Tile-Based Microscopic Image Processing For Malaria Screening Using Deep Learning Approach

Fetulhak Abdurahman (afetulhak@yahoo.com)
Jimma University
Kinde Anlay Fante
Jimma University

Research Article

Keywords: Deep Learning, Malaria, Object detection, Plasmodium falciparum, Thick smear microscopic image, Tile-based image processing, YOLOV4

Posted Date: March 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1460784/v1

RESEARCH

Tile-based microscopic image processing for malaria screening using deep learning approach

Fetulhak Abdurahman[1*] and Kinde Anlay Fante[2]

*Correspondence:
afetulhak@yahoo.com
[1]Faculty of Electrical and Computer Engineering, Jimma University, 378 Jimma, Ethiopia
Full list of author information is available at the end of the article

Abstract

Background: Manual microscopic examination remains the golden standard for malaria diagnosis. However, it is cumbersome and requires the experience of pathologists for accurate diagnosis. The shear workload and challenges involved in manual microscopy drives for alternative computer-aided diagnosis techniques. While the importance of computer-aided diagnosis is increasing at an enormous pace fostered by the advancement of deep learning algorithms, there are still challenges. Existing state-of-the-art (SOTA) deep learning-based object detection models suffer from effectively detecting small objects which are less represented on benchmark datasets and are affected by the loss of detailed spatial information due to in-network feature map downscaling. The problem even becomes harder when the input image is high resolution. This is due to the fact that existing SOTA models will not directly process high-resolution images limited by their low-resolution network input size.

Methods: In this study, an effective and robust tile-based image processing approach was proposed to enhance malaria parasite detection performance of SOTA object detection models. Three YOLOV4 based object detectors were adopted considering their detection speed and accuracy. The proposed detection models were developed by using tiles generated from 1,780 high-resolution P. falciparum-infected thick smear microscopic images. External validation was performed to verify the detection accuracy and generalization ability of the proposed models on three datasets acquired from a different domain.

Results: The best-performing model leveraging the proposed tile-based approach significantly outperforms its baseline method (Recall, [95.3%] vs [57%] and Average Precision, [87.1%] vs [76%]). In addition, the proposed method achieved state-of-the-art performance compared with previous research works which use different machine learning techniques on similar datasets.

Conclusions: The obtained qualitative and empirical results show that the proposed method reveals a fundamental performance improvement for the detection of P. falciparum from thick smear microscopic images while maintaining real-time speed. Furthermore, the proposed method could have the potential to assist and reduce the workload of laboratory technicians in malaria-endemic remote areas in developing countries where there is a critical skill gap and shortage of experts.

Keywords: Deep Learning; Malaria; Object detection; Plasmodium falciparum; Thick smear microscopic image; Tile-based image processing; YOLOV4

1 Background

Malaria is one of the most fatal diseases and the cause of high mortality rate in the world, which is transmitted from infected to healthy humans through the bites

of female Anopheles mosquitoes. It is caused by a unicellular parasite known as plasmodium. Once the parasite enters the human body it grows inside the liver and is released to the bloodstream to infect red blood cells (RBCs). Plasmodium parasite has five species including, plasmodium falciparum (P. falciparum), plasmodium vivax (P. vivax), plasmodium ovale (P.ovale), plasmodium Knowlesi (P. knowlesi), and plasmodium malariae (P. malariae); where P. falciparum and P. vivax are the most pathogenic and infect majority of the global population. According to statistics published by World Health Organization (WHO) malaria report 2020 an estimated 241 million malaria cases were reported in 2020, among which 95% of the cases are in WHO African region with only the remaining 5% outside African region. Globally, there were 627,000 malaria deaths in 2020 among which 96% of the cases occurred in 29 countries, and six countries – Nigeria (27%), the Democratic Republic of Congo (12%), Uganda (5%), Mozambique (4%), Angola (3.4%) and Burkina Faso (3.4%) – accounted for about 55% of all cases globally [1].

There are various approaches for malaria diagnosis including clinical diagnosis, microscopic diagnosis, rapid diagnostic test kits (RDTs) and polymerise chain reaction (PCR). Clinical diagnosis is based on various symptoms of malaria such as the history of fever and it has low specificity which leads to significant overuse of antimalarial drugs [2]. PCR is the most sensitive method but it is costly and complex, whereas RDTs are highly sensitive and are unable to quantify parasite density [3]. Parasitological confirmation by microscopy using thin/thick blood film remains the golden standard for malaria diagnosis, but it is a cumbersome method [4]. In the procedures of manual microscopy diagnosis, a thin or thick blood smear is prepared by spreading a drop of blood on a glass slide which is dried and stained before being visually examined by a microscopist for parasite identification. A thick smear has a large volume of blood and a large number of parasites per blood volume. It is usually used for determining whether the patient's blood contains malaria parasites or not. A thin blood smear has low blood volume and less number of parasites per blood volume and it is often used both for parasitemia detection and identification of parasite species and their life stage.

The accuracy of manual microscopic examination is severely affected by high intra/inter-observer variability which is supplemented by a large number of cases diagnosed per day in malaria-endemic regions with low resource settings [5]. Besides, visual examination using manual microscopy is tedious and time-consuming [6]. The shortage of trained experts and lack of a rigorous system to support expertise skill gaps at malaria-endemic regions leads to incorrect diagnosis results which contribute to inappropriate treatment [7], [8]. The challenges in the manual microscopy diagnosis procedure motivate the research community to develop automated computer-aided diagnostic (CAD) systems to improve malaria diagnosis accuracy and reduce the clinical challenges due to human error by empowering microscopists' diagnostic decisions for ultimate better patient treatment.

Most existing studies combine traditional image processing techniques with classical machine learning algorithms to develop automated malaria diagnosis systems. Traditional image processing approaches such as adaptive threshold techniques or other morphological operations have been used to segment the parasite candidates from the background of either thick or thin smear microscopic images [9],

[10],[11],[8]. However, such image processing approaches are very sensitive to variations in image quality since the segmentation techniques are determined empirically. On the other hand, conventional machine learning models using handcrafted features from segmented malaria parasite region of interest (ROI) are used for classifying malaria-infected and uninfected cells [11], [12],[13]. However, classical machine learning algorithms which are based on hand-engineered features struggle to generalize when their input data domain variation increases. Recently, the advancement of deep learning algorithms has celebrated a lot of achievements in its application in a range of medical imaging tasks including medical image segmentation and reconstruction [14], [15], classification tasks [16],[17], and object detection [18]. Furthermore, recent studies show deep learning-based algorithms outperformed conventional image processing and machine learning techniques for detection and identification of malaria parasites using microscopic images of thin and thick blood smear [19], [20].

1.1 Problem Statement and Motivation

Existing state-of-the-art (SOTA) deep learning-based object detection models are trained and evaluated on large-scale datasets such as ImageNet [21], Pascal VOC [22], and MS COCO [23] which are generic and largely represent medium and large scale objects. Although existing SOTA deep learning-based object detection models perform well for medium and large scale object detection tasks their direct application for specific object detection tasks, especially for small objects like malaria parasite screening, will not achieve a very good detection performance [24] [25]. When an input image passes through different layers of SOTA detectors it will lose too much spatial information which is crucial for small object localization. When the input image is high resolution (HR) another problem arises due to resizing the input HR image to detection networks input resolution. This is because existing SOTA detection models have specified low-resolution network input sizes to keep their computational demands low. These two problems 1) the in-network feature map downscaling leading to massive spatial information loss for small object localization and 2) detail information loss due to compressing the high-resolution image to low resolution network input size, exacerbate small object detection problem for SOTA deep learning-based object detection models.

