International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT2283111

Image Analysis for Detecting Malaria Cell Using Otsu Thresholding and Machine Learning Models

Miss. Spoorthi B[1], Dr. Aravinda C V[2]

*[1]Student, Department of Computer Science &Engineering, NMAM Institute of Technology affiliated to NITTE (Deemed to be University), Nitte, Karkala, Karnataka,India

[2]Associate Professor, Department of Computer Science &Engineering, NMAM Institute of Technology affiliated to NITTE (Deemed to be University), Nitte, Karkala, Karnataka, India

Article Info

Volume 8, Issue 3
Page Number : 453-470

Publication Issue :
May-June-2022

Article History
Accepted: 03 June 2022
Published: 15 June 2022

ABSTRACT

Motivation : Malaria is a dangerous disease that affects thousands of individuals each year all around the world. It can be fatal if not treated promptly. According to the most recent World Malaria Report from the World Health Organization, there would be 241 million malaria cases and 627 000 malaria deaths globally in 2020. Despite recent advances in malaria diagnosis, the microscopy approach remains the most widely used. Moreover, the efficiency of microscopic diagnosis is dependent on the expertise of the microscopist, which restricts malaria throughput. Distinguishing parasite development phases remains a very challenging task. **Goal:** The main aim is to develop a system to identify malaria stages in blood smears using machine learning models. This paper proposes a study of seven machine learning models and one ensemble model to foresee which model will better predict the malaria stage. **Results:** To avoid a large number of individuals from being infected with malaria, an early and precise diagnosis is essential. A web-based application is developed for the end-user using a flask, where the user can upload the sample images of the multi-stage malaria parasite and recognize the cell image. This will help the doctors to take the necessary steps to prevent the disease and choose the appropriate solution.

Keywords: Malaria, Infected Cell, Stages, Uninfected Cell

I. INTRODUCTION

Today, Cognitive computing replicates how people solve problems while artificial intelligence and machine learning approaches seek to provide new ways to address problems that people could be better served. In recent decades a lot of study has been done by means of cost-effective solutions utilizing machine learning algorithms to help health professionals in reducing diseases. Image processing techniques play a significant role in medical image interpretation and automated diagnosis.

Malaria is a usually dangerous parasite illness caused by female Anopheles mosquito bites. The majority of deaths occur in Africa, where a child dies from malaria nearly every minute, and malaria is a

leading cause of Neuro-disability among children. [14]. Infants, children under the age of five, pregnant women, HIV/AIDS patients, and others with weakened immune systems who go to malaria-endemic areas, such as migrant workers, marginalized groups, and tourists, are all at risk of transmitting malaria and becoming seriously ill.The WHO African Region anticipates 33.8 million pregnancies in 33 moderate and high transmission countries in 2020, with 11.6 million (34%) of them exposed to malaria infection during pregnancy. According to the World Health Organization's most current World Malaria Report, there would be 241 million malaria infections and 627 000 malaria deaths worldwide by 2020.Malaria symptoms include fever, tiredness, headaches, and, in severe cases, convulsions and coma, all of which can result in death.Malaria is a treatable illness, with medications and tablets available to help travellers visiting malaria-prone areas avoid infection. Malaria vaccines, on the other hand, are ineffective. Malaria is a quickly spreading infection that, once inflamed, has a significant chance of progressing to severe and cerebral malaria with neurologic symptoms in P. Falciparum infections. As a result, effective malaria diagnosis is critical[9].

Malaria is a mosquito-borne disease caused by Plasmodium unicellular protozoan parasites. Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, Plasmodium ovale, and Plasmodium knowlesi are the five Plasmodium species that cause malaria in humans. Most clinical cases of malaria in humans are caused by Plasmodium falciparum (Pf) and Plasmodium vivax [9]. Although less severe than Plasmodium falciparum, the deadliest of the five human malaria parasites, P. vivax malaria infections can result in severe sickness and death. In Asia and South America, Plasmodium vivax is responsible for 65 percent of malaria cases. Pf has a difficult life cycle, with the parasite's transmission between people via mosquitos undergoing an incredible sequence of morphological transformations[13]. During their lives (48 hours), every parasite species goes through many stages, each with its own characteristic visual aspect that may be observed under a microscope. The ring stage (R), trophozoite stage (T), schizont stage (S), and sexual development (gametocyte stage) are the phases in chronological sequence (G). For each species, Figure 1 depicts typical instances of all phases.

