Applied GIS

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

Deep Learning For Smartphone Based Malaria Parasite Detection In Thick Blood Smears

Sakruthi Jayasri, Nagireddy Mounika, Veerla Nithin, Rangaturam Srihind Reddy,
Dr. Vaka Murali Mohan(Principal and Professor),
Department of CSE,
MALLA REDDY INSTITUTE OF TECHNOLOGY AND SCIENCE, Telangana, Hyderabad.

Abstract—Objective: ***This work investigates the possibility of automated malaria parasite detection in thick blood smears with smartphones.Methods: Our deep learning algorithm is the first of its kind to identify malaria parasites in smartphone-optimized photos of thick blood smears. There are two stages to our procedure. To begin, we do a rapid screening of a thick smear picture using an intensity-based Iterative Global Minimum Screening (IGMS) to identify potential parasites. The next step is to use a tailored Convolutional Neural Network (CNN) to label each potential candidate as a background or parasite. In conjunction with this publication, we provide to the general public a dataset consisting of 1,819 thick smear photographs originating from 150 patients. In order to implement the deep learning strategy detailed in this article, we trained and tested it using this dataset.***

Using the following performance indicators: accuracy (93.46%±0.32%), AUC (98.39%±0.18%), sensitivity (92.59%±1.27%), specificity (94.33%±1.25%), precision (94.25%±1.13%), and negative predictive value (92.74%±1.09%), the customized CNN model was found to be effective in discriminating between positive (parasitic) and negative image patches in a patient-level five-fold cross-evaluation. Automatically discovered parasites and ground truth have very high correlation coefficients (>0.98) at both the picture and patient levels, proving that our technology is feasible.

Conclusion: A smartphone software that uses deep learning techniques to identify parasites in thick blood smears achieves promising results.

Importance: Automated parasite identification using cellphones offers a potential substitute for human parasite counting in the diagnosis of malaria, particularly in regions where trained parasitologists are few.

Glossary Terms– Advanced computer-assisted diagnostics; Deep learning; Malaria

I. INTRODUCTION

Malaria is a disease that may cause death all across the globe. Approximately 435 thousand people lost their lives to malaria in 2017, with 219 million cases recorded globally, as stated in the 2018 World Health Organization (WHO) malaria report [1]. The most reliable method for diagnosing malaria is to examine stained thick and thin blood smears under a microscope [2, 3]. Although it takes a lot of time and is inexpensive, microscopy inspection is readily accessible. The proficiency of the parasitologists doing the microscopy is also crucial to the accuracy of the results [4]. Parasitologists often operate in settings with limited resources and no formal structure to guarantee that their skills and diagnostic quality are maintained. This results in misdiagnoses and the wrong kind of therapy being administered [4]. For instance, if a malaria test comes back positive, people may take anti-malaria medication needlessly, which can cause side effects like nausea and abdominal pain. On the other hand, if a malaria test comes back negative, people may have to take antibiotics again, which can lead to unnecessary consultations and even the development of more severe malaria [5]. Therefore, it is an attractive research objective to build an automated method for malaria diagnosis in order to improve the treatment and management of individual patients. There are two major benefits to automated parasite detection: 1) it may lower diagnostic expenses and 2) it can provide a more accurate diagnosis, which is particularly useful in places with limited resources. In order to diagnose malaria and measure the severity of the condition, parasite numbers are crucial. In order to gauge medication efficacy and possible resistance, they are also useful for keeping tabs on patients. The purpose of this research is to examine the feasibility of using cellphones to automatically identify and

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

count malaria parasites in digital photographs of thick blood smears.

In order to identify if a blood sample contains malaria parasites, a thick blood smear is prepared. Its sensitivity is about eleven times more than that of a thin blood smear, allowing for more effective parasite identification [5]. Differentiating between distinct parasite species and phases of development is a common usage for the thin blood smear, which is made by spreading a drop of blood onto a glass slide. Figure 1 shows that various processing techniques are needed for parasite identification on thick and thin blood smears. Both red blood cells (RBCs) and white blood cells (WBCs) are easily discernible in thin blood smears. Segmenting red blood cells (RBCs) and then classifying each segmented RBC as infected or uninfected is a common procedure for automated parasite identification in thin smears [5]-[7]. Fig. 1(a) shows that in thick blood smears, the only visible components are white blood cells (WBCs) and red blood cell nuclei. Consequently, it is necessary to directly identify parasites. A common procedure involves preselecting potential parasites and then deciding whether they are real or just background noise. False parasite identification may result from artifacts created when white blood cell nuclei and other non-parasite components absorb dye.

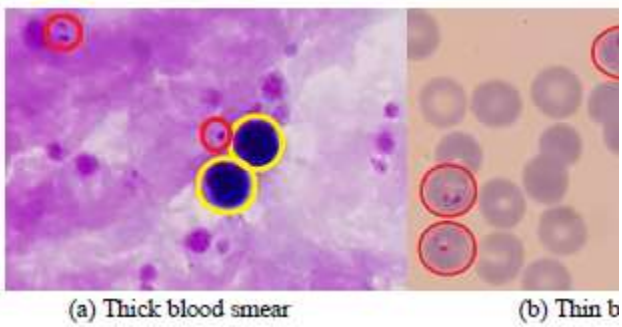


Fig. 1. Examples of thick and thin blood smears. Red circles are parasites and yellow circles are white blood cells.

A. Related work

Automated parasite identification has been the focus of several recent proposals for image processing and analysis on thin and thick blood smears. See also[5,8], [9] for reviews of the relevant literature. Here we provide a quick rundown of the methods that have been developed to detect malaria in thick blood smears.

Commonly used in traditional parasite detection methods are segmentation-based approaches that include thresholding and morphological procedures [10]-[13]. In order to separate parasites from white blood cells (WBCs) and potential parasites from the background, Kaewkamnerd et al. [10] suggest a technique that uses an adaptive threshold on the V-value histogram of the HSV picture. After testing it on 20 photos, the suggested approach reaches a 60% accuracy rate. To segment malaria parasites, Hanif et al. [11] first employ an intensity-stretching approach to improve the contrast of 255 thick blood smears, and then they use an empirical threshold. The authors provide qualitative findings on many photos, showcasing the use of various empirical criteria to get satisfactory segmentation outcomes. Parasites in thick blood smears may be identified using a combination of morphological segmentation and color information, as shown by Chakrabortya et al. [12]. After conducting experiments on 75 photos, the patch level assessment revealed a 95% success rate in detecting and a 10% false positive ratio. In order to identify red blood cells (RBCs) infected with malaria parasites in thin and thick blood smears, Dave et al. [13] use morphological procedures and adaptive thresholding based on histograms to denoised pictures. The approach discovers 533 parasites, compared to 484 parasites marked as ground truth, according to patch level assessment on 87 photos. While conventional methods for parasite identification are quick and easy to use, they struggle when faced with massive datasets. The reason for this is because conventional methods are very susceptible to changes in the images and their parameters are often found via empirical means. Evaluation of performance at the patch level on short datasets (ranging from twenty to three hundred pictures) might vary substantially when applied to large datasets, image level, or patient level evaluations.

