Comparative Analysis of Support Vector Machine and Convolutional Neural Network for Malaria Parasite Classification and Feature Extraction

Rika Rosnelly[1*], Bob Subhan Riza[2] and S. Suparni[3]

[1*] Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan, Indonesia. rikarosnelly@gmail.com, Orcid: https://orcid.org/0000-0002-0407-5160

[2] Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan, Indonesia. bob.potensi@gmail.com, Orcid: https://orcid.org/0000-0001-6358-9412

[3] Health Polytechnic, Kemenkes Medan, Indonesia. hajjahsuparni@gmail.com, Orcid:https://orcid.org/0009-0005-8152-9305

Received: June 15, 2023; Accepted: July 28, 2023; Published: September 30, 2023

Abstract

Malaria is a dangerous infectious disease because, if it is slow to handle, it can even cause death. Malaria is caused by a parasite called plasmodium, which is transmitted through the bite of a malaria mosquito called Anopheles. Parasites transmitted by mosquitoes attack human blood cells. The inspection method used to identify the type of malaria parasite is microscopic examination, whose accuracy and efficiency depend on human expertise. Examination methods using the Rapid Diagnostic Test (RDT) and Polymerase Chain Reaction (PCR) are not affordable, especially in underprivileged areas. This study compares the performance of classification methods, namely Support Vector Machine (SVM) and Convolutional Neural Network (CNN), to identify the type of malaria parasite and its stage and develop a feature extraction algorithm. The method of feature extraction is a decisive step to identifying the type of malaria parasite. The feature extraction process by developing a feature extraction algorithm is called the PEMA and KEHE feature tracking algorithm, or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy. The classifier uses a convolutional neural network (CNN) to divide the samples into 16 classes. The experiment used 446 images of malaria parasites. The outcome of identification showed that by tracking the PEMA and KEHE features with the SVM classifier, the best accuracy value was 85.08%, compared to CNN with an accuracy value of 61.40%.

Keywords: Types of Malaria Parasites, Shape Feature Extraction, Texture Feature Extraction, GLCM, SVM, CNN.

1 Introduction

According to Didik Budijianto, the Director of Prevention and Control of Vector-borne and Zoonotic Diseases of the Ministry of Health, malaria cases will increase in 2021, including in North Sumatra increasing by 1531 cases to 2529 cases, where in 2020 there were 998 cases. In the Lampung area, the number of cases increased from 424 in 2020 to 556 in 2021.NTB cases increased even further, rising

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 3 (September), pp. 194-217. DOI: *10.58346/JOWUA.2023.I3.015*

*Corresponding author: Faculty of Engineering and Computer Science, Universitas Potensi Utama, Medan, Indonesia.

from 247 in 2020 to 352 in 2021. A high increase in cases was experienced in the Papua region, which initially had 216,841 cases in 2020 but increased to 243,721 cases (Puspa, A., 2022).

Malaria is transmitted to humans by female mosquitoes of the genus Anopheles (Biernath, A., 2022). Malaria is a disease that can claim the human soul. WHO explained that there are five species of parasites, of which two types pose a threat, namely the plasmodium falciparum parasite and the plasmodium vivax parasite. Plasmodium falciparum is a malaria parasite that can kill humans and is most common on the African continent, while Plasmodium vivax is the dominant malaria parasite in most countries outside sub-Saharan Africa (Arnani, M., 2022). The conventional examination method for malaria is microscopic examination, where the accuracy and efficiency of this method is highly dependent on human expertise. Rapid diagnostic tests (RDTs) and Polymerase Chain Reaction (PCR) are not affordable especially in underprivileged areas of Bangladesh (Fuhad, K.F., 2020), so an automated system is needed to reduce costs and time for identification with good accuracy.

In their study, Garcia et al. (2019) implemented a mechanism for the automatic generation of slides during the image capturing process. The proposed approach involves the utilization of an SURF-based image merging technique, wherein several samples are amalgamated into a single image to enhance the zoom level. Certain researchers employ optical microscopes with a magnification of 400x or more to conduct various investigations pertaining to white blood cell quantification, examination of red blood cell structure, and facilitation of data transfer. The design of the product was achieved by the integration of technological advancements and extensive research on the optimal utilization of microscopes. The present investigation did not ascertain the specific species of parasite identified, but rather just detected the presence of malaria parasites. This technology offers a mobile solution for acquiring microscope pictures that may be afterwards subjected to analysis. The implementation of this solution will effectively alleviate the burden on healthcare professionals by establishing a means of conducting blood tests directly in the field. This facilitates more convenient diagnosis in geographically isolated areas and expedites the administration of therapy when necessary (Garcia, R.G., 2019).

The present study involves conducting research on the differentiation of infected and non-infected red blood cells through the utilization of a collection of extracted attributes. Feature-based datasets are generated with the purpose of providing training data for neural networks (Centers for Disease Control and Prevention., 2009) (Tek, F.B., 2006) (Tek, F.B., 2009) (Annaldas, S., 2015) (Savkare, S.S., 2011) (Liang, Z., 2016). The utilization of a Shape Descriptor feature is employed in order to accurately identify and quantify the roundness of objects that have been detected. The initial statistical characteristics at the first order: The assemblage of statistical measures encompassing the mean, standard deviation, variance, skewness, kurtosis, fifth and sixth moments, and entropy. One of the statistical features used in image analysis is the second order statistical feature, specifically the Gray-Level Co-occurrence Matrix (GLCM). The matrix that represents the co-occurrence of gray level, contrast, correlation, energy, and homogeneity is a set of high-level statistical variables that pertain to texture analysis (Kanojia, M., 2018).

Color normalization with K-nearest neighbors (KNN) and the Bayesian algorithm is described in a study by (Tek, F.B., 2006). This study used a stained pixel classifier, which outcomeed 88.5% true and 5.6% false detection; parasitic and non-parasitic classifiers achieved 74% sensitivity, 98% specificity, 88% positive prediction, and 95% negative prediction.

Research (Tek, F.B., 2009) used the watershed method and image processing color normalization at the segmentation stage, the classifier using Bayesian with a double threshold as the classifier of choice outcome in an accuracy of 96.72%. Research (Liang, Z., 2016) developed machine learning based on a

convolutional neural network to automatically classify single cells in diluted blood and output an accuracy of 97.37%.

The research conducted (Mehrjou, A., 2013) included an automated microscopic image collection system equipped with a GUI based image processing model to measure red blood cell counts. The use of the CNN model for cell image classification is appropriate, because CNN is able to recognize many features and label them correctly. This study uses a variant of the CNN model named VGG16 for recognition. To increase the robustness of this model, researchers have trained it with appropriately labeled images to cluster well even when images are used for classification. The results show that the truth level reaches 95.96% before adversarial training and reaches 95.79% after adversarial training (Huq, A., 2020). The research being carried out is to develop a feature extraction algorithm called the PEMA and KEHE feature tracking algorithm, or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy. Compared the performance of the classification method, namely the Support Vector Machine (SVM) and the Convolutional Neural Network (CNN), to identify the type of malaria parasite and its stage. The research results are expected to have better accuracy.

2 Research Method

This study uses several stages to get the best performance by comparing classification methods: SVM and CNN, to determine the type of malaria parasite and its stage (Salman, R., 2023). Malaria parasite image data were obtained from the North Sumatra Provincial Health Laboratory and photographed using a digital microscope at the North Sumatra Health Polytechnic. Image of the malaria parasite with four types of species and four stages, namely malaria falciparum, vivax, malariae, and ovale, with ring stages, trophozoites, schizonts, and gametocytes. There are 446 images of malaria parasites, with four species and four stages. The stages of the research conducted can be seen in Figure 1.

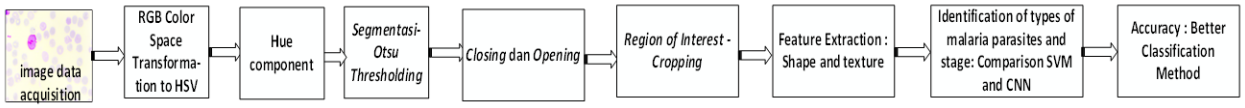


Figure 1: Stages of Feature Extraction Algorithm Development and Performance Comparison of the Classification Method

Image Data Collection

The initial phase of data collection pertaining to images of malaria parasites involves the examination of malaria parasite preparations, followed by the application of Giemsa staining. Subsequently, the application of emersion oil serves the purpose of enhancing the clarity and definition of an object. Subsequently, the digital microscope was employed to acquire the image at a magnification of 1000 times. The image is projected onto the monitor screen, and during the image capture process, health laboratory personnel are present to accurately identify the object, with a particular focus on determining the specific type of malaria parasite. The dimensions of the depicted malaria parasite image are 2592 x 1944. The study employed four distinct species of malaria parasites, including Plasmodium vivax, Plasmodium ovale, Plasmodium falciparum, and Plasmodium malariae. Each of these species exhibits four distinct stages in their life cycle, namely the ring stage, trophozoite stage, schizont stage, and gametocyte stage. Figure 2 displays visual representations of malaria parasites.

