A Review of Conventional and Machine Learning Techniques for Malaria Parasite Detection Using a Thick Blood Smear

Husnain Khalid[a*], Dr Nadeem Qaiser Mehmood[b], Bilawal Arif[c], Mubashar Mehmood[d]

[a,b,c]*University of Lahore Gujrat Campus, Near Chenab River, Gujrat and 50700, Pakistan*

[d]*Bahria University Islamabad, Sector E, Islamabad and 44210, Pakistan*

[a]*Email: husnainkhalid1@gmail.com*

[b]*Email: nadeem.mehmood@cs.uol.edu.pk*

[c]*Email: mbilawal333@gmail.com*

[d]*Email: mubasharsmart@gmail.com*

Abstract

Click here and insert your abstract text. Life-threatening malaria is caused by parasites that are lethally effective and harmful and are transmitted through the bite of female Anopheles mosquitoes. In 2015, WHO reported more than 200 million deaths occurred because of this. This makes malaria one of the most vulnerable diseases. The Plasmodium parasite needs to be detected at the early stages for the patient's survival. Microscopists over the years have been made such craftsmen that they through their expertise have been able to diagnose malaria, being followed by an area expansion support from computer-aided diagnosis. But the expertise required for feature extraction were questionable, which were later replaced by deep learning techniques through automatic feature extraction in CNN's. This paper provides a review of some such techniques and methods which were used for the said purposes.

Keywords: Computer-aided diagnosis; CNN; Deep learning; SVM; Malaria detection.

----

* Corresponding author.

1. Introduction

Transmission of malaria is caused by female anopheles mosquito bites. Mosquito habitats in tropical and subtropical areas make malaria as one of the health problems in the mentioned areas [1]. In 2015, World Malaria Report releases the data about the death cases which were caused by malaria. It stated that 438.000 people in 200 million cases are dead because of malaria disease [2]. Most of the death cases which were caused by malaria occurred in Africa. This is because of the favorable environmental conditions of some African regions which were supportive of mosquitoes nurturing. Moreover, the social and economic conditions of African people are still very poor that they have difficulties to receive proper treatment and equipment in countering malaria producing the high rate of death from malaria disease [2]. With the same region characteristics, eastern Indonesia also has the same problem, malaria is still one of the deathly infectious diseases which become one of the health issue concern in Indonesia. By utilizing a computer with special software developed to support the malaria diagnosis, it can be categorized as an easy tool that can be easily distributed in rural areas. Moreover, considering the variations of malaria cases in each region, by using a CAD technique with a machine learning approach, the development of CAD software for malaria diagnosis can be improved with the passage of time. Related research on this area has been widely done by several researchers. Our aim in writing this article is to review some techniques and methodologies used for automated malaria diagnosis based on microscopic thick blood smear images.

2. Malaria

Malaria is caused by Plasmodium parasite that is spread through the bites of the female Anopheles mosquitoes. Plasmodium parasites will settle and multiply in the system of the liver humans when it spreads throughout the human body while infecting red blood cells. There are 5 species of Plasmodium that can cause malaria in human which are P.falciparum, P.vivax, P.ovale, and P.malaria. In addition, the Plasmodium presence of malaria cases Plasmodium caused by P.knowlesi in the forests of Southeast Asia, especially the island of Borneo. Detection of P.Malaria using conventional methods. The Plasmodium parasite is a specie which lives in red blood cells. When it infects humans, it will multiply in red blood cells. Experts will need equipment such as a microscope, to find out the existence of parasite in the blood, to extract and enlarge blood images. These images are then analyzed in a laboratory. Detection using a microscope is an effective way of performing parasitic detection and still used until today.

- Clinical Malaria Diagnosis

  This is one of the traditional malaria detection techniques. The diagnosis is clinically done by observing the patient's body and identifying visible physical symptoms and characteristics. Although early symptoms happened are still common, medics will try to identify the disease such as fever, headache, diarrhea, cold, chills, abdominal pain, myalgia and vomiting [4].

  The number of symptoms that experts need to consider in clinical diagnosis is leading to the difficulty for accurate identification of the disease because the symptoms of the disease are likely still too general to

determine whether that the patient has malaria or not.

- Laboratory Malaria Diagnosis.

This kind of diagnosis requires special equipment, materials and expertise since most of these tools are only used by medical personnel. There are four commonly used methods for diagnosing malaria in the laboratory:

• Peripheral Blood Smear

The diagnosis of malaria using peripheral blood smear base on a blood smear. And the presence of such parasites in the blood smears was observed with a microscope. The blood smear can be thick and thin. The thick blood smear is used for detecting the presence of Plasmodium parasite and thin blood preparation is used for the detection of the type and phase of the parasite [6].

The results of sensitivity and specificity depend on the technique under use, expertise and the reagent of medical personnel. The required time to do the diagnosis is about 30-60 minutes.

• b) Serological Tests

The serologic examination is a test performed by recognizing human antibodies that work against parasites. The accuracy of the examination produces quite a good result [7]. A serologic examination can detect any type of Plasmodium and is very productive for epidemiological surveys. The disadvantage of this method is the required time to detect the parasite that is quite long, around 30-60 minutes.