Machine learning-based feature extraction and categorization are feature-based methodologies. From [14] to [18]. In order to identify parasites, Elter et al. [14] used a Support Vector Machine (SVM) classifier to extract 174 characteristics from candidates for plasmodia that had already been found. For 256 photos at the patch level, the authors state that the sensitivity is 97%. In their study, Purnama et al. [15] used Genetic Programming to determine the kind and stage of parasites by extracting characteristics from RGB channel histograms, H channel from HSV space, and HIS channel features. On average, their classification algorithm is able to identify parasites with 95.58% accuracy and non-parasites with 95.49%

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

accuracy when tested on 180 patches. The pre-segmented picture is processed by Yunda et al. [16] to extract color features, co-occurrence texture features, and wavelet-based texture features. Redundant data are reduced using Principal Component Analysis (PCA), and the final classification is performed using a neural network model. A sensitivity of 76.45% for parasite identification was determined by evaluation of 110 pictures. According to Quinn et al. [17], a randomized tree classifier should be used for classification after 475 randomly overlapping patches are extracted from each picture using linked component and moment features. The patches are created using downsampling and sliding window screening. Using 2903 pictures from 133 patients, the approach achieves a 90% accuracy rate with a 20% recall rate at the patch level. For parasite diagnosis, Rosado et al. [18] use an adaptive thresholding method. For WBC and parasite identification, they combine an RBF kernel based SVM classifier with geometry, color, and texture data. Their automated parasite prediction reached 91.8% accuracy, with sensitivity at 80.5% and specificity at 93.5%, when tested on 94 pictures from 6 patients; their white blood cell identification reached 98.2% sensitivity and 72.1% specificity. The effectiveness of the feature-based methods is assessed at the patch level. In other words, a single patch picture is used as the input sample and the assessment is usually the accuracy of patch categorization. The ultimate aim of malaria patient diagnosis, however, is to identify and categorize every patch—parasites and false positives alike. Having strong performance at the patch level is no guarantee of good performance at the picture level or patient level.

Due to its exceptional performance on massive data, deep learning has recently been the fad in machine learning. Many fields outside of medicine have already seen improved performance as a result. Deep learning has been making waves in computer-assisted diagnosis systems as of late. There were primarily two causes for this change: First, in contrast to more conventional approaches and feature-

TABLE I. EXISTING APPROACHES APPLIED TO PARASITE DETECTION IN THICK BLOOD SMEARS

Authors	Methods	Images	Patients	SE/SP(%)	Remarks
Kaewkamnerd et al., 2011	Adaptive threshold on the V-value histogram + size filtering	20	-	60% DR	
Hanif et al., 2011	Intensity-based contrast enhancement + threshold-based segmentation	255	-	Qualitative results	Threshold is empirical
Chakrabortya et al., 2015	Color information based morphological segmentation	75	-	95% DR + 10% FPR	Patch level evaluation
Dave et al., 2017	Histogram-based adaptive thresholding + morphological operations on denoised images	87	-	91% DR	Patch level evaluation
Elter et al., 2011	Feature extraction from pre-detected plasmodia candidates + SVM classifier	256	-	97% SE	Patch level evaluation
Purnama et al., 2013	Feature extraction from RGB histogram + Genetic Programming classification	-	-	96% ACC	Only 180 patch images were used.
Yunda et al., 2011	Feature extraction from color, co-occurrence matrix related texture and wavelet-based texture + PCA feature reduction + neural network classification	110	-	76% DR	Patch level evaluation
Quinn et al., 2014	Feature extraction from connected components and moment features + randomized tree classifier	2903	133	20% SE + 90% PR	Patch level evaluation
Rosado et al., 2016	Adaptive thresholding + feature extraction in color and texture + RBF kernel based SVM classifier	94	6	80.5% SE + 93.8% SP	Patch level evaluation
Delahunt et al., 2015	Parasite candidate localization + segmentation + feature extraction with CNN + linear SVM classification	-	143	LoD = 300 ~3000 p/μL + 92% SP	Patient level evaluation
Quinn et al., 2016	Small block splitting + CNN classifier	1182	-	97% PR	Patch level evaluation
Mehanian et al., 2017	Dynamic local thresholding for parasite candidate selection + SVM classifier + CNN classifier	1452	195	92% SE + 94% SP LoD =100 p/μL	Patch level and patient level evaluation
Torres et al., 2018	Same method as Mehanian et al., 2017		700	72% SE + 85% SP; 52% SE + 70% SP	Patient level evaluation

Note: DR indicates detection rate; FPR represents false positive rate; PR denotes precision; ACC means accuracy; SE indicates sensitivity; SP is specificity.

Unlike traditional methods that rely on segmentation and manually created features, deep learning provides a comprehensive solution and 2) can find data-derived hierarchical feature representations [23]-[25]. Automated feature extraction and parasite identification have been included into thick blood smears using deep learning algorithms throughout the last three years. After locating potential parasites, Delahunt et al. [19] suggest using a CNN in conjunction with a linear support vector machine (SVM) for classification. According to their reported results on 143 patients, the method used traditional features such as morphological, color, texture, and Harr-like features to predict a Limit of Detection

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

(LoD) of approximately 300 parasites/μL with a specificity of 92% at the patient level. However, they advocate for CNNs to be used for feature extraction. By using downsampled RGB pictures to train a four-layer convolutional neural network (CNN) model on overlapping patches, Quinn et al. [20] suggest a parasite identification method. On 1182 photos, they found an average accuracy of 97%. Nonetheless, rather of dividing the datasets according to patients, the scientists did it according to images. What this means is that they may use the same patient photos in both their training and testing datasets. Patch level is the only metric used to assess performance. In their study, Mehanian et al. [21] used a CNN model for feature extraction and classification after first detecting WBCs using a Gaussian-kernel SVM on thresholded candidates. They then trained a second Gaussian-kernel SVM on parasite candidates that were created using a dynamic local thresholding approach. According to the scientists, their approach is able to detect 94.1% of patches with a sensitivity of 91.6% and an accuracy of 89.7%, and 90.0% of patients with a LoD of around 100 parasites/μL and a specificity of 90.0%, using 1452 pictures from 195 patients. Having a parasite detection runtime of around 20 minutes (on a quad-core CPU) does not guarantee speed superior to human processing. Using regular microscopy and Polymerase Chain Reaction (PCR) as reference standards, Torres et al. [22] test Autoscope, a prototype digital microscope instrument with automated procedures suggested in [21]), at two peripheral health institutions. Assuming the slides have sufficient blood volume, the authors find that Autoscope's performance (sensitivity of 72% at a specificity of 85%) is comparable to standard microscopy (sensitivity of 68% at a specificity of 100%).

Table I provides a summary of the current methods used to identify parasites in thick blood smears. Low accuracy is a result of the use of tiny datasets in the conventional parasite identification methods described in [10]–[13]. The procedures have only been tested in qualitative or patch-level experiments, therefore we can't be sure that the outcomes will be the same when applied to patients. In comparison to more conventional methods, feature-based approaches [14]–[18] often improve the parasite identification rate. But all of the assessments are done at the patch level, and the majority of them [14–16], [18] employ datasets that include less than 256 pictures. There could be a significant performance hit if these methods are evaluated on a patient-level. In [17], a large dataset is used; nonetheless, the sensitivity that the authors attain is just 20%. Deep learning approaches have only been employed in four articles [19–22] to identify parasites in thick blood smears using large datasets; three of these papers [19, 21], [22] evaluated the system at the patient level. Thick smears captured by smartphone have only been the subject of three publications out of the aforementioned literature [17], [18], and [20]. With the goal of processing these pictures on a more robust platform for remote diagnosis, two of these articles, [17] and [20], rely only on the smartphone for data collection. The only study to far that has used a smartphone screening app is reference [18]. In this work, we set out to create a parasite detecting software for smartphones based on deep learning, which can provide a better performance compared to the traditional SVM, as implemented in [18].

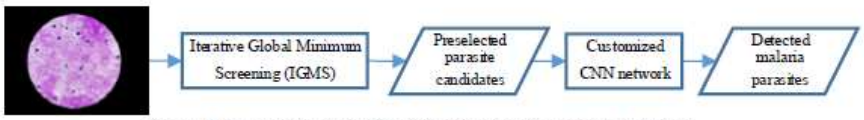


Fig. 2. Pipeline of the proposed system for automated parasite detection.