Nowadays, the advancement in digital data acquisition technologies in medical imaging has enabled pathologists to acquire HR images which are very important for various clinical management tasks [26]. However, HR images pose a formidable computational challenge for deep learning algorithms. Similarly, microscopic images captured through the eyepiece of a microscope by attaching a digital camera or smartphone camera have high resolution [7], [27]. Previous works [19], [28] for malaria parasite detection leverage such HR images directly to SOTA deep learning-based object detectors. However, directly feeding HR microscopic images to SOTA deep learning algorithms to detect very tiny objects - for example, P. falciparum, is almost impossible. The motivation of this study was to explore an effective and robust approach to process HR thick smear microscopic images in computationally-efficient SOTA deep learning-based object detection algorithms without degrading the detailed information found in the original HR images. In this study, tile-based

image processing was proposed to enhance small object detection accuracy of computationally efficient SOTA deep learning-based object detection models, which are limited by their network input resolution to process HR images. Even though there are similar research works that explore tile-based image processing for other problem domains such as remote sensing imagery [29], [30], to the best of our knowledge there is no other work that employed tile-based microscopic image processing for malaria parasite screening. The proposed approach does not apply dynamic tile processing since the parasites are sparsely located at any point in the microscopic image. The proposed method achieved a big performance improvement for SOTA object detection models and at the same time, it did not incur too much performance degradation on the models' inference speed.

The contributions of this study are as follows:

1 An effective and robust tile-based high-resolution thick smear microscopic image processing approach was explored for the detection of P. falciparum using SOTA deep learning-based object detection models. The proposed method significantly improves the detection performance of proposed SOTA object detection models with minimal effect on inference speed.
2 Extensive experimental analysis was performed with various combinations and evaluation strategies to investigate detection performance and generalization ability of SOTA object detection models by using datasets collected from different domains.
3 Detailed comparative analysis of the proposed method with baseline methods as well as with previously reported results was performed. The proposed method achieved state-of-the-art performance compared to baseline methods and previous works on three representative datasets for P. falciparum detection.

2 Related Work

In the past decade, attempts have been made to develop automated malaria parasite detection by using both thin and thick blood smear microscopic images. However, the majority of such automated diagnosis tools are developed using conventional image processing techniques and classical machine learning approaches which are based on handcrafted features, the reader is referred to the following literature review works done in a similar area [31],[32],[3], [33], [34].

Recently, the advancement in the field of machine learning with convolutional neural networks (CNNs) attracted much attention for medical image analysis including malaria parasite identification from thin and thick blood smear microscopic images [35]. This is due to their superpower performance compared to handcrafted feature extraction-based techniques by automatically learning robust feature representation from raw image pixels [36]. Several studies showcased the applicability of deep learning algorithms for the identification of malaria parasites from microscopic images of both thin and thick blood film. A study done by [7] proposed a pre-parasite candidate selection using intensity value of grayscale of thick blood film microscopic image and apply CNN for its final classification. The proposed parasite candidate selection technique will not be effective as it is evidenced for traditional image processing techniques when the technique is applied for images

obtained in various environmental conditions. In addition, the candidate parasite selection method will have a direct effect on the performance of the CNN classifier. In this work they have used a similar dataset with our study and their proposed system achieved a sensitivity of 82.73% and precision of 78.98% on test set data of 30 patients at the patch level.

Another study by [19] proposed a multi pipeline approach by applying Mask-RCNN as pre-candidate P. falciparum and P. vivax species detector followed by classifier head to filter out false positives. They evaluate their detection system using image-level and patient-level by using experimentally defined threshold scores. They achieved an accuracy of 90.8% on image level and 97.6 accuracy on patient-level evaluation. They have used a similar dataset to ours for P. faciparum detection but they did not evaluate their system at the patch level. Another study [20] proposed a dual deep learning framework for RBC segmentation using thin smear images. They have used U-Net as a pre-candidate RBC cluster segmentation and applied Faster-RCNN for the final detection of RBCs. However, they did not classify between malaria-infected and uninfected RBCs.

A mobile-based P. falciparum and white blood cell(WBC) localization using pre-trained deep learning models are proposed by [28]. In this study they prepared a new dataset of 903 fields stained thick blood smear microscopic images to train and evaluate their proposed models. Another study by [37] proposes an ensemble of pre-trained and custom CNN models for classification of infected and uninfected RBC cells segmented from thin blood smear microscopic images. In [38],[39] modified versions of YoloV3 and YoloV4 models are proposed to improve P. falciparum detection capability of SOTA deep learning models on thick smear microscopic images and to make detection models lightweight to be integrated with mobile phone-based diagnosis application.

Detection of malaria parasites using thick smear microscopic images is very difficult due to the very small size of the malaria parasite and most proposed deep learning models by previous research works have shown low performance for detection of parasites in thick smear images. However, thick blood film is the most recommended slide preparation technique for malaria diagnosis, and the development of robust automated tools will be very significant to reduce the challenges related to manual microscopy. Most of the existing studies on malaria parasite detection or classification are done using thin smear blood films [40], [41],[42],[43],[44]. This may be due to the easiness of detecting infected and uninfected RBCs due to their bigger size.

3 Methods

This study aims to improve detection performance and evaluate the effectiveness of SOTA deep learning-based object detection models for P. falciparum detection from thick smear microscopic images by solving the aforementioned challenges described in section 1.1. To this goal, a tile-based image processing was proposed to increase small object detection capability of SOTA object detection algorithms for operating on high-resolution thick smear microscopic images than their network input resolution allows. The general overview of the proposed scheme is illustrated in Figure 1.

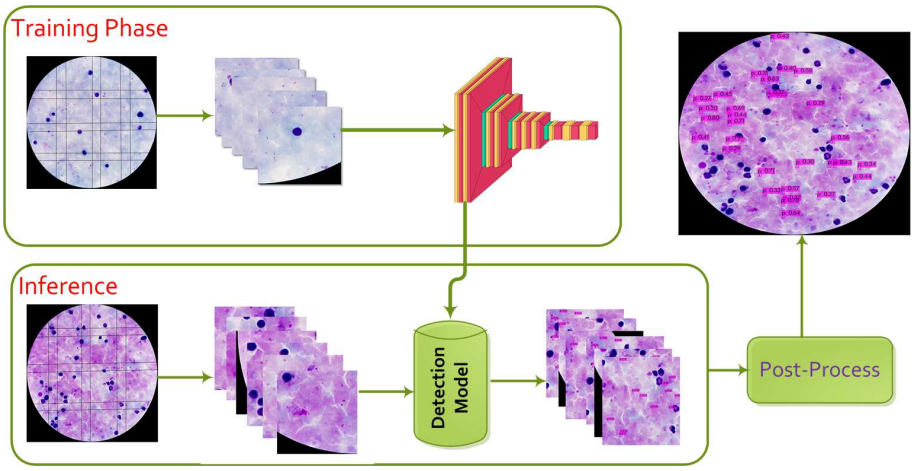

Figure 1: General overview of the proposed framework. First, the high-resolution image is sliced into overlapping tiles and fed as an input to the detection network for training. Then, the trained model is used during inference to predict parasite location at individual tiles. Finally, the initial prediction results are fused to generate refined detections to be superimposed on the input high-resolution image.

In the proposed approach the HR image was divided into overlapping small images called tiles so that each can then be fed to SOTA object detection networks without degrading the original image resolution. During model training tiles with their corresponding ground truth annotation mapped accordingly were used as an input and tiles which did not have ground truth annotation were excluded. Note that some objects may be cut at tile boundaries and the ratio of areas between partially segmented parts and the full object was taken to keep or discard the annotation during model training. Dividing HR images into tiles enabled to increase the relative area of small objects for the detection network input resolution.