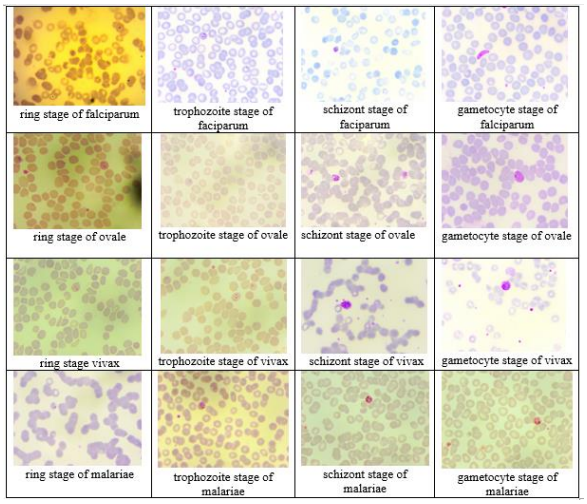


Figure 2: Image of Malaria Parasites with Stadium

The provided image depicts multiple strains of the malaria parasite, specifically falciparum, malariae, vivax, and ovale, each exhibiting four distinct developmental stages: ring, trophozoite, schizont, and gametocyte. Within the confines of the ring-shaped stadium, one may observe four distinct varieties of the malaria parasite. These parasites manifest as ring-shaped objects, each containing either one or two chromatin structures. Typically, infected red blood cells do not exhibit enlargement. However, in the case of ring-stadium vivax malaria, the infected red blood cells tend to manifest a bigger size compared to non-infected counterparts. Vacuoles are frequently observed in the Plasmodium ovale parasite, which causes malaria.

The trophozoite stage of the Plasmodium falciparum parasite exhibits a consistent ring-like morphology, characterized by a robust and compact structure. The quantity of chromatin pigment may also exhibit an augmentation, accompanied by a compact, amoeboid morphology. A Schuffner point is identifiable by its characteristic smooth and dark colour. Typically, red blood cells that have been infected tend to have a bigger size in comparison to their uninfected counterparts. The malaria parasite in question, during the trophozoite stage, has distinct traits such as a ribbon-like morphology, a cytoplasm that typically lacks vacuoles, and spherical chromatin. The Plasmodium ovale parasite exhibits numerous distinctive James spots.

The schizont stage of the malaria parasite Plasmodium malariae is characterized by the presence of merozoites, which are frequently found in irregular clusters. The malaria ovale parasite exhibits a slightly enlarged nucleus, with subsequent cytoplasm following each nucleus. The Plasmodium falciparum parasite harbors merozoites. The mature schizont stage typically encompasses around two-thirds of the red blood cells that have been infected. The schizont stage of vivax malaria is characterized by the presence of merozoites, each of which contains chromatin dots. During this phase, the red blood cells that have been infected exhibit an increase in size.

The gametocyte stage of the Plasmodium falciparum malaria parasite typically exhibits a crescent-shaped morphology. The cytoplasm of a female macrogametocyte has a darker and bluer hue, in contrast to the paler cytoplasm observed in a male microgametocyte. Vivax malaria often exhibits a spherical morphology and invades the host cell. Typically, red blood cells that are infected tend to have a bigger size in comparison to their normal counterparts. The cytoplasm exhibits a dark blue hue and is characterized by the presence of brown pigment. Scuffner points are observable by the application of

appropriate coloration techniques. Malaria caused by Plasmodium malariae does not result in the expansion of infected red blood cells. The cytoplasm exhibits a blue hue, while the chromatin displays a red coloration. The presence of ovale-shaped, round patches in cases of malaria has been observed by James.

RGB to HSV Color Space Transformation and Hue Components

HSV (Hue Saturation Value) colors are derived from RGB (Red Green Blue) colors, necessitating a color conversion procedure from RGB to HSV. The transformation from the RGB color system to the HSV color space involves converting the coordinates R, G, and B, which range from 0 to 1, into the corresponding values for hue, saturation, and value. In the RGB color space, the order of the primary colors is red, green, and blue, and the maximum value is represented as "max". Among the values, namely red, green, and blue, the minimal value, referred to as "min," is obtained by evaluating the color values associated with red, green, and blue. In order to obtain the accurate hue angle [0.360] within the HSV color space, it is recommended to refer to equations 1 through 3 (Kurniastuti, I., 2021).

$$H = \begin{cases} 0, \ if\ max = min \\ 60°x\left(\frac{(G-B)}{max-min} + 0\right), if\ max = R\ and\ G \geq B \\ 60°x\left(\frac{(G-B)}{max-min} + 360\right), if\ max = R\ and\ G < B \\ 60°x\left(\frac{(B-R)}{max-min} + 120\right), if\ max = G \\ 60°x\left(\frac{(R-G)}{max-min} + 240\right), if\ max = B \end{cases} \quad (1)$$

The values for S and V in HSV can be seen in equations 2 and 3.

$$S = \begin{cases} 0, \ if\ max = 0 \\ 1 - \frac{min}{max}, \ otherwise \end{cases} \quad (2)$$

$$V = max \quad (3)$$

The coordinates [0, 1] represent the values for red, green, and blue in the RGB color space. The variable "max" denotes the highest value among red, green, and blue, while "min" represents the lowest value among these colors. The aforementioned calculation yields values and saturation levels within the RGB range of [0.1] by initially multiplying by 255 to generate a value within the RGB range of [0.255].

Segmentation – Otsu Thresholding

The general approaches to segmentation can be grouped into three classes: pixel, regional and edge. Image segmentation is the process of separating objects where objects that are needed become the foreground and objects that are not needed become the background. The segmentation stage in this study was by dividing the network image into sixteen regions representing the types of malaria parasites and stages. The area of the type of malaria parasite and its stage refers to the object that characterizes the type of malaria parasite and its stage, while the area of type of non-malarial parasite refers to the background. In research conducted using the Otsu Thresholding algorithm for the segmentation process, the Otsu Thresholding equation can be seen in equation 4 (Suradi, S.H., 2021).

$$\begin{aligned} \sigma^2 &= X_{nw}(Y_{nw} - Y)^2 + X_w(Y_w - Y)^2 \\ Y &= X_{nw}.Y_{nw} + X_w.Y_w \\ X_{nw} &+ X_w = 1 \\ t^* &= \underset{a \leq t \leq b}{\text{Arg Max}}\{X_{nw}(Y_{nw} - Y)^2 + X_w.(Y_w - Y)^2 \end{aligned} \quad (4)$$

where,

σ^2 : Variants in class of malaria parasite species and stages

X_{nw} : Probability pixel values for non-malarial parasite classes

X_w : the probability of the pixel value for the malaria parasite class and its stage

Y_{nw} : the average pixel value of the non-malarial parasite class

Y_w : the average pixel value of the malaria parasite class and its stage

Y : the average pixel value of the image

t^* : threshold value

Otsu thresholding is used because it is able to produce an optimum threshold so that it can distinguish objects of the type of malaria parasite and its stages in the foreground from the background.

Morphology-Closing and Opening Operations

The image of the malaria parasite that has been obtained and analyzed can be seen; in fact, there are several image objects of the type of malaria parasite after segmentation; there are missing pixels, so a closing operation is carried out, where the method of closing is to widen the object by performing a dilation operation, then smoothing the edges of the object, namely with erosion surgery. Because the shape of the malaria parasite tends to be round and some are elongated, the structural elements are square in shape.

In this study, an opening morphology operation was performed to remove small objects that were not of the malaria parasite type so that the image could smooth the boundaries of the malaria parasite object. This stage begins with calling the image resulting from the OTSU thresholding segmentation, then the morphological process is carried out using closing and opening. Then it displayed images of the results of morphological operations: closing and opening. Finally, save the image of the results of morphological operations: closing and opening.

Region of Interest (ROI)

The ROI stage of the closing and opening images is carried out by taking the closing and opening images and cropping them using the tools available in the Matlab programming language to obtain image objects of the type of malaria parasite and its stage. The results of the ROI process are resized to a size of 512x512 pixels. The ROI results are saved in a file with the bmp extension.

Feature Extraction

Feature extraction is featuring retrieval, where the value obtained will be analyzed for further processing. This study combines shape and texture feature extraction algorithms called "feature tracking" (PEMA and KEHE algorithms) with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy.

a. The perimeter is the length of the malaria parasite object and its stage. The perimeter equation can be seen in equation 5.

$$P_i = \sum_{x=0}^{R_i-1} \sum_{y=0}^{S_i-1} O_i(x,y), \forall O_i\,(\mathrm{x,y}) \in 4 - \mathrm{neighbour}(O_i(\mathrm{x,y})) = 0 \quad (5)$$

where R and S_i are based on the number of rows and columns, respectively. x,y image boundaries of a malaria parasite object, as well as the closing and opening stages .

b. Eccentricity is the comparison of minor elliptical foci with major elliptical foci in malaria parasite species. The eccentricity equation can be seen in equation 6.

$$E_i = \sqrt{1 - \frac{d_i}{c_i}} \quad (6)$$

where $c_i, dan\ d_i$ re the sizes of the major and minor axes of the O_i(x,y) object.

c. Metric is the ratio between the area and the circumference of an object's area. The metric equation can be seen in equation 7

$$M = \frac{4\mu xA}{Cf^2} \quad (7)$$

where M=metric, A=area, Cf = *circumference*

d. Area: the number of pixels in object boundary area. The area equation can be seen in equation 8.

$$A_i = \sum_{x=0}^{P_i-1} \sum_{y=0}^{Q_i-1} O_i(x,y) \quad (8)$$

where $P_i, dan\ Q_i$ each based on the number of rows and columns – i.

