae*, *P. ovale*, and *P. vivax* respectively, with 99.97% for the *P. falciparum* classification. This shows that their model is more biased towards the classification of *P. falciparum* and the high accuracy is mainly contributed by the *P. falciparum* class only.

TABLE 11. Performance comparison of the proposed model with the other published works for malaria species classification.

References	Techniques	Species classification	Significant findings
Malihi, et al. [20]	SVM, NM, KNN, 1-NN, Fisher	<i>P.falciparum</i> <i>P.vivax</i> <i>P.ovale</i> <i>P.malariae</i>	KNN attained best performance with accuracy of 91%
Krishnadas, et al. [30]	YOLOv5 and Scaled YOLOv4	<i>P.falciparum</i> <i>P.vivax</i> <i>P.ovale</i> <i>P.malariae</i>	Parasite classification based on object detection model: Precision -74.1% Recall - 79.7% mAP - 78%
Garba et al. [55]	Customized CNN	<i>P.falciparum</i> <i>P.vivax</i> <i>P.malariae</i> <i>P.ovale</i>	Stage classification: Precision – 44.7% Recall – 55.6% mAP – 39.9%
Guemas et al. [56]	Real-time detection transformer (RT-DETR_)	<i>P.falciparum</i> <i>P.vivax</i> <i>P.malariae</i> <i>P.ovale</i>	mAP: 63.8%
Proposed model	Two-stage deep learning technique First Stage: YOLOv5 Second Stage: DenseNet-121	<i>P.falciparum</i> <i>P.vivax</i> <i>P.ovale</i> <i>P.malariae</i>	First stage classification (infected & non-infected cells) mAP – 97.7% Precision – 95.3% Recall – 93.6% Second-stage human malaria classification Accuracy - 95.5%

In the meantime, we proposed a two-stage deep learning framework to address the shortcomings of previously published works. This methodology was developed and evaluated on the public dataset of MP-IDB, and the supplementary samples gathered from MRC-UNIMAS, Sarawak. The first stage of the proposed work focused on identifying and localizing malaria-infected cells using YOLOv4 architecture, while the second stage classified malaria-infected cells into their four species using DenseNet-121 architecture. Our work achieved a mAP of 89% for the first stage of identifying malaria-infected cells and a classification accuracy of 95.5% for malaria species.

V. CONCLUSION

In malaria diagnosis, the automated identification of malaria-infected cells and classification of human malaria parasites using a two-stage deep learning technique marks a significant milestone. This innovative framework, which incorporates a two-stage method of accurate localization and species categorization, has demonstrated its potential, bringing in a new era of precision and dependability in malaria detection, especially in the face of challenging scenarios, including overlapping cells. The implications for clinical practice are significant since this unique methodology provides healthcare providers with a powerful tool for making personalized treatment decisions, resulting in

improved patient outcomes. It becomes increasingly evident that the malaria landscape is dynamic and continuously evolving. Emerging malaria species, such as *P. knowlesi*, highlight the urgent need for adaptable diagnostic strategies. The capacity of this framework to adapt and accommodate newly emergent malaria strains is a crucial area that requires further investigation. In addition, classifying parasite stages, such as early trophozoites, schizonts, and gametocytes, bears tremendous promise for a more exhaustive disease assessment. Findings from this research have achieved good performance in detecting malaria-infected cells with mAP of 89% using an optimized YOLOv4. Meanwhile, regarding species classification, the proposed DenseNet-121 architecture has successfully identified complex patterns of malaria species with an overall accuracy of 95.5%. This study serves as both evidence of the capabilities of deep learning in the healthcare field and as a catalyst for continued advancements in malaria diagnosis. Through a continuous process of improving our methodologies and embracing rising obstacles, it is possible to advance the field of malaria detection and treatment, making significant contributions to the worldwide endeavor of eradicating this very destructive illness and saving innumerable lives.