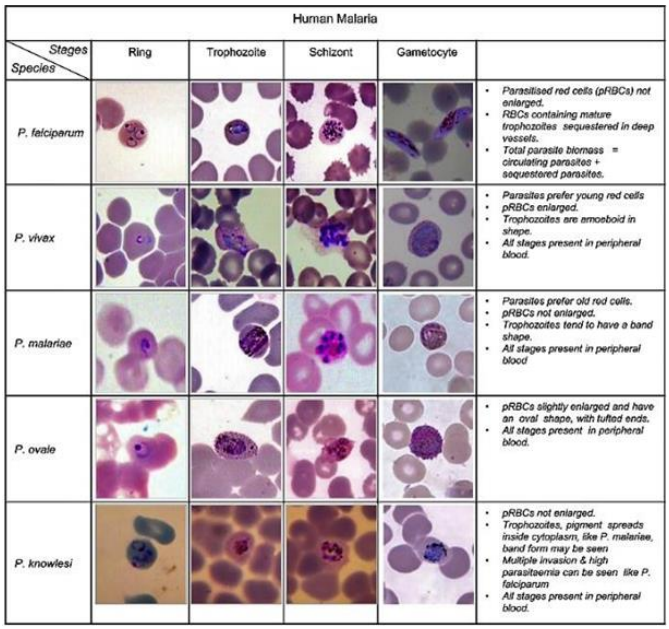

Figure1: Five different human malaria Plasmodium species and their life stages in thin blood film[9]

Around the end of the first week of parasitemia, gametocytes of P. vivax are typically identified in human peripheral blood. The gametocytes are spherical to oval in form, with brown pigment scattered throughout, and can almost entirely fill the RBC.P. vivax rings have numerous chromatin spots and ameboid cytoplasm. P. vivax schizonts are big, with 12 to 24 merozoites, yellowish-brown aggregated pigment, and the ability to completely fill the RBC. Trophozoites of P. vivax exhibit amoebic cytoplasm, big chromatin spots, and fine yellowish-brown pigment. Thin and thick blood smears are often examined by microscopists to diagnose the condition and evaluate parasitemia. Using a normal light microscope with a 100-oil objective, a sample of the blood of the patient is placed on a glass slide, which is then stained to detect parasites. Blood smears of two types are commonly used to diagnose malaria: thick and thin smears. A thick smear is used to indicate the existence of parasites in a drop of blood.

Thick smears detect parasites more effectively than thin smears due to an 11-fold increase in insensitivity. Smears that are thin, on the other hand, are generated by dispersing a drop of blood on a glass slide and have additional benefits. They make it easier for the

examiner to tell the difference between malaria species and parasite stages. An experienced microscopist will examine a single blood slide under the microscope for 15–30 minutes in order to find quantitative parasites and identify species. Their accuracy, on the other hand, is dependent on smear consistency and prior knowledge of parasitized and uninfected cells in classifying and counting parasitized and uninfected cells [9].

Multiple procedures are still used, including polymerase chain reaction (PCR) and rapid diagnostic testing (RDT). In disease-endemic areas, however, PCR tests have limits in terms of performance, and RDTs are less cost-effective [9]. Malaria screening may be made more efficient and successful with the use of an autonomous malaria detection system. To measure parasitemia and its stages in each smear, digital pictures of thin blood smears are analyzed using image recognition and machine learning technologies. The phases of malaria in blood smears are depicted in Figure 2.

Healthcare providers can make better judgments about patient diagnosis and treatment based on the early prediction results. Humans may also use machine learning algorithms to help them analyze vast, complex medical databases and analyze them for therapeutic insights. As a result, machine learning-based automated malaria screening has proven to be a great testing tool[9].

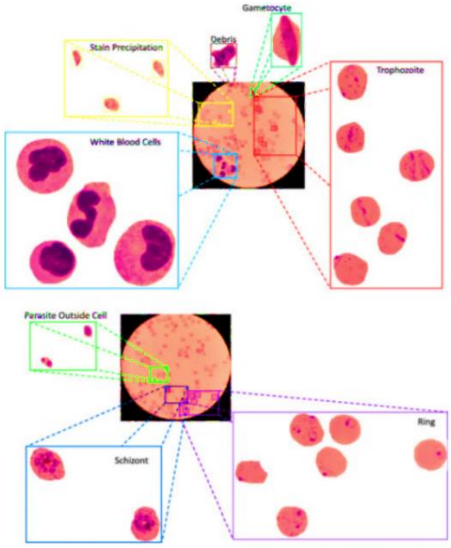

Figure 2:stages of malaria in blood smears[10]

This paper proposes a study of seven machine learning models namely Random Forest(RF), SVM, Naïve Bayes(NB), LDA, Decision tree(DT), logistic regression(LR), KNN, and one ensemble model.

In this paper, we analyze the evaluation parameters namely classification accuracy y, precision, recall and f1-score for these machine learning models to foresee which model is best suitable for predicting the stage of malaria.

