Tropical Biomedicine 40(2): 208-219 (2023)
https://doi.org/10.47665/tb.40.2.013

RESEARCH ARTICLE

An automated malaria cells detection from thin blood smear images using deep learning

Sukumarran, D.[1], Hasikin, K.[1,5*], Mohd Khairuddin, A.S.[2,5], Ngui, R.[3*], Wan Sulaiman, W.Y.[3], Vythilingam, I.[3], Divis, P.C.S.[4]

[1]Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[2]Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[3]Department of Para-Clinical Sciences, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak
[4]Malaria Research Centre, Faculty of Medicine and Health Sciences, Universiti Malaysia Sarawak, Sarawak, Malaysia
[5]Center of Intelligent Systems for Emerging Technology (CISET), Faculty of Engineering, Universiti Malaya, 50603 Kuala Lumpur, Malaysia
*Corresponding author: khairunnisa@um.edu.my; nromano@unimas.my

ARTICLE HISTORY

Received: 13 January 2023
Revised: 15 March 2023
Accepted: 20 March 2023
Published: 30 June 2023

ABSTRACT

Timely and rapid diagnosis is crucial for faster and proper malaria treatment planning. Microscopic examination is the gold standard for malaria diagnosis, where hundreds of millions of blood films are examined annually. However, this method's effectiveness depends on the trained microscopist's skills. With the increasing interest in applying deep learning in malaria diagnosis, this study aims to determine the most suitable deep-learning object detection architecture and their applicability to detect and distinguish red blood cells as either malaria-infected or non-infected cells. The object detectors Yolov4, Faster R-CNN, and SSD 300 are trained with images infected by all five malaria parasites and from four stages of infection with 80/20 train and test data partition. The performance of object detectors is evaluated, and hyperparameters are optimized to select the best-performing model. The best-performing model was also assessed with an independent dataset to verify the models' ability to generalize in different domains. The results show that upon training, the Yolov4 model achieves a precision of 83%, recall of 95%, F1-score of 89%, and mean average precision of 93.87% at a threshold of 0.5. Conclusively, Yolov4 can act as an alternative in detecting the infected cells from whole thin blood smear images. Object detectors can complement a deep learning classification model in detecting infected cells since they eliminate the need to train on single-cell images and have been demonstrated to be more feasible for a different target domain.

Keywords: Malaria; Yolov4; Faster R-CNN; SSD300; deep learning.

INTRODUCTION

Malaria is a severe and occasionally fatal disease caused by a *Plasmodium* parasite that commonly infects the *Anopheles* mosquito, which feeds on human blood. World Health Organization (WHO) has been committed to malaria eradication since 1955. However, in 2020 alone, 241 million cases and 627,000 deaths were recorded. According to the World malaria report for 2021, the number of reported malaria cases remained the same for 2000 and 2020. This demonstrates that malaria remains a serious global health issue 67 years after the eradication pledge. As one of the tropical countries, Malaysia is also affected by this disease. During the past decade, the incidence of zoonotic *P. knowlesi* cases in Southeast Asia has been increasing, with Malaysia reporting the highest number of infections (WHO, 2021a). Based on WHO (2021b) harnessing innovation and expanding research is one of the supporting elements in the fight against malaria. Substantial research and development to create new tools and strategies in medicines, diagnostics, vector control, and vaccines can contribute to malaria elimination and, eventually, its global eradication.

Microscopic examination remains the gold standard for malaria diagnosis. The results of the examination highly depend on the microscopist's interpretation. Diagnosing the disease is challenging in non-endemic countries as the disease is rarely seen, and expertise in malaria diagnosis needs to be better maintained. However, in malaria-endemic countries, a lack of resources is a significant barrier to reliable and timely diagnosis of diseases (Tangpukdee *et al*., 2009; Capela *et al*., 2019). Therefore, alternate use of artificial intelligence for decision-making can play a crucial role in facilitating the rapid detection and prompt diagnosis of diseases.

With current advances, deep learning, a subset of artificial intelligence (AI), plays a prominent role in healthcare by improving the reliability and efficiency of diagnosis and treatment across various specializations (Ahuja, 2019). In previous studies, different deep-learning methods have been applied to malaria diagnosis. A Convolutional Neural Network (CNN) is a type of artificial neural network mainly used in image recognition and processing and is famously used to detect and classify infected cells. However, one main challenge in applying deep learning in malaria diagnosis is the need for a comprehensive dataset. To learn the infected cells'

features, single-cell images are needed to train the CNN classification model. However, they usually exist in a whole blood smear image with other cells, such as healthy red blood cells (RBC), white blood cells, etc. In previous works, traditional segmentation methods have been used to segment the cells of interest. Liang *et al*. (2016) used active contour to segment the cells, while Dong *et al*. (2017) segmented cells using thresholding techniques and morphological operations. Both rely on Hough Circle transform for their cell separation. Rajaraman *et al*. (2019) applied a level-set algorithm for RBC segmentation. Morphological operations for cell segmentation have been widely reported (Dong *et al*., 2017; Gopakumar *et al*., 2018; Molina *et al*., 2021; Arshad *et al*., 2022). Furthermore, in several recent studies, readily segmented single-cell images from the National Library of Medicine were used to binary classify cells as infected or non-infected using deep learning models (Kassim *et al*., 2021; Maqsood *et al*., 2021; Imran *et al*., 2022).

Several researchers have employed object detection algorithms, including Zhao *et al*. (2020), who used a deep learning object detection algorithm known as SSD300 to detect all the infected and non-infected red blood cells from thin blood smear images. On thin blood smear images, the effectiveness of object detection algorithms such as Faster R-CNN (Hung *et al*., 2017), Yolov2 (Yang *et al*., 2020) and SSD300 (Zhao *et al*., 2020), Mask R-CNN (Loh *et al*., 2021), have been evaluated. Besides these architectures, Yolov4 was mainly used in previous studies to detect malaria parasites from thick blood smear images (Abdurahman *et al*., 2021; Koirala *et al*., 2022). The modified Yolov4 model achieved a higher mean average precision of 96.32% than Faster R-CNN and SSD in detecting infected cells in thick blood smear images (Abdurahman *et al*., 2021), but their performance is yet to be compared on thin blood smear images. So far, to our knowledge, only one study has experimented with and compared the performance of scaled Yolov4 and Yolov5 on malaria thin blood smear images for the detection and classification of malarial cells according to their parasite (Krishnadas *et al*., 2022). Based on their experiment, the performance of scaled Yolov4 surpasses the performance of Yolov5. There is limited research on the application of Yolov4 for the detection of malarial infected cells. However, the application of Yolov4 was recently studied on other blood smear images with satisfying results, such as for blast cell detection for acute lymphoblastic leukaemia diagnosis (Khandekar *et al*., 2021; Akalin & Yumu÷ak, 2022).

In this study, we aim to develop an end-to-end identification and localization of the malaria-infected cells using a deep learning approach. Object detectors have an advantage in classifying and localizing objects. They can be trained to recognize the specific morphological characteristic of the infected cells using full blood smear images rather than single-cell images. Despite being trained on the entire blood smear images, the model will recognise which features to be learnt since, in the training images, the bounding boxes must be drawn on the objects of interest. Following that, during prediction, an object detector can identify the location of the detected cells (i.e., the coordinates of the detected cells). With the detected localised coordinates, it is possible to do automated cropping and processing of the single-cell image. Overall, an object detector reduces the need to generate single-cell images and permits the generation of single-bounded cell images for other classification applications. This technology will increase the viability of end-to-end deep-learning malaria diagnosis.

In this study, several deep-learning object detectors are evaluated to determine the best-performing object detector in distinguishing and detecting only infectious cells from thin blood smear images. In contrast to the prior research, the images used in this study reflect infections caused by all malaria species at various stages of infection. This study aims to explore the capability of object detection architecture in recognizing the infected cells regardless of morphological differences between species and the infection phases. In addition, a specific algorithm will be integrated into the final model to automatically locate, crop, and create single-cell images of the detections. The crop-infected cells can be employed in future applications as an input to the second stage of the classification model in identifying the malaria species or the infection stages.

MATERIALS AND METHOD

In this study, the object detector algorithms are trained to determine the infected cells from the whole thin blood smear images. The idea of implementing object detectors is to automatically localize the objects of interest with minimal human intervention. Figure 1 shows the overview of the proposed object detection algorithm conducted in this study. The three object detectors trained for this problem are Yolov4, Faster R-CNN and SSD-300.