DATA AVAILABILITY

The datasets generated during and/or analyzed during the current study are available in the Malaria Parasite Image Database (MP-IDB) GitHub repository, <https://github.com/andrealoddo/MP-IDB-The-Malaria-Parasite-Image-Database-for-Image-Processing-and-Analysis>. The dataset from MRC-UNIMAS, Sarawak that supports the findings of this study are available from the corresponding author, upon reasonable request.

REFERENCES

- [1] A. Z. Chin, M. C. M. Maluda, J. Jelip, M. S. B. Jeffree, R. Culleton, and K. Ahmed, "Malaria elimination in Malaysia and the rising threat of plasmodium knowlesi," *J. Physiological Anthropology*, vol. 39, no. 1, p. 36, Nov. 2020, doi: [10.1186/s40101-020-00247-5](https://doi.org/10.1186/s40101-020-00247-5).
- [2] Y. W. Lee, J. W. Choi, and E.-H. Shin, "Machine learning model for predicting malaria using clinical information," *Comput. Biol. Med.*, vol. 129, Feb. 2021, Art. no. 104151, doi: [10.1016/j.compbiomed.2020.104151](https://doi.org/10.1016/j.compbiomed.2020.104151).
- [3] J. Nadia and F. Lu, "Historical experiences on mass drug administration for malaria control and elimination, its challenges and China's experience: A narrative review," *Acta Tropica*, vol. 225, Jan. 2022, Art. no. 106209, doi: [10.1016/j.actatropica.2021.106209](https://doi.org/10.1016/j.actatropica.2021.106209).
- [4] World Health Organization. (2021). *World Malaria Report*. Accessed: Jun. 4, 2024. [Online]. Available: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>
- [5] E. Charpentier, E. Benichou, A. Pages, P. Cahuvin, J. Filliaux, A. Valentin, H. Guegan, E. Guemas, A. S. Salabert, C. Armengol, S. Menard, S. Cassaing, A. Berry, and X. Iriart, "Performance evaluation of different strategies based on microscopy techniques, rapid diagnostic test and molecular loop-mediated isothermal amplification assay for the diagnosis of imported malaria," *Clin. Microbiol. Infection*, vol. 26, no. 1, pp. 115–121, Jan. 2020, doi: [10.1016/j.cmi.2019.05.010](https://doi.org/10.1016/j.cmi.2019.05.010).
- [6] World Health Organization. (2011). *Guidelines for the Treatment of Malaria*. Accessed: Jun. 4, 2024. [Online]. Available: <https://iris.who.int/handle/10665/162441>
- [7] World Health Organization. (2016). *Malaria Microscopy Quality Assurance Manual*. Accessed: Jun. 4, 2024. [Online]. Available: <https://www.who.int/publications/item/9789241549394>