II. LITERATURE SURVEY

Due to an increase in the number of patients and a shortage of diagnostic facilities and competent medical workers, there is a need for novel diagnostic techniques to enhance existing approaches.As a result, Abubakar, Aliyu, Mohammed Ajuji, and Ibrahim Usman Yahya [2] suggested that machine-learning models be used to detect the malaria parasite in blood smear images in their study. VGG16, VGG19, ResNet50, ResNet101, DenseNet121, and DenseNet201 models were used to extract six distinct characteristics. These six characteristics were then used to train DT, SVM, NB, and K-NN classifiers. Precision, recall, f-1score, accuracy, and computing time are all included in the comprehensive performance analysis. The results show that automating the process recognizes the malaria parasite in blood samples with more accuracy (around 94%) and less complexity than previous approaches.

Malarial fever disease is substantially caused by the Plasmodium parasite that's contagious to red blood cells. Manual blood cell counting is a time-consuming procedure that yields a poor diagnostic method. These procedures have a significant influence on the whole screening process. Aravinda C V, Meng Lin, Udaya Kumar Reddy K R, and Amar Prabhu G[6] suggested a Gabor Filter-based technique for analyzing this malarial disease, followed by a comparison of the XG-Boost classifier, SVM, and Neural Network Classifier algorithms as the preferred architecture for finding and categorizing these malarial blood cells. S.V.M obtained 94 percent in the studies, whereas XG-Boost reached 90 percent and a neural network classifier achieved 80 percent. To increase decision-making accuracy, S.V.M performed well in identifying and separating parasitized and uninfected blood cells.

Malaria is a fatal disease carried by female anopheles mosquito bites that are endemic in so many parts of the world. This study [7] intends to enhance malaria diagnosis by segmenting patches from microscopic pictures of red blood cell smears using a deep convolutional neural network. Unlike earlier systems, which relied on time-consuming manual feature extraction, the proposed technique employs deep learning in an end-to-end setting that does both feature extraction and classification straight from raw segmented patches of red blood smears. The NIH Malaria Dataset was used in this research. To assess and choose the best performing architecture, the evaluation measure accuracy and loss, as well as 5-fold cross-validation, were utilized. Existing traditional pre-processing procedures from the literature were also examined for performance improvement. To determine which model works best, several more sophisticated structures have been designed and tested. A holdout test was performed to see how well the suggested model generalizes data. The top model proposed has an accuracy of roughly 97.77 percent.

Md. Khayrul Bashar[8] proposed a supervised approach for recognizing malaria parasite stages from microscope pictures in the study. This approach combines color and texture information with a support vector machine (SVM) classifier to achieve the goal. Three texture characteristics were evaluated: an oriented pattern's histogram (HOG), a local binary pattern (LBP), and the Grey-level Co-occurrence Matrix (GLCM), as well as four color features: local color moments (StatMom) and ccolorhistograms (HSV, LAB, and YCrCb). An experiment was run with an imbalanced dataset of 46,978 single-cell thin blood smear photographs, and the cocolorharacteristics outperformed the texture features. Employing the SVM classifier, the suggested color texture feature (YCrCb HOG) appears to have the greatest accuracy rate (96.9%) on average, outperforming a previously reported strategy using the HOG LBP feature (87.1 percent).

Image recognition algorithms and machine learning approaches were used to measure malaria parasites in microscopic blood slides in order to enhance diagnosis. This study provides an understanding of current improvements in image processing and machine learning for microscopic malaria detection, as well as a description of these methodologies. Imaging techniques, image filtering, parasite detection and cell segmentation, attribute calculation, and automated cell categorization are all discussed. Future advancements in deep learning and mobile technology for malaria detection are also highlighted.

An automated approach for detecting malaria parasites in blood pictures was described by Vinayak K. Bairagi and Kshipra C. Charpe[11]. Image processing methods are utilised to diagnose and detect the stages of the malaria parasite. In blood images, factors such as statistical features and textural aspects of malaria parasites are used to diagnose parasite stages. This article compares how textural based elements are utilised separately and how they are used in groups. The comparison is based on the accuracy, sensitivity, and specificity of the characteristics for identical photos in the database.

Machine learning approaches are widely used in sentiment analysis, chatbot designs, pattern recognition, and many more [17][18].

III. RESULTS AND DISCUSSION

A .Dataset Collection

The photos of the multi-stage malaria-infected cells were taken from blood smear samples stained with Giemsa reagent. The Broad Bioimage Benchmark Collection (BBBC) website [1] hosts this image collection, which comprises 1,364 pictures at 1,000 magnification. All of the photos were taken by hand under 1,000 magnification from Plasmodium vivax–infected people in Manaus, Brazil, and Thailand, and annotated by three separate specialists from around the world. [1]. Figure 3 illustrates Giemsa staining of two types of uninfected cells (RBCs and leukocytes) and four types of parasitized cells (gametocytes, rings, trophozoites, and schizonts). The total number of microscopic pictures utilized in this study is 6,856 microscopic images out of which 6,170 images for training and 686 images for testing. This dataset is released on April 21, 2021.