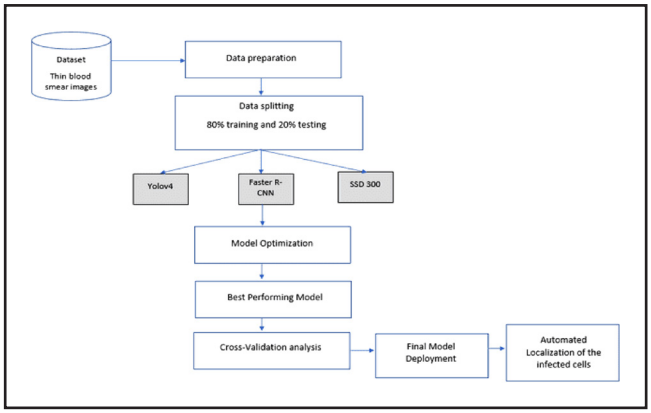

Figure 1. Overall research methodology.

Object detection architecture

Faster Region-based Convolutional Neural Network (Faster R-CNN)

Faster R-CNN is a two-stage object detector introduced by Ren *et al*. (2015). It consists of two stages: the region proposal stage for object localization and the detector layer to assign labels to these objects. It is known as a two-stage detector, as the architecture first extracts the features from the images and then calculates the probability of the region containing the object. In Faster R-CNN, the image input is sent to the region proposal network (RPN). In the RPN, proposals are created in the input images. Region proposals are the parts that likely contain the object of interest. First, the features from the images are extracted, and feature maps are created to create region proposals. Once the feature maps are obtained, anchor boxes, as shown in Figure 2b, are placed all over the feature map. The anchor boxes represent where the RPN will search for the objects of interest. Finally, the RPN eliminates the anchor boxes that have lesser than a certain threshold of intersection over union (IoU) with the ground truth annotation (Figure 2a). If the intersection is high, the anchor box likely contains the object and is kept; otherwise, it is discarded. At the end of RPN, only a few anchor boxes are used as region proposals.

In the region of interest (ROI) pooling, the feature maps and the region proposals are fed into the fully connected layers. In ROI pooling, fully connected layers are used to reduce the feature maps to the same size, as they are all in different sizes. The output of the ROI pooling is fed into two other fully connected layers and a softmax (classifier) which classifies the region proposals. At the same time, the regressor refines the location and dimension of the bounding boxes to increase localization accuracy.

Single Shot Multibox Detector (SSD)

As compared to the Faster R-CNN, SSD has no region proposals networks. It predicts the boundary boxes and the classes of the

objects directly from feature maps in a single pass. Single Shot detector (SSD) consists mainly of extracting feature maps and applying a convolutional filter for object detection. Upon image input, a convolutional neural network such as VGG-16 is used as the backbone for feature extraction and Conv layers for object detection. In SSD, multi boxes of different sizes, like anchor boxes in Faster R-CNN, are created across the image. A total of 8732 boxes will be created to find the default box that overlaps the most with the ground truth bounding boxes containing the object.

During training, a matching strategy is used to match the default boxes over aspect ratio, location, and scale to the ground truth boxes. The boxes with the highest overlap with the ground truth over a certain threshold are selected. Sometimes there is more than one overlapping box on the ground truth. Therefore, non-max suppression is conducted where the predicted box with the maximum overlap is chosen as the final object location.

You Only Look Once, version 4 (Yolov4)

You Only Look Once (Yolo) is a single-stage object detector as SSD. The main improvement in Yolo is the integration of object detection and classification in a single pass. Yolo does not perform feature extraction and region separately; instead, it executes them in a single pass. The approach of Yolov4 for bounding box regression sets it apart from the other object detection architectures. The detection of the object consists mainly of grid-cell, bounding box, and IoU. During training, the input image is divided into grid cells SxS. The model predicts N bounding boxes in every cell to search if the cell contains the centre point of the object of interest based on the ground truth box. When more than one bounding box contains the centre point, the model applies IoU overlap (Figure 2a) to determine and keep only the boxes exceeding a certain threshold. If there is more than one bounding box with IoU more than the threshold, non-max suppression is used to select the box with the maximum overlap with the ground truth.

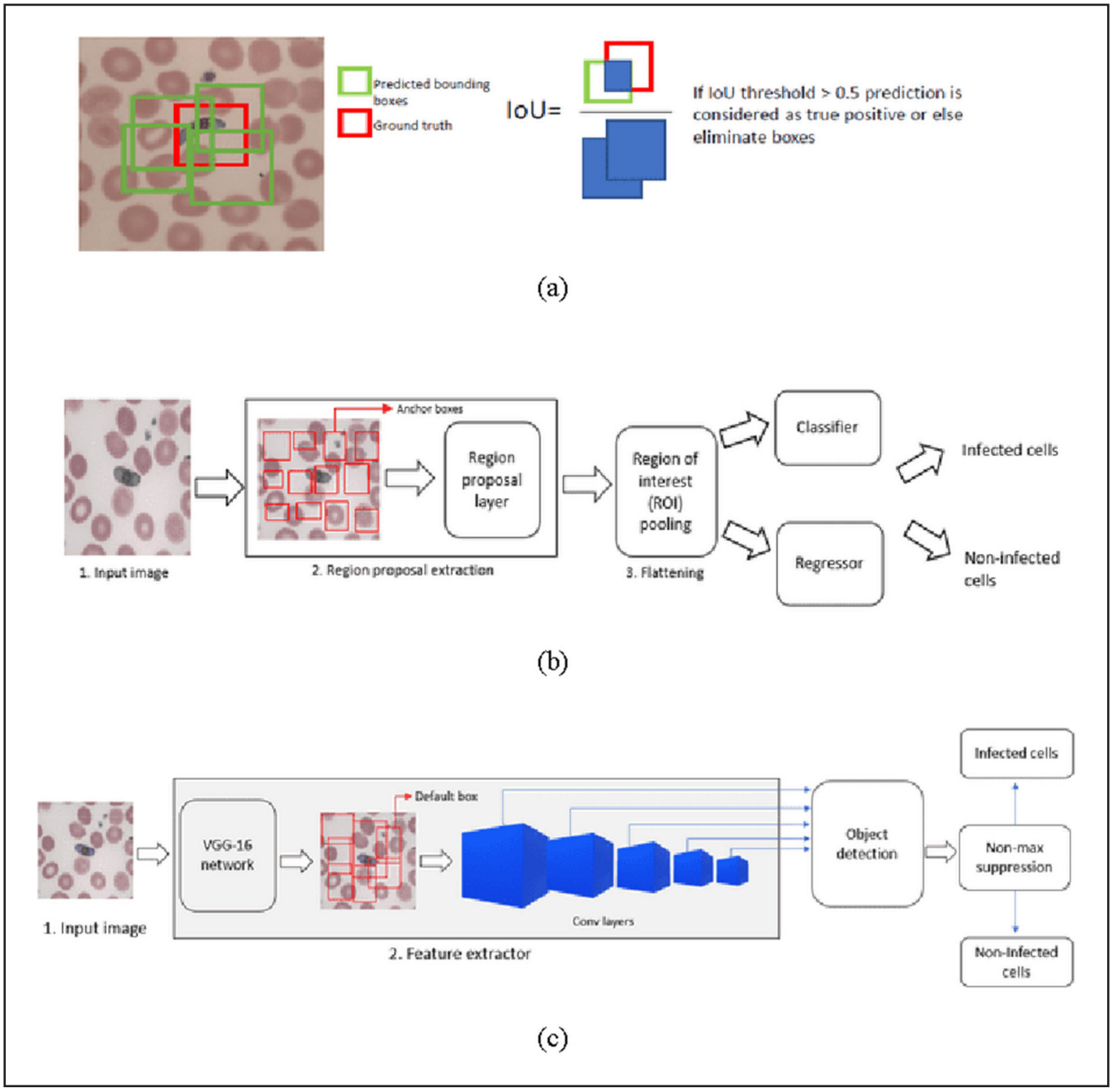

Figure 2. Bounding box selection using intersection over union (IoU)(a), Overview of model predictions Faster R-CNN(b), SSD300(c), Yolov4(d).