- [8] M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. Thoma, "Image analysis and machine learning for detecting malaria," *Transl. Res.*, vol. 194, pp. 36–55, Apr. 2018, doi: [10.1016/j.trsl.2017.12.004](https://doi.org/10.1016/j.trsl.2017.12.004).
- [9] J. Frean, "Microscopic determination of malaria parasite load: Role of image analysis," *Microscopy, Sci., Technol., Appl. Educ.*, vol. 3, pp. 862–866, Jan. 2010.
- [10] M. Maity, A. K. Maity, P. K. Dutta, and C. Chakraborty, "A web-accessible framework for automated storage with compression and textural classification of malaria parasite images," *Int. J. Comput. Appl.*, vol. 52, no. 15, pp. 31–39, Aug. 2012.
- [11] S. Moon, S. Lee, H. Kim, L. H. Freitas-Junior, M. Kang, L. Ayong, and M. A. E. Hansen, "An image analysis algorithm for malaria parasite stage classification and viability quantification," *PLoS ONE*, vol. 8, no. 4, Apr. 2013, Art. no. e61812, doi: [10.1371/journal.pone.0061812](https://doi.org/10.1371/journal.pone.0061812).
- [12] N. Ahirwar, S. Pattnaik, and B. Acharya, "Advanced image analysis based system for automatic detection and classification malarial parasite in blood images," *Int. J. Inf. Technol. Knowl. Manage.*, vol. 5, no. 1, pp. 59–64, Jun. 2012.
- [13] J. Hung, S. C. P. Lopes, O. A. Nery, F. Nosten, M. U. Ferreira, M. T. Duraisingh, M. Marti, D. Ravel, G. Rangel, B. Malleret, M. V. G. Lacerda, L. Rénia, F. T. M. Costa, and A. E. G. Carpenter, "Applying faster R-CNN for object detection on malaria images," in *Proc. Comput. Vis. Pattern Recognit. Workshops*, Honolulu, HI, USA, Jul. 2017, pp. 808–813.
- [14] F. B. Tek, A. G. Dempster, and I. Kale, "Parasite detection and identification for automated thin blood film malaria diagnosis," *Comput. Vis. Image Understand.*, vol. 114, no. 1, pp. 21–32, Jan. 2010, doi: [10.1016/j.cviu.2009.08.003](https://doi.org/10.1016/j.cviu.2009.08.003).
- [15] M. J. Mammoottil, L. J. Kulangara, A. S. Cherian, P. Mohandas, K. Hasikin, and M. Mahmud, "Detection of breast cancer from five-view thermal images using convolutional neural networks," *J. Healthcare Eng.*, vol. 2022, pp. 1–15, Feb. 2022, doi: [10.1155/2022/4295221](https://doi.org/10.1155/2022/4295221).
- [16] Y. S. Leong, K. Hasikin, K. W. Lai, N. Mohd Zain, and M. M. Azizan, "Microcalcification discrimination in mammography using deep convolutional neural network: Towards rapid and early breast cancer diagnosis," *Frontiers Public Health*, vol. 10, Apr. 2022, Art. no. 875305, doi: [10.3389/fpubh.2022.875305](https://doi.org/10.3389/fpubh.2022.875305).
- [17] A. Nanoti, S. Jain, C. Gupta, and G. Vyas, "Detection of malaria parasite species and life cycle stages using microscopic images of thin blood smear," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, vol. 1, Coimbatore, India, Aug. 2016, pp. 1–6.
- [18] H. Telang and K. Sonawane, "COVID-19 and malaria parasite detection and classification by bins approach with statistical moments using machine learning," *Int. J. Image, Graph. Signal Process.*, vol. 15, no. 3, pp. 1–13, Jun. 2023, doi: [10.5815/ijigsp.2023.03.01](https://doi.org/10.5815/ijigsp.2023.03.01).