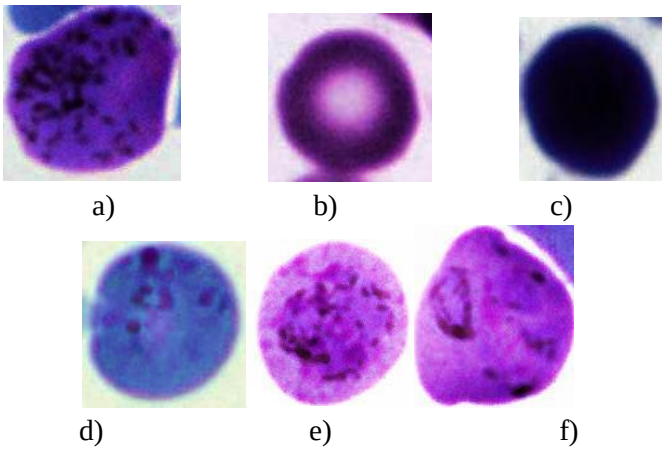

Figure 3. Sample images from dataset a) gametocyte b) leukocyte c) red blood cell d) Ring e) schizont f) trophozoite.

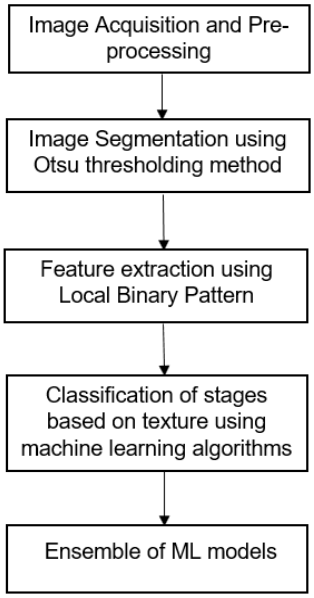

Figure 4. Block diagram of a Proposed method.

B. Proposed work

The methodology used in predicting the stage of malaria is shown in figure 4.The steps are as follows.

1. Import necessary packages.
2. Read the image. The input image is a Giemsa-stained blood sample of multi-stage malaria cells.
3. Convert images to grayscale and then apply Otsu segmentation
4. From segmented images extract the features using LBP. These characteristics can help the classification unit to decide which human host is affected by malaria and in what stage it is in.
5. Divide the data into training and testing sets using the train_test_split function with test size=0.1, such that the training set has roughly 6170 photos and the testing set has 686 images.
6. Define and train effective machine learning classifiers for classifying parasite stages into distinct classes.
7. Apply K-fold cross-validation(k=10) .Finally, based on the feature classes discovered by the classifier, a decision is made about the information conveyed by the image. Choose the best models based on accuracy.
8. Ensemble different machine learning models
9. Predict or classify the stages, check on the training dataset.

C. Segmentation

Image segmentation is an important part in image processing. In this work, the Otsu thresholding method is used for image segmentation. The OTSU technique is an image segmentation methodology that uses a global adaptive binarization threshold. It is named after Nobuyuki Otsu and is used to perform picture thresholding automatically. It will be easier to do processing on those segmented images.

Segmentation is necessary because -

- It transforms an image's representation into something more meaningful and easier to examine.
- After segmentation, it is much easier to properly detect the image's objects and borders.
- It is the process of labelling every pixel in a picture so that pixels with the same label have similar features.

Iterating over all possible threshold values and computing a measure of spread for the pixel levels on each side of the threshold, i.e. pixels that are either in the foreground or background, is Otsu's thresholding technique. The aim is to find the value at which the overall spread of the foreground and background is the least. The histogram of pictures is used in Otsu segmentation. If the intensity of a pixel in the input image exceeds a threshold, the output pixel position is marked as white (foreground), and if it is equal to or below the threshold, it is marked as black (background).

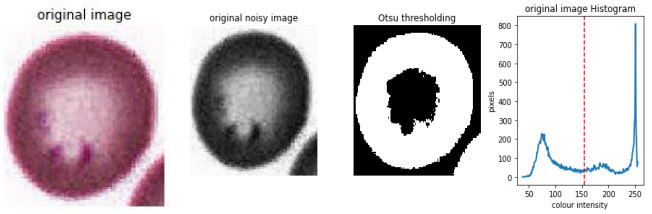

Figure 5 Example of Otsu segmentation

The resolution of the original image as shown in figure 5 was 96*96 dpi and the image dimensions are 114 x 128 pixels.