At inference time a similar approach was followed by dividing the high-resolution inference image into smaller overlapping tiles and initial detection results were obtained for individual tiles. Then, the initial detection results at tile level were merged and non-maximum suppression (NMS) algorithm was applied to avoid duplicate detections at overlapping regions with an intersection over union (IOU) threshold of 30%. Finally, the refined detection results were stitched on the input high-resolution images. The proposed tile-based approach demonstrated that computationally efficient SOTA deep learning models, which are limited by their network input resolution to process HR images, detect small objects with increased detection accuracy, and at the same time the proposed approach does not increase too much computational overhead during inference time.

3.1 Datasets

To verify the effectiveness of the proposed approach in this work experiments were run by using two kinds of datasets. The first one, which is called model development dataset, is used for training, validation, and testing the proposed SOTA detection

models and finally to select the best model considering its detection accuracy and computation speed. This dataset consists of high-resolution thick smear microscopic images infected with P. falciparum which is acquired from previous researchers [7]. It is collected from 150 patients at Chittagong Medical College Hospital, Bangladesh, and manually annotated by experienced experts. The dataset contains an average number of 12 images per patient and 47 parasites per image. The images are in RGB color and have a high resolution of 4032 x 3024 pixels. Sample images from the development dataset with their corresponding annotations are shown in figure 2.

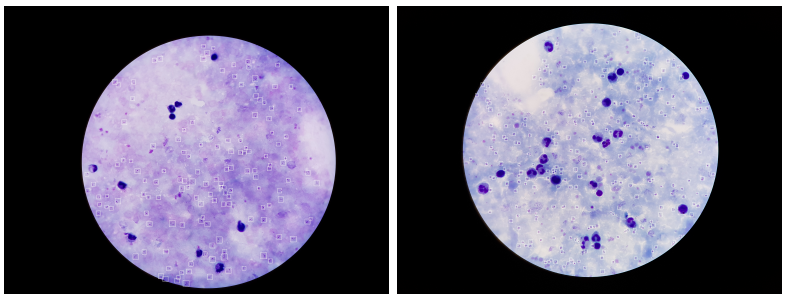

Figure 2: Sample images taken from the model development dataset. The boxes indicate ground truth bounding box locations of P. falciparum.

Table 1: Dataset Description

	Training	Validation	Testing	Total
Model Development Dataset				
Number of patients	96	24	30	150
Number of images	1,140	266	374	1,780
Number of parasites	49,520	11,174	22,783	83,477
External Dataset 1				
Number of images	958	106	118	1,182
Number of parasites	6,214	694	719	7,628
External Dataset 2				
Number of images	664	166	100	930
Number of parasites	7,734	470	1,035	9,239
External Dataset 3				
Number of images	-	-	-	1,141

The second dataset, which is obtained from a different domain, is used as an external dataset to further evaluate the generalization capability of the proposed detection model which is selected as the best model based on the model development dataset. In addition, the external dataset was integrated with model development dataset to fine-tune the selected model. The external dataset consists of three different datasets. The first external dataset, which is released by previous research work [10], consists of a low resolution (750 x 750 pixels) thick smear microscopic images infected with P. falciparum collected from 133 individuals. The second external dataset is also a collection of thick smear microscopic images obtained from Mulago National referral hospital in Uganda and released by previous research work [28]. This dataset consists of 903 high-resolution images of size 3264 x 2448 pixels. The third external dataset is used to test performance of the proposed detection model on negative images. This dataset consists of 1141 thick smear microscopic

images obtained from 50 uninfected patients and it was released by previous research work [19]. All the datasets are captured by attaching a mobile phone camera to the eyepiece of microscope and they differ in their staining style and imaging characteristics. A detailed description of the dataset is given in Table 1.

3.2 Proposed Detection Networks

The proposed P. falciparum detection algorithms were based on the YOLOV4 object detection model [45] which is chosen considering its high detection performance and inference speed for downstream object detection tasks compared to other single-stage and two-stage detectors [45]. In this study, three different YOLOV4 based object detection models were evaluated to identify the best model with an optimal trade-off between detection performance and inference speed. The first detection network, which is called YOLOV4-MOD, is based on previous work [39] which is modified to improve small object detection performance of the original YOLOV4 model with minimal computation cost. This model has a large number of trainable parameters and more convolution layers. The second two detection networks are based on lightweight versions of YOLOV4 known as tiny YOLOV4 models, which are designed for improved inference speed with compromise on their detection accuracy. Among these lightweight models, the first model, which is called YOLOV4-tiny, has two detection heads while the second detection model, which is called YOLOV4-tiny-3l, has three detection heads. The addition of more detection heads on YOLO-based models enabled better detection performance for small objects [39]. These two lightweight networks have a small number of trainable parameters and fewer convolution and max-pooling layers compared to the large size model.

3.3 **Evaluation Metrics**

In this study, two widely used evaluation metrics for object detection tasks, namely *average precision (AP)* and *recall (R)*, were used to evaluate performance of the proposed models. The average precision is based on the area under interpolated Precision-Recall Curve (PRC), where precision is calculated as the ratio of the number of true positive detections to all detected objects. The recall measures the fraction of detections that are true positive. The formulas for the evaluation metrics are given below.

$$Precsion = \frac{|TP|}{|TP + FP|} \qquad Recall = \frac{|TP|}{|TP + FN|}$$

$$AP = \frac{1}{11} \sum_{r \epsilon \{0.0, \ldots 1.0\}} p_{interp}(r)$$

where

$$p_{interp}(r) = \max_{\widetilde{r} \geq r} p(\widetilde{r})$$

where True Positive (TP) indicates the number of correctly detected objects, False Positive (FP) indicates the number of incorrectly predicted suspicious objects, and False Negative (FN) indicates the number of undetected objects. p_{interp} represents the interpolated precision (p) over a given recall (r) values in an ascending order from 0.0 to 1.0 into 11 points — 0, 0.1, 0.2, . . . , 0.9 and 1.0.

4 Experiments

4.1 Experimental Setup

In this study, several experimental settings were followed to develop and evaluate the effectiveness of the proposed tile-based thick smear microscopic image processing approach for malaria parasite screening. First, the three proposed YOLOV4 based detection models described in section 3.2 were trained using tiles with different tile sizes generated from the model development dataset described in section 3.1. Afterward, the best model is selected considering its trade-off between computation speed and detection performance on the model development test data. The dataset is divided into training, validation, and testing at patient level where 96 patients were used for training, 24 patients for validation, and 30 patients for testing. In this experimental setting, optimal network hyper-parameters of the detection models were also selected by using the model development validation dataset.

After selection of the best model among the three proposed detection models, its P. falciparum detection performance was tested on the first and second external datasets described in section 3.1. When the selected model was evaluated using the first external dataset tile-based processing was not applied since resolution of the dataset is not high compared to the proposed detection network's input resolution. The second external dataset consists of high-resolution thick smear microscopic images and the tile-based approach was applied when it was evaluated using this dataset. Furthermore, two experimental setups were followed during evaluation of the selected model using these two external datasets. In the first experimental setting, the selected model was evaluated by utilizing the whole dataset as test data. In the second experimental setting, the datasets were partitioned into training, validation, and test as shown in Table 1 and combined with the model development dataset to fine-tune the selected model.

In addition, the selected model was evaluated on 1141 uninfected images collected from 50 uninfected patients [19] which have not been used during model training. This dataset is a high-resolution thick smear microscopic image and has a resolution of 4032 x 3024 pixels. Thus, the proposed tile-based approach was applied to investigate performance of the selected model on identifying uninfected images.