- [19] F. Yang, N. Quizon, H. Yu, K. Silamut, R. J. Maude, S. T. Jaeger, and S. Antani, "Cascading YOLO: Automated malaria parasite detection for Plasmodium vivax in thin blood smears," *Proc. SPIE*, vol. 11314, pp. 404–410, Feb. 2020.
- [20] L. Malih, K. Ansari-Asl, and A. Behbahani, "Malaria parasite detection in giemsa-stained blood cell images," in *Proc. 8th Iranian Conf. Mach. Vis. Image Process. (MVIP)*, Zanjan, Iran, Sep. 2013, pp. 360–365.
- [21] G. Díaz, F. A. González, and E. Romero, "A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images," *J. Biomed. Informat.*, vol. 42, no. 2, pp. 296–307, Apr. 2009, doi: [10.1016/j.jbi.2008.1.1005](https://doi.org/10.1016/j.jbi.2008.1.1005).
- [22] N. Hallfors, A. Shanti, J. Sapudom, J. Teo, G. Petroianu, S. Lee, L. Planelles, and C. Stefanini, "Multi-compartment lymph-node-on-a-chip enables measurement of immune cell motility in response to drugs," *Bioengineering*, vol. 8, no. 2, p. 19, Jan. 2021.
- [23] D. Romano, G. Rossetti, and C. Stefanini, "Learning on a chip: Towards the development of trainable biohybrid sensors by investigating cognitive processes in non-marine ostracoda via a miniaturised analytical system," *Biosyst. Eng.*, vol. 213, pp. 162–174, Jan. 2022, doi: [10.1016/j.biosystemseng.2021.11.004](https://doi.org/10.1016/j.biosystemseng.2021.11.004).
- [24] S. J. Harper, L. I. Ward, and G. R. G. Clover, "Development of LAMP and real-time PCR methods for the rapid detection of *Xylella fastidiosa* for quarantine and field applications," *Phytopathology*, vol. 100, no. 12, pp. 1282–1288, Dec. 2010, doi: [10.1094/phyto-06-10-0168](https://doi.org/10.1094/phyto-06-10-0168).
- [25] Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. A. Hossain, A. Sameer, R. J. Maude, J. X. Huang, S. Jaeger, and G. Thoma, "CNN-based image analysis for malaria diagnosis," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, Shenzhen, China, Dec. 2016, pp. 493–496.
- [26] Q. A. Arshad, M. Ali, S.-U. Hassan, C. Chen, A. Imran, G. Rasul, and W. Sultani, "A dataset and benchmark for malaria life-cycle classification in thin blood smear images," *Neural Comput. Appl.*, vol. 34, no. 6, pp. 4473–4485, Nov. 2021, doi: [10.1007/s00521-021-06602-6](https://doi.org/10.1007/s00521-021-06602-6).
- [27] S. Rajaraman, S. K. Antani, M. Poostchi, K. Silamut, M. A. Hossain, R. J. Maude, S. Jaeger, and G. R. Thoma, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images," *PeerJ*, vol. 6, p. e4568, Apr. 2018, doi: [10.7717/peerj.4568](https://doi.org/10.7717/peerj.4568).
- [28] O. S. Zhao, N. Kolluri, A. Anand, N. Chu, R. Bhavaraju, A. Ojha, S. Tiku, D. Nguyen, R. Chen, A. Morales, D. Valliappan, J. P. Patel, and K. Nguyen, "Convolutional neural networks to automate the screening of malaria in low-resource countries," *PeerJ*, vol. 8, p. e9674, Aug. 2020, doi: [10.7717/peerj.9674](https://doi.org/10.7717/peerj.9674).
- [29] N. K. Chauhan and K. Singh, "A review on conventional machine learning vs deep learning," in *Proc. Int. Conf. Comput., Power Commun. Technol. (GUCCON)*, Uttar Pradesh, India, Sep. 2018, pp. 347–352.
- [30] P. Krishnadas, K. Chadaga, N. Sampathila, S. Rao, and S. Prabhu, "Classification of malaria using object detection models," *Informatics*, vol. 9, no. 4, p. 76, Sep. 2022, doi: [10.3390/informatics9040076](https://doi.org/10.3390/informatics9040076).