Morphological feature extraction begins with reading the image resulting from closing and opening, then calling the image resulting from closing and opening. Next calculate the perimeter, eccentricity, metric, area. After that, get the results of perimeter, eccentricity, metric, area. Finally, the results of perimeter, eccentricity, metric, area are stored.

The process of feature extraction from texture commences by invoking the resultant grayscale image and constructing a framework with dimensions of 256x256. Next, proceed to initialize the variable representing the texture feature. In addition, the co-occurrence matrix is constructed by populating the matrix with the count of spatial associations that are present. The matrix is subsequently subjected to a transposition operation to achieve a symmetrical angle of 180°. Subsequently, the co-occurrence matrices are aggregated to ensure symmetry, followed by normalization through summation of all symmetrical matrices and division by the total number of pixels within those matrices. In addition, the computation of contrast, energy, homogeneity, and entropy is carried out. The utilization of second-order statistical feature extraction demonstrates a high level of effectiveness in the representation of visual texture. This is achieved by the incorporation of measurable factors such as contrast, energy, homogeneity, and entropy (Althubiti, S.A., 2022). The utilized texture features encompass:

a. Contrast is a measure of the spatial frequency of an image and the difference in GLCM (Gray Level Co-Occurrence Matrix) moments. The difference in question is the difference in the high and low of a pixel. The contrast equation can be seen in equation 9.

$$C = \sum_{a=1}^{m} \sum_{b=1}^{n} P_{a,b}(a-b)^2 \quad (9)$$

b. Energy is a measure of textural uniformity. The energy value is high if the pixel values are similar to each other; otherwise, the value will be small, indicating that the value of the normalized GLCM is heterogeneous. The energy equation can be seen in equation 10.

$$E_1 = \sum_{c}^{m} \sum_{d}^{n} \frac{p\,(c,d)}{1+(c-d)^2} \quad (10)$$

where p is the normalized GLCM matrix, c is the row index of the E1 matrix, and d is the column index of the E1 matrix.

c. Homogeneity is measured by measuring image homogeneity. The sensitivity of this value is significantly influenced by the values near the major diagonal. A high value is observed when all pixels possess an identical and uniform value. The homogeneity equation can be seen in equation 11.

$$H = \sum_{c}^{m} \sum_{d}^{n} P(c,d)^2 \quad (11)$$

where p is the normalized GLCM matrix, c is the row index of the H matrix, and d is the column index of the H matrix.

d. Entropy is a measure of the complexity of an image. Entropy is high if the image is not uniform. The entropy equation can be seen in equation 12.

$$E_2 = -\sum_{c}^{m} \sum_{d}^{n} p(c,d)\log\{p(i,j)\} \quad (12)$$

where p is the normalized GLCM matrix, c is the row index of the E_2 matrix, and d is the column index of the E^2 matrix.

Finally, save the results of contrast, energy, homogeneity, and entropy.

Identification of Types of Malaria Parasites Along with Stage-comparison Classification Methods Using CNN and Multi-class SVM

The present study employs a neural network for the purpose of image classification. The neural network will analyze photos depicting various types of malaria parasites and their respective stages in order to determine the feasibility of identifying these parasites and stages using image-based testing. The purpose of training is to obtain weight values from the network. The weight acquired will be recorded, transformed into a numerical representation, and subsequently utilized in the testing phase of the test image. This study employed a neural network for the purpose of categorizing the test photos into several categories, namely pertaining to the type of malaria parasite and its corresponding stage.

Convolutional neural networks (CNNs) are a prominent category of neural networks that exhibit significant prowess in various computer vision applications, hence garnering considerable interest across multiple domains. CNNs, short for convolutional neural networks, are a mathematical framework that is frequently employed for the purpose of classification. The Convolutional Neural Network (CNN) is composed of interconnected neurons that have been organized in a manner that results in the creation of a filter with specific dimensions, namely length and height (Sultan, S., 2021). The Convolutional Neural Network (CNN) is an advancement of the multilayer perceptron that is designed for the processing of data using a neuron configuration that consists of three dimensions: width, height, and depth. The dimensions of the layers are represented by width and height, while the number of layers is denoted as depth (Matinkia, M., 2022). In a general context, layers can be categorized into two distinct categories, specifically (Niu, S., 2022):

1. The initial layer of the design is dedicated to extracting picture features. This layer consists of many sublayers, each comprising interconnected neurons that are receptive to the local region of the preceding layer. The initial layer is referred to as the convolutional layer, whereas the subsequent layer is denoted as the pooling layer. An activation function is assigned to each layer.
2. The categorization layer consists of many layers, with each layer including completely interconnected neurons. The responsibility of this layer is to receive input in the form of a vector from the output of the picture feature extraction layer. The architecture of the Convolutional Neural Network (CNN) is depicted in Figure 3 (Lina, Q., 2018).

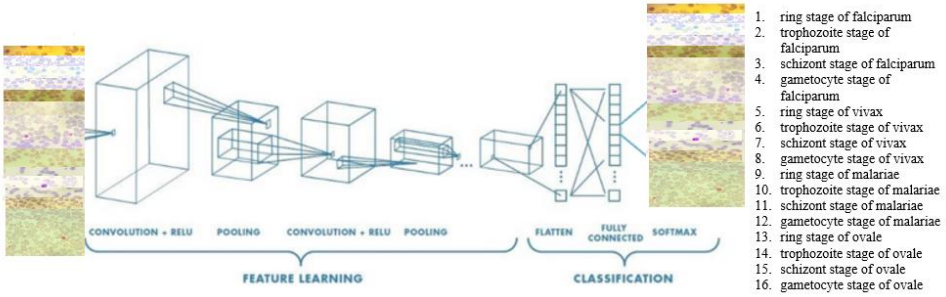


Figure 3: CNN Architecture – Performance Comparison of Classification Method Determining Malaria Parasites and Stages

CNN comprises 4 parts: input, feature extraction, classification, and output. The extraction phase of Convolutional Neural Networks (CNNs) encompasses multiple layers that are concealed, including convolution layers, rectification linear units (activation functions), and pooling. The output generated by the initial convolution layer serves as the input for the subsequent convolution layer. The classification process involves a set of interconnected nodes and an activation function, specifically the softmax function, which produces a classification outcome as the output (Prajapati, R., 2022).

The primary function of the convolution layer is to apply a filtering process that enables the extraction of objects that correspond to malaria parasites and their respective phases from the input image. Filters containing weighted values are employed for the purpose of identifying object attributes, specifically edges, curves, or colors. The process of convolution results in a linear transformation of the input image, which is determined by the spatial characteristics included in the data.

ReLU (Rectification Linear Unit) is an operation to introduce nonlinearity and improve the model representation. The ReLU activation function is f(x) = max (0, x) (Liu et al., 2022). The output value of the neuron is declared 0 if the input is negative. If the input value is positive, then the output of the neuron is the activation input value itself (Banihashemi, S.B., 2022).

There are two separate types of pooling that are commonly used, namely average pooling and maximal pooling. As stated in the referenced literature (Nandi, U., 2023), average pooling involves calculating the mean of the input values, while maximum pooling involves selecting the highest value. The present layer obtains input from the previous stage in order to determine the attributes that have the highest connection with a particular class. The fundamental objective of this layer is to facilitate the integration of distinct nodes into a unified dimension (Xiao, C., 2022).

The fully connected layer comprises a set of convolutional operations, wherein this layer receives input derived from preceding operations in order to ascertain the features that exhibit the highest correlation with a specific class. The purpose of this layer is to integrate all nodes into a single dimension (Liu, Z., 2022). The Softmax Activation function is employed to derive classification outcomes that yield unnormalized probabilities for each class, which are subsequently interpreted as probabilities. The initial phases of CNN operations commence with reference to (Hari, M. S., 2022):

1. Breaking the image into smaller images with the same convolution.
2. Subsequently, the feature representation is generated by utilizing each individual little image obtained through the convolution process as input. This allows CNN to accurately identify items, namely objects belonging to the malaria parasite category, and determine the specific stage of the object's location inside an image. The aforementioned procedure is executed on every component of each miniature representation, with an identical filter, so ensuring that each image is subjected to an equivalent scaling factor. If an object of importance is present in each picture, it is duly identified and tagged as such.
3. Save the results from each thumbnail into a new array.
4. During stage 3, the array continues to possess a substantial size. In order to decrease the dimensions of the array, a technique known as downsampling is employed. This downsampling technique, referred to as "max pooling," involves selecting the maximum pixel value from each pooling kernel. Although the number of parameters is reduced, the crucial information from that area is still retrieved.
5. In the context of making predictions, an array can be defined as a collection of integers. By employing this limited-sized array, it can be incorporated as an input for another neural network.

The ultimate neural network will determine the appropriateness of the image. At this juncture, the network is referred to as a "fully connected network".