• c) Quantitative Buffy Coat

It is designed to improve malaria detection based on microscopic tests, QBC also makes easier malaria detection [7]. It has better sensitivity and specificity compared to PBS and is able to detect parasites in less than 15 minutes. The QBC method is only able to detect specific types of Plasmodium parasites.

• d) Rapid Diagnostic Tests (RDT)

To solve the problems inflicted by microscopic tests, experts designed a simple method that uses antigens from parasites, which identified blood flow runs through membranes containing malarial antibodies. It does not require electricity or special expertise and is able to provide high accuracy detection results. However, it cannot determine the type of Plasmodium parasite.

• Staining Methods.

Giemsa's blood staining was first applied for the diagnosis of malaria in 1902. It is recognized for its low cost, high sensitivity and specificity. It is widely used in microscopic malaria examinations [8]. But it

requires multiple reagents and experienced individuals. It is labor intensive and time-consuming. Typically, it takes at least 45 minutes to stain one slide [8]. Other staining methods have been used to overcome the staining time.

- Computer-based Malaria Diagnosis

The use of digital image processing in Computer Aided Diagnosis (CAD) support the diagnosis procedure with the help from the computer. Usually, the initial step of CAD is the acquisition of digital images of blood smears. This initial step breaks down the different approaches for the different types of microscopy, thin or thick blood slides and staining [14].

In general, the flow of digital image processing is as follows: (1) image preprocessing, (2) image segmentation, (3) feature extraction, (4) classification based on the similarity of object features.
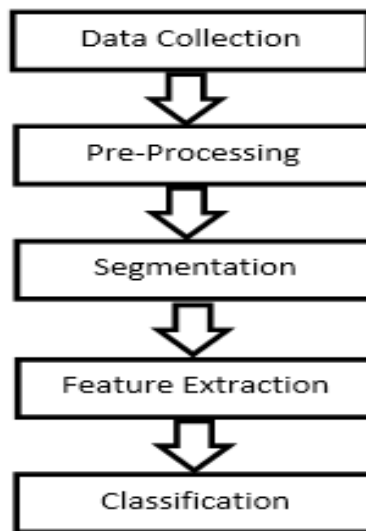
3. Methodology

Figure 1: Typical Classification Procedure

3.1 Pre-processing

In image processing, the initial stage that needs to be done is the pre-processing stage. Improving image quality and image size changes are done at this stage. The image that is being the object often has poor quality, for example, the image has noises during the transmission through a transmission line, the image brightness level, and the obscurity of the object in the image. Through this initial processing operation, the image quality is improved so that the image can be used for further applications. Frean and his colleagues [15] used a rolling ball algorithm to clean and repair the background image. The parameter used in this algorithm was a light background which produces a brighter background. Basically, this algorithm refines the image by reducing the

noise in each area per 3x3 pixels. After the background became clearer, edge detection was done using a Sobel edge detector by highlighting the intensity changes that occur between objects and the background. The final stage of preprocessing in this research was to convert the image into a black and white binary image and filled the holes that may appear on objects by using mathematical morphology. Elter and his colleagues [16] performed two major stages to detect malaria parasites. Pre-processing was in the first stage. According to this research, the proportion of red and green components of a blood smear image is the best feature for identifying objects which contain chromatin. Thus, in the preprocessing stage, this research tried to transform the original blood smear image into the monochrome image, where the object containing chromatin turned into dark grey, and objects that did not contain chromatin became light grey. Arco and his colleagues [17] performed pre-processing which consist of 2 operations; image filtering to reduce noise, and improving image quality. At the noise reduction stage, a Gaussian low pass filter was used to get the signal region more clearly and suppress the effect of noise. Once the Gaussian filter applied, the image became brighter, thus, the intensity of the pixels became higher than before. The next stage in the preprocessing was the histogram equalization. This operation produced images with high contrast. This method will be useful when the image is represented by adjacent contrast values. In this research, the adaptive histogram equalization method was used to calculate multiple histograms, in which each histogram was corresponding to different image sections, and used it to redistribute the brightness value of an image. The optimal value of the size of each image sub-division has been determined, 64 pixels (8x8) and bilinear interpolation of the pixel edge from 2 subdivisions has been calculated. Then, to prevent over-amplification of noise, contrast limits were used. After going through the image improvement process, an h-minima transformation operation was performed. This operation effectively suppressed all the intensity of minima in the image. With this operation, the difference between background pixels could be reduced . Abidin and his colleagues [18] did 2 processes of image quality improvement on the image of ROI cutting result; noise filtering and contrast increases. Noise filtering involved median filtering, Gaussian filtering, and low pass filtering. Then, contrast enhancement was involving contrast stretching. Dave and his colleagues [19] converted RGB image to HSV. It was based on the assumption that the particles in the color space of HSV had significant differences and the image in the HSV color space could decrease brightness variation and blood staining concentration, which was very influential on the image with RGB color space. The channel used in the HSV color space is a saturation channel. Median filter with a size of 3x3 then performed on the channel to reduce noise. Mehanian and his colleagues [20] used white balancing techniques to compensate for some color variations. Basically, this white balancing technique is scaling red, green and blue (RGB) pixel values based on the mean color of the brightest pixels in each image individually [20]. The result of this technique is pixels from all fields of view pooled and global color balance affine transform for each blood sample is computed. Rosado and his colleagues [21] uses an optical magnification prototype that can be easily adapted to a smartphone. The image processing software then installed in a smartphone application. A whole slide of thick blood smear images that acquired using this methodology has circular shape field of view with black background in each corner. In order to remove the circular shape, optical circle detection is used in the first place. The second step of this study is to detect the white blood cell (WBC). Mean shift filtering was applied, this method can preserve edges [21]. After all, WBC detected, the last step is trophozoites detection, which is done by comparing the dimensions and stain contrast of trophozoites cytoplasm with WBC objects. Salamah and her colleagues proposed two main steps image preprocessing. The first step is the contrasting correction. Contrast correction is