- [31] A. Loddio, C. Fadda, and C. Di Ruberto, "An empirical evaluation of convolutional networks for malaria diagnosis," *J. Imag.*, vol. 8, no. 3, p. 66, Mar. 2022, doi: [10.3390/jimaging8030066](https://doi.org/10.3390/jimaging8030066).
- [32] C. Daneshvar, T. M. Davis, J. Cox-Singh, M. Z. Rafa'ee, S. K. Zakaria, P. C. Divis, and B. Singh, "Clinical and parasitological response to oral chloroquine and primaquine in uncomplicated human plasmodium knowlesi infections," *Malaria J.*, vol. 9, no. 1, p. 238, Aug. 2010, doi: [10.1186/1475-2875-9-238](https://doi.org/10.1186/1475-2875-9-238).
- [33] N. Linder, R. Turkki, M. Walliander, A. Mårtensson, V. Diwan, E. Rahtu, M. Pietikäinen, M. Lundin, and J. Lundin, "A malaria diagnostic tool based on computer vision screening and visualization of plasmodium falciparum candidate areas in digitized blood smears," *PLoS ONE*, vol. 9, no. 8, Aug. 2014, Art. no. e104855, doi: [10.1371/journal.pone.0104855](https://doi.org/10.1371/journal.pone.0104855).
- [34] H. Telang and K. Sonawane, "Effective performance of bins approach for classification of malaria parasite using machine learning," in *Proc. IEEE 5th Int. Conf. Comput. Commun. Autom. (ICCCA)*, Greater Noida, India, Oct. 2020, pp. 427–432.
- [35] W. Kudisthalert, K. Pasupa, and S. Tongsim, "Counting and classification of malarial parasite from giemsa-stained thin film images," *IEEE Access*, vol. 8, pp. 78663–78682, 2020, doi: [10.1109/ACCESS.2020.2990497](https://doi.org/10.1109/ACCESS.2020.2990497).
- [36] K. M. F. Fuhad, J. F. Tuba, M. R. A. Sarker, S. Momen, N. Mohammed, and T. Rahman, "Deep learning based automatic malaria parasite detection from blood smear and its smartphone based application," *Diagnostics*, vol. 10, no. 5, p. 329, May 2020.
- [37] S. Rajaraman, S. Jaeger, and S. K. Antani, "Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images," *PeerJ*, vol. 7, p. e6977, May 2019, doi: [10.7717/peerj.6977](https://doi.org/10.7717/peerj.6977).
- [38] M. Umer, S. Sadiq, M. Ahmad, S. Ullah, G. S. Choi, and A. Mehmood, "A novel stacked CNN for malarial parasite detection in thin blood smear images," *IEEE Access*, vol. 8, pp. 93782–93792, 2020, doi: [10.1109/ACCESS.2020.2994810](https://doi.org/10.1109/ACCESS.2020.2994810).
- [39] A. Maqsood, M. S. Farid, M. H. Khan, and M. Grzegorzek, "Deep malaria parasite detection in thin blood smear microscopic images," *Appl. Sci.*, vol. 11, no. 5, p. 2284, Mar. 2021, doi: [10.3390/app11052284](https://doi.org/10.3390/app11052284).
- [40] P. A. Pattanaik, M. Mittal, and M. Z. Khan, "Unsupervised deep learning CAD scheme for the detection of malaria in blood smear microscopic images," *IEEE Access*, vol. 8, pp. 94936–94946, 2020.
- [41] E. Irmak, "A novel implementation of deep-learning approach on malaria parasite detection from thin blood cell images," *Electrica*, vol. 21, no. 2, pp. 216–224, May 2021.
- [42] D. Sukumarran, K. Hasikin, A. S. M. Khairuddin, R. Ngui, W. Y. W. Sulaiman, I. Vytilingam, and P. C. S. Divis, "An automated malaria cells detection from thin blood smear images using deep learning," *Tropical Biomed.*, vol. 40, no. 2, pp. 208–219, Jun. 2023.