Baseline Method: The proposed tile-based approach was compared with a baseline method in which a full high-resolution image is downscaled to the proposed SOTA detection network's input resolution both during model training and inference. This technique offers fast training and inference speed but it leads to detailed information loss found in the original high-resolution image which resulted in massive detection performance degradation.

4.2 Proposed Detection Network Training and Hyper-parameter Optimization

During proposed model training, hyperparameter optimization, and selection of best model a publicly available high-resolution thick smear microscopic image dataset

collected from [7] was used. Different tile sizes relative to the detection network's input resolution were utilized to prevent loss of detailed information in HR images and to keep the computation cost optimal as well. To overcome the problem of a limited dataset in all experimental settings pre-trained models using MS COCO dataset [23] were leveraged and fine-tuning was applied by using the target dataset.

In the training process, default configurations in original versions of the proposed YOLOV4 based models were used unless specified. Anchor box sizes were modified based on the network input resolution and ground truth bounding box information of the specific dataset used for training. A batch size of 2 for large size YOLOV4 based model (YOLOV4-MOD) and a batch size of 8 for lightweight models (YOLOV4-tiny and YOLOV4-tiny-3l) was used. All the models were trained for 4000 iterations using the default settings for data augmentation, optimizer, and loss functions. The initial learning rate was 0.001 for large size model and 0.00261 for lightweight models and decreased by a factor of 10 at 80% and 90% of the training iteration. The detection network input sizes were set as 416 x 416 and 512 x 512 for large size model due to computational constraint and 416 x 416, 512 x 512 and 608 x 608 for lightweight models. All the experiments were conducted by using Google Colaboratory with NVIDIA TESLA K80 processor and 12 GB of RAM.

The proposed tile-based image processing approach introduces additional parameters such as tile size and overlapping ratio, that must be tuned during model validation. Tile sizes were selected relative to the proposed detection networks' input size. Too large or small tile sizes were not selected compared with detection networks' input resolution since large tile sizes contribute to loss of detail information due to resizing. On the other hand, the selection of too small tile sizes will add too much computation time and affect detection accuracy due to upsampling to network input resolution.

Table 2: Comparisons of detection performance and inference speed for YOLOV4-MOD by using the proposed and baseline approach on model development test set data. The highlighted rows in blue indicate the best-performing model.

Model	Input Res.	Tile size		AP (%)	R (%)	sec/img
		train	inference			
YOLOV4-MOD@416	4032 x 3024	1088 x 1088	1088 x 1088	83.1	94.3	3
YOLOV4-MOD@416	4032 x 3024	1088 x 1088	832 x 832	85.5	95.1	4
YOLOV4-MOD@416	4032 x 3024	1088 x 1088	608 x 608	82.8	91.3	8
YOLOV4-MOD@512	4032 x 3024	1088 x 1088	1088 x 1088	81.8	90.5	4
YOLOV4-MOD@512	4032 x 3024	1088 x 1088	832 x 832	82.5	92.6	5
YOLOV4-MOD@512	4032 x 3024	1088 x 1088	608 x 608	78.2	91.7	9
YOLOV4-MOD@416	4032 x 3024	832 x 832	1088 x 1088	80.6	90.8	3
YOLOV4-MOD@416	4032 x 3024	832 x 832	832 x 832	85.0	93.5	4
YOLOV4-MOD@416	4032 x 3024	832 x 832	608 x 608	75.0	86.1	7
YOLOV4-MOD@512	4032 x 3024	832 x 832	1088 x 1088	66.1	82.3	4
YOLOV4-MOD@512	4032 x 3024	832 x 832	832 x 832	82.5	93.2	5
YOLOV4-MOD@512	4032 x 3024	832 x 832	608 x 608	79.6	92.7	10
YOLOV4-MOD@416	4032 x 3024	608 x 608	1088 x 1088	43.5	63.3	3
YOLOV4-MOD@416	4032 x 3024	608 x 608	832 x 832	66.7	82.5	3.5
YOLOV4-MOD@416	4032 x 3024	608 x 608	608 x 608	76.8	90.6	6
YOLOV4-MOD@512	4032 x 3024	608 x 608	1088 x 1088	55.3	85.7	3
YOLOV4-MOD@512	4032 x 3024	608 x 608	832 x 832	69.0	85.6	5
YOLOV4-MOD@512	4032 x 3024	608 x 608	608 x 608	77.9	90.1	8.5
with out tiling						
YoloV4-MOD@512	4032 x 3024	-	-	79.78	80	0.25
YoloV4-MOD@416	4032 x 3024	-	-	72.8	76	0.2

Based on experimental analysis tiles of size 1088 x 1088, 832 x 832 and 608 x 608 provide optimal detection performance for proposed models. Considering the selected tile sizes and the proposed models' network input resolution, which is 416 x 416, 512 x 512, and 608 x 608, the loss of detailed information due to resizing is reduced compared to directly resizing the full high resolution (4032 x 3024 pixels) image. As shown in Table 2, the proposed large size model performs well on large tile sizes whereas decreasing the tile size becomes detrimental for the model to differentiate objects at a similar scale to P. falciparum which resulted in a large number of false-positive detections. However, lightweight models perform better on small size tiles as shown in Tables 3 and 4. The overlapping ratio between tiles were selected based on experimental results on the validation dataset. The overlap between tiles prevents missing objects due to image partitioning at tile boundaries. An overlap ratio of 0.2 achieved optimal detection accuracy for the proposed detection models.

During inference time the tile size can be different from the one used during model training. Thus, in all experimental settings, the proposed detection models trained in one of the selected tile sizes were evaluated on three different tile sizes at inference time, to analyze the effect of varying tile sizes at train and inference time. There is a relationship between training and inference tile sizes such that the proposed detection models perform better when the inference tile sizes are equal or less than the training tile sizes. This may be due to the lack of CNNs ability for strong generalization across scales [46] and the effect of input image downscaling to the network's input resolution which contributes to loss of some detailed information. When the network input resolution is increased for the large size model it does not improve the detection performance of the proposed models. This may be due to the deep network architecture of the large size model which contributes to equivalent localization feature for objects on a similar scale with P. falciparum.

5 Results

In Tables 2, 3, and 4 detail experimental results are illustrated for the different experimental settings described in section 4.1. The experimental results demonstrate the effects of various techniques used in this study, including different detection networks proposed, variation in network input resolution, and tile size variation both during training and inference stage. By using the model development test data, which consists of 30 patients and 374 images, among the proposed three detection models, YOLOV4-tiny with 512 x 512 input resolution and tile size of 608 x 608 both at train and inference time performs well with a maximum recall of 95.3% and maximum average precision of 87.1%. In a similar test set data, YOLOV4-tiny-3l with 608 x 608 input resolution and tile size of 608 x 608 both at train and inference time achieved a maximum recall of 95.1% and maximum average precision of 87.4%. The large size model (YOLOV4-MOD) with input resolution of 416 x 416 achieved a maximum recall of 95.1% and maximum average precision of 85.5% by using a training tile size of 1088 x 1088 and inference tile size of 832 x 832. YOLOV4-tiny and YOLOV4-tiny-3l perform 2.6× and 2× faster compared to YOLOV4-MOD with better recall and average precision respectively. Surprisingly, lightweight models achieve a better trade-off between detection performance and inference speed for

Table 3: Comparisons of detection performance and inference speed for YOLOV4-tiny by using the proposed and baseline approach on model development test set data. The highlighted rows in blue indicate the best-performing model.