The CNN model has the ability to classify image data with very good accuracy, so in this study the CNN model will be used in the identification of malaria parasites and their stages. The level of accuracy of the model produced by the switching process with SVM is very dependent on the kernel functions and parameters used. SVM consists of two types: linear and non-linear. Linear SVM is data that has been separated linearly, i.e., by separating the two classes with a soft margin on the hyperplane. Non-linear SVM applies the trick kernel function to a high-dimensional space.

Provide evidence or arguments in favor of a particular claim or position. The Vector Machine is a machine learning system that utilizes a hypothetical space represented by a linear function in a feature space with a high number of dimensions. It is learned using a learning method rooted in optimization theory. The initial Support Vector Machine (SVM) proposed by Vapnik is limited to binary classification, where data may only be categorized into two distinct classes. Additional investigation is required to advance the development of Support Vector Machines (SVM) in order to enhance its capability to categorize data across several classes, beyond the current limitation of binary classification. There exist two approaches for developing multi-class Support Vector Machines (SVMs): the first involves integrating multiple binary SVMs, while the second involves aggregating all data points from multiple classes into a unified optimization problem. The multi-class SVM method employs the one-against-all approach, wherein k binary SVM models are constructed, with k representing the total number of classes. The classification model for each individual classification (indexed by i) is trained using the entire dataset in order to identify effective solutions for the given challenges. For example, research conducted with 16 classes can be seen in Figure 4 (Heaton, J., 2015).

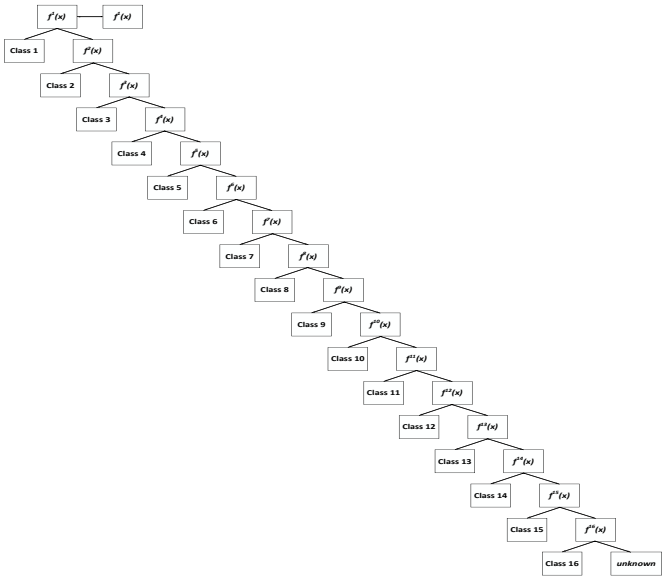


Figure 4: Example of "One Against All" Classification for 16 Classes

Accuracy

Accuracy is the amount of data that is correctly predicted by the model, both negative and positive (Kaur, D., 2020). The calculation is calculated by equation 13.

$$\frac{TP+TN}{TP+TN+FP+FN}\% \quad (13)$$

where,

A True Positive (TP) refers to positive data that is accurately identified by the model.

A false positive (FP) refers to the occurrence of positive data being mistakenly identified by a model.

A true negative (TN) refers to negative data that has been accurately classified as negative by the model.

A false negative (FN) refers to data that is classified as negative by a model, but is actually positive.

3 Result and Discussion

The model that has been designed in point 2 is the research method; the following are the results of the method and discussion at each step.

Image Collection

This stage uses image data on the types of malaria parasites and their stages. The number of images obtained so far is 446 images consisting of 332 images of malaria parasites and the stages of the training process and 114 images of malaria parasites and the stages of the testing process. examination using a digital microscope. Immersion oil is useful for clarifying an object. Furthermore, the preparations were examined using a digital microscope. Then they labeled several types of malaria parasites. Data for malaria parasite preparations were sourced from the Health Laboratory of North Sumatra Province and Giemsa staining was carried out on the preparations and then dripped with immersion oil, where the immersion oil functions to sharpen or clarify an object. The image is then captured with a digital microscope, magnification 1000 times. Images are resized to speed up computing.

Results of the Transformation of the RGB Color Space to HSV and the Hue Component

The current step exhibits the resultant output of the conversion process from the RGB color space to the HSV color space, as well as the specific Hue component, subsequent to the acquisition of parasitic photos. The hue represents the inherent color of an object, while saturation refers to the degree of purity or intensity of that color. On the other hand, value pertains to the level of brightness exhibited by the color. One of the benefits associated with utilizing the HSV color space is its ability to represent colors that closely resemble those perceived by the human visual system. Figure 5 displays the outcomes of converting the RGB color space into HSV and hue components.

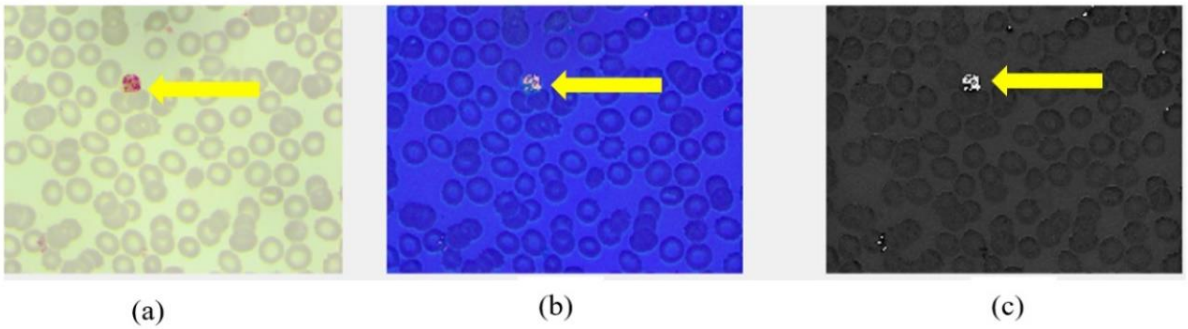
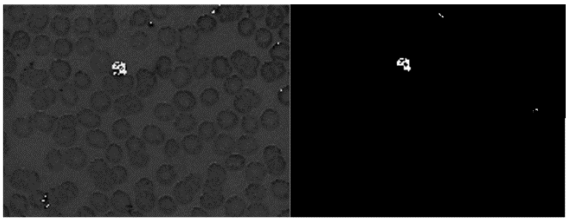


Figure 5: Color Space Transformation Results (a) An RGB Image; (b) An Image Converted from RGB to HSV; (c) An Image Resulting from the Hue Component

Figure 5 (a) illustrates an RGB image that has undergone a conversion process to the HSV color space, as seen in Figure 5 (b). In the process of converting the RGB color space to HSV, it is reasonable to assume that the coordinates R, G, and B (with values of either 0 or 1) represent the red, green, and blue components in the RGB color space. The variable "max" represents the maximum value among the red, green, and blue components, while "min" represents the minimum value among these components. The hue component of the image obtained after converting the color space to HSV is utilized because to its ability to visually differentiate between the object and the backdrop, among the three components of HSV. The hue component is derived from the HSV image. The outcome of the conversion of the HSV color model, specifically the hue component, is observable as it enables differentiation between objects and backgrounds. The color component of the retrieved image will be kept and then utilized for the segmentation stage.

Segmentation Results-Otsu Thresholding

The segmentation stage is responsible for the separation of the necessary objects from the backdrop. The segmentation process involves the differentiation of various things, including malaria parasites of different types and stages, as well as other entities such as platelets, normal red blood cells, white blood cells, and artifacts. The first step involves retrieving the image data obtained from the Hue component and subsequently applying Otsu thresholding for further processing. The utilization of Otsu thresholding is motivated by its ability to provide an optimal threshold that effectively discriminates between foreground items of the malaria parasite and its respective stage, and the background. Figure 6 displays the outcomes of segmentation achieved through the utilization of OTSU thresholding in conjunction with the HSV (Hue, Saturation, Value) staining approach.



(a) (b)

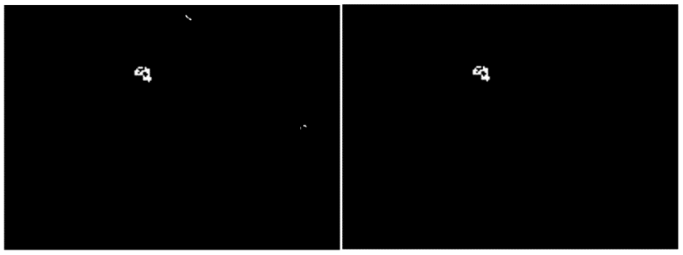
Figure 6: (a) Image Resulting from the Hue Component; (b) Image Resulting from Threshold Segmentation

The outcome of the hue component is depicted in Figure 6(a). The outcome of segmentation using OTSU thresholding is illustrated in Figure 6.b). One notable benefit of employing Otsu thresholding is its ability to automatically compute the optimal threshold value. This approach effectively maximizes the variance between distinct classes that are separated by the threshold. The primary objective of this technique is to accurately distinguish items belonging to the category of malaria parasites and their respective stages from the background (Li, L., 2022).

Results of Morphological Operations: Closing and Opening

The morphological analysis stage involves the application of closing and opening operations to the obtained image. Figure 7 displays the resulting image, wherein the objective of the morphological operation is to alter the shape of the object from its original form and eliminate extraneous elements present in the vicinity of the malaria parasite image object.