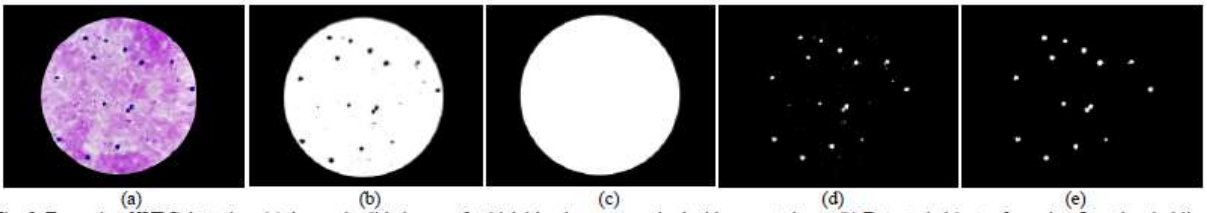


Fig. 3. Example of WBC detection. (a) A sample slide image of a thick blood smear acquired with a smartphone. (b) Detected objects after using Otsu thresholding. (c) Detected field of view ROI mask. (d) Detected WBCs including small areas of noise. (e) Detected WBCs after filtering noise in (e).

B. Contributions

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

Our contributions to thick blood smear processing are as follows, in comparison to previous work: We begin by creating a smartphone app that can automatically detect parasites in thick blood smears. This app will use our customized CNN model [23]-[28] to classify parasite candidates as either parasites or background, and our proposed intensity-based Iterative Global Minimum Screening (IGMS) method to quickly automatically preselect parasite candidates. As far as we are aware, this is the first effort to develop a deep learning-based system for parasite identification in cellphones for thick blood smears. We also have a quick method. On a standard Android smartphone, detecting parasites in a 3024×4032 photograph takes around 10 seconds. Third, we openly share a much bigger picture collection that we use to evaluate our technique. This set includes 1819 thick smear images and 84,961 tagged parasites, and it was gathered from 150 patients.

Here is how the remainder of the paper is structured: In Section II, we lay forth our plan for an automated system to identify parasites. In Section III, we provide the dataset, methodologies, and outcomes of the experiment. Our findings are detailed in Section IV, and Section V provides a summary and last thoughts.

II. METHODS

Splitting our problem into a screening and classification step allows faster processing because we only need to predict on a relatively small number of pixel patches, which reduces the overall processing cost. We illustrate the pipeline of our method in Fig. 2.

A. Parasite Candidate Screening

The screening stage reduces the size of the initial search space and preselects a subset of most likely parasite candidates. Parasite candidates are selected according to the lowest intensities in grayscale based on a histogram analysis, exploiting that the nuclei of parasites and WBCs have darker intensities than the background (Fig. 3(a)). To eliminate WBC

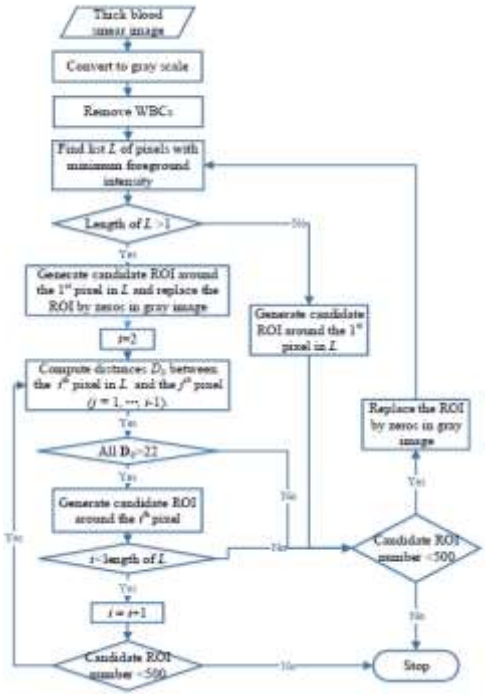


Fig. 4. Flowchart of parasite candidate screening (IGMS method). *i* is the iterator through *L*.

distraction, we filter out WBCs before performing the parasite candidate screening. Therefore, our intensity-based screening method for parasite candidate preselection consists of WBC detection and parasite candidate generation. The WBC detection first filters all WBCs present in the image. Then, the parasite candidate generation produces regions of interest by localizing the lowest intensities across a thick blood smear image.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

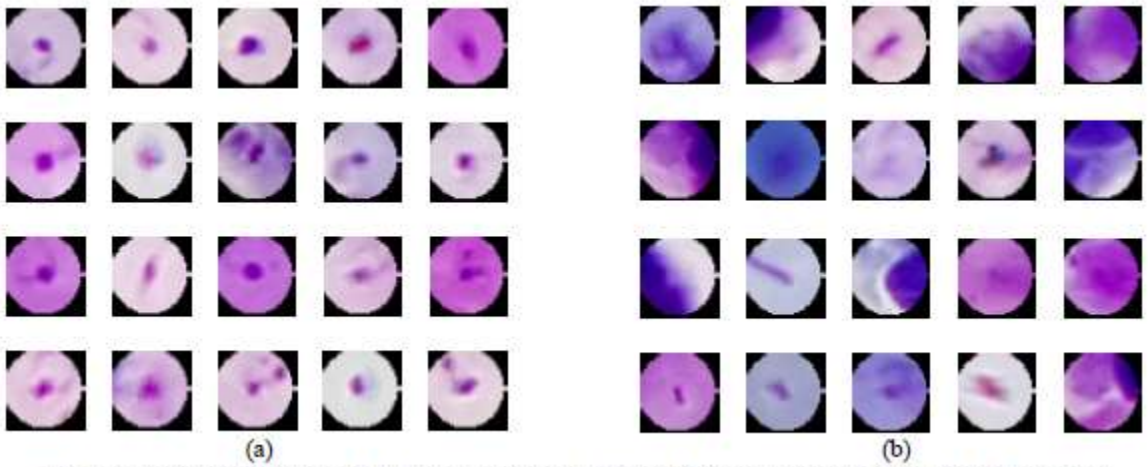


Fig. 5. Parasite candidates generated by the IGMS method. (a) 20 positive patches. (b) 20 negative patches.

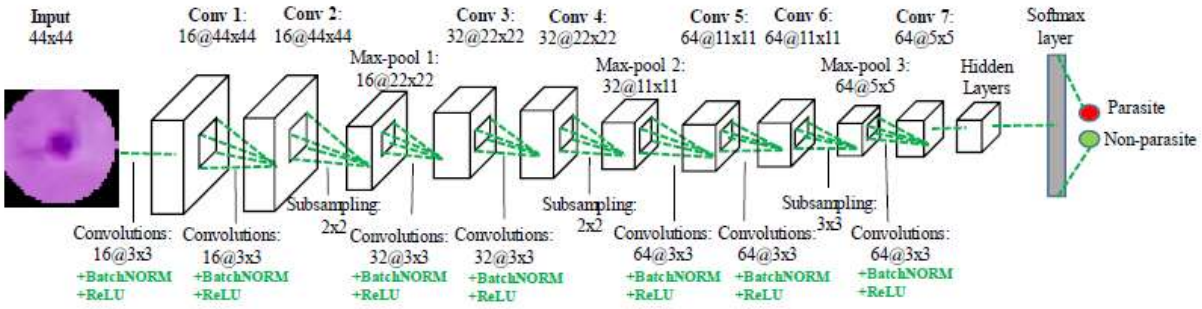


Fig. 6. Architecture of the customized CNN model for parasite classification. Conv, Max-pool and BatchNorm

denote convolution, max-pooling and batch normalization, respectively. The numbers above the cuboids indicate the dimensions of the feature maps. The numbers below the green dotted line represent the convolutional kernel sizes and the sizes of the max-pooling regions. The hidden layers include three fully connected layers and two dropout layers with a dropout ratio of 0.5. The output softmax layer computes the probabilities of the input image being either a parasite or background.