Model	Input Res.	Tile size		AP(%)	R(%)	sec/img
		train	inference			
YOLOV4-tiny@416	4032 x 3024	1088 x 1088	1088 x 1088	56	60	1
YOLOV4-tiny@416	4032 x 3024	1088 x 1088	832 x 832	67	73	1
YOLOV4-tiny@416	4032 x 3024	1088 x 1088	608 x 608	78	86	1.4
YOLOV4-tiny@416	4032 x 3024	1088 x 1088	416 x 416	65	75.8	1.4
YOLOV4-tiny@512	4032 x 3024	1088 x 1088	1088 x 1088	70	75.9	1
YOLOV4-tiny@512	4032 x 3024	1088 x 1088	832 x 832	80.4	86.1	1.2
YOLOV4-tiny@512	4032 x 3024	1088 x 1088	608 x 608	82.6	91.4	1.6
YOLOV4-tiny@608	4032 x 3024	1088 x 1088	1088 x 1088	77.9	83.4	1
YOLOV4-tiny@608	4032 x 3024	1088 x 1088	832 x 832	85.9	92.7	1.3
YOLOV4-tiny@608	4032 x 3024	1088 x 1088	608 x 608	81.9	89.3	1.7
YOLOV4-tiny@416	4032 x 3024	832 x 832	1088 x 1088	53.3	57.6	1
YOLOV4-tiny@416	4032 x 3024	832 x 832	832 x 832	66	71.9	1.1
YOLOV4-tiny@416	4032 x 3024	832 x 832	608 x 608	81.4	88.1	1.6
YOLOV4-tiny@512	4032 x 3024	832 x 832	1088 x 1088	66.7	72.8	1
YOLOV4-tiny@512	4032 x 3024	832 x 832	832 x 832	79.1	85.0	1.1
YOLOV4-tiny@512	4032 x 3024	832 x 832	608 x 608	86.0	94.1	1.6
YOLOV4-tiny@512	4032 x 3024	832 x 832	416 x 416	78.8	89.2	2.7
YOLOV4-tiny@608	4032 x 3024	832 x 832	1088 x 1088	76.7	82.6	1
YOLOV4-tiny@608	4032 x 3024	832 x 832	832 x 832	86	92.9	1.3
YOLOV4-tiny@608	4032 x 3024	832 x 832	608 x 608	84.4	91.5	1.6
YOLOV4-tiny@416	4032 x 3024	608 x 608	1088 x 1088	48	57.1	1
YOLOV4-tiny@416	4032 x 3024	608 x 608	832 x 832	60.9	67.1	1
YOLOV4-tiny@416	4032 x 3024	608 x 608	608 x 608	83.2	90.1	1.3
YOLOV4-tiny@416	4032 x 3024	608 x 608	416 x 416	84.0	91.1	2
YOLOV4-tiny@512	4032 x 3024	608 x 608	1088 x 1088	54.3	60.8	1
YOLOV4-tiny@512	4032 x 3024	608 x 608	832 x 832	76.4	79.8	1
YOLOV4-tiny@512	4032 x 3024	608 x 608	608 x 608	87.1	95.3	1.5
YOLOV4-tiny@512	4032 x 3024	608 x 608	416 x 416	84.0	94.9	2.6
YOLOV4-tiny@608	4032 x 3024	608 x 608	1088 x 1088	69.4	85.7	1
YOLOV4-tiny@608	4032 x 3024	608 x 608	832 x 832	84.3	92.1	1.3
YOLOV4-tiny@608	4032 x 3024	608 x 608	608 x 608	87.0	95.3	1.9
with out tiling						
YoloV4-tiny@416	4032 x 3024	-	-	54	21	0.15
YoloV4-tiny@512	4032 x 3024	-	-	69	48	0.15
YoloV4-tiny@608	4032 x 3024	-	-	76	57	0.15

P. falciparum detection compared to the large YOLOV4-MOD model by leveraging the proposed tile-based approach. Among the lightweight models, YOLOV4-tiny was selected as the best model with comparable performance with YOLOV4-tiny-3l but has high computation speed.

Comparison to Baseline method: The proposed detection models' performance was also compared with the baseline method which directly applies high-resolution images during model training and inference. As shown in Table 2, 3 and 4 detection performance of models using the proposed tile-based approach shows significant performance improvement compared to their baseline counterparts. Compared to their baseline models, lightweight models perform 10× slower in computation speed but their detection performance increase with a significant amount, which is about 17% in terms of recall for yolov4-tiny-3l and about 38% for the yolov4-tiny model. The detection performance improvement due to the proposed approach for YOLOV4-tiny-3l is lower than YOLOV4-tiny since the additional detection head enabled better detection accuracy on small objects.

Assessment of the selected model on external datasets: Furthermore, performance of the selected model (YOLOV4-tiny) was evaluated on two external

Table 4: Comparisons of detection performance and inference speed for YOLOV4-tiny-3l by using the proposed and baseline approach on model development test set data. The highlighted rows in blue indicate the best-performing model.

Model	Input Res.	Tile size		AP(%)	R(%)	sec/img
		train	inference			
YOLOV4-tiny-3l@416	4032 x 3024	1088 x 1088	1088 x 1088	55.4	59.8	1
YOLOV4-tiny-3l@416	4032 x 3024	1088 x 1088	832 x 832	65.9	71.2	1
YOLOV4-tiny-3l@416	4032 x 3024	1088 x 1088	608 x 608	76.9	84.6	1.4
YOLOV4-tiny-3l@512	4032 x 3024	1088 x 1088	1088 x 1088	67.2	72.8	1.3
YOLOV4-tiny-3l@512	4032 x 3024	1088 x 1088	832 x 832	78.0	84.3	1.4
YOLOV4-tiny-3l@512	4032 x 3024	1088 x 1088	608 x 608	83.4	92.3	2
YOLOV4-tiny-3l@608	4032 x 3024	1088 x 1088	1088 x 1088	78.4	84.3	1.2
YOLOV4-tiny-3l@608	4032 x 3024	1088 x 1088	832 x 832	86.1	93.0	1.6
YOLOV4-tiny-3l@608	4032 x 3024	1088 x 1088	608 x 608	81.6	89.3	2
YOLOV4-tiny-3l@416	4032 x 3024	832 x 832	1088 x 1088	50	53.7	1
YOLOV4-tiny-3l@416	4032 x 3024	832 x 832	832 x 832	64.3	70	1
YOLOV4-tiny-3l@416	4032 x 3024	832 x 832	608 x 608	80.5	87.7	1.5
YOLOV4-tiny-3l@512	4032 x 3024	832 x 832	1088 x 1088	64.9	71.4	1.1
YOLOV4-tiny-3l@512	4032 x 3024	832 x 832	832 x 832	78.1	84.5	1.4
YOLOV4-tiny-3l@512	4032 x 3024	832 x 832	608 x 608	85.7	94.2	1.9
YOLOV4-tiny-3l@608	4032 x 3024	832 x 832	1088 x 1088	74.5	81.0	1.3
YOLOV4-tiny-3l@608	4032 x 3024	832 x 832	832 x 832	85.4	92.6	1.6
YOLOV4-tiny-3l@608	4032 x 3024	832 x 832	608 x 608	84	91	2
YOLOV4-tiny-3l@416	4032 x 3024	608 x 608	1088 x 1088	41.8	57.1	0.8
YOLOV4-tiny-3l@416	4032 x 3024	608 x 608	832 x 832	64.4	71.4	1
YOLOV4-tiny-3l@416	4032 x 3024	608 x 608	608 x 608	81.7	89.7	1.4
YOLOV4-tiny-3l@512	4032 x 3024	608 x 608	1088 x 1088	51.8	57.2	1
YOLOV4-tiny-3l@512	4032 x 3024	608 x 608	832 x 832	74.6	81.1	1.3
YOLOV4-tiny-3l@512	4032 x 3024	608 x 608	608 x 608	87.1	95.0	1.9
YOLOV4-tiny-3l@608	4032 x 3024	608 x 608	1088 x 1088	69.2	85.7	1
YOLOV4-tiny-3l@608	4032 x 3024	608 x 608	832 x 832	83.8	91.9	1.4
YOLOV4-tiny-3l@608	4032 x 3024	608 x 608	608 x 608	87.4	95.1	1.9
with out tiling						
YoloV4-tiny-3l@416	4032 x 3024	-	-	71.46	71	0.2
YoloV4-tiny-3l@512	4032 x 3024	-	-	79.4	78	0.2
YoloV4-tiny-3l@608	4032 x 3024	-	-	78.73	76	0.2

datasets collected from a different domain. By utilizing the first external dataset which is a low-resolution image obtained from [10], an average precision of 57.8% and recall of 75.1% was achieved by using the whole dataset as test data. Similarly, by using the second external data obtained from [28] an average precision of 71.1% and recall of 86.3% was achieved.