- [43] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
- [44] S. Sarkar, R. Sharma, and K. Shah, “Malaria detection from RBC images using shallow convolutional neural networks,” 2020, *arXiv:2010.11521*.
- [45] A. Çınar and M. Yıldırım, “Classification of malaria cell images with deep learning architectures,” *Ingénierie des systèmes d’Inf.*, vol. 25, no. 1, pp. 35–39, Feb. 2020, doi: [10.18280/isi.250105](https://doi.org/10.18280/isi.250105).
- [46] K. Sriport, C.-F. Tsai, C.-E. Tsai, and P. Wang, “Analyzing malaria disease using effective deep learning approach,” *Diagnostics*, vol. 10, no. 10, p. 744, Sep. 2020, doi: [10.3390/diagnostics10100744](https://doi.org/10.3390/diagnostics10100744).
- [47] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, *arXiv:1409.1556*.
- [48] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Lake Tahoe, NV, USA, Dec. 2012, pp. 1–8.
- [49] J. Zhang, J. Zhang, S. Ghosh, D. Li, S. Tasci, L. Heck, H. Zhang, and C. C. J. Kuo, “Class-incremental learning via deep model consolidation,” in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, 2020, pp. 1131–1140.
- [50] L. Zedda, A. Loddo, and C. Di Ruberto, “A deep learning based framework for malaria diagnosis on high variation data set,” in *Proc. ICIAP*, Jul. 2022, pp. 358–370.
- [51] R. Liu, T. Liu, T. Dan, S. Yang, Y. Li, B. Luo, Y. Zhuang, X. Fan, X. Zhang, H. Cai, and Y. Teng, “AIDMAN: An AI-based object detection system for malaria diagnosis from smartphone thin-blood-smear images,” *Patterns*, vol. 4, no. 9, Sep. 2023, Art. no. 100806, doi: [10.1016/j.patter.2023.100806](https://doi.org/10.1016/j.patter.2023.100806).
- [52] L. Zedda, A. Loddo, and C. Di Ruberto, “YOLO-PAM: Parasite-attention-based model for efficient malaria detection,” *J. Imaging*, vol. 9, no. 12, p. 266, Nov. 2023, doi: [10.3390/jimaging9120266](https://doi.org/10.3390/jimaging9120266).
- [53] E. Özbilge, E. Güler, and E. Özbilge, “Ensembling object detection models for robust and reliable malaria parasite detection in thin blood smear microscopic images,” *IEEE Access*, vol. 12, pp. 60747–60764, 2024, doi: [10.1109/ACCESS.2024.3393410](https://doi.org/10.1109/ACCESS.2024.3393410).
- [54] Z. Xiong and J. Wu, “Multi-level attention split network: A novel malaria cell detection algorithm,” *Information*, vol. 15, no. 3, p. 166, Mar. 2024, doi: [10.3390/info15030166](https://doi.org/10.3390/info15030166).
- [55] S. Garba, M. B. Abdullahi, S. A. Bashir, and O. A. Abisoye, “Implementation of malaria parasite detection and species classification using dilated convolutional neural network,” in *Proc. 5th Inf. Technol. Educ. Develop. (ITED)*, Abuja, Nigeria, Nov. 2022, pp. 1–6.
- [56] E. Guemas, B. Routier, T. Ghelfenstein-Ferreira, C. Cordier, S. Harthuis, B. Marion, S. Bertout, E. Varlet-Marie, D. Costa, and G. Pasquier, “Automatic patient-level recognition of four plasmodium species on thin blood smear by a real-time detection transformer (RT-DETR) object detection algorithm: A proof-of-concept and evaluation,” *Microbiol. Spectr.*, vol. 12, no. 2, Feb. 2024, Art. no. e01440, doi: [10.1128/spectrum.01440-23](https://doi.org/10.1128/spectrum.01440-23).