1) WBC detection

A sample smear image is shown in Fig. 3(a). We first convert the RGB image into a grayscale image. Then, we convert the grayscale image into a binary mask $M1$ using Otsu's method [30]. In this binary mask $M1$, the large ROI area corresponding to the field of view is shown as foreground (white) while WBCs are shown as background (dark); see Fig. 3(b). By filling the holes inside the large field of view ROI area, we obtain the field of view mask $M2$, shown in Fig. 3(c). WBCs can then be separated out by subtracting the binary mask $M1$ from the ROI mask $M2$ (see Fig. 3(d)). Clean WBCs are finally obtained by filtering small noisy areas. Fig. 3(e) demonstrates the result of this step. The pixels of WBCs are set to zeros for the following parasite detection steps.

2) Parasite preselection using *Iterative Global Minimum Screening (IGMS)*

IGMS generates RGB parasite candidates by localizing the minimum intensity values in a grayscale image. If only one pixel is localized, a circular region centered at this pixel location with a pre-defined radius of 22 pixels (average parasite radius) is cropped from the original RGB image and is selected as a parasite candidate (Fig. 5 (a)). If more than one pixel is localized, a new parasite candidate centered at the ith pixel is added when all the distances between the ith pixel and previously

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

selected pixels are larger than 22. Once a parasite candidate is selected, the intensity values inside this region of the grayscale image will be replace by zeros to guarantee the convergence of the IGMS method. The screening stage stops when the number of parasite candidates reaches a given number. In our experiments, we select 500 parasite candidates for each image to cover the true parasites as much as possible. Experiments on our dataset of 150 patients show that we can achieve a sensitivity above 97% on patch level, image level, and patient level when using this number. Each parasite candidate is a 44×44×3 RGB patch image, with pixels having a distance greater than 22 to the center set to zero. Fig. 4 shows the processing flowchart for IGMS and Fig. 5 shows examples of positive and negative patches extracted by IGMS.

B. Parasite Classification

Once the parasite candidates are extracted, we use a CNN model to classify them either as true parasites or background. In this work, we customize a CNN model consisting of seven convolutional layers, three max-pooling layers, three fully connected layers, and a softmax layer as shown in Fig. 6. A batch normalization layer is used after every convolution layer to allow a higher learning rate and to be less sensitive to the initialization parameters [30], followed by a rectified linear unit (ReLU) as the activation function [23]. Max-pooling layers are introduced after every two successive convolutional layers to select feature subsets. The last convolutional feature map is connected to three fully connected layers with 512, 50, and 2 hidden units, respectively. Between the three fully connected layers, two dropout layers [31] with a dropout ratio of 0.5 are applied to reduce model overfitting. The network is derived from VGG19 [27] by selecting the first six convolutional layers and three corresponding max-pooling layers from the VGG19 architecture to stop the feature maps at 64@5×5, followed directly by the fully connected and dropout layers. This shorter network structure provides similar performance while being faster and requiring less memory, which is important for smartphone applications. The output of the CNN model is a score vector, which gives the probabilities of the input image patch being either a parasite or background. We can obtain a larger or smaller number of predicted parasites by applying an adaptive probability threshold to the score vector.

Compared with pre-trained networks such as VGG [27], GoogLeNet [28], ResNet-50 [26], our customized CNN model has several advantages: 1) runtime is reduced by using a smaller set of customizable parameters, with the input size of the model being determined by the average parasite size in thick smear images (44×44×3), which is much smaller than the input size used by the other networks (224×224×3); 2) our smaller network structure with fewer layers is more suitable for smartphones. Since the input size is smaller, our network should in fact be less deep to avoid feature maps that are too small. A smaller network structure with less parameters also avoids over-training on the smaller input space. Compared to the pre-trained networks mentioned above, our customized CNN model achieves a better accuracy, despite having less network layers, and a shorter runtime. For an input image of 4032×3024×3 pixels, our system can complete the parasite detection within ten seconds (about eight seconds for candidate screening and two seconds for classification) on a standard Android smartphone. Both the smaller set of parameters and the smaller network structure contribute to the reduced runtime.

C. Smartphone-Based Application

Based on the IGMS method and customized CNN model for parasite detection, we develop a smartphone-supported automated system to diagnose malaria in thick blood smears. We implement this system as a smartphone app for the Android mobile operating system. When using this app, the camera of the smartphone is attached to the eyepiece of the microscope. The user adjusts the microscope to find the target field in the blood smear and takes pictures with the app. The app then detects and counts parasites, records parasite numbers in a patient database, and displays the results in the user interface. Users will take several images until they have collected enough data to meet the requirements of their local protocols. The app will aggregate the parasite counts across all images taken. We implemented all algorithms using the OpenCV4Android SDK library.

After the image acquisition and processing stage, the app will go through a series of input masks for the user to fill in the information associated with the current patient and smear. This information is saved in the local database of the app, which we build with the SQLite API provided by Android. The app offers a user interface to the database where the user can view the data and images of previous smears, allowing hospital staff to monitor the condition of patients. Fig. 7 shows a smartphone running our app

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

connected to a microscope (left-hand side) and a sample screenshot displaying a thick smear image with parasite counts (right-hand side).

III. DATA PREPARATION AND EXPERIMENTAL RESULTS

A. Dataset

We photographed Giemsa-stained thick blood smear slides from 150 *P. falciparum* infected patients at Chittagong Medical College Hospital, Bangladesh, using a smartphone camera for the different microscopic field of views. Fig. 7 shows the smartphone-microscope setup and a screenshot of the phone displaying a thick smear image. Images are captured with 100x magnification in RGB color space with a 3024×4032 pixel resolution. An expert slide reader manually annotated each image at the Mahidol-Oxford Tropical Medicine Research Unit (MORU), Bangkok, Thailand. We de-identified all images and their annotations, and archived them at the National Library of Medicine (IRB#12972). In this work, we use 1819 thick blood smear images from these 150 patients. We publish the data here:

Fig. 7. Smartphone-based malaria data acquisition and parasite detection.

B. Statistics of the Dataset

We first perform a statistical analysis on the whole dataset of 150 patients. There are in total 84,961 annotated parasites, whose radius varies from two to 96 pixels, with an average radius of 22 pixels (Fig. 8(a)). Each image includes one to 341 parasites, with an average number of 47 parasites (Fig. 8(b)). Each patient set contains three to 22 images with an average number of 12 images (Fig. 8(c)), and contains eight to 3,130 parasites with an average number of 522 parasites (Fig. 8(d)).

C. Data partitioning

We divide the data on patient level into two sets: Set A and Set B, by a ratio of 4:1. Our data division strategy is shown in Fig. 9. Set A includes 120 patients with 1443 images and 72,184 parasites, and is used for the CNN model training and evaluation. Set B includes 30 patients including 375 images and 12,777 parasites, and is used for the performance evaluation of our method for automated parasite detection using screening and classification. We further split Set A into training sets and test sets on patient level, and evaluate the CNN model performance based on five-fold cross evaluation. To achieve a better performance for the CNN model, we use a balanced training set with an equal number of positive and negative patches. For each image in Set B, we generate 500 patches using IGMS.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