Steps of Otsu thresholding:

1. Read image: First, the image is loaded and is converted into the grayscale mode, and a histogram is generated. We can clearly detect expressed mono peak and its close area in the above picture of the histogram, as well as a slightly expressed peak at the beginning of the scale.
2. Calculate the Otsu's threshold with the OpenCV threshold function with the THRESH OTSU and THRESH BINARY flags set.

   • THRESH BINARY: If the pixel intensity exceeds the given threshold, the value is set to 255; otherwise, it is set to 0. (black).

   • THRESH OTSU: The best threshold value will be automatically returned. The threshold value obtained for this example shown in figure 5 is 155.0.
3. final binarized image is obtained after the application of Otsu's method. So, from this, we can easily see that the background and main elements in the image have been separated

D.Feature Extraction

Feature extraction is the process of generating a new and smaller collection of features that captures the majority of the important information in raw data. It has something to do with dimensionality reduction. When an algorithm's input data is too vast to analyze and appears to be redundant, it can be reduced to a small number of traits. This is known as feature extraction. The extracted features should contain useful information from the input data, allowing the task to be completed using this reduced representation rather than the entire initial data set.

In this study, the Local Binary Pattern (LBP) is employed to extract features. It's a wellknown method for representing and categorising images. The first LBP operator was created by Ojala et al. [18]. The LBP texture operator is a basic yet powerful pattern extractor that classifies pixels in an image by thresholding their surrounds and interpreting the result as a binary integer [16]. LBP is used to classify and recognize textures and patterns in photos automatically.

Working of LBP*:*

LBP may be used to characterize the texture and form of a digital picture. This is accomplished by segmenting a picture into tiny sections from which the characteristics are retrieved (figure 6).

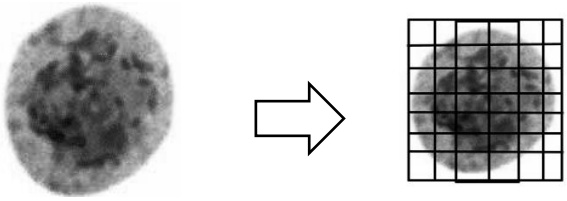

Figure 6 Preprocessed Image divided into several small regions

LBP has four parameters: radius, neighbors, grid x (number of cells in the horizontal direction), and grid y. (number of cells in the vertical direction). This technique employs a sliding window idea depending on factors such as radius and neighbors.

Figure 7 shows the example of the LBP operator. The technique entails thresholding the window's center pixel with its surrounding pixels using window mean, window median, or the actual center pixel as thresholds. It uses a histogram of the patterns for texture classification.

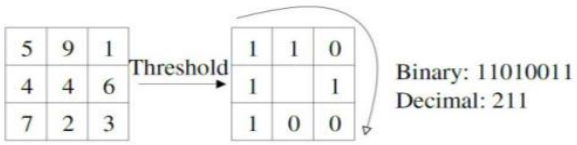

Figure 7 Example of LBP Operator

LBP steps[16]:

- we have a multi-stage malaria cell image converted in to grayscale.
- Because this approach employs the sliding window principle, each 3x3 window in the picture must be analysed in order to extract LBP information.

- It can also be represented as a 3x3 matrix containing the intensity of each pixel (0~255).
- Then, take the central value of the matrix to be used as the threshold.
- The traditional LBP operator compares the grey value of the centre pixel with the grey values of its eight neighbours pixels. All neighbour pixels with values more than or equal to those of the centre pixels are given a value of one, while those with values lower than those of the central pixels are given a value of zero. After that, the binary values corresponding to eight neighbors are acquired in clockwise order. Then we transform this binary value to a decimal number and assign it to the matrix's center value, which is a pixel from the original picture.
- At the completion of this operation (LBP technique), we obtain a new image that better depicts the properties of the original image.

The key advantage of using LBP is that we can capture incredibly fine-grained features in the image. The algorithm's capacity to collect features on such a tiny scale is also one of its primary drawbacks. Ojala et al. [17][18] proposed two factors to address this.

1. number of points p in the circularly symmetric neighborhood to consider

2. radius of circle r, which allows us to account for different scales.

Given a number of points p in LBP there are p+1 uniform patterns. Figure 8 shows the circular neighbor-sets for three different values of P and R. A pixel's neighborhood is given in the form of P nneighbors within a radius of R. It is an extremely strong descriptor that identifies all of the image's possible edges. In a Local Binary Pattern, the number of uniform prototypes is entirely determined by the number of points p. In this proposed work we chose p as 24 and r as 8.