done globally and locally. For a global contrast correction, a dark stretching is performed, and for a local contrast correction, contrast limited adaptive histogram equalization is used. [22]. Hanif and his colleagues performed a dark stretching technique to enhance and to segment the parasite objects from thick blood smear image slide [23]. To enhance the brightness and contrast level of the image, the auto scaling method which is a linear mapping function is used in dark stretching process [24]. This technique will be based on the original brightness and contrast level of the image to do the adjustment [24]. To enhance the contrast of the image, Kaewkamnerd and his colleagues controlled the vertical movement of the motorized unit so that the system is able to capture images in different depths of the field [25]. To detect edges for in-focus pixel positions of each depth of field, Laplacian spatial filter [26] is used because of its high accuracy and high speed.

Table 1: Summary of pre-processing techniques

Challenges	Pre-processing methods	Remarks
Noise reduction	Rolling ball algorithm [15].	Remove noises while preserving the edges.
	Gaussian low-pass filter [17, 18]. Adaptive histogram equalization [17]. Median filtering [18, 19]. Mean shift filtering [21] Laplacian spatial filter [26].	Very sensitive to noises. Requires image smoothing techniques so that noises are reduced.
Contrast enhancement	Dark stretching [18, 22].	Create a brighter image, objects easily recognized visually.
Color variation	White balancing [20]	Compensate color variation.

3.2 Segmentation

Segmentation is the process of dividing the image into sections based on certain criteria. The main focus at this

stage is the separation between the objects (white blood cells and Plasmodium) and the background. Proper segmentation techniques will greatly impact the next stage and affect diagnosis accuracy. The separation of the object against the background was done by Frean and his colleagues through particle analysis [15]. Initially, the particles considered as parasitic objects are selected manually by using point pickers. Point picker is a feature of the ImageJ software used in research by Frean and his colleagues [15]. Not only parasites which contain chromatin, but also leukocytes, platelets, and artefacts. Because of that, Elter and his colleagues [16] separated plasmodium bacteria with leukocytes based on the characteristic of its former shape, while for platelets was using shape characteristics and intensity of staining. Blacktop-hat morphological operators and threshold count operation was used to separate plasmodium bacteria with other objects. Basically, a morphological operator uses a general form that represents the target object to separate it with other objects. This research used the object with a paraboloid shape with a radius of 7 pixels, which was obtained from the size of plasmodium objects in general. The threshold value was specified with a fixed value and performed globally. To combine the blobs contained in the binary image resulting from the threshold operation, the method of dilation with the circular form of element structure was used. The radius of this circular element structure was chosen to match the radius of the elemental structure used in black-top-hat operator. The adaptive technique was again performed by Arco and his colleagues [17] at the segmentation stage. In the previous process, this research produced an image histogram with unimodal characteristics, with several small peaks formed a large peak and unclear valleys. This made difficulties to determine the threshold value. To solve this problem, an adaptive thresholding process was used. First, image median was calculated with the mean filter (15x15 mask).Before performing feature extraction, Arco and his colleagues [17] performed mathematical morphological operations to improve the shape of objects in a blood smear. There are 2 operations performed; dilation and erosion. Dilation operation was performed to close the hole, while erosion was to remove the irrelevant details of the object. Then, object analysis based on interrelated components was done. This process aimed to select the objects of plasmodium candidates. Abidin and his colleagues [18] used image resulted from the selection of malaria parasite candidate in form of an image with improved quality at the pre-processing stage, to perform segmentation afterwards where active contour without an edge (ACWE) is one of the main methods of segmentation. The process of segmentation itself consists of 6 stages, they are:

1) *Complement*

It is the process of turning the image into a negative image form. The image used in this research consists of dark-colored objects or parasites, and background with brighter colors. Once the complement process is done, the object or parasite will turn brighter, and the background becomes darker.

2) ACWE

This algorithm is designed to segment objects that have an obscure edge. In this research, it produced binary images with white objects and black backgrounds.

3) *Erosion + Dilation*

This stage aims to remove irrelevant small objects and refine the edges of the object.

4) *Masking* 5) *Contrast stretching*

This stage aims to make the background darker and the object becomes lighter, resulting in more contrasting images.