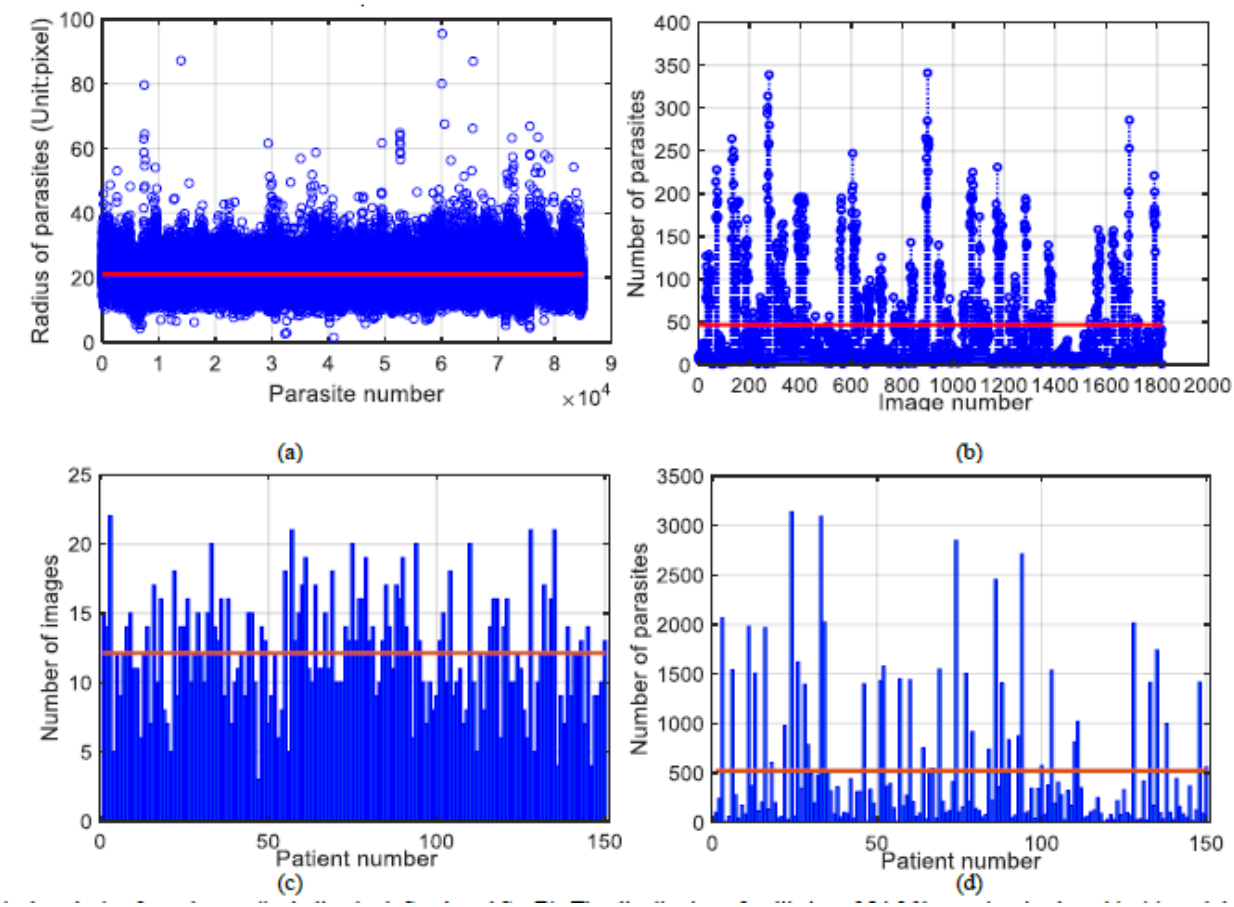


Fig. 8. Statistical analysis of our dataset (including both Set A and Set B). The distribution of radii size of 84,961 parasites is plotted in (a), and the distribution of the number of parasites in 1819 images is illustrated in (b). The number of images and parasites in each patient is illustrated in (c) and (d) respectively. The red lines in the four subfigures indicate the average parasite radius, the average number of parasites in each image, the average number of images for each patient, and the average number of parasites for each patient, respectively

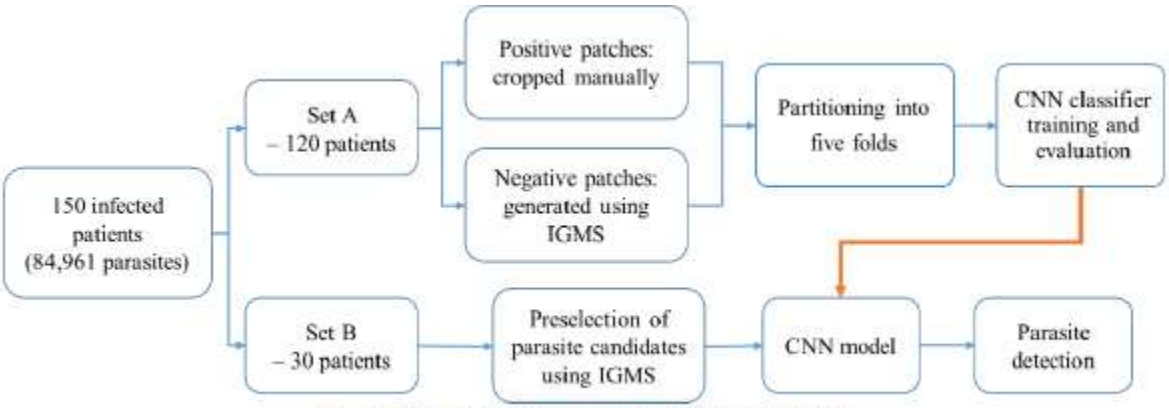


Fig. 9. Data division strategy used in the experiments.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

D. Preselection Performance

We evaluate the performance of IGMS as follows: We consider a parasite candidate generated by IGMS as a truly identified parasite if the overlap between it and the corresponding manually annotated parasite is larger than 50%. This overlap ratio is chosen empirically based on the balance of preselection sensitivity and classification accuracy. Then, we compute the sensitivity of IGMS as the ratio of the number of truly identified parasites to the total number of annotated parasites. Fig. 10 presents the sensitivity of IGMS on both image level and patient level for Set B. For parasite preselection, the proposed IGMS method achieves a sensitivity of 97.04% on patch level, 97.49%±5.40% on image level (Fig. 10(a)), and 96.59%±5.52% on patient level (Fig. 10(b)), respectively.

E. Performance of the Customized CNN model

We evaluate the performance of the customized CNN model on Set A using five-fold cross evaluation. Each fold contains 24 patients. Table II and Fig. 11 present the classification performance and receiver operating characteristic (ROC). According to Fig. 11, our customized CNN model achieves an average AUC score of 98.39%, and a standard deviation of 0.18%, showing its robustness and effectiveness. The average accuracy, F-score, specificity, sensitivity, precision, and negative predictive values for our customized CNN model are 93.46%, 93.40%, 94.33%, 92.59%, 94.25%, and 92.74%, respectively.

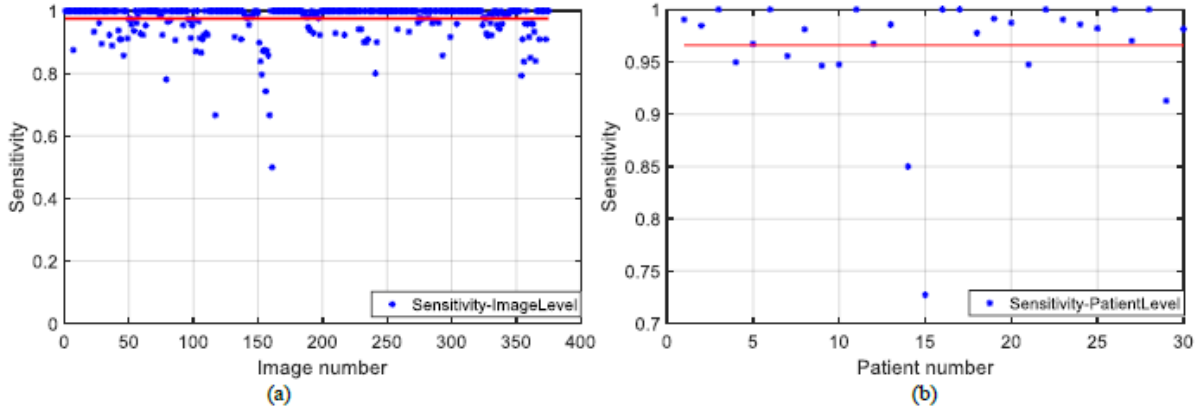


Fig. 10. IGMS sensitivity of parasite preselection for 120 patients in Set B on image level (a) and patient level (b).