The benefit of LBP is that monatomic grayscale changes do not alter, computing complexity is minimal, small rotation invariant, illumination invariant, and easy multiscale expansions.

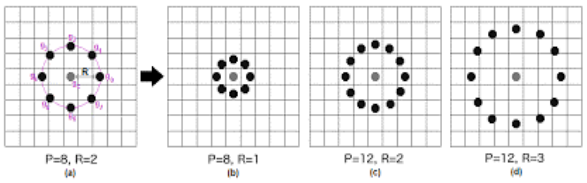

Figure 8. Circularly neighbour-sets for three different values of P and R [17].

The LBP matrix may be expressed as a histogram, which will be used as the feature vector of the original picture. Figure 9 shows an example of deriving the LBP representation and histogram (Right) from the original schizont input picture (Left).

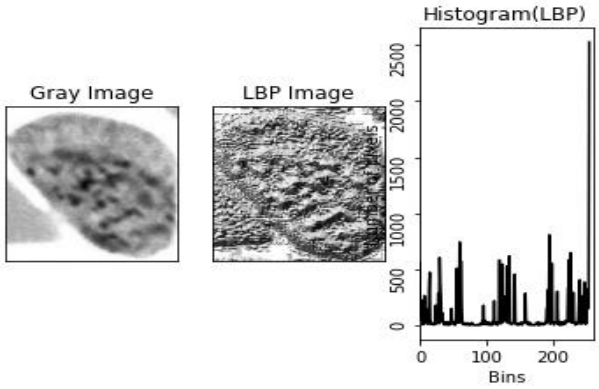

Figure 9 An example of computation of LBP image and histogram from the original image

E.Machine Learning Algorithms

1. **Random Forest:** Random Forest is a comprehensive machine learning (ML) classification approach. An ensemble of decision trees enables RF. Each tree expects a classification and "votes" for the required parameters independently, with the majority of the votes deciding the overall RF predictions. The use of several trees reduces the possibility of overfitting.

2**. Decision Tree:** A decision Tree is an algorithm that generates a classification tree using the training dataset. It's a tree-like structure in which an internal node represents a characteristic (or attribute), a branch represents a decision rule, and each leaf node reflects the outcome. It is based on the technique of "divide and conquer." It is divided into three sections: partitioning the nodes, identifying terminal nodes, and assigning class labels to terminal nodes. It is organised according to a hierarchical system. It can handle high-dimensional data and communicate knowledge in a tree format that is simple for humans to understand. Here we train the dataset using a decision tree classifier in scikit-learn.

3.**KNN:** The acronym KNN stands for "K-Nearest Neighbors." It's also a common algorithm for multi-class classification. KNeighborsClassifier is used to implement it in sklearn. This method is entirely dependent on the distance between feature vectors, which in this case are the raw RGB pixel intensities of the images.The distance metric used for the tree is Minkowski**.** Here I have chosen p=2, so that standard Euclidian distance in the Minkowski is obtained. The k-NN approach categorizes unknown data points by locating the most frequent class among the k nearest examples. Each of the k nearest data points votes, and the class with the most votes wins. we have chosen neighbors randomly so the algorithm will take into account the nearest neighboring data points for determining the class of malaria. As the value of K increases the prediction curve becomes smoother.

4. **Support Vector Classifier:** SVM (Support vector machine) is an efficient classification strategy when the feature vector has a high dimension. In sci-kit learn, we can specify the kernel function. The kernel trick is a technique for calculating relationships between data points. SVM employs a variety of kernels, the most common of which are polynomial, radial, and linear kernels. The linear kernel is used in this work. With a C value of 1.0, linear SVM is utilised. In this work, the random state is selected as 9. It manages the generation of pseudo-random numbers used to shuffle data for probability estimations.

5. **Naive Bayes:** The Bayes Theorem is used by Nave Bayes, a classifier. It computes the likelihood of a data item belonging to each class and assigns the label with the highest probability to that class.
The Bayes rule is defined as:

$$P(A/B) = P(B/A).P(A) / P(B) \quad (1)$$

Where A and B denote the class and the characteristics, respectively. P(A/B) denotes the likelihood of belonging to class A with all given attributes of B. P(B/A) denotes the probability of belonging to class B with all of A's specified characteristics. P(A) denotes the probability of all classes. P(B) represents the probability of all features.. Naive Bayes is a fast and basic classification technique that is commonly used as a preliminary step for classification tasks. Gaussian naïve Bayes is used here to classify the stages of malaria. It posits that classifiers must be unbiased of one another. It accepts continuous values and assumes that each class is distributed uniformly.

6. **Logistic Regression**: Multinomial logistic regression is a logistic regression variant that supports multi-class classification issues natively. Lbfgs is used as a solver here. The abbreviation stands for limited-memory BFGS. This solver just computes an approximation to the Hessian based on the gradient, making it more computationally efficient.