The image obtained after the segmentation procedure is subsequently subjected to the morphological analysis stage. Figure 7(a) illustrates the disparity observed between the image of the malaria parasite and the stadium when employing OTSU thresholding. Notably, the outcomes of this thresholding technique exhibit the presence of noise in the form of small white spots. Consequently, to mitigate this issue, noise elimination is performed through the utilization of mathematical morphology techniques such as closing and opening operations. Upon the completion of the mathematical morphological operations of closure and opening, the noise is effectively eliminated, as depicted in Figure 7 (b). An illustration of the outcomes of morphological examination for different classes and stages of malaria parasites may be observed in Figure 7 (b).



(a) (b)

Figure 7: (a) Image of the Segmentation Result of OTSU Thresholding for the Type of Malaria Parasite and Its Stage (b) Examples of Results of Closing and Opening Operations for Types of Malaria Parasites and their Stages

Furthermore, automatic cutting is carried out at the boundary of the type of malaria parasite and its stage, where it is cut from 446 image data points of the type of malaria parasite and its stage. Figure 7 (b) is an example of the results of automatic cutting samples on the class of malaria parasite species and their stages.

Region of Interest Results

The image of the type of malaria parasite and the stage resulting from the morphological operation that has gone through the process of removing noise will then proceed to the region of interest process, namely cropping the image of the type of malaria parasite and the stage. Images of the type of malaria parasite and its stage were taken for use in the next process. The cropped image of the type of malaria parasite and its stage can be seen in Figure 8.

(a) (b)

Figure 8: (a) Depicts the Results of Morphological Analysis of Different Types of Malaria Parasite Classes and Stages. (b) Cropping Result Sample of Malaria Parasite Class and Stage

Feature Extraction Results

The image of the type of malaria parasite and the stages that have been carried out through the cropping process are then subjected to feature extraction on the image of the type of malaria parasite and the stage. The process of extracting texture features through the utilization of the Gray Level Co-Occurrence Matrix (GLCM) involves the computation of the likelihood of adjacent associations within a specified proximity. The selection of distance is a crucial component in the calculation of the Gray-Level Co-occurrence Matrix (GLCM). The present work used a spatial distance of 1 unit in conjunction with straightforward computational methods.

The feature extraction process for the type of malaria parasite and its stage by combining shape and texture feature extraction algorithms is called the PEMA and KEHE feature tracking algorithm, or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy. Examples of feature value results obtained from the results of the PEMA and KEHE algorithms for feature tracking or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy can be seen in Table 1. An example of the results of image feature values for malaria parasite types along with stadium.

Table 1: Example of Feature Value Results from the Results of the PEMA and KEHE Algorithms for Feature Tracking or Feature Tracking with Perimeter, Eccentricity, Metric, Area, Contrast, Energy, Homogeneity, and Entropy

No	Perimeter	Eccentricity	Metric	Area	Contrast	Energy	Homogeneity	Entropy
1	1,73286E+14	0.80373322	0.108806	260	0.072413(	0.631279087	0.9637931958	0.786853954
2	1,12914E+14	0.83174639	0.113347	115	0.097025[	0.502854800	0.9514871252	1,062E+14
3	3,1303E+14	0.90190418	0.718167	56	0.238464[	0.276799740	0.8807679063	1,67566E+14
4	4,7777E+14	0.89071656	0.429404	78	0.213110[	0.324214213	0.8934449404	1,46737E+14
5	2,5282E+14	0.65476375	1E+14	51	0.443303[	0.169249135	0.7878720238	1,96381E+14
6	1,11185E+14	0.82840081	0.341551	336	0.306019[	0.139057538	0.8557419334	2,2331E+14
7	7,0999E+14	0.83865443	0.416314	167	0.180382(	0.205687078	0.9098086734	1,7839E+14
8	1,1217E+14	0.92201659	0.098876	99	0.063592[	0.629047635	0.9682037193	0.7734749103
9	4,2699E+14	0.65204753	0.992514	144	0.666472(	0.103127943	0.7581101190	2,6067E+14
10	9,8614E+14	0.95025442	0.286870	222	0.301648(	0.139389316	0.8605719768	2,32768E+14
11	3,0906E+14	0.56895230	1,1E+14	82	0.645745(	0.128183728	0.7613187160	2,29732E+14
12	2,38811E+14	0.78354533	0.221225	1004	0.105562(	0.294131418	0.9472189658	1,63399E+14
13	8,9219E+14	0.17625204	0.704092	446	0.284698[	0.156287300	0.8665452244	2,104E+14
14	7,1769E+14	0.88073785	0.380592	156	0.327115(	0.163292758	0.8497000455	2,24969E+14
15	2,7432E+14	0.66702578	1,2E+14	71	0.399621[	0.209137071	0.8077651515	1,89785E+14
16	1,12293E+14	0.90123607	0.296975	298	0.395849[	0.131971767	0.8331856577	2,41885E+14

Table 1 shows an example of the results of image feature values for the type of malaria parasite and its stage using a shape and texture feature extraction algorithm called the PEMA and KEHE feature tracking algorithm or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy. The number of images is 446 with details of 332 images of the training process and 114 images of the testing process.

The Results of Identifying the Type of Malaria Parasite and Its Stage by Developing a Feature Extraction Algorithm and a Comparison of Classification Methods Using CNN and Multi-class SVM

In the identification stage of the malaria parasite and its stadium, there are two processes carried out, namely the training and testing processes. The training phase uses processed data from feature extraction. training to identify types of malaria parasites and their stages using the convolutional neural network (CNN). CNN was applied to 446 images, consisting of 332 images identifying the type of malaria parasite and its stages for the training process and 114 images identifying the type of malaria

parasite and its stages for the testing process. Before starting training on the CNN network, the parameters and components that will be used in building the network must first be determined. The parameters used in network development include a batch size of 128, a learning rate of 0.1, and a momentum of 0.9. The results of the training can be seen in Table 2.

Table 2: Shows the Training Results for the PEMA and KEHE Algorithm Feature Tracking or Feature Tracking with Perimeter, Eccentricity, Metric, Area, Contrast, Energy, Homogeneity, and Entropy using CNN

Epoch	Classification Method	Number of types of identification of the type of malaria parasite along with its stage	Feature extraction type	Accuracy (%)
100	CNN	16	perimeter, entropy, metric, area, contrast, energy, homogeneity, eccentricity.	92,36

Table 2 shows the results of the training to develop a feature extraction algorithm called the PEMA and KEHE algorithm feature tracking or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy using CNN as a classification method and obtained using epoch 100 to get an accuracy value of 92.36%.

Training to identify types of malaria parasites and their stages using the Support Vector Machine (SVM) SVM was applied to 446 images, consisting of 332 images for the identification of malaria parasite types and stages for the training process and 114 images for the identification of malaria parasite types and stages for the testing process. Before starting training on the SVM network, the parameters and components that will be used in building the network must first be determined. The SVM model uses a polynomial kernel. The results of the training can be seen in Table 3.

Table 3: Shows the Training Results for the PEMA & KEHE Algorithm Feature Tracking or Feature Tracking with Perimeter, Eccentricity, Metric, Area, Contrast, Energy, Homogeneity, and Entropy using SVM

No	Kernel Type	Classification Method	Number of Types of Identification of the Type of Malaria Parasite along with its Stage	Features Extraction Type	Accuration (%)
1	Polynomial	SVM	16	Perimeter, Entropy, Metric, Area, Contrast, Energy, Homogeneity, Eccentricity.	93,61

The test aims to test the results of the training because the CNN network is smarter at training. CNN is a very powerful extractor for extracting features with weights that meet the requirements that have been trained so that if the network is given certain inputs, it will produce the expected output.

In the testing process, the data used amounted to 114 test results. The parameters used were a batch size of 128; a learning rate of 0.1; and a momentum of 0.9. The results of testing 114 types of malaria parasites and their stages can be seen in Table 4.