Data acquisition

This study used 446 thin blood smear images (examples as shown in Figure 3). The images were acquired from the publicly available dataset MP-IDB (https://github.com/andrealoddo/MP-IDB-The-Malaria-Parasite-Image-Database-for-Image-Processing-and-Analysis). Besides that, archived thin blood films of malaria infections from Kapit Hospital, Sarawak was obtained from previous studies (Daneshvar *et al*., 2010; Divis *et al*., 2015, 2018; Hu *et al*., 2021; Yunos *et al*., 2022). These blood films were fixed with absolute ethanol (BDH, England) for 10 seconds and stained with 10% Giemsa (BDH, England) in Gurrâ buffered water, pH 7.2 (BDH, England). Blood films were examined using a light microscope (Olympus model BX53) at x 1,000 magnification with immersion oil, and images were taken using the Cell^B software version 3 (Olympus). The images obtained are infected by all malaria parasites from various stages of infection. There is a total of 3038 parasitic cells annotated from all the images. For the images from the MP-IDB database, every image has its ground truth of infected cells in the form of a binary mask, and every file name indicates the stages of infection. The ground truth helps identify the infected cells' location in every image (Figure 3).

Data Preparation

The main data preparation step done on the dataset was image augmentation. Augmentation is done to generate and increase the amount of the training images from the existing ones, such as by rotating the image and shifting and flipping the image horizontally and vertically. Data augmentation can only be done on images kept to train the model and not on the test images. Augmenting the training images provides more data to train the model and improves the model's accuracy by enhancing its ability to recognize new variants of the training data.

The object detectors perform their data augmentations. Such as one primary augmentation conducted by the Yolov4 model is mosaic data augmentation, where the training images are combined into a certain ratio, Faster RCNN performs random flips, and SSD 300 performs random crop. Besides these augmentations, a separate geometric augmentation was performed on the training images. As shown in Figure 4, the MP-IDB dataset is used to train and test the model, whereas the dataset from Kapit Hospital, Sarawak, is used to cross-validate the model's performance. As shown in Table 1, firstly, the 210 images from MP-IDB dataset are split into training and testing images. From the dataset, 80% of the images are used to train the model and 20% to test the model's performance. The train images are then divided according to parasite and stages of infection. The augmentation was done to achieve an equal number of images in all phases of the infection of parasites. Upon augmentation, 1000 training images were obtained (Figure 5). Besides that, the 236 images from Kapit Hospital, Sarawak were used to cross-validate the performance of the object detector.

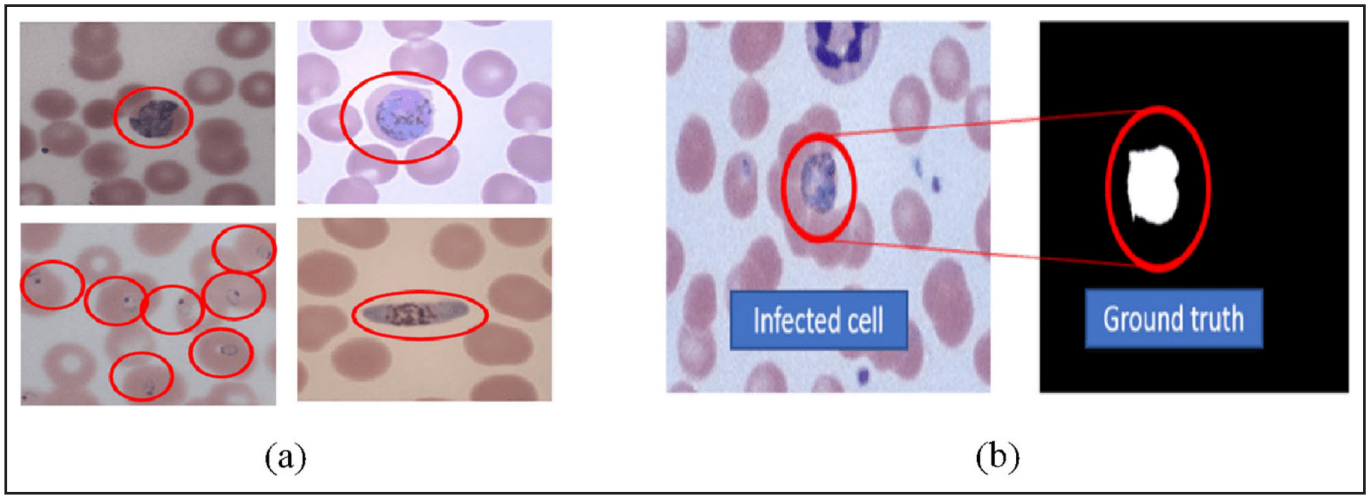

Figure 3. Examples of thin blood smears, red circles enclose the parasites(a) Example of image and its ground truth from the public dataset(b).

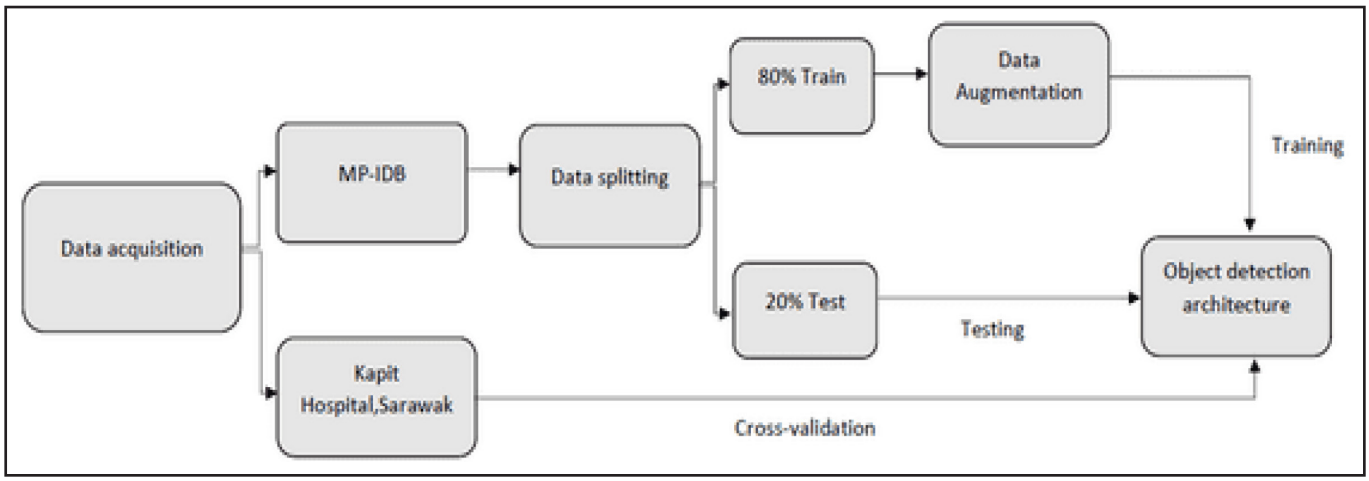

Figure 4. Overview of data preparation.

Table 1. Data distribution for training and testing the model

Database	Total images before augmentation	Training images		Testing images
		Training images before augmentation	Training images after augmentation	No augmentation
MP-IDB	210	168	1000	42
Kapit Hospital, Sarawak	236	–	–	236