TABLE II
CLASSIFICATION PERFORMANCE ON FIVE FOLDS FOR SET A.

Accuracy	F-score	Specificity	Sensitivity	Precision	Neg-pred
93.46% ±0.32%	93.40% ±0.33%	94.33% ±1.25%	92.59% ±1.27%	94.25% ±1.13%	92.74% ±1.09%

Note: Neg_pred is the negative predictive value.

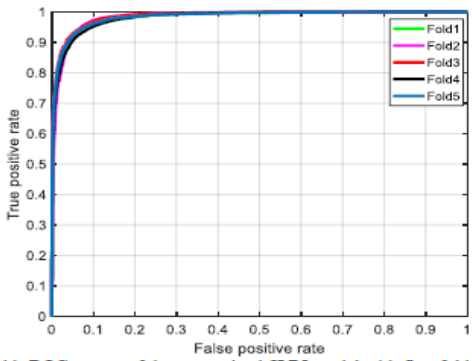


Fig. 11. ROC curves of the customized CNN model with five-fold cross evaluation for Set A on patch level (AUC=98.39%±0.18%).

TABLE III
CONFUSION MATRIX FOR SET B ON PATCH LEVEL.

	Predicted Positive	Predicted Negative
Parasites	5.8%	1.2%
Non-parasites	1.5%	91.5%

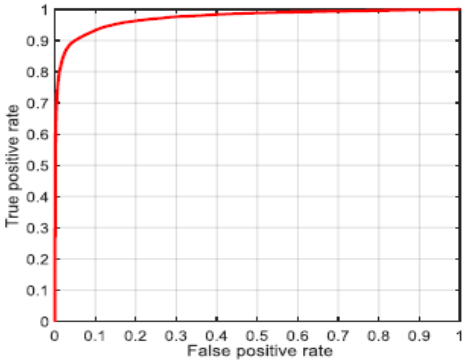
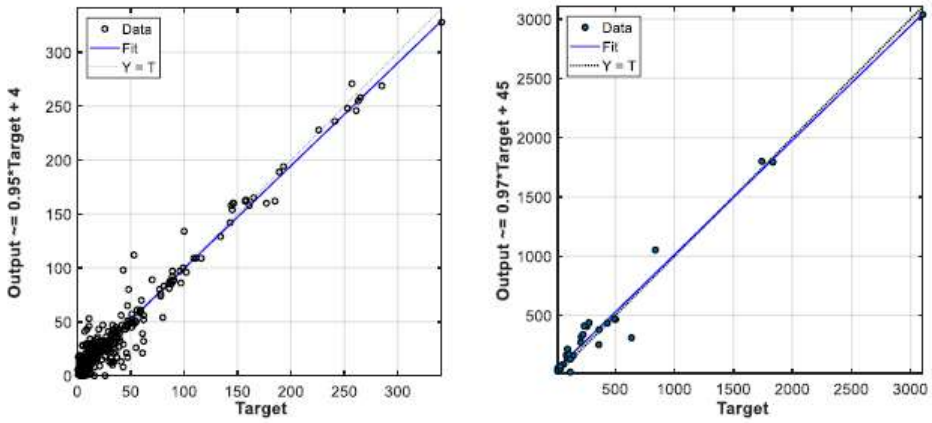


Fig. 12. ROC curve of the proposed method on automatic parasite detection for Set B on patch level (AUC=97.34%).

Applied GIS

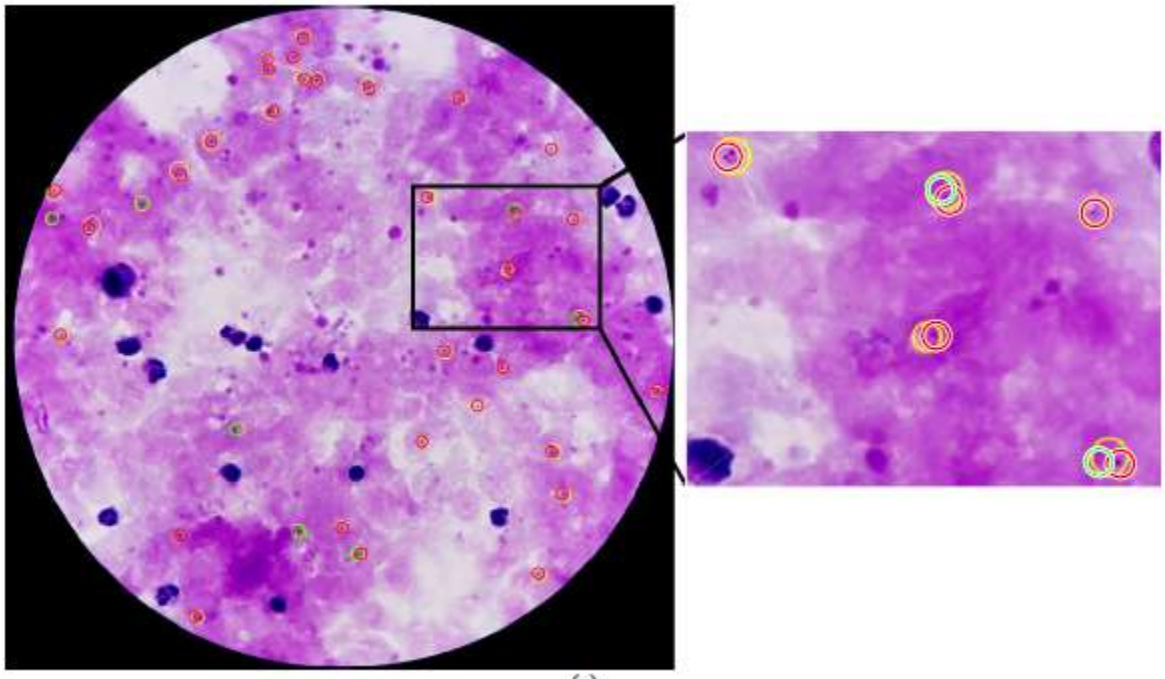
ISSN: 1832-5505

Vol-11 Issue-04 Nov 2023



(a) (b)

Fig. 13. Linear regression between the number of automatically detected parasites (y-axis) and the number of parasites in the ground truth (x-axis) on image level (R=0.98) (a) and patient level (R=0.99) (b).



(a)