Table 4: Show the Test Results for Developing the PEMA and KEHE Algorithm Feature Tracking or Feature Tracking with Perimeter, Eccentricity, Metric, Area, Contrast, Energy, Homogeneity, and Entropy Using CNN

No	Epoch	File Name	Identification Result	Test Result (✓or ×)
1	100	gamet-falci-thin54	gametocyte stage of falciparum	✓
2	100	gamet-falci-thin55	gametocyte stage of falciparum	✓
3	100	gamet-falci-thin56	gametocyte stage of falciparum	✓
4	100	gamet-falci-thin57	gametocyte stage of falciparum	✓
5	100	gamet-falci-thin58	gametocyte stage of falciparum	✓
6	100	gamet-falci-thin59	gametocyte stage of falciparum	✓
7	100	gamet-falci-thin60	gametocyte stage of falciparum	✓
8	100	gamet-falci-thin61	gametocyte stage of falciparum	✓
9	100	gamet-falci-thin62	gametocyte stage of falciparum	✓
10	100	gamet-falci-thin63	gametocyte stage of falciparum	✓
11	100	gamet-falci-thin64	gametocyte stage of falciparum	✓
12	100	gamet-falci-thin65	gametocyte stage of falciparum	✓
13	100	gamet-falci-thin66	gametocyte stage of falciparum	✓
14	100	gamet-falci-thin67	gametocyte stage of falciparum	✓
15	100	gamet-falci-thin68	gametocyte stage of falciparum	✓
16	100	gamet-falci-thin69	gametocyte stage of falciparum	✓
17	100	gamet-falci-thin70	gametocyte stage of falciparum	✓
18	100	gamet-falci-thin71	gametocyte stage of falciparum	✓
19	100	gamet-falci-thin72	gametocyte stage of falciparum	✓
20	100	gamet-falci-thin73	gametocyte stage of falciparum	✓
21	100	gamet-falci-thin74	gametocyte stage of falciparum	✓
22	100	gamet-falci-thin75	gametocyte stage of falciparum	✓
23	100	gamet-falci-thin76	gametocyte stage of falciparum	✓
24	100	gamet-falci-thin77	gametocyte stage of falciparum	✓
25	100	gamet-falci-thin78	gametocyte stage of falciparum	✓
26	100	gamet-falci-thin79	gametocyte stage of falciparum	✓
27	100	gamet-falci-thin80	gametocyte stage of falciparum	✓
28	100	gamet-falci-thin81	gametocyte stage of falciparum	✓
29	100	ring-falci-thin15	gametocyte stage of falciparum	×
30	100	ring-falci-thin19	ring stage of vivax	×
31	100	ring-falci-thin20	gametocyte stage of malariae	×
32	100	schizon falci 15	trophozoite stage of ovale	×
33	100	schizon falci 16	trophozoite stage of ovale	×
34	100	troporing-falci-thin24	gametocyte stage of falciparum	×
35	100	troporing-falci-thin29	gametocyte stage of falciparum	×
36	100	gamet-malariae-thin19	gametocyte stage of malariae	✓
37	100	gamet-malariae-thin20	gametocyte stage of malariae	✓
38	100	gamet-malariae-thin21	gametocyte stage of malariae	✓
39	100	gamet-malariae-thin22	gametocyte stage of malariae	✓
40	100	malariae ring 7	gametocyte stage of malariae	×
41	100	malariae ring 12	gametocyte stage of falciparum	×
42	100	malariae schizon17	schizont stage of malariae	✓
43	100	malariae schizon18	schizont stage of malariae	✓

44	100	malariae schizon19	gametocyte stage of falciparum	×
45	100	malariae schizon20	gametocyte stage of falciparum	×
46	100	malariae schizon21	gametocyte stage of falciparum	×
47	100	tropo-malariae-thin33	gametocyte stage of falciparum	×
48	100	tropo-malariae-thin40	gametocyte stage of malariae	×
49	100	tropo-malariae-thin41	gametocyte stage of falciparum	×
50	100	tropo-malariae-thin44	gametocyte stage of falciparum	×
51	100	tropo-malariae-thin46	gametocyte stage of falciparum	×
52	100	gametosit6 ovale	trophozoite stage of ovale	×
53	100	gametosit8 ovale	gametocyte stage of ovale	✓
54	100	gametosit9 ovale	gametocyte stage of ovale	✓
55	100	gametosit10 ovale	gametocyte stage of ovale	✓
56	100	gametosit11 ovale	ring stage of vivax	×
57	100	ovale-ring1	gametocyte stage of falciparum	×
58	100	ovale-ring2	gametocyte stage of falciparum	×
59	100	ovale-ring3	gametocyte stage of falciparum	×
60	100	schizon10 ovale	trophozoite stage of ovale	×
61	100	schizon11 ovale	trophozoite stage of ovale	×
62	100	tropozoit9 ovale	trophozoite stage of ovale	✓
63	100	tropozoit10 ovale	trophozoite stage of ovale	✓
64	100	tropozoit11 ovale	trophozoite stage of ovale	✓
65	100	tropozoit13 ovale	trophozoite stage of ovale	✓
66	100	tropozoit14 ovale	trophozoite stage of ovale	✓
67	100	tropozoit15 ovale	trophozoite stage of ovale	✓
68	100	tropozoit16 ovale	trophozoite stage of ovale	✓
69	100	tropozoit18 ovale	trophozoite stage of ovale	✓
70	100	tropozoit19 ovale	trophozoite stage of ovale	✓
71	100	tropozoit20 ovale	trophozoite stage of ovale	✓
72	100	tropozoit21 ovale	trophozoite stage of ovale	✓
73	100	tropozoit28 ovale	ring stage of vivax	×
74	100	tropozoit29 ovale	gametocyte stage of ovale	×
75	100	gamet-vivax-thin7	gametocyte stage of ovale	×
76	100	gamet-vivax-thin8	schizont stage of falciparum	×
77	100	gamet-vivax-thin10	trophozoite stage of ovale	×
78	100	gamet-vivax-thin13	Malariae Gametosit	×
79	100	gamet-vivax-thin17	gametocyte stage of vivax	✓
80	100	gamet-vivax-thin19	trophozoite stage of ovale	×
81	100	gamet-vivax-thin21	gametocyte stage of vivax	✓
82	100	gamet-vivax-thin23	gametocyte stage of malariae	×
83	100	Vivax cincin2	ring stage of vivax	✓
84	100	Vivax cincin3	ring stage of vivax	✓
85	100	Vivax cincin4	ring stage of vivax	✓
86	100	Vivax cincin5	ring stage of vivax	✓
87	100	Vivax cincin6	ring stage of vivax	✓
88	100	Vivax cincin7	ring stage of vivax	✓
89	100	Vivax cincin8	ring stage of vivax	✓
90	100	Vivax cincin9-10	ring stage of vivax	✓
91	100	Vivax cincin11-12	ring stage of vivax	✓

92	100	Vivax cincin13-14	ring stage of vivax	✓
93	100	Vivax cincin16	ring stage of vivax	✓
94	100	Vivax cincin17	ring stage of vivax	✓
95	100	Vivax cincin18	ring stage of vivax	✓
96	100	schizon-vivax-thin7	trophozoite stage of ovale	×
97	100	schizon-vivax-thin12	trophozoite stage of ovale	×
98	100	schizon-vivax-thin14	gametocyte stage of vivax	×
99	100	schizon-vivax-thin15	trophozoite stage of ovale	×
100	100	schizon-vivax-thin16	trophozoite stage of ovale	×
101	100	shizon-vivax-24-thin13	trophozoite stage of ovale	×
102	100	vivax tropozoit2	ring stage of vivax	×
103	100	vivax tropozoit3	schizont stage of falciparum	×
104	100	vivax tropozoit4	trophozoite stage of vivax	✓
105	100	vivax tropozoit5-6	trophozoite stage of vivax	✓
106	100	vivax tropozoit9-10	trophozoite stage of vivax	✓
107	100	vivax tropozoit11	trophozoite stage of vivax	✓
108	100	vivax tropozoit13	trophozoite stage of vivax	×
109	100	vivax tropozoit14	trophozoite stage of vivax	✓
110	100	vivax tropozoit16	ring stage of vivax	×
111	100	vivax tropozoit17	Falciparum Skizon	×
112	100	vivax tropozoit19	trophozoite stage of vivax	✓
113	100	vivax tropozoit21	ring stage of vivax	×
114	100	vivax tropozoit22	schizont stage of malariae	×

Table 4. This can be seen from the results of identifying types of malaria parasites and their stages (vivax, falciparum, malariae, and ovale with ring stages, trophozoites, schizonts, and gametocytes). Using CNN in the testing process, 70 images can be correctly grouped from 114 images with an accuracy value of 61.40%. Testing using SVM aims to test the results of training because SVM is better at training. In the testing process, the data used amounted to 114 types of malaria parasites and their stages, as seen in Table 5.

Table 5: Test Results for Developing a Feature Extraction Algorithm Called the PEMA & KEHE Algorithm Feature Tracking or Feature Tracking with Perimeter, Eccentricity, Metric, Area, Contrast, Energy, Homogeneity, and Entropy using SVM

No	Epoch	File Name	Identification Result	Test Result (✓or ×)
1	100	gamet-falci-thin54	gametocyte stage of falciparum	✓
2	100	gamet-falci-thin55	gametocyte stage of falciparum	✓
3	100	gamet-falci-thin56	gametocyte stage of falciparum	✓
4	100	gamet-falci-thin57	gametocyte stage of falciparum	✓
5	100	gamet-falci-thin58	gametocyte stage of falciparum	✓
6	100	gamet-falci-thin59	gametocyte stage of falciparum	✓
7	100	gamet-falci-thin60	gametocyte stage of falciparum	✓
8	100	gamet-falci-thin61	gametocyte stage of falciparum	✓
9	100	gamet-falci-thin62	gametocyte stage of falciparum	✓
10	100	gamet-falci-thin63	gametocyte stage of falciparum	✓
11	100	gamet-falci-thin64	gametocyte stage of falciparum	✓
12	100	gamet-falci-thin65	gametocyte stage of falciparum	✓
13	100	gamet-falci-thin66	gametocyte stage of falciparum	✓