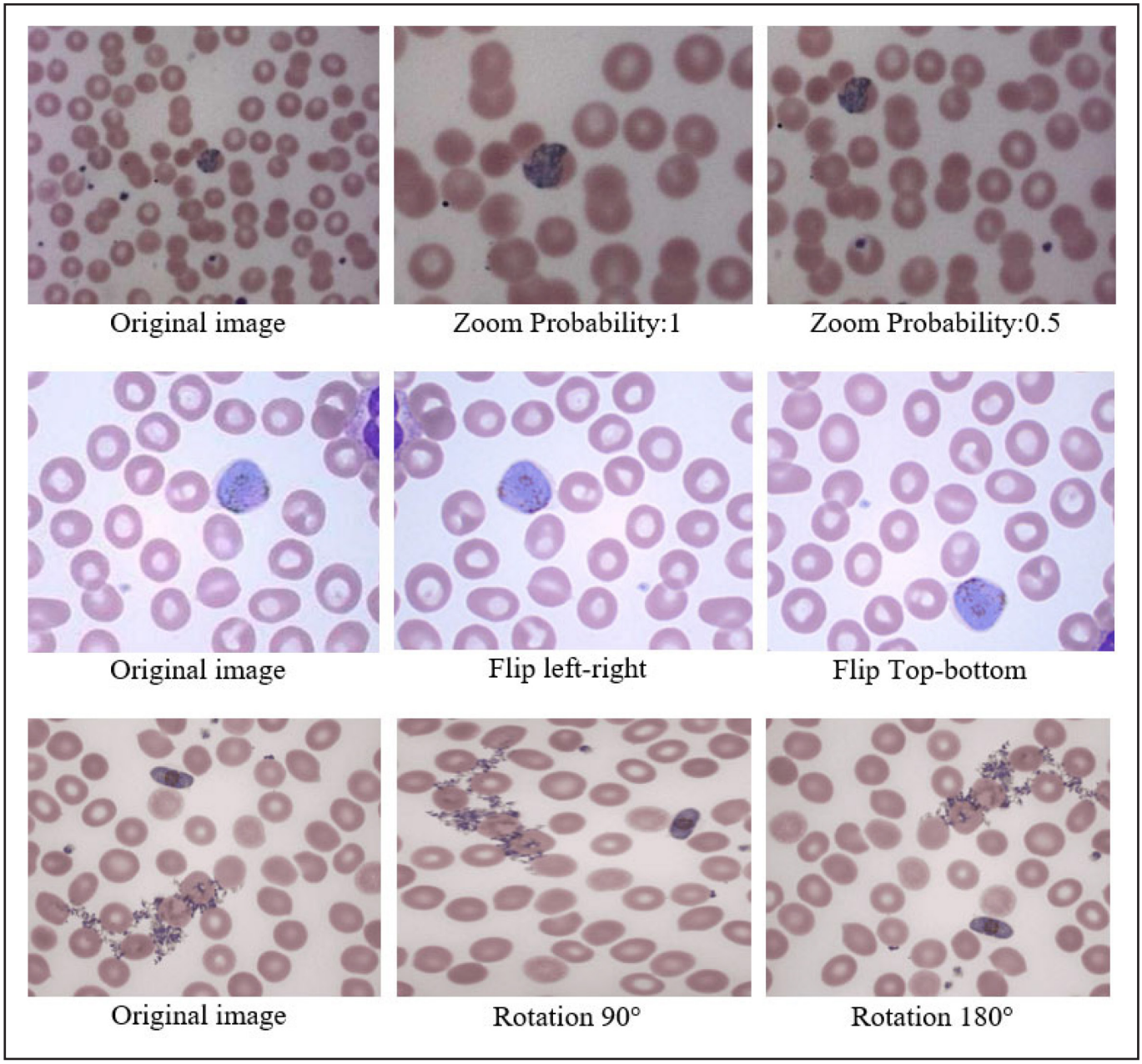

Figure 5. Examples of images before and after augmentations.

Upon image augmentation, the train and test images were annotated using LabelImg. As shown in Figure 6, bounding boxes are manually drawn around the infected cells for all images. The bounding boxes provide the minimum and maximum *x* and *y* coordinates (xmin, ymin,xmax,ymax) where the objects of interest are located for the model to learn their features. Besides drawing the bounding boxes, the infected cells are labelled under the same class despite their parasite and stage of infection. For Yolov4, the bounding boxes' coordinates are saved in Yolo format, whereas coordinates for Faster R-CNN and SSD300 are saved in XML format. These files with the coordinates of the bounding boxes will be used throughout the training and testing of the models.

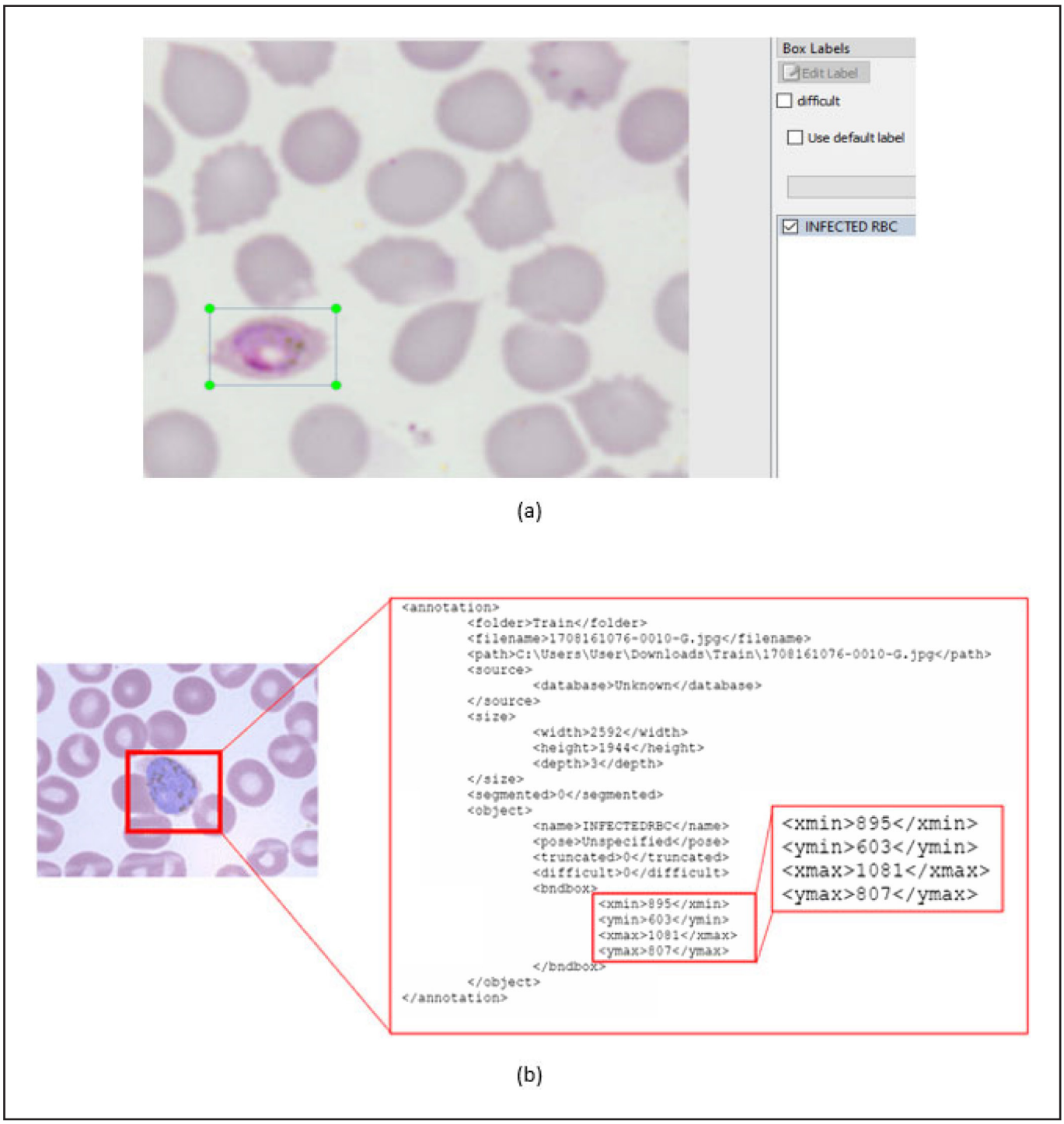

Figure 6. Annotation of infected cell in labelImg(a), Example of XML file created upon annotation with coordinates of bounding boxes(b).

Model Implementation

A custom model configuration of Yolov4, Faster RCNN, and SSD300 was performed according to the number of classes to be detected. In this study, the predictive model of binary classes of "infected cells" and "non-infected cells" are constructed, and experiments were conducted in Google Colab. The models were trained with 80% of the original MP-IDB dataset augmented and tested with 20% of the dataset.

For Yolov4, a learning rate of 0.01 and 0.001 was used to train the model. The training of the Yolov4 model was conducted twice at a learning rate of 0.001. Firstly, it was conducted with distortion as one of the augmentation methods. Next, the training was undertaken again after evaluating the results and making a few changes to the augmentation and annotation of the images to improve the model's prediction. The distortion augmentation method was eliminated, and the cells blurred or cut off in the training images were not included in the annotation. For training Faster R-CNN, models with different backbones are selected, such as Inceptionv2 and Resnet-50. Inceptionv2 and Resnet-50 are convolutional neural networks with different architectures and act as feature extractors to create feature maps in Faster R-CNN. Similarly, SSD-300 with MobileNetv2 backbone was trained. Faster R-CNN Inceptionv2, Resnet-50, and SSD300 models were first trained with their default learning rate of 0.0002,0.0003,0.0003, respectively, followed by the optimization process with a learning rate of 0.001 as tabulated in Table 2.