EE SAM LOH received the Bachelor of Biomedical Engineering degree from Universiti Malaya, Malaysia. He is currently a Software Development Engineer. He has a keen interest in artificial intelligence, and his involvement in the automated malaria diagnosis project showcases his commitment to applying technology to address real-health challenges.



ANIS SALWA MOHD KHAIRUDDIN (Member, IEEE) received the Bachelor of Electrical and Electronics degree from Universiti Teknologi Malaysia, Malaysia, the master’s degree in computer engineering from Universiti Malaya, Malaysia, the M.Eng. degree in computer engineering from Royal Melbourne Institute of Technology, Australia, and the Ph.D. degree in electrical engineering from Universiti Teknologi Malaysia. She is currently a Senior Lecturer with the Department of Electrical Engineering, Universiti Malaya, specializing in autonomous systems, image processing, and computer vision and expert systems. She has authored or co-authored many peer-reviewed articles in national and international journals. She has supervised many Ph.D. and master’s students by research and has completed many projects on deep learning, machine learning, and intelligent systems.



ROMANO NGUI received the Bachelor of Biomedical Science and Master of Medical Science degrees in parasitology, and the Ph.D. degree in medical parasitology from Universiti Malaya, Malaysia. He was a Senior Lecturer with the Department of Parasitology, University Malaya. He is currently a Senior Lecturer with the Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak (UNIMAS), Sarawak, Malaysia. He has supervised many Ph.D. and master’s students by research and has more than 100 publications. His research interests include the epidemiological distribution of infectious diseases particularly parasitology and tropical medicine, which developed and applied molecular and statistical models/tools to support operational research activities relating to infectious disease control. He is a member of the Malaria Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak. He has been an Active Member of the Malaysian Society of Parasitology and Tropical Medicine (MSPTM), since 2008. He has served as an Assistant Honorary Secretary and the Council Member of the Society.



WAN YUSOFF WAN SULAIMAN was the Deputy Dean of the Faculty of Medicine, Universiti Malaya, Malaysia. He is currently a Senior Lecturer with the Department of Parasitology, Faculty of Medicine, Universiti Malaya. He has supervised many Ph.D. and master’s students by research and has more than 70 publications. His research emphasizes parasitology and infectious disease. He is actively involved in the development and application of molecular and statistical models and tools to enhance operational research endeavors.



DHEVISHA SUKUMARRAN received the Bachelor of Biomedical Engineering degree from Universiti Malaya, Malaysia, where she is currently pursuing the Ph.D. degree. Her research interests include the application of artificial intelligence, namely deep learning for automated malaria diagnosis, and species identification.



INDRA VYTHILINGAM received the Bachelor of Science degree in zoology from Madras University, India, the M.Sc. degree in zoology from Victoria University of Wellington, New Zealand, and the Ph.D. degree in zoology from Universiti Malaya, Malaysia. Currently, she is a Professor with Universiti Malaya, specializing in parasitology. She is the Assistant Director in parasites and vectors, a Committee Member with the Ministry of Health, Malaysia, and an Examiner with Universiti Malaya. She has authored or co-authored many peer-reviewed articles in national and international journals. She has supervised many Ph.D. and master's students by research and has completed many national and international projects related to parasitology.



KHAIRUNNISA HASIKIN received the Bachelor of Electrical Engineering and M.Eng.Sc. degrees in biomedical microelectromechanical systems from Universiti Malaya, Malaysia, and the Ph.D. degree in engineering from Universiti Sains Malaysia, Malaysia. She is currently an Associate Professor and a Researcher with the Department of Biomedical Engineering, Universiti Malaya, specializing in medical image analysis, medical image processing, artificial intelligence, and sustainability management. She has supervised many Ph.D. and master's degree students in research and has more than 90 publications. She is a member of the Center of Intelligent Systems for Emerging Technology (CISET), Faculty of Engineering, Universiti Malaya. She has led many research projects more than the years. She has authored or co-authored many peer-reviewed articles in national and international journals and has completed many national and international projects on artificial intelligence, computer science, and medicine.



PAUL CLIFF SIMON DIVIS received the Bachelor of Science degree from Universiti Kebangsaan Malaysia, the Master of Science degree in molecular parasitology from Universiti Malaysia Sarawak, Malaysia, and the Ph.D. degree in applied parasitology from London School of Hygiene and Tropical Biomedicine, U.K. He is currently a Senior Lecturer with the Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak, specializing in population genetics, phylogenetics, parasitology, epidemiology, and bioinformatics. He is a member of the Malaria Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak. He has supervised many Ph.D. and master's students in research and has more than 30 publications.

• • •