6) *Thresholding*

By using a pre-determined thresholding value, this method is segmenting the image into 2 areas, objects and background. Dave and his colleagues [19] used an adaptive histogram thresholding technique to separate parasitic candidate objects. The threshold value was derived from the search of the cumulative distribution function value. At this stage, not all detected objects were malaria parasites. Several techniques are implemented by Mehanian and his colleagues [20]. Before performing a segmentation step, Mehanian and his colleagues [20], applied a dark threshold to a grayscale intensity image as an initial detector for malaria parasites candidates. The initial detector resulting in high sensitivity while its precision is low, a large amount of distractor are also detected. These distractor objects cause a high number of false positive detections, which degrades system performance. To overcome that problem, adaptive grayscale intensity and dynamic local thresholding are performed. The segmentation step which performed by Mehanian and his colleagues [20] consists of two passes, first passes, white blood cell(WBC) candidates are segmented using dark threshold tied to grayscale intensity statistics. Individual WBC candidates are then filtered using morphological and clustering operations in order to classify each individual WBC using Gaussian-kernel SVM. The result of WBC segmentation is the RGB color collection statistics and a random sampling of background pixels. Another machine learning technique is implemented to compute the optimal projection in RGB space in order to separate WBC pixels from background pixels. Segmenting the WBC objects is an important part of the study by Rosado and his colleagues [21]. The segmentation technique used in this study is adaptive thresholding from the mean intensity value of the square region centered on the pixel location. After the pre-processing step is done, Kaewkamnerd and his colleagues [25] convert the whole in-focus image into HSV (Hue-saturation value) color format. For segmentation purposes, the value component from the HSV color channel is employed. The segmentation is done in three steps, first, construct a histogram of value components and extract non-background objects using adaptive threshold [27, 28] based on the histogram value. The second step is to divide the image into small windows of 300 by 300 pixels for the efficient use of resources in the searching process . The connected component analysis then performed in order to label each object for future reference. The last step is filtering malaria parasites from other objects according to their difference in sizes. The second and third steps are repeated until all the parasites are discovered and labelled [25].

Table 2: Summary of segmentation techniques

egmentation Techniques	Remarks
Black-top-hat morphology [16]	With correct parameters, these morphological operators can remove WBC objects. However, in a certain case, it can increase the false positive number.
Global Threshold [18]	Threshold value applied to all slide image without considering the image characteristics.
Adaptive thresholding [17, 19, 21, 25]	The threshold value is determined based on image calculation, such as histogram, mean intensity value.
Dilation + Erosion [17]	Holes filling and remove small objects.
Active Contour [18]	Very effective with the uniform lighting condition, the curve evolution enables whole object capture with a similar shape. Heavy calculation.

3.3 Feature extraction

After objects are separated from the background, feature extraction is applied to those objects to recognize objects based on feature similarities. In a research conducted by Frean and his colleagues [15], after the object coordinates were obtained, the analysis of the object was done by using particle analysis to obtain object characteristics, so the calculation can be done automatically. This particle analysis feature requires an object's size parameter and level of the circle. Particle analysis generates the size of the object that most likely is the characteristic of malaria parasite object. Elter and his colleagues [16] performed feature extraction in the first (detection) and second (error detection reduction) stage. The feature extraction in the first stage was performed to extract the positions of plasmodium candidates using a simple connected component labelling algorithm. The extracted objects were determined as the positions of plasmodium candidates based on their centroid. In the second stage, feature extraction was performed by cutting the region of interest (ROI) in 80x80 pixels on each plasmodium candidate from the input image. Based on these ROI image pieces; statistical features, texture analysis features, and color textures were used as a set of features to be extracted. It generated a large set of features. These features were then selected to produce an optimized feature subset using univariate rankings to select the top 60 subsets that have a great correlation with objects and genetic algorithms, to minimize features generated by univariate ratings. After the objects of the candidate of Plasmodium are filtered, Arco and his colleagues [17] performed object property measurements. The main difference between white blood cells and malaria parasites is their size, thus, the size became one of the properties used. Then, the entire area i.e. the number of pixels forming the area) was checked, resulting in 2 parts; parasites or white blood cells, which in this way, the higher the value of the label given to a field indicated the number of fields (parasites and white blood cells) in the image. Cutting of objects detected as parasitic candidates is done by Dave and his colleagues [19], to perform feature extraction based on the result of the cutting of object afterwards. The feature extraction

of each object produced 123 features that each object had. Features possessed by each object were 56 statistical features, 16 geometry textures, 24 intensity textures, 22 color features, and 1 frequency domain feature. Based on Mehanian and his colleagues [20] paperwork, objects that survived the distractor filter are then assumed as malaria parasite candidates. Because of the proposed method is mainly about Convolutional Neural Network (CNN), it performed some experiments using deep learning framework, including Alexnet, VGG and GoogleNet. After subsequent experiments, it comes to the conclusion that VGG feature extractor is better than Alexnet and GoogleNet. To avoid overfitting when training the CNN, this paperwork performs an augmentation of data by randomizing gamma correction of individual color channels. This process gave more realistic blood smear microscopy image colors as well as improved performance [20]. Rosado and his colleagues [21] Extract WBC and trophozoites candidates. For each WBC candidate, a total of 152 image features were extracted and grouped into 3 major groups: Geometry, Color and Texture features. Purnama and his colleagues [29] perform a manual crop of parasite objects as a region of interest. Extracted features from cropped images are mean, standard deviation, kurtosis, skewness and entropy of the red-green-blue channel color histogram, hue channel of the HSV histogram, and hue channel of HSI histogram.