Table 2. Hyperparameters setting of object detectors

Parameter	Yolov4	Faster R-CNN Inception-v2	Faster R-CNN Resnet-50	SSD-Mobilenetv2
Learning rate	0.001	0.001	0.001	0.001
	0.01	0.0002	0.0003	0.0003
Batch size	32	32	32	32
Input image size	416x416	600x1024	600x1024	300x300
IOU threshold	0.5	0.5	0.5	0.5

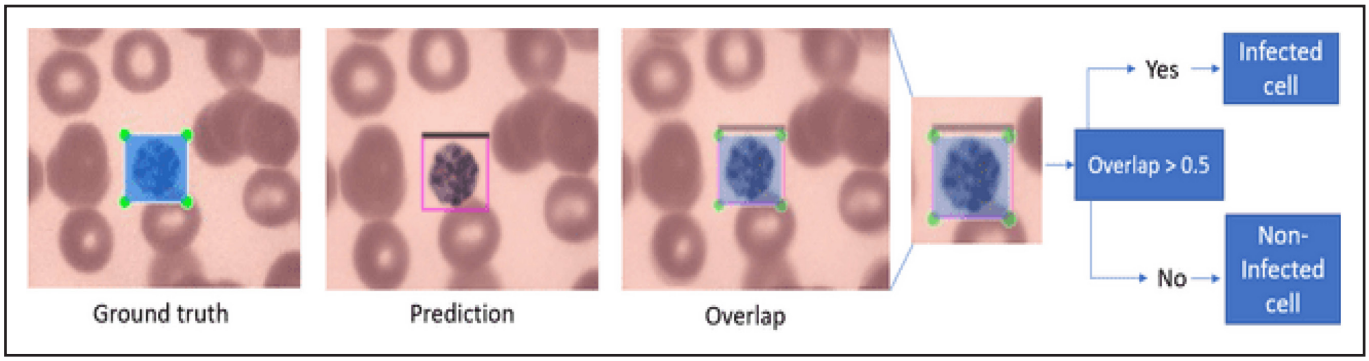

Figure 7. Infected cells prediction strategy using IoU threshold.

Performance Evaluation and best model selection

The results were evaluated with four performance metrics, namely recall rate, precision, F1-score, and mean average precision (mAP). The performance metric of the model depends on the IoU threshold allocated for predictions. In this study, an IoU threshold of 0.5 was set. If the bounding boxes predicted by the model overlap with the ground truths for more than a threshold of 0.5, the predictions are considered true positives; otherwise, they are false positives (Figure 7). The IoU determines the true and false positive predictions; with this, we get the precision and recall rate of the model. Precision is the intolerance of the model towards false positives (1), The recall rate represents the intolerance of the models towards false negatives (2), and the F1-score (3) is the harmonic mean of recall and precision. F1-score gives equal weight to recall and precision. Mean average precision (4) is the typical metric used to compare the performances of object detectors. It is the mean of the average precision of all the classes detected by the object detection model. The model with the highest mAP is chosen as the best-performing model.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

$$F1-score = \frac{2[precision(class=1)\ .recall(class=1)]}{precision(class=1)+recall(class=1)} \quad (3)$$

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \quad (4)$$

Cross-dataset validation and Optimal IoU determination

Upon selecting the best-performing model, cross-dataset validation of the model was performed on an independent dataset from Kapit Hospital, Sarawak. Cross-dataset validation tests the model's generalizability on an independent dataset acquired under different staining conditions. A total of 236 images containing all malaria parasites were used to cross-test the model's performance, and the model's performance was recorded. Besides that, the model's performance was further evaluated at different IoU thresholds besides 0.5, such as at 0.3 and 0.7, to determine the optimal IoU threshold for this problem.

Automated localization algorithm

In yolov4, localization and labelling of classes are done during predictions. In this study, we integrated a separate algorithm to automatically crop and save the predictions by the model to obtain single-cell images. The final best model was first saved as a TensorFlow model to achieve automatic cropping. Upon that, a separate function file is created to crop the detections. The function file performs the cropping by reading the coordinates of every bounding box (left_x, top_y, width, and height) predicted by the model on an image and saves each predicted cell separately (Figure 8).

RESULTS

Performance metric of object detectors

From Table 3, the Yolov4 training was conducted twice with a learning rate of 0.001. Firstly, it obtained a mAP of 89.93% (Model 1) with distortion as one of the augmentation methods and achieved a mAP of 93.87% (Model 2) upon excluding distortion augmentation. The main performance metric used to compare the model's performances is Mean average precision(mAP).

Yolov4 model 2, with a learning rate of 0.001, gives the highest mAP at the IoU threshold of 0.5 (Table 3). The predictions were compared from before (Model 1) and after (Model 2) few adjustments were made in the training images' labelling and augmentation methods. From the comparison, a few of the true positive cells not detected by model 1 were caught by model 2. Therefore, the Yolov4 model 2 with the highest mAP is chosen as the best-performing model.

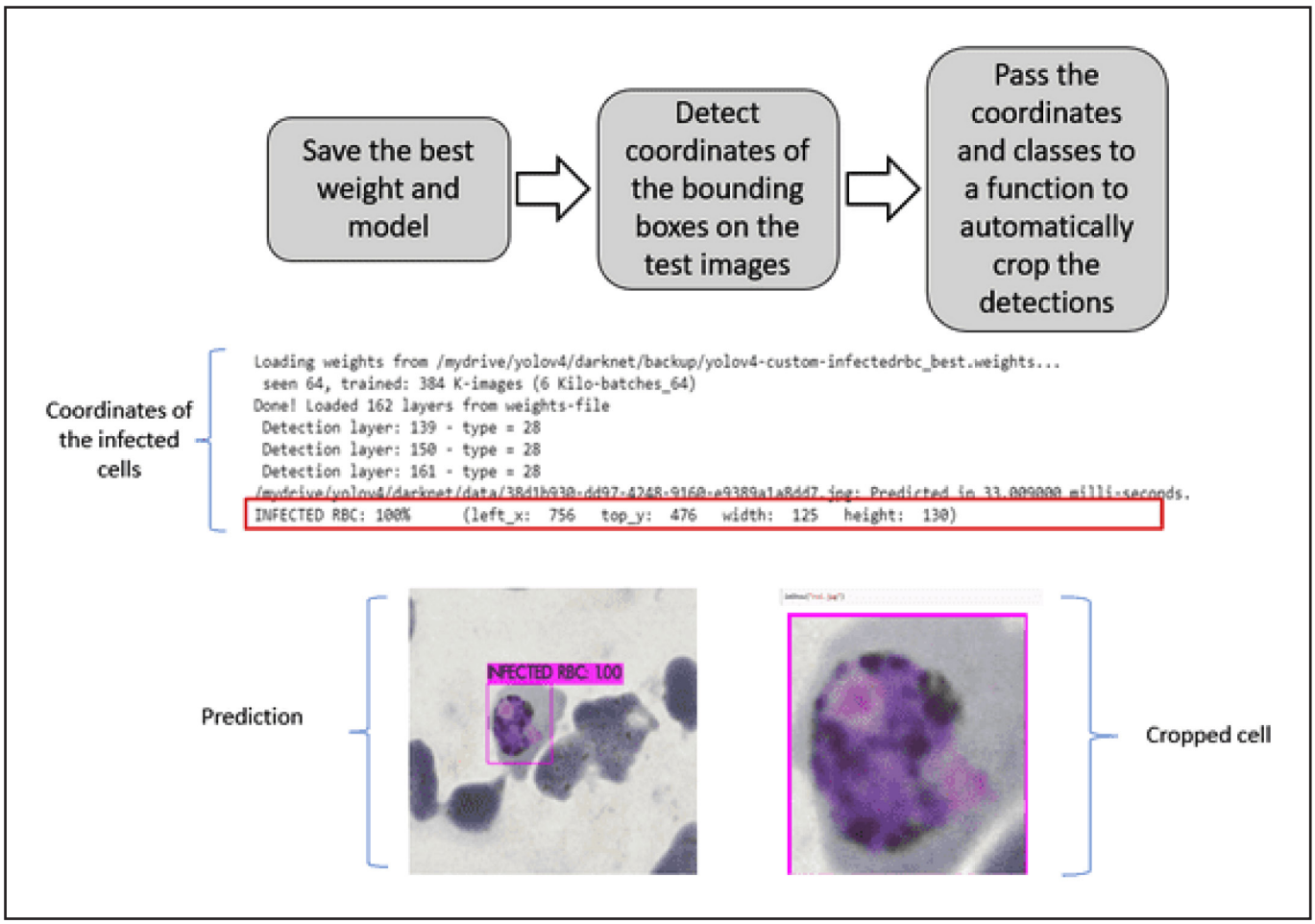

Figure 8. Overview of the automated cropping of cells.