7**. LDA:** Linear discriminant analysis is a technique for reducing dimensionality. To overcome the dimensionality curse, characteristics from higher-dimensional space are projected onto lower-dimensional space. It estimates probability using the Bayes theorem. They produce predictions based on the likelihood that each new input dataset corresponds to one of the classes. The class with the highest probability is considered as the output class, and predictions are made.

F.Ensemble of Machine Learning Models.

To create superior predictions on a dataset, ensemble learning uses many machine learning models. Ensemble learning refers to the ML models that combine the predictions from two or more models. An ensemble model is created by training many models on the same dataset. The predictions of these models are then merged in an ensemble model to make a final prediction.. The maximum voting approach is commonly utilized for classification problems. This technique use many models to provide predictions for each data point. The predictions of each model are viewed as a 'vote.' The final prediction is based on the majority of the models' projections. Its primary purpose is to enhance categorization.

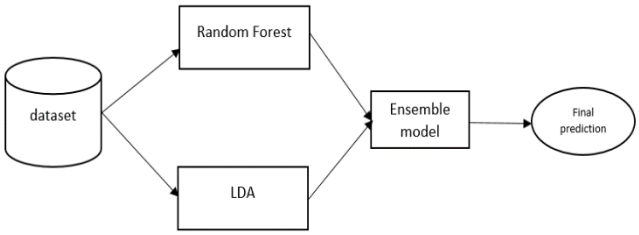

Figure 10. Ensemble of machine learning algorithms

In this task, two different types of machine learning models are used as weak learners to build an ensemble learning model as shown in figure 10. A homogenous collection of weak learners is used in this model. The models are Random Forest and LDA.

G.Classification Report

The classification report provides a comprehensive summary of precision, recall, F1score, and support for each class. It enables us to have a better understanding of the overall performance of our trained model. We used a variety of machine learning algorithms to train the dataset, including random forest, logistic regression, LDA, decision tree, and Naive Bayes.

Precision is defined as the proportion of true positives to the total of true and false positives. It represents the fraction of positive recognition that were right. A model with a precision of 1.0 yields no false positives.

- The ratio of true positives to the total of true positives and false negatives is known as recall. It represents the fraction of true positives that were accurately categorized. A model with a recall of 1.0 delivers no false negatives.
- The weighted harmonic mean of accuracy and recall is F1. The closer the F1 score number is to 1.0, the better the expected performance of the model is.
- Support is the number of actual occurrences of samples in each class in the dataset
- Accuracy is the proportion of correct classifications from the overall number of cases.
- Macro Avg is the average precision, recall and F1 score between classes
- Weighted avg is the weighted average precision, recall and F1 score between classes. Weighted means that each measure is determined in proportion to the number of samples in each class.

IV.RESULTS AND DISCUSSION

H.Classification report of Different Machine Learning Models.

Classification Report of Decision Tree

Table 1 shows the classification report of DecisionTree. The model identified red blood cell, ring, schizont and trophozoite stages. for red blood cells, the precision score is 85%, recall score is 87% and f1 score is 86%.For schizont the precision, recall and f1 score is 20%.for trophozoite stage the precision score is 57%, recall score is 52% and f1 score is 54%. The precision score is high for red blood cell so it shows that red blood cell is more correctly identified by the model when compared to other stages. The overall accuracy of the model is 76%.

Table 1 Classification Report of Decision Tree

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.85	0.87	0.86	508
ring	0.15	0.14	0.15	21
Schizont	0.20	0.20	0.20	5
trophozoite	0.57	0.52	0.54	139
accuracy			0.76	686
Macro avg	0.29	0.29	0.29	686
Weighted avg	0.75	0.76	0.75	686

However, due to the additional complexity and time required for training, it is quite expensive. If the data set is vast, a single tree may become complicated, leading to overfitting. In this scenario, Random Forest should be used instead of a single Decision Tree. To solve the Decision Tree's limitations, we must employ Random Forest, which does not reliant on a single tree. It grows a forest of trees before making a choice depending on the number of votes received.

Classification Report of Random Forest model

Since Random forest is very stable and helps in automatically handling of missing data we first trained our dataset using this algorithm. Table 2 shows the classification report of the random forest model.