In addition, the selected model was fine-tuned and evaluated by partitioning the two external datasets into training, validation, and testing and merging them with the model development dataset. In that way, an average precision of 83.4% and a recall of 94.7% was achieved on test data of the first external dataset and an average precision of 73.1% and recall of 96.3% was achieved on test data of the second external dataset.

5.1 Qualitative Results

The qualitative comparison, which shows the profound effect of the tile-based approach on SOTA deep learning-based object detectors for P. falciparum detection from high-resolution thick smear microscopic images, is shown in Figure 3. As shown in the figure detection models using the proposed method show very good detection results. The sample visualization results in the first row are detection results based on the selected YOLOV4-tiny model using the proposed tile-based approach.

The second row contains detection results for YOLOV4-tiny using the baseline approach. Ground truth bounding boxes are in green and predicted boxes are in red. Comparing visualization results for image (a) and image (d) it is noticeable that there are more false positives predictions (red boxes without green boxes) and false negatives predictions (green boxes without red boxes) using the baseline approach. The same is true for the remaining images in columns 2 and 3. Even though, the proposed method performs better in detecting P. falciparum with high sensitivity in high-resolution input images it still has a lot of false positives with low precision due to dark distractors which have a very similar shape and color with the P. falciparum parasite. In the future, the research will be continued to reduce the number of false positives by applying hard negative mining technique and by adding a classifier head in front of the detection network.

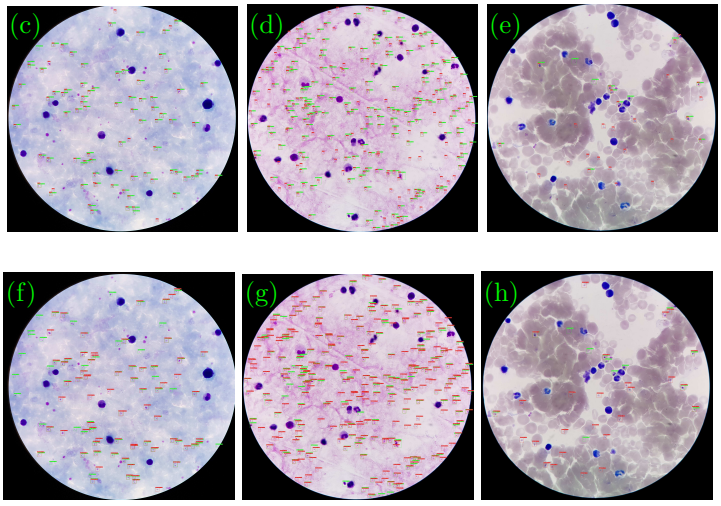

Figure 3: Sample visualization results of best-performing model (YOLOV4-tiny). The top row shows detection results for three test images using the proposed method and the bottom row shows detection results using the baseline approach. Ground truth bounding boxes are in green and predicted boxes are in red. The figure also shows how the images vary in color and infection rate.

6 Discussion

The experimental results on the external datasets indicate that the proposed model struggles to generalize in a new domain dataset. This is because the model development dataset was collected from one domain. The experimental results obtained by fine-tuning the selected model using external datasets indicate a significant increase in detection accuracy with improved generalization ability of the model. Therefore, this study shows that training SOTA deep learning models by using datasets collected from different health centers and geographic locations by considering the real-time clinical procedures during manual microscopy diagnosis enabled improvement of model generalization and detection accuracy.

The Performance of the selected model utilizing the proposed tile-based approach was also compared with existing work [7] which uses a similar dataset with this work.

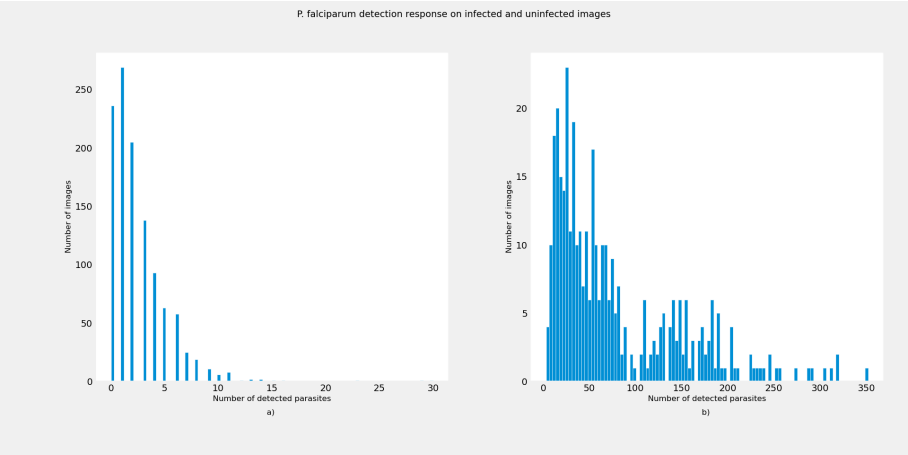

Figure 4: Detection response of the proposed method for a) uninfected images and b) infected images. From the histograms, it is evidenced that the proposed method is very effective in identifying P. falciparum on infected images and it is also effective on uninfected images with very few false positives due to distractors.

In their work, parasite locations were pre-segmented by using an intensity-based threshold technique and utilize a custom CNN classifier on the segmented candidate suspicious regions for final detection of P. falciparum. Compared with their work the proposed method surpasses their performance by 7% in terms of precision and 12% in terms of recall. The results obtained in terms of detection performance were also compared for the two external datasets obtained from a different domain. For the external dataset obtained from [28] they obtained a precision of 67% and recall of 80% which is very less compared to the proposed model's detection performance; 73.1% average precision and 96.3% of recall. For the other external dataset which is obtained from [10] in their proposed model, they achieved high precision of 97% but the recall is very bad with 22% whereas by using the proposed method an average precision of 83.4% and a recall of 94.7% was achieved.

The extensive validation experiments for the proposed method using datasets collected from a different domain demonstrate that the proposed method can be used effectively for malaria parasite screening by recommending exact and suspicious P. falciparum parasite locations and this will greatly reduce workload of laboratory technicians in malaria-endemic remote areas where there is critical skill gap and shortage of experts. Surprisingly, this study demonstrated that a significant performance improvement is achieved by utilizing the proposed method on lightweight SOTA YOLOV4 based object detection models compared to large size complex models. Thus, the proposed tile-based approach can even be deployed on low end-devices such as smartphones which can be integrated with the microscope without the need for high computing resources and memory.

To demonstrate the effectiveness of the proposed approach for P. falciparum detection analysis was done for detection responses of uninfected and infected images. A total of 1141 uninfected images collected from 50 uninfected patients and infected images from the model development test dataset were used. Due to artifacts,

which are suspicious due to their similarity to P. falciparum parasites and occurred because of staining and imaging procedures, the proposed model generates one or two false positive detections in negative images. As shown in Figure 4 the best performing model shows an effective detection response for uninfected images.