Performance analysis and cross-validation of best performing Yolov4 model

The yolov4 model achieved the highest mAP of 93.87% at an IoU threshold of 0.5. Table 4 shows the performance metric of Yolov4 model 2; epochs are the number of iterations the training data are passed through the model during training. The mAP is derived from the precision and recall rate of the model. When the precision value is higher, the model is more confident when classifying an object as positive. Whereas the recall rate is higher, more objects are classified correctly as positive as the model recalls the features better. For this problem, the model with a higher recall rate is preferred to flag as many infected cells as possible. Since both precision and recall are important, the precision-recall curve gives a better trade-off between these two metrics to select the threshold that maximizes both metrics. However, a more convenient way is to evaluate the F1-score of the model. The higher the F1- score, the higher the balance between the model's precision and recall rate. From Table 4, the highest recall rate achieved by the model is 95% at 4000 epochs. However, the precision decreases from 85 to 83%, yet the model achieved the highest F1-score of 89%, showing a balance between precision and recall.

The performance of the best-performing Yolov4 model was evaluated with an independent dataset. Table 5 shows the confusion matrix obtained on the source and target domains containing the predictions' true positive, false positive and false negative. The source domain indicates the datasets used to train and test the model, and the target domain indicates the independent dataset used to cross-validate the model. Initially, the model's performance on the source and target dataset was evaluated at an IoU threshold of 0.5 to ensure a balance in predictions with not too many false

Table 3. Mean average precision (%) obtained by object detector models on test images from the source domain

Object detectors		Learning rate	mAP(%)
Yolov4	Model 1	0.01	88.18
	Model 1	0.001	89.93
	Model 2	0.001	93.87
Faster R-CNN InceptionV2		0.0002	67.3
		0.001	76.7
Faster R-CNN Resnet50		0.0003	43.8
		0.001	66.7
SSD300 MobileNetV2		0.0003	69.3
		0.001	63.9

Table 4. Performance metric of Yolov4 model 2

Epoch	Precision (%)	Recall (%)	F1-score (%)	mAP (%)
1000	74.00	92.00	82.00	89.08
2000	82.00	93.00	88.00	92.08
3000	85.00	87.00	86.00	87.63
4000	83.00	95.00	89.00	93.87

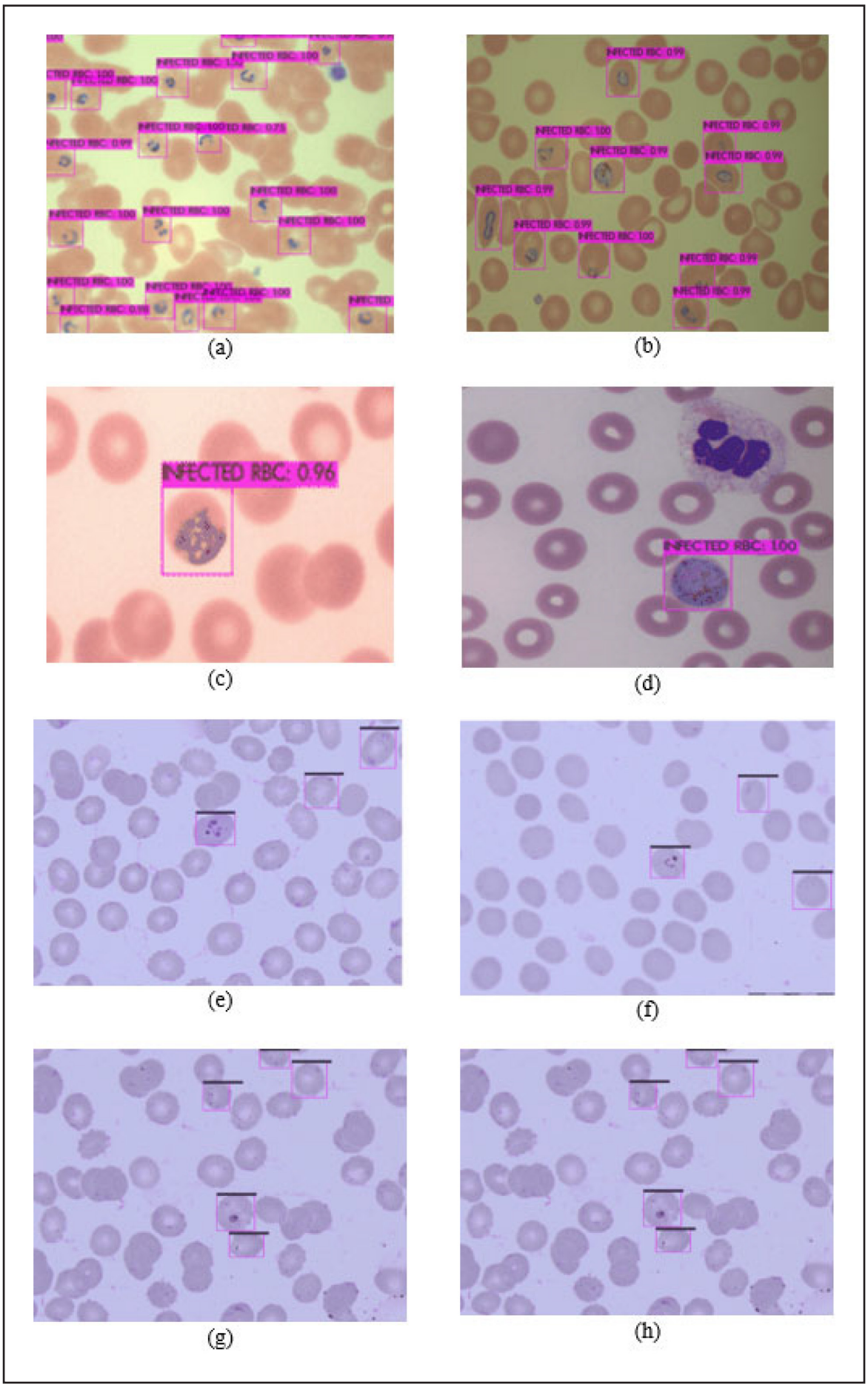

Figure 9. Examples of predictions of Yolov4 model 2 in ring(a)Trophozoite(b), schizont(c) and gametocyte(d) stages of infection, Examples of false positives results predicted by the Yolov4 model 2(e,f,g,h).

Table 5. Confusion matrix of Yolov4 model 2 in source and target domain

Confusion matrix of Yolov4 model on source domain

	Predicted class			Predicted class			Predicted class	
	Yolov4			Yolov4			Yolov4	
	1	0		1	0		1	0
1	346	54	1	337	63	1	238	162
0	9	TN	0	18	TN	0	117	TN
Threshold:0.3			Threshold:0.5			Threshold:0.7		

Confusion matrix of Yolov4 model on Target domain

	Predicted class			Predicted class			Predicted class	
	Yolov4			Yolov4			Yolov4	
	1	0		1	0		1	0
1	372	174	1	360	186	1	283	263
0	65	TN	0	77	TN	0	154	TN
Threshold:0.3				Threshold:0.5		Threshold:0.7		

positives and false negatives values. The predictions were then evaluated at different IoU threshold values to determine the optimal IoU threshold for detecting the infected cells. During the experiment, it was determined that the Yolov4 model achieved the highest mAP, recall rate, and true positives prediction on the malaria dataset with an IoU threshold of 0.3. From Table 4, more true positive infected cells are correctly detected at IoU threshold 0.3, contributing to a higher recall rate and mAP.

DISCUSSION

This study uses deep learning algorithms as object detectors to detect infected cells from thin blood smear images. The images that are infected with malaria parasites from various infection stages are used in this study to measure the capability of each model in detecting infected cells despite their morphological differences. The images were augmented in the training stage so that there was the same number of images infected by all parasites and stages of infection. This is to reduce the model's bias towards a certain parasite or stage of infection.