3.4 Classification

The similarity of features between objects obtained in the previous process then classified to recognize objects based on feature similarity, so that whether the object is a malaria parasite or other artefact in the blood can be determined. Frean and his colleagues [15] again performed particle analysis with parameters that already obtained in the previous feature extraction process. After the median size of the malaria parasite was determined, this study again conducted a particle analysis to separate malaria parasite with other artefacts.

Table 3: Summary of extracted feature

Author	Extracted Feature
Frean [15]	Object size and circular level.
Elter [16]	Statistical features, texture analysis features, and color textures
Arco [17]	Object size
Dave [19]	Statistical, Geometry, Intensity, Color, Frequency domain.
Mehanian [20]	Using CNN framework such as Alexnet, GoogleNet, and VGG.
Rosado [21]	Geometry, Color, and Texture features.
Purnama [29]	Mean, standard deviation, kurtosis, skewness and entropy of color channel combination.

Elter and his colleagues [16] used Support Vector Machine (SVM) algorithm to perform the classification. Once a feature subset was found, the SVM algorithm was used to study it, then classified it by 2 classes; Plasmodium or artefacts. Arco and his colleagues [17] used the opening operation to remove white blood cells, resulting in images containing only malaria parasites. The structural size of the element of average white blood cell size had been determined so that all elements smaller than the specified value will be removed. This step produced an image that contains only white blood cells. By calculating the difference between the two images, the original image with the image containing the white blood cells produced an image containing only malaria parasites. Dave and his colleagues [19] classified it into 5 classes; ring form, ET form, late trophozoite form, leucocytes, and non-parasites like artefacts, platelets and other particles. The LT class consists of two parasitic phases; schizont and gametocyte because these two parasitic phases are difficult to be recognized. SVM was used to classify, extracted 123 features from 2323 ROI's tested results from 43 image data trained, and tested in 2290 ROI's results of cutting object from 44 images data test. In this research, the more features used will improve accuracy. Parasitic calculations by proposed algorithms were able to detect parasites in ring, trophozoite, schizont and gametocyte phases. The proposed algorithm error rate was 7.14% and still in the WHO quality control limit. Mehanian and his colleagues [20] implemented a logistic regression as the external classifier. The reasons behind the use of logistic regression as external classifier are mimicked the CNN fully-connected plus Softmax output and the software package implements a robust, large-scale learning algorithm for logistic regression based on SGD [20]. Rosado and his colleagues [21] performed 2 classifications. First, to classify WBC and second, trophozoites classification. WBC and trophozoites classification used the same classification method. A two class SVM classifier with an RBF kernel was used to create a classification model, using a grid search approach to obtain the best parameters [21]. Kaewkamnerd and his colleagues [25] classify the malaria parasites into two species, *Plasmodium falciparum* and *Plasmodium vivax* based on chromatin size feature. The observation of the chromatin size shows that Pf parasites have a smaller size of chromatin than those of PV. To verify this argument, chromatin size of a total of 4.000 samples of both parasite species were investigated. The decision process was then performed by evaluating the distribution of chromatin size as in the following criteria [25]: •Number of parasite = 0: classified as no infection

- Chromatin size< 64.5 nm: classified as unknown object
- 64.5 nm < chromatin size< 258 nm: classified as Pf parasite
- 64.5 nm <chromatin size < 688 nm, and the amount of chromatin size of > 258 nm greater than 20 %: classified as PV parasite.

Purnama and his colleagues [29] proposed genetic programming (GP) method to classify parasite and non parasite objects. GP is a method to evolve computer programs [30]. Automatic programming requires developing a computer program that can produce the desired output for a given set of inputs [31]. To represent the evolutionary computation, GP used tree and graph. This method is usually used for machine learning.

Table 4: Summary of classification methods

Author	Classification method
Frean [15]	Particle analysis.
Elter [16]	SVM
Arco [17]	Connected component labelling filter.
Dave [19]	SVM.
Mehanian [20]	CNN.
Rosado [21]	SVM.
Purnama [29]	Genetic programming.