7 Conclusion

Recent development of microscopy techniques accompanied by improvements in computer vision technologies hold enormous potential to aid medical diagnosis in developing countries where there is a critical shortage of resources. The challenges in manual microscopy for malaria parasite screening motivates researchers to explore an alternative computer vision-based diagnostic approach. Although existing SOTA deep learning-based models for natural object detection show promising results, their application for a specific domain such as malaria parasite screening has its own challenges. Existing SOTA object detection models perform poorly on small object detection tasks such as malaria parasite detection. The other challenge is related to high-resolution microscopic images benefited from the advancement in digital data acquisition technologies such as high-resolution cameras. Directly applying high-resolution images by downscaling to SOTA detection network's low resolution input degrades small object detection performance due to the detail information loss.

In this study, performance of SOTA deep learning-based object detection models were increased for P. falciparum detection in high-resolution thick smear microscopic images. To achieve this, effectiveness of tile-based image processing which relies on training deep learning-based object detection models using tiles generated from an input high-resolution image was systematically evaluated. Besides, the proposed models were evaluated using datasets obtained from a different domain to validate their generalization ability and detection accuracy. Based on the extensive experimental analysis, lightweight YOLOV4 based models achieved a significant performance improvement using the proposed tile-based approach with 38% performance boost concerning their baseline method, while requiring only minimal additional computation cost. In addition, the proposed method outperforms detection results obtained by previous research works using similar datasets. The proposed malaria parasite screening technique has the potential to reduce workload of laboratory technicians by providing exact parasite locations or suspicious regions so that it will support the doctors to make their final decision. In the future work, we will focus on reducing the number of false-positive to improve precision of the proposed models by applying hard negative mining techniques and adding a classifier head that will be used to filter the detected objects.

Abbreviations

AP: Average Precision; CAD: Computer-Aided Diagnostic; CNN: Convolutional Neural Network; FP: False Positive; FN: False Negative; GPU: Graphics Processing Unit; HR: High Resolution; IOU: Intersection Over Union; NMS: Non-maximum suppression; PCR: Polymerise Chain Reaction; PRC: Precision-Recall Curve; RBCs: Red Blood Cells; RDT: Rapid Diagnostic Test; RGB: Red, Green and Blue; ROI: Region of Interest; SOTA: State of the art; SSD: Single Shot multibox Detector; TP: True Positive; WBC: White Blood Cell; WHO: World Health Organization; YOLOV4: You Only Look Once Version 4

Declarations

Ethics approval and consent to participate

No informed consent or ethical approval is required for the purpose of this study.

Consent for publication
Not applicable.

Availability of data and materials
The model development dataset used in this study which was released by research work of [7] is available on this link https://data.lhncbc.nlm.nih.gov/public/Malaria/Thick_Smears_150/index.html (last accessed: 03/10/2022). The first external dataset used in this study released by research work of [10] is available here http://air.ug/downloads/plasmodium-phonecamera.zip (last accessed: 03/10/2022). The second external dataset used in this study released by research work of [28] is available on this link https://drive.google.com/drive/folders/1p45Dt-BJy8hhoI-rYnhcaL6IMl5FsFL-?usp=sharing (last accessed: 03/10/2022). The third external dataset for uninfected images released by research work of [19] is available on this link https://data.lhncbc.nlm.nih.gov/public/Malaria/NIH-NLM-ThickBloodSmearsU/NIH-NLM-ThickBloodSmearsU.zip (last accessed: 03/10/2022).

Competing interests
The authors declare that they have no competing interests.

Funding
Not applicable.

Authors' contributions
FA: Conceptualization of the study, study methodology, experimental design and analysis, wrote and revised the manuscript. **KA**: Contributed to study methodology, writing result and discussion, revision of the manuscript.

Acknowledgements
Not applicable.

Author details
[1]Faculty of Electrical and Computer Engineering, Jimma University, 378 Jimma, Ethiopia. [2]Faculty of Electrical and Computer Engineering, Jimma University, , 378 Jimma, Ethiopia.

References

1. WHO. World malaria report 2021. Geneva: World Health Organization; 2021. Licence: CC BY-NC-SA 3.0 IGO:. https://www.mmv.org/newsroom/publications/world-malaria-report-2021?gclid=Cj0KCQiAraSPBhDuARIsAM3Js4r35mcOuyexpOjmcQ1Sl_6rPon5hDZfKsJPQgkqKEm9vE7kDVhFTVQaAtiREALw_wcB, 2021. [Online; accessed 21-January-2022].
2. Samuel Shillcutt, Chantal Morel, Catherine Goodman, Paul Coleman, David Bell, Christopher Whitty, and Anne Mills. Cost-effectiveness of malaria diagnostic methods in sub-saharan africa in an era of combination therapy. *Bulletin of the World Health Organization*, 86:101–10, 03 2008. .
3. Mahdieh Poostchi, Kamolrat Silamut, Richard Maude, Stefan Jaeger, and George Thoma. Image analysis and machine learning for detecting malaria. *Translational Research*, 194, 01 2018. .
4. Karan Makhija, Samuel Maloney, and Robert E Norton. The utility of serial blood film testing for the diagnosis of malaria. *Pathology*, 47:68–70, 2015.
5. Kassahun Desta, G. Mengistu, and Beteley Gelaw. The reliability of blood film examination for malaria at the peripheral health unit. *Ethiop J Health Dev*, 17, 01 2004.
6. Fatima A. Merchant and Kenneth R. Castleman. Chapter 27 - computer-assisted microscopy. In Al Bovik, editor, *The Essential Guide to Image Processing*, pages 777–831. Academic Press, Boston, 2009. ISBN 978-0-12-374457-9. . URL https://www.sciencedirect.com/science/article/pii/B9780123744579000275.
7. Feng Yang, Mahdieh Poostchi, Hang Yu, Zhou Zhou, Kamolrat Silamut, Jian Yu, Richard James Maude, Stefan Jaeger, and S. Antani. Deep learning for smartphone-based malaria parasite detection in thick blood smears. *IEEE Journal of Biomedical and Health Informatics*, 24:1427–1438, 2020.
8. Courosh Mehanian, Mayoore Jaiswal, Charles Delahunt, Clay Thompson, Matt Horning, Liming Hu, Travis Ostbye, Shawn McGuire, Martha Mehanian, Cary Champlin, et al. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 116–125, 2017.
9. Kaushik Chakrabortya, Arnab Chattopadhyayb, A. Chakrabarti, Tinku Acharyad, and Anjan Kr Dasguptae. A combined algorithm for malaria detection from thick smear blood slides. *Journal of Health and Medical Informatics*, 6:1–6, 2015.
10. John Quinn, Alfred Andama, Ian Munabi, and Fred Kiwanuka. *Automated Blood Smear Analysis for Mobile Malaria Diagnosis*, page Pages 115–132. 09 2014. ISBN Print ISBN: 978-1-4665-8929-2.
11. Luís Rosado, José Manuel Correia da Costa, Dirk Elias, and Jaime Cardoso. Automated detection of malaria parasites on thick blood smears via mobile devices. *Procedia Computer Science*, 90:138–144, 12 2016. .
12. Han Sang Park, Matthew T. Rinehart, Katelyn A. Walzer, Jen-Tsan Ashley Chi, and Adam Wax. Automated detection of p. falciparum using machine learning algorithms with quantitative phase images of unstained cells. *PLOS ONE*, 11(9):1–19, 09 2016. . URL https://doi.org/10.1371/journal.pone.0163045.
13. F. Boray Tek, Andrew Graham Dempster, and Izzet Kale. Parasite detection and identification for automated thin blood film malaria diagnosis. *Computer Vision and Image Understanding*, 114(1):21 – 32, 2010. ISSN 1077-3142. . URL http://www.sciencedirect.com/science/article/pii/S1077314209001234.
14. Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo. Deep learning for undersampled MRI reconstruction. *Physics in Medicine & Biology*, 63(13):135007, jun 2018. . URL https://doi.org/10.1088/1361-6560/aac71a.

15. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. volume 9351, pages 234–241, 10 2015. ISBN 978-3-319-24573-7. .
16. Samreen Naeem, Aqib Ali, Salman Qadri, Wali Khan Mashwani, Nasser Tairan, Habib Shah, Muhammad Fayaz, Farrukh Jamal, Christophe Chesneau, and Sania Anam. Machine-learning based hybrid-feature analysis for liver cancer classification using fused (mr and ct) images. *Applied Sciences*, 10(9), 2020. ISSN 2076-3417. . URL https://www.mdpi.com/2076-3417/10/9/3134.
17. Ghulam Murtaza, Liyana Shuib, Ainuddin Wahid, Ghulam Mujtaba, Henry Nweke, Mohammed Al-Garadi, Fariha Zulfiqar, Ghulam Raza, and Nor Azmi. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. *Artificial Intelligence Review*, 53, 03 2020. .
18. Arthur Marka, Joi Carter, Ermal Toto, and Saeed Hassanpour. Automated detection of nonmelanoma skin cancer using digital images: a systematic review. *BMC Medical Imaging*, 19, 2019.
19. Yasmin Kassim, Feng Yang, Hang Yu, Richard Maude, and Stefan Jaeger. Diagnosing malaria patients with plasmodium falciparum and vivax using deep learning for thick smear images. *Diagnostics*, 11:1994, 10 2021. .
20. Yasmin M. Kassim, Kannappan Palaniappan, Feng Yang, Mahdieh Poostchi, Nila Palaniappan, Richard James Maude, S. Antani, and Stefan Jaeger. Clustering-based dual deep learning architecture for detecting red blood cells in malaria diagnostic smears. *IEEE Journal of Biomedical and Health Informatics*, 25:1735–1746, 2021.
21. Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael S. Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *CoRR*, abs/1409.0575, 2014. URL http://arxiv.org/abs/1409.0575.
22. Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International Journal of Computer Vision*, 88:303–308, September 2009. URL https://www.microsoft.com/en-us/research/publication/the-pascal-visual-object-classes-voc-challenge/. Printed version publication date: June 2010.
23. Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312, 2014. URL http://arxiv.org/abs/1405.0312.
24. Kang Tong, Yiquan Wu, and Fei Zhou. Recent advances in small object detection based on deep learning: A review. *Image and Vision Computing*, 97:103910, 03 2020. .
25. Yanwei Pang, Jiale Cao, Yazhao Li, Jin Xie, Hanqing Sun, and Jinfeng Gong. TJU-DHD: A diverse high-resolution dataset for object detection. *CoRR*, abs/2011.09170, 2020. URL https://arxiv.org/abs/2011.09170.
26. Hao Zhang, Chunyu Fang, Xinlin Xie, Yicong Yang, Wei Mei, Di Jin, and Peng Fei. High-throughput, high-resolution deep learning microscopy based on registration-free generative adversarial network. *Biomedical Optics Express*, 10:1044, 03 2019. .
27. Mingfei Gao, Ruichi Yu, Ang Li, Vlad I. Morariu, and Larry S. Davis. Dynamic zoom-in network for fast object detection in large images. *CoRR*, abs/1711.05187, 2017. URL http://arxiv.org/abs/1711.05187.
28. Rose Nakasi, Ernest Mwebaze, and Aminah Zawedde. Mobile-aware deep learning algorithms for malaria parasites and white blood cells localization in thick blood smears. *Algorithms*, 14:17, 01 2021. .
29. Zhong Chen, Ting Zhang, and Chao Ouyang. End-to-end airplane detection using transfer learning in remote sensing images. *Remote Sensing*, 10(1), 2018. ISSN 2072-4292. . URL https://www.mdpi.com/2072-4292/10/1/139.
30. Fan Yang, Heng Fan, Peng Chu, Erik Blasch, and Haibin Ling. Clustered object detection in aerial images. *CoRR*, abs/1904.08008, 2019. URL http://arxiv.org/abs/1904.08008.
31. Salam Shuleenda Devi, Ngangbam Herojit Singh, and Rabul Hussain Laskar. Performance analysis of various feature sets for malaria-infected erythrocyte detection. In Kedar Nath Das, Jagdish Chand Bansal, Kusum Deep, Atulya K. Nagar, Ponnambalam Pathipooranam, and Rani Chinnappa Naidu, editors, *Soft Computing for Problem Solving*, pages 275–283, Singapore, 2019. Springer Singapore. ISBN 978-981-15-0184-5.
32. Andrea Loddo, Cecilia Di Ruberto, and Michel Kocher. Recent advances of malaria parasites detection systems based on mathematical morphology. *Sensors (Basel, Switzerland)*, 18, 02 2018. .
33. Amin Siddiq Sumi, Hanung Adi Nugroho, and Rudy Hartanto. A systematic review on automatic detection of plasmodium parasite. *International Journal of Engineering and Technology Innovation*, 11:103–121, 2021.
34. F. Tek, Andrew Dempster, and Izzet Kale. Computer vision for microscopy diagnosis of malaria. *Malaria journal*, 8:153, 02 2009. .
35. Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017. ISSN 1361-8415. . URL https://www.sciencedirect.com/science/article/pii/S1361841517301135.
36. Yann LeCun, Y. Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521:436–44, 05 2015. .
37. Sivaramakrishnan Rajaraman, Stefan Jaeger, and S. Antani. Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images. *PeerJ*, 7, 2019.
38. Samson Chibuta and Aybar Acar. Real-time malaria parasite screening in thick blood smears for low-resource setting. *Journal of Digital Imaging*, 33, 01 2020. .
39. Fetulhak Abdurahman, Kinde Anlay Fante, and Mohammed Aliy. Malaria parasite detection in thick blood smear microscopic images using modified yolov3 and yolov4 models. *BMC Bioinformatics*, 22, 2021.
40. Muhammad Umer, Saima Sadiq, Muhammad Ahmad, Saleem Ullah, Gyu Sang Choi, and Arif Mehmood. A novel stacked cnn for malarial parasite detection in thin blood smear images. *IEEE Access*, 8:93782–93792, 2020. .
41. Feng Yang, Nicolas Quizon, Hang Yu, Kamolrat Silamut, Richard James Maude, Stefan Jaeger, and S. Antani. Cascading yolo: automated malaria parasite detection for plasmodium vivax in thin blood smears. In *Medical Imaging*, 2020.
42. Aimon Rahman, Hasib Zunair, Tamanna Rahman Reme, M. Sohel Rahman, and Mahdy Rahman Chowdhury Mahdy. A comparative analysis of deep learning architectures on high variation malaria parasite classification

dataset. *Tissue & cell*, 69:101473, 2020.
43. De Loh, Yong Xin, Jullian Yapeter, Karupppasamy Subburaj, and Rajesh Chandramohanadas. 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, 01 2021. .
44. Poonguzhali Elangovan and Malaya Nath. A novel shallow convnet-18 for malaria parasite detection in thin blood smear images: Cnn based malaria parasite detection. *SN Computer Science*, 2, 09 2021. .
45. Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.
46. Tsung-Yi Lin, Piotr Dollár, Ross B. Girshick, Kaiming He, Bharath Hariharan, and Serge J. Belongie. Feature pyramid networks for object detection. *CoRR*, abs/1612.03144, 2016. URL http://arxiv.org/abs/1612.03144.