Initially, the Yolov4 model 1 obtained a mAP of 89.93%. This was achieved when the model was trained with the augmented dataset. However, one of the augmentation methods used was distortion. After evaluating the results, the distortion augmentation was eliminated to not further complicate the learning process of the model as there is already no other pre-processing of images, such as noise removal. Besides that, on the training images, the cells that are blurred or cut off from the images were not included in the annotation to facilitate the cell features learning by the model. The model's performance (model 2) improved to 93.87%, as shown in Table 3.

Yolov4 outperforms Faster R-CNN and SSD 300 in detecting the infected red blood cells, as shown in Table 3. According to the previous studies, it can be concluded that there is no superior object detector, as their effectiveness changes based on the type of detection problem and dataset. The SSD300 Mobilenet is a short architecture notable for its accuracy-to-speed ratio. However, the model needs help producing an accurate mAP for the malaria dataset. The degree of accuracy of SSD300 may have been reduced compared to Faster R-CNN as it needed to identify small objects, mainly cells from the ring stage of infection. Although the test images contain the same number of images for each stage of infection, typically, there are more infected cells in the ring stage in a single blood smear image which can impact the model's performance. Previous studies indicate that Faster R-CNN performs relatively better in multi-class than single-class classification (Hung *et al*., 2017). Faster R-CNN with Inceptionv2 as a feature extractor gives better mAP than Resnet 50. Faster R-CNN inceptionv2 is deeper than Faster RCNN Resnet50; however, from the results deeper network should not perform worse than a shallow network given proper optimization. With a change to the learning rate of Inceptionv2, its performance was significantly enhanced.

In previous studies, Yolov4 was mostly employed to identify infected cells in thick blood smear images. Meanwhile, this study concluded that Yolov4 could efficiently execute regression and classification on images of thin blood smear cells. As depicted in Figure 9, although distinct morphologies from different stages of infected malaria, the proposed Yolov4 can recognize them. Given the importance of finding infected cells, the recall rate attained by the model to identify as many infected or true positive cells as feasible is given greater weight. Consequently, the performance of the model was also assessed at varying thresholds. Although the model initially produces a reasonable result with an IoU threshold of 0.5, it was discovered that with an IoU threshold of 0.3, the model could identify more true positive cells and has a higher recall rate.

Although few measures were taken during the training to reduce the model's bias, it is crucial to perform cross-data validation to observe the model's performance and robustness on a dataset from a different domain. In most previous studies, cross-dataset validation of the deep learning models is not performed. Consequently, these questions the model's robustness on an independent dataset. Among the few studies that performed cross-dataset validation, Zhao *et al*. (2020) and Rahman *et al*. (2021) reported a decrease in models' classification performance upon testing. The results of the classification model on an independent dataset show that it is difficult to directly apply models trained on the source domain to the target domain when they differ excessively (Loddo *et al*., 2022).

In previous studies, the object detectors were not cross-validated in an independent malaria dataset. In this study, the Yolov4 model was validated with an independent dataset and achieved a mAP of 84.04% at the optimal threshold. From Figure 9 (e, f, g, h), the model mainly mistakes small materials or faint stains and smears as infected cells leading to false positives. Besides that, in some infections, there are typically an average of 3 or more cells in a single blood smear image, mainly in the ring, schizont, and trophozoites stages of *P. falciparum* infection. However, based on

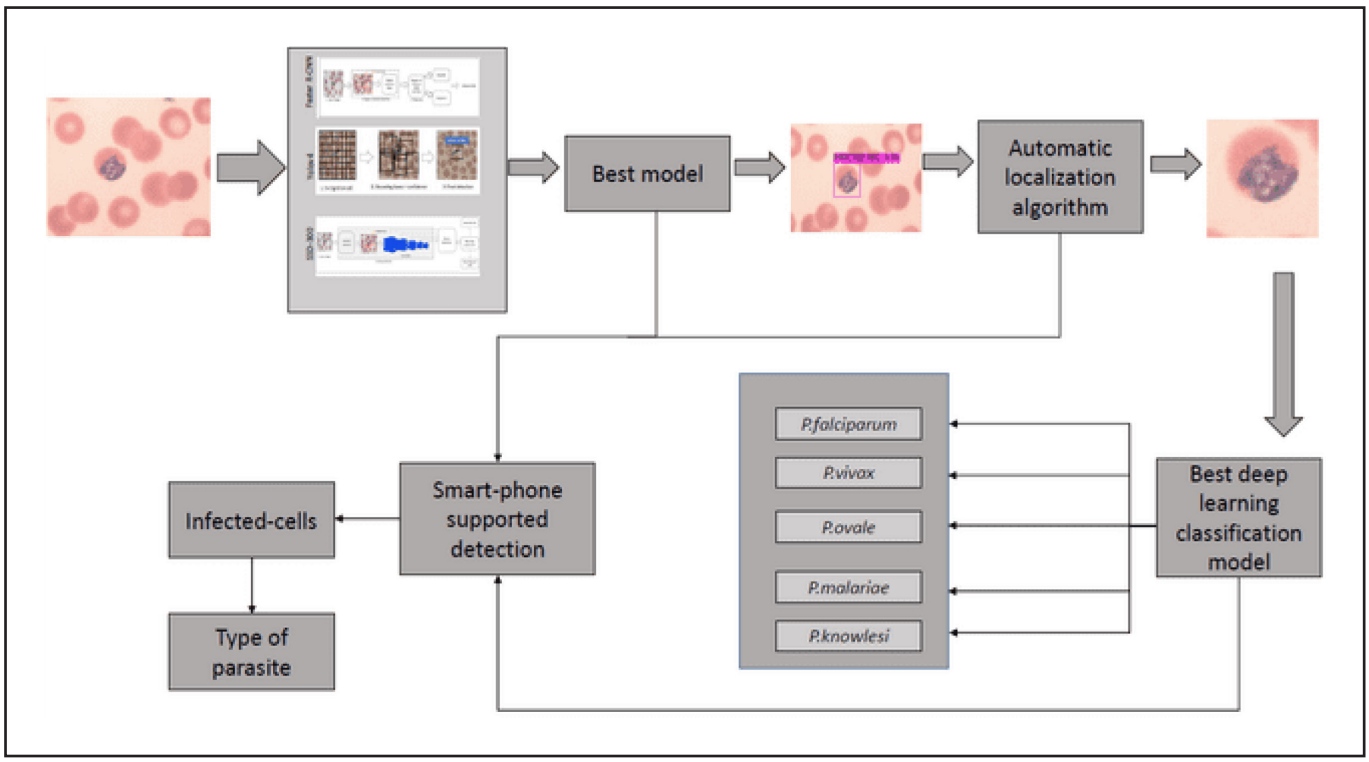

Figure 10. Overall proposed framework.

the predictions, the model can still have at least predicted two infected cells correctly in a single blood smear image, flagging that the person is infected. As mentioned earlier, for this problem, false positives are more acceptable compared to false negatives to flag as many infected cells as possible. The performance of the Yolov4 object detector shows that it is feasible to be used on a cross-dataset without fine-tuning. As for future applications, the cropped cells can facilitate further applications of deep learning in mobile malaria diagnoses, such as classifying the cells according to the species that causes the infection, as shown in the proposed framework in Figure 10.