4. Discussion

In the early stages of Frean and his colleague's experiment [15], to obtain a particle analysis algorithm that can be used in thin blood smear, user intervention was needed in adjusting the various algorithmic variables when analyzing blood preparations with low parasitic density. Despite the high accuracy and reliable results, Frean and his colleague's research had not been fully automated. In this research, Frean and his colleagues [15] also stated that the staining process of high-quality blood contributed to result in high accuracy. With the conditions in malaria-endemic areas where limited medical infrastructure, the image of the good quality blood smear is still difficult to be obtained. In blood smear with low Plasmodium density, the approach used by Elter e and his colleagues 16] was capable of providing high detection sensitivity (97%) with low false positive detection in each image (8%). This research used the image of blood preparation in above standard quality, with a little amount of parasite density in the blood smear (less than 5). Separation of Plasmodium with artefacts was also not performed and could lead to failure in detecting parasites and artefacts. The used of black-top-hat morphological surgery was able to distinguish between malaria parasites and background well. This research used an image with a low parasitic density (3 parasites in 1 image for average). A high degree of sensitivity which leads to failure in malaria parasite detection was completed by the second stage; reduced sensitivity by classifying detected region of interest based on statistical features, textures and colors. The feature extraction based on 3 object characteristics highly depends on the segmentation stage (separation of object and background). If in the segmentation stage does not produce object's separation in the good result (for example, the actual form of a parasite does not resemble its original form), feature extraction will not produce good features for classification, thus impacting classification accuracy. According to Arco and his colleagues [17], to calculate the number of the parasites quickly in the digital images, this research can use adaptive histogram equalization and the adaptive thresholding approach. Certain traits of the image will lead to misinterpretation of parasites and increase the calculation mismatch between manual and digital counting. The application of the proposed method yields an average accuracy rate of up to 96.46% with low processing time (2 seconds of each

image on a medium computing platform). The malaria parasite objects in the image used in this study have the same shape when converted into binary (white dots), therefore the connected component labelling algorithm provides good results for object classification. Implementation of a connected component labelling algorithm can give unfavorable results if the form of the object varies. Variations in the shape of an object usually form at a low parasite density. The research conducted by Abidin and his colleagues [18] tried to use the ACWE segmentation approach that matched the characteristics of the to-be-processed low-quality image with unclear parasite form. In this research, the combination of image quality improvement between low pass filtering and contrast stretching provides the best quality image results. The accuracy of segmentation in this research reached 97.57% with false-negative rate is 12.04%. This study cuts the ROI object from a thick blood smear image, for pre-processing and segmentation afterwards. The implementation of ACWE segmentation is not performed on the whole field of thick blood smear images. This raises the question of whether this segmentation technique would be effective when used in the whole field of blood smear image, taking the characteristics of the active contour method into account, which requires masking initiation at first, author's assumption is that this masking initiation will impact the 2 segmentation stages; the first stage is the segmentation to get ROI, and the second stage is to improve the shape of previously segmented objects. To think that the computation time of ACWE method is quite high, the addition of segmentation in the early stages will provide additional computing time. The classification of the parasite into 5 classes performed by Dave and his colleagues [19] did not publish the points of accuracy. However, this study opens the possibility of identifying the phase of the malaria parasite in the thick blood smear that is usually performed on a thin blood smear. The promising results of computer-aided diagnosis for malaria in thick blood smear images is done by Mehanian and his colleagues [20]. The proposed method is working well because of the use of a large amount of thick blood smear image data which suitable for the neural-network-based method. In this case, CNN is used to classify between parasites and non-parasite objects. Evaluation of the proposed method is done in two ways, patient level and object level. The results indicate that the system has achieved malaria diagnosis accuracy sufficient to attain competence level 1 in the WHO external competency assessment of malaria microscopists for Plasmodium falciparum, which means that it performs a par with well trained microscopists for this species [20]. However, the system is working well with a high density of parasite. Malaria parasite quantitation is sufficiently accurate for greater than1000 p/µL parasitemia and supported by a large amount of data. Exploration of the proposed method in low parasite density and the use of limited data is still open for further research. Moreover, the high quality of the images also has a big impact on the proposed method to give good results. A mobile based framework for malaria parasites detection has been developed by Rosado and his colleagues [21]. The classification results are presented in terms of three metrics: sensitivity, specificity and accuracy. Automatic detection of WBC in thick blood smears achieved 98.2% of sensitivity and 72.1% specificity, while the Plasmodium falciparum trophozoite detection achieved a sensitivity of 80.5% and a specificity of 93.8%. In this study, the focus of the detection is WBC and trophozoite, which can be easily segregated based on its characteristics. Moreover, the blood staining process can emerge artefact object that has similar characteristics with trophozoites which are not considered to be an object to be detected. This issue can lead to false detection of trophozoites and reduce the specificity. The adaptive local contrast correction to enhance image quality by Salamah and her colleagues [22] is very sensitive to noise. In other words, the proposed scheme is dependent on the image smoothing technique. The smoothing technique should be as good as possible to remove noises so that the adaptive local contrast

correction can be performed optimally. The evaluation of the enhanced image is done using peak signal to noise ratio and minimum square error. Their values are better if the range of the image histogram is short [22]. Based on this finding, Salamah and his colleagues [22] stated that the proposed method is effective if the contrast of the original image is low. The result of the dark stretching technique proposed by Hanif and his colleagues [23] is the stretched dark areas of the image whereas bright areas are compressed. In the thick blood smear images, dark area is referred to the parasites, thus the parasites become clearer due to the stretching step in dark stretching technique [23]. In order to find the suitable threshold value, the threshold parameter value had been applied in a different value, meanwhile, the dark stretching factor only used the same parameter for all images, which is 255. The evaluation of the dark stretched image is done by visual inspection. The darker objects and the brighter background show that the proposed image enhancement technique is well performed. The proposed approach by Kaewkamnerd and his colleagues [25] to detect and classify malaria parasites is done on 60 Giemsa-stained thick blood films in which 40 blood films contain infected red blood cells and 20 control blood films of normal red blood cells. The proposed method detected positive and negative parasite blood films at a rate of 95% and 68.5% accuracy, respectively. The performance of the proposed classification was correctly classified with the success rate of 75% while the accuracy of Pf classification was 90%. Images that are acquired by motorized units for controlling the movements of objective lens and microscope stage can produce high-quality image slide, while, sometimes it is quite difficult to have such motorized unit to produce a high-quality image. With the good quality of the image, the end result of parasite detection or classification will have a promising result. Purnama and his colleagues [29] classified all detected object into two and six classes. Classification process performed using the GP method for 120 training data and 60 testing data. The result of the classification shows that the two-class classification has higher accuracy than six class classification. The combination of color channels which performed in this study is an interesting idea and it needs to do more exploration.