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

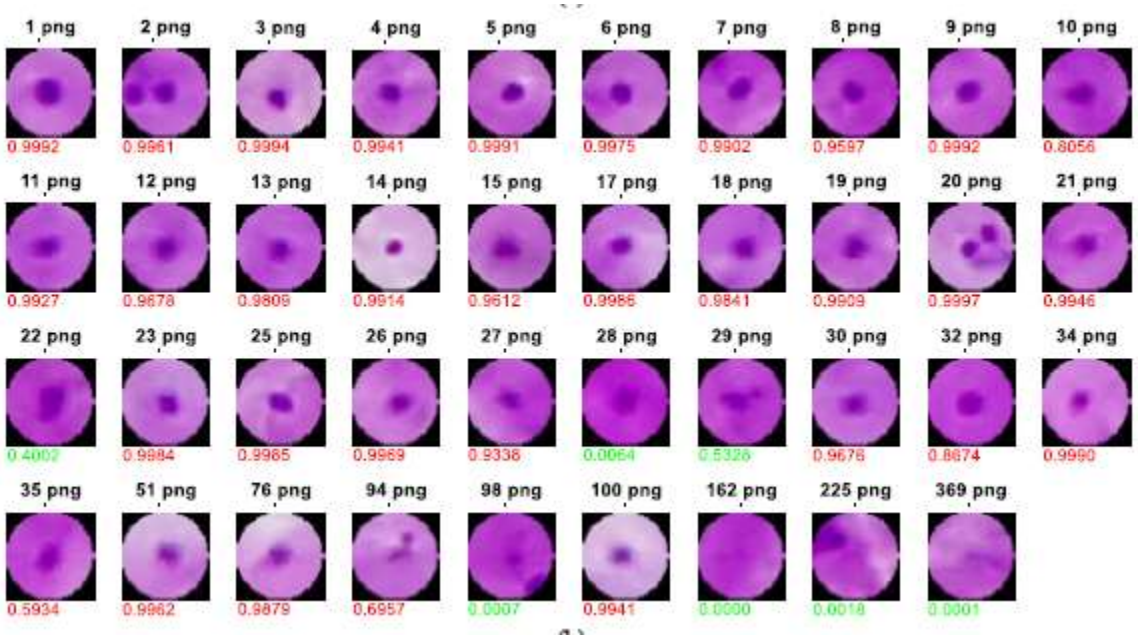


Fig. 14. Parasite detection on an example image using our proposed method. (a) Parasites annotated in the ground truth (yellow circles) and screened parasite candidates that overlap more than 50% with the parasites in the ground truth (red and green circles). Red circles indicate candidates that are finally predicted as parasites (true preselected parasites), and green circles indicates those that are predicted as non-parasites (false preselected parasites). (b) Probabilities of parasite candidates that overlap more than 50% with parasites in the ground truth. The number under each patch denotes the output probability of the CNN. Red and green numbers indicate probabilities larger than 0.6 and smaller than 0.6, respectively.

F. Evaluation on Patch, Image and Patient Level

Set B contains 30 patients, 375 pictures, and 12,777 parasites; to evaluate our automated parasite identification approach, we apply IGMS and CNN classifier to this set. Out of the 187,500 patches generated by IGMS, 13,066 are deemed positive due to their overlap of more than 50% with the ground truth annotations. Using a classifier score cutoff of 0.6, we apply the modified CNN model to the 187,500 patches and forecast 13,687 patches as parasites. After five rounds of cross-validation on Set A, this cutoff showed the best accuracy based on the ROC curve. Accuracy 97.26%, area under the curve 97.34%, sensitivity 82.73%, specificity 98.39%, precision 79.99%, and F-score 80.81% were the performance metrics obtained on the patch level for this threshold. Figure 12 and Table III show the relevant ROC curve and confusion matrix. Reducing the classifier score threshold allows us to get a sensitivity of 93% for a specificity of 90%, as seen in the ROC curve in Fig. 12.

We also use linear regression to assess our technique on both the image and patient levels, as seen in Figure 13. We anticipate an average of 35 parasites per picture at the image level, but the ground truth images typically include 34 parasites per image. The significant connection between the projected number of parasites and the ground truth Fig. 13(a) is shown by the high correlation coefficient of R=0.98. Predicted parasite loads at the patient level are 456 on average, compared to 426 in the ground reality. As shown in Figure 13(b), we also find a substantial link at the patient level, with a correlation value of R=0.99.

Parasite detection, including screening and classification, on an RGB picture with dimensions 4032 × 3024 × 3 pixels, when executed using TensorFlow Mobile on a Samsung Galaxy S6 (powered by an Exynos 7 Octa 7420 Processor and running Android version 7.0), takes around 10 seconds.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

The suggested strategy is shown in Fig. 14 by a realistic example of parasite identification. The ground truth parasites are represented by yellow circles in Figure 14 (a), whereas the IGMS-screened parasite candidates that overlap more than 50% with the ground truth parasites are shown by red and green circles, respectively. In the modified CNN model, red circles represent the screened candidates that are ultimately projected as parasites (positives), while green circles represent those that are predicted as non-parasites (negatives). Figure 14's magnified rectangle area shows seven potential parasites (red and green circles) and five identified parasites (yellow circles). The modified CNN model has predicted that five of the parasite candidates are, in fact, parasites. The CNN-provided probabilities for each candidate are shown in Figure 14(b).

G. Comparison with pre-trained networks

We perform evaluations on Set B to compare the performance between our method and pre-trained networks, such as AlexNet, VGG19, and ResNet50. First, we extract patch candidates using IGMS, and then apply different models to detect the true parasites. We compare performances in terms of accuracy, sensitivity, specificity, precision, F-score, AUC, and sensitivity for a given specificity. As listed in Table IV, the accuracy of our customized CNN model is about 1% higher than AlexNet, and about 4% higher than VGG19 and ResNet50. The F-score of our customized CNN model is about 5%, 15%, and 16% higher than the F-score of AlexNet, VGG19, and ResNet50, respectively. For a specificity of 98.39%, the sensitivity of our customized CNN is about 5%, 15%, and 23% higher than the sensitivity of AlexNet, VGG19, and ResNet50, respectively.

TABLE IV. PERFORMANCE OF DIFFERENT NETWORKS ON SET B.

Network	Accu	Sensi	Speci	Preci	F-score	AUC	Sensi1
IGMS+ ResNet50	93.88%	81.34%	94.82%	54.04%	64.94%	95.48%	59.97%
IGMS+ VGG19	93.72%	87.31%	94.20%	52.99%	65.95%	96.99%	67.26%
IGMS+ AlexNet	96.33%	82.15%	97.39%	70.23%	75.73%	96.97%	77.77%
IGMS+ Our CNN	97.26%	82.73%	98.39%	78.98%	80.81%	97.34%	82.73%

Note: Accu, Sensi, Speci, and Preci indicate accuracy, sensitivity, specificity and precision, respectively. Sensi1 is the sensitivity for a given specificity of 98.39%.

IV. DISCUSSION

Here, we build a parasite detection app for smartphones using deep learning and our IGMS approach. As shown in Section III, our program attains an area under the curve (AUC) of 97.34% and a patch level accuracy of 97.26%, while achieving correlation coefficients over 98% on both the picture level and the patient level. There are primarily two reasons for this: To start, IGMS does a good job of selecting candidates for parasites based on the ground reality. Secondly, our convolutional neural network (CNN) model is able to accurately categorize the pre-selected candidates thanks to its customizable input size and network layers.

Our CNN model learns to minimize false positives by producing fake positive patches (negative patches) that closely resemble parasites (positive patches). This is achieved by our IGMS approach. Negative patches picked at random from the backdrop have also been the subject of our research. On the other hand, Set B saw an accuracy drop below 75%. This is due to the fact that an excessive number of clean negative patches are produced by the random picking of negative patches. Consequently, a large number of false positives are produced when the CNN model is trained on such patches.

Using three distinct input patch sizes—36×36×3, 44×44×3, and 52×52×3—we have evaluated how well our technique performs. We find that a large number of false positives are picked up when the patch size is 36×36×3. Because there isn't enough data to reliably detect parasites with this patch size, the procedure fails. The AUC value on the patch level is 97.30% when the patch size is raised to 52×52×3, which is very similar to the findings we published for a 44×44×3 patch size. But at the patient and picture levels, the correlation values drop to 0.97 and 0.96, respectively. This is because a larger patch results in a greater amount of ambient noise.

When compared to a conventional support vector machine (SVM) classifier trained on HOG features, our tailored CNN model achieves 6% better accuracy, 8% better sensitivity, 4% better specificity, 5% better precision, 6% better negative prediction, and 6% better F-score.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

TABLE V. PERFORMANCE COMPARISON BETWEEN THE CUST MODEL AND DL MODELS IN THE LITERATURE ON SE

Network	CPU or GPU	Learning rate	Training time
ResNet50	CPU	0.001	47382
VGG19	GPU	0.001	13698
AlexNet	GPU	0.001	1613
Our CNN	GPU	0.001	1487

Based on the comparison between our customized CNN model and the three pre-trained networks Alexnet, VGG19, and ResNet50 (on a CPU), in Tables IV and V, we find: 1) our customized CNN model is more than ten times faster than VGG19 and ResNet50 (see Table V); 2) the accuracy of our customized CNN network is significantly better on Set A, between one and two percent, than the accuracy of a pre-trained VGG19 ($p<0.001$) and AlexNet ($p<0.01$), with a larger difference on Set B (Table IV). ResNet50 achieves an accuracy around 92.50% on Set A. However, ResNet50 is too big and too slow for our smartphone application; 3) according to the ROC curve, our customized CNN outperforms AlexNet, VGG19, and ResNet50 from 5% to 23% in terms of sensitivity for the given specificity. We have also applied object detection networks, such as faster-RCNN [32] and YOLO [33], to detect parasite candidates. However, these object detection networks do not work well for very small objects like parasites, with an average size of 44x44 pixels in an image of 4032×3024 pixels, resulting in many false negatives.