CONCLUSION

This study investigated the capability of deep learning object detectors such as Yolov4, Faster R-CNN, and SSD300 to detect the infected cells from various malaria species and stages of infection from thin blood smear images. The Yolov4 model outperforms the other object detectors on the malaria dataset, and several experiments were conducted to determine the optimal IoU threshold for the malaria dataset. The model's robustness was tested with cross-dataset validation. This was done to test the feasibility of the model trained on the source domain on an independent dataset without fine-tuning. Deep-learning object detectors can act as an alternative to deep-learning classification models in determining infected cells from thin blood smear images more effectively, eliminating the need for single-cell images. Furthermore, the best-performing object detection architecture was integrated with a separate algorithm to crop the cells detected to create single-cell images. These single-cell images can facilitate further applications of deep learning in malaria diagnosis, such as for classification according to species or stages of infection.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

Abdurahman, F., Fante, K.A. & Aliy, M. (2021). Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models. *BMC Bioinformatics* **22**: 112. https://doi.org/10.1186/s12859-021-04036-4

Ahuja, A.S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ* **7**: e7702. https://doi.org/10.7717/peerj.7702

Akalin, F. & Yumuşak, N. (2022). Early detection of ALL disease using YOLOv4 algorithm on peripheral blood smear images. Innovations in Intelligent Systems and Applications Conference (ASYU), Antalya, Turkey. https://doi.org/10.1109/ASYU56188.2022.9925427

Arshad, Q.A., Ali, M., Hassan, S.U., Chen, C., Imran, A., Rasul, G. & Sultani, W. (2022). A dataset and benchmark for malaria life-cycle classification in thin blood smear images. *Neural Computing and Applications* **34**: 4473-4485. https://doi.org/10.1007/s00521-021-06602-6

Capela, R., Moreira, R. & Lopes, F. (2019). An overview of drug resistance in protozoal diseases. *International Journal of Molecular Sciences* **20**: 5748. https://doi.org/10.3390/ijms20225748

Daneshvar, C., Davis, T.M., Cox-Singh, J., Rafa'ee, M.Z., Zakaria, S.K., Divis, P.C. & Singh, B. (2010). Clinical and parasitological response to oral chloroquine and primaquine in uncomplicated human *Plasmodium knowlesi* infections. *Malaria Journal* **9**: 238. https://doi.org/10.1186/1475-2875-9-238

Divis, P.C.S., Singh, B., Anderios, F., Hisam, S., Matusop, A., Kocken, C.H., Assefa, S.A., Duffy, C.W. & Conway, D.J. (2015). Admixture in humans of two divergent *Plasmodium knowlesi* populations associated with different macaque host species. *PLOS Pathogens* **11**: e1004888. https://doi.org/10.1371/journal.ppat.1004888

Divis, P.C.S., Duffy, C.W., Kadir, K.A., Singh, B. & Conway, D.J. (2018). Genome-wide mosaicism in divergence between zoonotic malaria parasite subpopulations with separate sympatric transmission cycles. *Molecular Ecology* **27**: 860-870. https://doi.org/10.1111/mec.14477

Dong, Y., Jiang, Z., Shen, H., David Pan, W., Williams, L.A., Reddy, V.V.B., Benjamin, W.H. & Bryan, A.W. (2017). Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2017, Orlando, Florida, USA. https://doi.org/10.1109/BHI.2017.7897215

Gopakumar, G., Swetha, M., Siva, S.G., & Subrahmanyam, G.R.K.S. (2018). CNN based malaria diagnosis from focus-stack of blood smear images acquired using custom-built slide scanner. *Journal of Biophotonics* **11**(5): e201700003. https://doi.org/10.1002/jbio.201700003

Hung, J., Lopes, S.C.P., Nery, O.A., Nosten, F., Ferreira, M.U., Duraisingh, M.T., Marti, M., Ravel, D., Rangel, G., Malleret, B. *et al*. (2017). Applying Faster R-CNN for bbject detection on malaria images. IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

Hu, T.H., Rosli, N., Mohamad, D.S.A., Kadir, K.A., Ching, Z.H., Chai, Y.H., Ideris, N.N., Ting, L.S.C., Dihom, A.A., Kong, S.L. *et al*. (2021). A comparison of the clinical, laboratory and epidemiological features of two divergent subpopulations of *Plasmodium knowlesi*. *Scientific Reports* **11**: 20117. https://doi.org/10.1038/s41598-021-99644-8

Imran, T., Khan, M.A., Sharif, M., Tariq, U., Zhang, Y.D., Nam, Y., Nam, Y. & Kang, B.G. (2022). Malaria blood smear classification using deep learning and best features selection. *Computers, Materials & Continua* **70**: 1875-1891. https://doi.org/10.32604/cmc.2022.018946

Kassim, Y.M., Yang, F., Yu, H., Maude, R.J. & Jaeger, S. (2021). Diagnosing malaria patients with *Plasmodium falciparum* and *vivax* using deep learning for thick smear images. *Diagnostics* **11**: 1994. https://doi.org/10.3390/diagnostics11111994

Khandekar, R., Shastry, P., Jaishankar, S., Faust, O. & Sampathila, N. (2021). Automated blast cell detection for Acute Lymphoblastic Leukemia diagnosis. *Biomedical Signal Processing and Control* **68**: 102690. https://doi.org/10.1016/j.bspc.2021.102690

Koirala, A., Jha, M., Bodapati, S., Mishra, A., Chetty, G., Sahu, P.K., Mohanty, S., Padhan, K.T., Mattoo, J. & Hukkoo, A. (2022). Deep learning for real-time malaria parasite detection and counting using YOLO-mp. *IEEE Access* **10**: 102157-102172. https://doi.org/10.1109/ACCESS.2022.3208270

Krishnadas, P., Chadaga, K., Sampathila, N., Rao, S., Swathi, S.K. & Prabhu, S. (2022). Classification of malaria using object detection models. *Informatics* **9**: 76. https://doi.org/10.3390/informatics9040076

Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K., Guo, P., Hossain, Md.A., Sameer, A., Maude, R.J. *et al*. (2016). CNN-based image analysis for malaria diagnosis. IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China.

Loh, D.R., Yong, W.X., Yapeter, J., Subburaj, K. & Chandramohanadas, R. (2021). A deep learning approach to the screening of malaria infection: automated and rapid cell counting, object detection and instance segmentation using Mask R-CNN. *Computerized Medical Imaging and Graphics* **88**: 101845. https://doi.org/10.1016/j.compmedimag.2020.101845

Loddo, A., Fadda, C. & Di Ruberto, C. (2022). An empirical evaluation of convolutional networks for malaria diagnosis. *Journal of Imaging* **8**: 66. https://doi.org/10.3390/jimaging8030066

Maqsood, A., Farid, M.S., Khan, M.H. & Grzegorzek, M. (2021). Deep malaria parasite detection in thin blood smear microscopic images. *Applied Sciences* **11**: 2284. https://doi.org/10.3390/app11052284

Molina, A., Rodellar, J., Boldú, L., Acevedo, A., Alférez, S. & Merino, A. (2021). Automatic identification of malaria and other red blood cell inclusions using convolutional neural networks. *Computers in Biology and Medicine* **136**: 104680. https://doi.org/10.1016/j.compbiomed.2021.104680

Rajaraman, S., Jaeger, S. & Antani, S.K. (2019). Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PeerJ* **7**: e6977. https://doi.org/10.7717/peerj.6977

Rahman, A., Zunair, H., Reme, T.R., Rahman, M.S. & Mahdy, M.R.C. (2021). A comparative analysis of deep learning architectures on high variation malaria parasite classification dataset. *Tissue and Cell* **69**: 101473. https://doi.org/10.1016/j.tice.2020.101473

Ren, S., He, K., Girshick, R. & Sun, J. (2015). Faster R-CNN: towards real-time object detection with region proposal networks. *NeurIPS Proceedings*. https://doi.org/10.48550/arXiv.1506.01497

Tangpukdee, N., Duangdee, C., Wilairatana, P. & Krudsood, S. (2009). Malaria diagnosis: a brief review. *Korean Journal of Parasitology* **47**: 93-102. https://doi.org/10.3347/kjp.2009.47.2.93

World Health Organization (WHO). (2021a). World malaria report, 2021. World Health Organization. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021. Accessed 6 December 2021.

World Health Organization (WHO). (2021b). Global technical strategy for malaria 2016–2030. World Health Organization. https://www.who.int/publications/i/item/9789240031357. Accessed 19 July 2021.

Yang, F., Quizon, N., Yu, H., Silamut, K., Maude, R.J., Jaeger, S. & Antani, S. (2020). Cascading YOLO: automated malaria parasite detection for *Plasmodium vivax* in thin blood smears. *Medical Imaging 2020: Computer-Aided Diagnosis* **11314**, Houston, Texas, United States. https://doi.org/10.1117/12.2549701

Yunos, N.E., Sharkawi, H.M., Hii, K.C., Hu, T.H., Mohamad, D.S.A., Rosli, N., Masron, T., Singh, B. & Divis, P.C.S. (2022). Spatio-temporal distribution and hotspots of *Plasmodium knowlesi* infections in Sarawak, Malaysian Borneo. *Scientific Report* **12**: 17284. https://doi.org/10.1038/s41598-022-21439-2

Zhao, O.S., Kolluri, N., Anand, A., Chu, N., Bhavaraju, R., Ojha, A., Tiku, S., Nguyen, D., Chen, R., Morales, A. *et al*. (2020). Convolutional neural networks to automate the screening of malaria in low-resource countries. *PeerJ* **8**: e9674. https://doi.org/10.7717/peerj.9674