5. Conclusion

Research to detect malaria parasites in blood smear has been widely practiced. Before determining what method to use, previous researches identified image characteristics which are going to be processed first. Stages of sequential image processing have an impact on the importance of pre-processing as the initial stage that will affect the next stages. The adaptive technique proposed by Arco and his colleagues [17] provides promising results, both in the preprocessing stage and in the segmentation stage. However, the characteristics of the image used will greatly affect the detection result. Moreover, in their research, Arco and his colleagues [17] assumed that all objects with area value smaller than detected white blood cells are parasites, in fact, there is a possibility that some objects that have similar area value as parasites are artefacts. The use of images with different characteristics in each research has an impact on the difficulties to determine which method is best to use. This suggests that the proposed method in this field of research highly depends on the characteristics of the image used. The exploration of method or scheme of an optimal image processing sequence in a particular image characteristic can still be deeply explored for further research. Moreover, the majority of the publication in thick blood smear does not separate malaria parasite with artefacts that might be appearing as a result of the staining process. The limited data of the medical image in each study is one factor that restricts the researchers in conducting further exploration. When medical image data can be easily obtained, it opens the possibility to

explore with a deep learning algorithm. Several studies [31, 32, and 33] in the topic of malaria identification began to use deep learning, but the image used was still from the thin blood smear.

Acknowledgements

References

[1] World Health Organization. Investing to overcome the global impact of neglected tropical diseases: third WHO report on neglected tropical diseases 2015. World Health Organization; 2015 Aug 5.

[2] Nugroho HA, Akbar SA, Murhandarwati EE. Feature extraction and classification for detection malaria parasites in thin blood smear. In2015 2nd International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) 2015 Oct 16 (pp. 197-201). IEEE.

[3] Nasir AA, Mashor MY, Mohamed Z. Segmentation based approach for detection of malaria parasites using moving k-means clustering. In2012 IEEE-EMBS Conference on Biomedical Engineering and Sciences 2012 Dec 17 (pp. 653-658). IEEE.

[4] Saputra WA, Nugroho HA, Permanasari AE. Toward development of automated plasmodium detection for Malaria diagnosis in thin blood smear image: An overview. In2016 International Conference on Information Technology Systems and Innovation (ICITSI) 2016 Oct 24 (pp. 1-6). IEEE.

[5] Somasekar J, Reddy BE. Segmentation of erythrocytes infected with malaria parasites for the diagnosis using microscopy imaging. Computers & Electrical Engineering. 2015 Jul 1; 45: 336-51.

[6] She RC, Rawlins ML, Mohl R, Perkins SL, Hill HR, Litwin CM. Comparison of immunofluorescence antibody testing and two enzyme immunoassays in the serologic diagnosis of malaria. Journal of travel medicine. 2007 Mar 1; 14(2):105-11.

[7] Clendennen III TE, Long GW, Baird JK. QBC® and Giemsa-stained thick blood films: diagnostic performance of laboratory technologists. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1995 Mar 1; 89(2):183-4.

[8] Keiser J, Utzinger J, Premji Z, Yamagata Y, Singer BH. Acridine Orange for malaria diagnosis: its diagnostic performance, its promotion and implementation in Tanzania, and the implications for malaria control. Annals of Tropical Medicine & Parasitology. 2002 Oct 1; 96(7):643-54.

[9] Houwen B. Blood film preparation and staining procedures. Clinics in laboratory medicine. 2002 Mar 1; 22(1):1-4.

[10] Shute GT, Sodeman TM. Identification of malaria parasites by fluorescence microscopy and acridine orange staining. Bulletin of the World Health Organization. 1973; 48(5):591.

[11] Suwalka I, Sanadhya A, Mathur A, Chouhan MS. Identify malaria parasite using pattern recognition technique. In2012 International Conference on Computing, Communication and Applications 2012 Feb 22 (pp. 1-4). IEEE.

[12] Azif FM, Nugroho HA, Wibirama S. Detection of malaria parasites in thick blood smear: A review. Communications in Science and Technology. 2018 Jun 14; 3(1):27-35.

[13] Wahab IH. *Deteksi Parasit Malaria dalam Sampel Darah Untuk Menunjang Keperluan Diagnosis Berbasis Data Ciri Tekstur Warna* (Doctoral dissertation, Universitas Gadjah Mada).