V. CONCLUSION

In this research, we develop a deep learning program for mobile devices that can identify malaria parasites in thick smear pictures. Two steps make up our processing pipeline for automated parasite detection: screening for parasites and classifying them. A rapid screening of a whole thick smear picture is conducted via an intensity-based Iterative Global Minimum Screening (IGMS) to identify potential parasites. Next, each candidate is categorized as either a background or a parasite using a modified CNN model. The feasibility of our approach for automated identification of malaria parasites is shown by our experimental findings. We believe ours is the first study to use deep learning methods to identify parasites in thick blood smears on smartphones, and the second to create a smartphone app for this purpose [18]. We also evaluated our app on a patient-level. To help researchers address the scarcity of thick blood smear training data for automated malaria detection, we are making accessible our dataset of 1819 pictures from 150 individuals. Improving the performance and runtime of our automated parasite identification system on smartphones is our next effort. We will be leveraging network ensemble methods for this purpose.

REFERENCES

[*1] WHO, World Malaria Report 2018. 2018.*

[2] WHO, Guidelines For The Treatment of Malaria, Third edition. World Health Organization, 2015.

[3] K. S. Makhija, S. Maloney, and R. Norton, "The utility of serial blood film testing for the diagnosis of malaria," Pathology, vol. 47, no. 1, pp. 68–70, 2015.

[4] WHO, Malaria micropscopy quality assurance manual, Version 2. World Health Organization, 2016.

[5] M. Poostchi, K. Silamut, R. J. Maude, S. Jaeger, and G. Thoma, "Image analysis and machine learning for detecting malaria," Transl. Res., vol. 194, pp. 36–55, Apr. 2018.

[6] Z. Liang, A. Powell, I. Ersoy, M. Poostchi, K. Silamut, K. Palaniappan, P. Guo, M. A. Hossain, A. Sameer, R. J. Maude, J. X. Huang, S. Jaeger, and G. Thoma, "CNN-based image analysis for malaria diagnosis," in Proc. BIBM, ShenZhen, China, 2017, pp. 493–496.

[7] S. Rajaraman K. Silamut; M. A. Hossain, I. Ersoy, R. J. Maude, S. Jaeger, G. R. Thoma, and S. K. Antani, "Understanding the learned behavior of customized convolutional neural networks toward malaria parasite detection in thin blood smear images," J. Med. Imaging, vol. 5, no. 3, p. 034501, July 2018.

[8] L. Rosado, J. M. Correia da Costa, D. Elias, and J. S. Cardoso, "A Review of Automatic Malaria Parasites Detection and Segmentation in Microscopic Images," Anti-Infective Agents, vol. 14, no. 1, pp. 11–22, Mar. 2016.

[9] P. A. Pattanaik and T. Swarnkar, "Comparative analysis of morphological techniques for malaria detection," Int. J. Healthc. Inf. Syst. Informatics, vol. 13, no. 4, pp. 49-65, Oct. 2018.

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

[10]S. Kaewkamnerd, A. Intarapanich, M. Pannarat, S. Chaotheing, C. Uthaipibull, and S. Tongsima, "Detection and classification device for malaria parasites in thick-blood films," in Proc. IDAACS, Prague, Czech Republic, 2011, pp. 435–438.

[11]N. S. M. M. Hanif, M. Y. Mashor, and Z. Mohamed, "Image enhancement and segmentation using dark stretching technique for Plasmodium Falciparum for thick blood smear," in Proc. CSPA, Penang, Malaysia, 2011, pp. 257–260.

[12]K. Chakrabortya, "A Combined Algorithm for Malaria Detection from Thick Smear Blood Slides," J. Heal. Med. Informatics, vol. 6, no. 1, pp. 179-186, Jan. 2015.

[13]I. R. Dave and K. P. Upla, "Computer Aided Diagnosis of Malaria Disease for Thin and Thick Blood Smear Microscopic Images," in Proc. SPIN, Noida, India, 2017, pp. 4–8.

[14]M. Elter, E. Hasslmeyer, and T. Zerfass, "Detection of malaria parasites in thick blood films," in Proc. EMBS, Boston, MA, USA, 2011, pp. 5140–5144.

[15]I. K. E. Purnama, F. Z. Rahmanti, and M. H. Purnomo, "Malaria parasite identification on thick blood film using genetic programming," Proc. ICICI-BME, Bandung, Indonesia. pp. 194–198, 2013.

[16]L. Yunda, "Automated Image Analysis Method for p-vivax Malaria Parasite Detection in Thick Film Blood Images," Rev. S&T, vol. 10, no. 20, pp. 9–25, Mar. 2011.

[17]J. A. Quinn, A. Andama, I. Munabi, and F. N. Kiwanuka, "Automated Blood Smear Analysis for Mobile Malaria Diagnosis," in Mobile Point-of-Care Monitors and Diagnostic Device Design, W. Karlen and K. Iniewski, Eds. CRC Press, 2014, pp. 1–20.

[18]L. Rosado, J. M. C. Da Costa, D. Elias, and J. S. Cardoso, "Automated Detection of Malaria Parasites on Thick Blood Smears via Mobile Devices," Procedia Comput. Sci., vol. 90, no. July, pp. 138–144, Dec. 2016.

[19]C. B. Delahunt et al., "Automated microscopy and machine learning for expert-level malaria field diagnosis," in Proc. GHTC, Seattle, WA, USA, 2015, pp. 393–399.

[20]J. A. Quinn, R. Nakasi, P. K. B. Mugagga, P. Byanyima, W. Lubega, and A. Andama, "Deep Convolutional Neural Networks for Microscopy-Based Point of Care Diagnostics," in Proc. ICMLHC, Los Angeles, CA, USA, 2016, pp. 271–281.

[21]C. Mehanian, M. Jaiswal, C. Delahunt, and C. THOMPson, "Computer-Automated Malaria Diagnosis and Quantitation Using Convolutional Neural Networks," in Proc. ICCVW, Venice, Italy, 2017, pp. 116–125.

[22]K. Torres et al., "Automated microscopy for routine malaria diagnosis: A field comparison on Giemsa-stained blood films in Peru," Malar. J., vol. 17, no. 1, pp. 339–50, Sept. 2018.

[23]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., pp. 1–9, Dec. 2012.

[24]Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in Proc. CVPR, Columbus, OH, USA, 2014, pp. 1891–1898.

[25]S. Ren et al., "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proc. CVPR, Boston, USA, 2015, vol. 794, pp. 1–15.

[26]K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. CVPR, Las Vegas, NA, USA, 2016, pp. 770–778.

[27]K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv Prepr. arXiv1409.1556, p. [Online] Available: https://arxiv.org/abs/1409.155, 2014.

[28]C. Szegedy et al., "Going deeper with convolutions," in Proc. CVPR, Boston, USA, 2015, pp. 1–9.

[29]N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," IEEE Trans. Syst. Man. Cybern., vol. 9, no. 1, pp. 62–66, Jan. 1979.

[30]S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing

ISSN: 1832-5505
Vol-11 Issue-04 Nov 2023

Internal Covariate Shift," in Proc. ICML, Lille, France, 2015, pp. 81–87.