14	100	gamet-falci-thin67	gametocyte stage of falciparum	✓
15	100	gamet-falci-thin68	gametocyte stage of falciparum	✓
16	100	gamet-falci-thin69	gametocyte stage of falciparum	✓
17	100	gamet-falci-thin70	gametocyte stage of falciparum	✓
18	100	gamet-falci-thin71	gametocyte stage of falciparum	✓
19	100	gamet-falci-thin72	gametocyte stage of falciparum	✓
20	100	gamet-falci-thin73	gametocyte stage of falciparum	✓
21	100	gamet-falci-thin74	gametocyte stage of falciparum	✓
22	100	gamet-falci-thin75	gametocyte stage of falciparum	✓
23	100	gamet-falci-thin76	gametocyte stage of falciparum	✓
24	100	gamet-falci-thin77	gametocyte stage of falciparum	✓
25	100	gamet-falci-thin78	gametocyte stage of falciparum	✓
26	100	gamet-falci-thin79	gametocyte stage of falciparum	✓
27	100	gamet-falci-thin80	gametocyte stage of falciparum	✓
28	100	gamet-falci-thin81	gametocyte stage of falciparum	✓
29	100	ring-falci-thin15	ring stage of falciparum	✓
30	100	ring-falci-thin19	ring stage of vivax	×
31	100	ring-falci-thin20	gametocyte stage of malariae	×
32	100	schizon falci 15	gametocyte stage of vivax	×
33	100	schizon falci 16	schizont stage of vivax	×
34	100	troporing-falci-thin24	trophozoite stage of falciparum	✓
35	100	troporing-falci-thin29	gametocyte stage of falciparum	×
36	100	gamet-malariae-thin19	gametocyte stage of malariae	✓
37	100	gamet-malariae-thin20	gametocyte stage of malariae	✓
38	100	gamet-malariae-thin21	gametocyte stage of malariae	✓
39	100	gamet-malariae-thin22	gametocyte stage of malariae	✓
40	100	malariae ring 7	gametocyte stage of malariae	×
41	100	malariae ring 12	gametocyte stage of falciparum	×
42	100	malariae schizon17	schizont stage of malariae	✓
43	100	malariae schizon18	schizont stage of malariae	✓
44	100	malariae schizon19	schizont stage of malariae	✓
45	100	malariae schizon20	schizont stage of malariae	✓
46	100	malariae schizon21	schizont stage of malariae	✓
47	100	tropo-malariae-thin33	trophozoite stage of malariae	✓
48	100	tropo-malariae-thin40	trophozoite stage of malariae	✓
49	100	tropo-malariae-thin41	gametocyte stage of falciparum	×
50	100	tropo-malariae-thin44	gametocyte stage of falciparum	×
51	100	tropo-malariae-thin46	trophozoite stage of malariae	✓
52	100	gametosit6 ovale	gametocyte stage of ovale	✓
53	100	gametosit8 ovale	gametocyte stage of ovale	✓
54	100	gametosit9 ovale	gametocyte stage of ovale	✓
55	100	gametosit10 ovale	gametocyte stage of ovale	✓
56	100	gametosit11 ovale	gametocyte stage of ovale	✓
57	100	ovale-ring1	gametocyte stage of malariae	×
58	100	ovale-ring2	gametocyte stage of malariae	×
59	100	ovale-ring3	gametocyte stage of malariae	×
60	100	schizon10 ovale	gametocyte stage of ovale	×
61	100	schizon11 ovale	trophozoite stage of ovale	×

62	100	tropozoit9 ovale	trophozoite stage of ovale	✓
63	100	tropozoit10 ovale	trophozoite stage of ovale	✓
64	100	tropozoit11 ovale	trophozoite stage of ovale	✓
65	100	tropozoit13 ovale	trophozoite stage of ovale	✓
66	100	tropozoit14 ovale	trophozoite stage of ovale	✓
67	100	tropozoit15 ovale	trophozoite stage of ovale	✓
68	100	tropozoit16 ovale	trophozoite stage of ovale	✓
69	100	tropozoit18 ovale	trophozoite stage of ovale	✓
70	100	tropozoit19 ovale	trophozoite stage of ovale	✓
71	100	tropozoit20 ovale	trophozoite stage of ovale	✓
72	100	tropozoit21 ovale	trophozoite stage of ovale	✓
73	100	tropozoit28 ovale	trophozoite stage of ovale	✓
74	100	tropozoit29 ovale	trophozoite stage of ovale	✓
75	100	gamet-vivax-thin7	gametocyte stage of vivax	✓
76	100	gamet-vivax-thin8	ring stage of vivax	×
77	100	gamet-vivax-thin10	gametocyte stage of vivax	✓
78	100	gamet-vivax-thin13	gametocyte stage of vivax	✓
79	100	gamet-vivax-thin17	gametocyte stage of vivax	✓
80	100	gamet-vivax-thin19	gametocyte stage of vivax	✓
81	100	gamet-vivax-thin21	gametocyte stage of vivax	✓
82	100	gamet-vivax-thin23	gametocyte stage of vivax	✓
83	100	Vivax cincin2	ring stage of vivax	✓
84	100	Vivax cincin3	ring stage of vivax	✓
85	100	Vivax cincin4	ring stage of vivax	✓
86	100	Vivax cincin5	ring stage of vivax	✓
87	100	Vivax cincin6	ring stage of vivax	✓
88	100	Vivax cincin7	ring stage of vivax	✓
89	100	Vivax cincin8	ring stage of vivax	✓
90	100	Vivax cincin9-10	ring stage of vivax	✓
91	100	Vivax cincin11-12	ring stage of vivax	✓
92	100	Vivax cincin13-14	ring stage of vivax	✓
93	100	Vivax cincin16	ring stage of vivax	✓
94	100	Vivax cincin17	ring stage of vivax	✓
95	100	Vivax cincin18	ring stage of vivax	✓
96	100	schizon-vivax-thin7	schizont stage of vivax	✓
97	100	schizon-vivax-thin12	gametocyte stage of vivax	×
98	100	schizon-vivax-thin14	schizont stage of vivax	✓
99	100	schizon-vivax-thin15	schizont stage of vivax	✓
100	100	schizon-vivax-thin16	schizont stage of vivax	✓
101	100	shizon-vivax-24-thin13	schizont stage of vivax	✓
102	100	vivax tropozoit2	ring stage of vivax	×
103	100	vivax tropozoit3	trophozoite stage of vivax	✓
104	100	vivax tropozoit4	trophozoite stage of vivax	✓
105	100	vivax tropozoit5-6	trophozoite stage of vivax	✓
106	100	vivax tropozoit9-10	trophozoite stage of vivax	✓
107	100	vivax tropozoit11	trophozoite stage of vivax	✓
108	100	vivax tropozoit13	trophozoite stage of vivax	✓
109	100	vivax tropozoit14	trophozoite stage of vivax	✓

110	100	vivax tropozoit16	trophozoite stage of vivax	✓
111	100	vivax tropozoit17	trophozoite stage of vivax	✓
112	100	vivax tropozoit19	trophozoite stage of vivax	✓
113	100	vivax tropozoit21	trophozoite stage of vivax	✓
114	100	vivax tropozoit22	trophozoite stage of vivax	✓

Table 5 It can be seen that the results of identifying the types of malaria parasites and their stages show that 97 images can be correctly grouped from 114 images with an accuracy value of 85.08% using SVM in the testing process. The results of the comparison of values in the training and testing processes in this study using two classification methods and the parameters used to identify the type of malaria parasite and its stage can be seen in Table 6.

Table 6: The Results of the Accuracy Values at the Training and Testing Stages with the CNN and SVM Methods

No	Epoch	Classification Method	Kernel Type	Training Result (%)	Testing Result (%)
1	100	CNN	-	92,36	61,40
2	-	SVM	Polynomial	93,61	85,08

Table 6 shows that the two classification methods used, namely CNN and SVM, for the identification of malaria parasites and their stages produced the best accuracy value using the SVM method with an accuracy value of 93.61% in the training process and 85.08% in the testing process, compared to using the CNN method with an accuracy value of 92.36% in the training process and 61.40% in the testing process.

4 Conclusion

Based on the previous explanation, it is concluded that the identification of malaria parasites and their stages; developing a feature extraction method, namely, a feature extraction algorithm called the PEMA and KEHE feature tracking algorithm, or feature tracking with perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy. Then get the results in the form of values from the features perimeter, eccentricity, metric, area, contrast, energy, homogeneity, and entropy.

Malaria parasite and stage identification using CNN and SVM classification methods, specifically comparing the two methods to produce a higher accuracy value. The results of identifying malaria parasites and their stages by developing a feature extraction method, namely developing a feature extraction algorithm called the PEMA and KEHE feature tracking algorithm, or feature tracking with perimeter, eccentricity, metric, area and contrast, energy, homogeneity, and entropy using SVM, get accuracy values better with a value of 93.61% in the training process and 85.08% in the testing process, while CNN gets accuracy values of 92.36% in the training process and 61.40% in the testing process.

The next stage is to develop feature extraction by analyzing the methods used on shape, texture, and color texture features in order to obtain a high accuracy value.

References

[1] Althubiti, S.A., Paul, S., Mohanty, R., Mohanty, S.N., Alenezi, F., & Polat, K. (2022). Ensemble learning framework with GLCM texture extraction for early detection of lung cancer on CT images. *Computational and Mathematical Methods in Medicine*, *2022.*

[2] Annaldas, S., & Shirgan, S.S. (2015). Automatic diagnosis of Malaria parasites using neural network and support vector machine. *International Journal of Advanced Foundation in Computer (IJAFRC)*, *2*, 62-65.