[14] Poostchi M, Silamut K, Maude RJ, Jaeger S, Thoma G. Image analysis and machine learning for detecting malaria. Translational Research. 2018 Apr 1; 194:36-55.

[15] Frean JA. Reliable enumeration of malaria parasites in thick blood films using digital image analysis. Malaria journal. 2009 Dec; 8(1):218.

[16] Elter M, Haßlmeyer E, Zerfaß T. Detection of malaria parasites in thick blood films. In2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2011 Aug 30 (pp. 5140-5144). IEEE.

[17] Arco JE, Górriz JM, Ramírez J, Álvarez I, Puntonet CG. Digital image analysis for automatic enumeration of malaria parasites using morphological operations. Expert Systems with Applications. 2015 Apr 15; 42(6):3041-7.

[18] Abidin SR, Salamah U, Nugroho AS. Segmentation of malaria parasite candidates from thick blood smear microphotographs image using active contour without edge. In2016 1st International Conference on Biomedical Engineering (IBIOMED) 2016 Oct 5 (pp. 1-6). IEEE.

[19] Dave IR. Image analysis for malaria parasite detection from microscopic images of thick blood smear. In2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) 2017 Mar 22 (pp. 1303-1307). IEEE.

[20] Mehanian C, Jaiswal M, Delahunt C, Thompson C, Horning M, Hu L, Ostbye T, McGuire S, Mehanian M, Champlin C, Wilson B. Computer-automated malaria diagnosis and quantitation using convolutional neural networks. InProceedings of the IEEE International Conference on Computer Vision 2017 (pp. 116-125).

[21] Rosado L, da Costa JM, Elias D, Cardoso JS. Automated detection of malaria parasites on thick blood smears via mobile devices. Procedia Computer Science. 2016 Jan 1; 90:138-44.

[22] Salamah U, Sarno R, Arifin AZ, Nugroho AS, Gunawan M, Pragesjvara V, Rozi E, Asih PB. Enhancement of low quality thick blood smear microscopic images of malaria patients using contrast and edge corrections. In2016 International Conference on Knowledge Creation and Intelligent Computing (KCIC) 2016 Nov 15 (pp. 219-225). IEEE.

[23] Hanif NS, Mashor MY, Mohamed Z. Image enhancement and segmentation using dark stretching technique for Plasmodium Falciparum for thick blood smear. In2011 IEEE 7th International Colloquium on Signal Processing and its Applications 2011 Mar 4 (pp. 257-260). IEEE.

[24] Isa NA, Mashor MY, Othman NH. Contrast enhancement image processing technique on segmented Pap smear cytology images in World Congress on Medical Physics and Biomedical Engineering (WC2003) 2003.

[25] Kaewkamnerd S, Uthaipibull C, Intarapanich A, Pannarut M, Chaotheing S, Tongsima S. An automatic device for detection and classification of malaria parasite species in thick blood film. InBmc

Bioinformatics 2012 Dec (Vol. 13, No. 17, p. S18). BioMed Central.

[26] Peli E. Contrast in complex images. J Opt Soc Am A. 1990; 7:2032– 2040. doi: 10.1364/JOSAA.7.002032.

[27] Haralick R. M., Shapiro, LG: Computer and Robot Vision.

[28] Shapiro L, George C. Stockman g: computer vision. InPrentice Hall 2002.

[29] Purnama IK, Rahmanti FZ, Purnomo MH. Malaria parasite identification on thick blood film using genetic programming. In2013 3rd International Conference on Instrumentation, Communications, Information Technology and Biomedical Engineering (ICICI-BME) 2013 Nov 7 (pp. 194-198). IEEE.

[30] Koza JR, Koza JR. Genetic programming: on the programming of computers by means of natural selection. MIT press; 1992.

[31] Koza JR. Hierarchical Genetic Algorithms Operating on Populations of Computer Programs. InIJCAI 1989 Aug 20 (Vol. 89, pp. 768-774).

[32] Liang Z, Powell A, Ersoy I, Poostchi M, Silamut K, Palaniappan K, Guo P, Hossain MA, Sameer A, Maude RJ, Huang JX. CNN-based image analysis for malaria diagnosis. In2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2016 Dec 15 (pp. 493-496). IEEE.

[33] Dong Y, Jiang Z, Shen H, Pan WD, Williams LA, Reddy VV, Benjamin WH, Bryan AW. Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. In2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) 2017 Feb 16 (pp. 101-104). IEEE.

[34] Gopakumar GP, Swetha M, Sai Siva G, Sai Subrahmanyam GR. Convolutional neural network-based malaria diagnosis from focus stack of blood smear images acquired using custom-built slide scanner. Journal of biophotonics. 2018 Mar;11(3):e201700003.

6. Appendix

- ML = Machine Learning
- DL = Deep learning
- WHO = World Health Organization
- RBC = Red Blood Cell
- MSE = Minimum Square Error
- CV = Computer Vision
- DM = Data Mining
- RDT = Rapid Diagnostic Tests
- SVM = Support Vector Machine
- ANN = Artificial Neural Network
- RNN = Recurrent Neural Network
- CNN = Convolutional Neural Network
- NN = Neural Network
- PSNR = Peak Signal To Noise Ratio
- MSE = Minimum Square Error