[3] Arnani, M., (2022). Penyakit Malaria: Spesies Parasit, Gejala dan Pencegahannya, https://www.kompas.com/sains/read/2022/02/12/090100223/penyakit-malaria--spesies-parasit-gejala-dan-pencegahannya?page=all, tanggal akses 23 April 2022
[4] Banihashemi, S.B., & Akhtarkavan, E. (2022). Encrypted Network Traffic Classification Using Deep Learning Method. *In IEEE 8th International Conference on Web Research (ICWR)*, 1-8.
[5] Biernath, A., (2022). Covid: Apakah dari Pandemi, virus corona akan menjadi endemic, dan apa artinya?, https://www.bbc.com/indonesia/majalah-60414985, tanggal akses 23 April 2022
[6] Centers for Disease Control and Prevention. (2009). CDC health information for international travel 2010. Atlanta: US Department of Health and Human Services. *Public Health Service*, 1-4.
[7] Fuhad, K.F., Tuba, J.F., Rahman, T., & Mohammed, N. (2020). CNN Based Model for Malaria Diagnosis with Knowledge Distillation. *In Proceedings of the 4th International Conference on Digital Signal Processing*, 131-135.
[8] Garcia, R.G., Paglinawan, C.C., Paglinawan, A.C., Gatus, H.M.F., Mallari, J.E.P., Maniacup, M. A.G., & Saguiguit, A.A. (2019). Modification of Optical Compound Microscope to Detect Malaria Blood Parasites Using Wright-Giemsa Staining Process with SURF Based Stitching. *In Proceedings of the 9th International Conference on Biomedical Engineering and Technology*, 125-129.
[9] Hari, M. S., Manimegalai, R., & Sheeba, A. (2022). Optical Character Recognition using Deep Learning Techniques. In *Applications of Machine intelligence in Engineering* (pp. 117-129). CRC Press.
[10] Heaton, J. (2015). *AIFH, volume 3: deep learning and neural networks*, 228-238. Heaton Research.
[11] Huq, A., & Pervin, M.T. (2020). Robust deep neural network model for identification of malaria parasites in cell images. *In IEEE Region 10 Symposium (TENSYMP)*, 1456-1459.
[12] Kanojia, M., Gandhi, N., Armstrong, L.J., & Pednekar, P. (2018). Automatic identification of malaria using image processing and artificial neural network. *In Intelligent Systems Design and Applications: 17th International Conference on Intelligent Systems Design and Applications (ISDA) held in Delhi, India,* 846-857. Springer International Publishing.
[13] Kaur, D., & Walia, G.K. (2020). A hybrid aco-svm approach for detecting and classifying malaria parasites. *Computational Network Application Tools for Performance Management*, 139-152.
[14] Kurniastuti, I., Wulan, T.D., & Andini, A. (2021). Color Feature Extraction of Fingernail Image based on HSV Color Space as Early Detection Risk of Diabetes Mellitus. *In IEEE International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE)*, 51-55.
[15] Li, L., Qiao, J., Yu, G., Wang, L., Li, H.Y., Liao, C., & Zhu, Z. (2022). Interpretable tree-based ensemble model for predicting beach water quality. *Water Research*, *211*.
[16] Liang, Z., Powell, A., Ersoy, I., Poostchi, M., Silamut, K., Palaniappan, K., & Thoma, G. (2016). CNN-based image analysis for malaria diagnosis. *In IEEE international conference on bioinformatics and biomedicine (BIBM)*, 493-496.
[17] Lina, Q., (2018). Implementasi Deep Learning Menggunakan Convolutional Neural Network untuk Klasifikasi Gambar (Mata Juling dan Mata Normal), https://medium.com/@16611110/implementasi-deep-learning-menggunakan-convolutional-neural-network-untuk-klasifikasi-gambar-mata-87dcc0ad26e0, tanggal akses 3 Januari 2023
[18] Liu, Z., Dong, A., Yu, J., Han, Y., Zhou, Y., & Zhao, K. (2022). Scene classification for remote sensing images with self-attention augmented CNN. *IET Image Processing*, *16*(11), 3085-3096.
[19] Matinkia, M., Sheykhinasab, A., Shojaei, S., Vojdani Tazeh Kand, A., Elmi, A., Bajolvand, M., & Mehrad, M. (2022). Developing a new model for drilling rate of penetration prediction using

convolutional neural network. *Arabian Journal for Science and Engineering*, *47*(9), 11953-11985.

[20] Mehrjou, A., Abbasian, T., & Izadi, M. (2013). Automatic malaria diagnosis system. *In IEEE First RSI/ISM International Conference on Robotics and Mechatronics (ICRoM)*, 205-211.

[21] Nandi, U., Ghorai, A., Singh, M.M., Changdar, C., Bhakta, S., & Kumar Pal, R. (2023). Indian sign language alphabet recognition system using CNN with diffGrad optimizer and stochastic pooling. *Multimedia Tools and Applications*, *82*(7), 9627-9648.

[22] Niu, S., & Srivastava, V. (2022). Simulation trained CNN for accurate embedded crack length, location, and orientation prediction from ultrasound measurements. *International Journal of Solids and Structures*, *242*.

[23] Phatak, S. (2021). *Morphological Classification of Subtypes of Volumetric Projection Neurons from Mouse Brain Scans* (Doctoral dissertation, Johns Hopkins University).

[24] Prajapati, R., & Kwon, G.R. (2022). A binary classifier using fully connected neural network for Alzheimer's disease classification. *Journal of Multimedia Information System*, *9*(1), 21-32.

[25] Puspa, A., (2022). Kasus Malaria di sejumlah Wilayah Alami Peningkatan, https://mediaindonesia.com/humaniora/469628/kasus-malaria-di-sejumlah-wilayah-alami-peningkatan, tanggal akses 23 April 2022

[26] Salman, R., & Banu, A.A. (2023). DeepQ Residue Analysis of Computer Vision Dataset using Support Vector Machine. *Journal of Internet Services and Information Security, 13*(1), 78-84.

[27] Savkare, S.S., & Narote, S.P. (2011). Automatic detection of malaria parasites for estimating parasitemia. *International Journal of Computer Science and Security (IJCSS)*, *5*(3), 310-315.

[28] Sultan, S., & Bekeneva, Y. (2021). A comparative analysis of a designed CNN and AlexNet for image classification on small datasets. In *International Symposium on Intelligent and Distributed Computing*, 441-446. Cham: Springer International Publishing.

[29] Suradi, S.H., Abdullah, K.A., & Isa, N.A.M. (2021). Breast lesions detection using FADHECAL and Multilevel Otsu Thresholding Segmentation in digital mammograms. *In International Conference on Medical and Biological Engineering*, 751-759. Cham: Springer International Publishing.

[30] Tek, F.B., Dempster, A.G., & Kale, I. (2006). Malaria parasite detection in peripheral blood images. BMVA.

[31] Tek, F.B., Dempster, A.G., & Kale, I. (2009). Computer vision for microscopy diagnosis of malaria. *Malaria journal*, *8*, 1-14.

[32] Xiao, C., Wang, X., Dou, H., Li, H., Lv, R., Wu, Y., & Zhai, R. (2022). Non-uniform synthetic aperture radiometer image reconstruction based on deep convolutional neural network. *Remote Sensing*, *14*(10), 1-16.

Authors Biography

Rika Rosnelly

Dr. Rika Rosnelly, S. Kom, M. Kom was born in Medan, North Sumatra on September 19, 1975. As for my Educational History, S1 at the STMIK Logika Medan graduated in 2005. S2 at Putera Indonesia University YPTK Padang graduated in 2008. S3 Computer Science, at Gadjah Mada University Yogyakarta graduated in 2017. Currently I am serving as a Permanent Lecturer at Universitas Potensi Utama. My research study in the field of Image Processing, Artificial Intelligence, Decision Support System. Email: rikarosnelly@gmail.com; rika@potensi-utama.ac.id, Orcid: https://orcid.org/0000-0002-0407-5160

Bob Subhan Riza

Dr. Bob Subhan Riza, ST, M. Kom was born in Medan, North Sumatra on August 5, 1970. As for my Educational History, - D3 at USU Medan Polytechnic graduated in 1993 - S1 at the University of North Sumatra, Medan graduated in 2000 - S2 at Putera Indonesia University YPTK Padang graduated in 2014. S3 Information Technology at Putera Indonesia University YPTK Padang graduated in 2023. Currently I am serving as a Permanent Lecturer at Universitas Potensi Utama. My research study in the field of Image Processing. Email: bob.potensi@gmail.com Orcid: https://orcid.org/0000-0001-6358-9412

S. Suparni

Suparni was born in Medan, North Sumatra, as for my Educational History, - S2 at USU Medan, 2009, Currently I am serving as a Permanent Lecturer at Health Polytechnic, Kemenkes Medan, North Sumatra Indonesia. Email: hajjahsuparni@gmail.com, Orcid:https://orcid.org/0009-0005-8152-9305