Table 2 Classification Report of Random Forest Model

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.86	0.95	0.90	508
ring	0.50	0.05	0.09	21
Schizont	0.00	0.00	0.00	5
trophozoite	0.70	0.59	0.64	139
accuracy			0.83	686
Macro avg	0.34	0.27	0.27	686
Weighted avg	0.79	0.83	0.80	686

When it comes to precision and recall, red blood cells have a precision of 86% and a recall of 95% and an f1 score is 90%. Precision is less because of false positives. Similarly, the ring has a precision of 50%, recall of 5%

and f1 score of 9%. trophozoite have a precision of 70%, recall of 59% and f1 score of 64%. Here in trophozoite and ring recall is less because of false negative. Along with red blood cell ,ring and trophozoite stage is also been identified by the model. Random forest decreases overfitting in decision trees, as well as variation, and hence enhances accuracy. Although this model has an overall accuracy of 83%, the main restriction of random forest is that a high number of trees might cause the algorithm to slow down and become inefficient for real-time predictions. To solve this limitation, we must train and test the data using the Logistic Regression technique, which is readily extensible to several classes (multinomial regression) and provides a natural probabilistic perspective of class predictions.

Classification of Logistic Regression

Table 3 shows the classification report of the logistic regression model. When it comes to precision and recall, red blood cells have a precision of 74% and a recall of 100%. Because no other species of red blood cell has been identified.

Table 3. Classification Report of Logistic Regression

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.74	1.00	0.85	508
ring	0.00	0.00	0.00	21
Schizont	0.00	0.00	0.00	5
trophozoite	0.00	0.00	0.00	139
accuracy			0.74	686
Macro avg	0.12	0.17	0.14	686
Weighted avg	0.55	0.74	0.63	686

Because none of the red blood cells has been classified as some other species. Precision is less because of false positives. Similarly, none of the other species have been incorrectly identified as Redbloodcell. F1 score is 85%. A low number of false positives and false negatives suggests an excellent F1 score. The overall average accuracy is 74%. The assumption of linearity between the dependent and independent variables is a fundamental disadvantage of Logistic Regression. On high-dimensional datasets, this may cause the model to be over-fit on the training set, implying that the accuracy of predictions on the training set is exaggerated, and so the model is unable to predict correct outcomes on the test set. To address this limitation, we are analyzing the dataset using LDA, a simple, rapid, and portable approach. It outperforms some algorithms even when its assumptions are satisfied (logistic regression).

Classification Report of LDA:

Table 4 shows the classification report of Linear Discriminant Analysis. If we take a look at precision ,f1-score and recall score, red blood cell has a precision of 83%, recall of 97% and f1-score of 89%. Along with red blood cells, trophozoites have also been identified by the model. Trophozoite has a precision of 77%, recall of 47% and f1 score of 58%.In the case of trophozoites recall is less because of more number of false negatives. A greater precision indicates that an algorithm produces more relevant results than irrelevant outcomes, but a high recall indicates that an algorithm produces the majority of relevant results. LDA assumes that all classes are linearly separable and hence the accuracy and recall of this model is 81%.But it is prone to overfitting.we further train and test our dataset by using KNN algorithm as it requires no training before making predictions and is easy to implement.

Table 4.Classification Report of LDA

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.83	0.97	0.89	508
ring	0.00	0.00	0.00	21
Schizont	0.00	0.00	0.00	5
trophozoite	0.77	0.47	0.58	139
accuracy			0.81	686
Macro avg	0.27	0.24	0.25	686
Weighted avg	0.77	0.81	0.78	686

Classification Report of KNN

Table 5. Classification Report of KNN

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.81	0.94	0.87	508
ring	0.00	0.00	0.00	21
Schizont	1.00	0.20	0.33	5
trophozoite	0.61	0.43	0.51	139
accuracy			0.78	686
Macro avg	0.40	0.26	0.29	686
Weighted avg	0.73	0.78	0.75	686

Table 5 shows the classification report of KNN. For red blood cells precision is 81%, recall is 94% and f1-score is 87%. %. Whereas schizont has a 100% precision score .Because there are no false positives and has of 20% of recall score and 33% of F1score. For trophozoites the precision is 61%, recall is 43% and f1-score is 51%. The recall of the model evaluates its ability to recognize positive samples. The greater the recall, the more positive samples are found. The overall accuracy and recall of the model is 78%. KNN does not operate well with bigger datasets because the cost of computing the distance between the new and existing points is considerable and may reduce the algorithm's performance, even though the overall accuracy of this model is 78 percent. To address this restriction, we apply gaussian naive bayes, which is extremely scalable in terms of predictors and data points and is not sensitive to irrelevant characteristics.

Classification Report of Gaussian Naïve Bayes

Table 6 shows the classification report of Naïve Bayes. The precision, recall and f1 score of red blood cells is more when compared to other stages. Hence, we can say that red blood cell is more highly predicted correctly by the Naive Bayes model.

Table 6: Classification Report of Gaussian Naïve Bayes

	Precision	Recall	F1-Score	Support
gametocyte	0.01	0.17	0.02	6
leukocyte	0.02	0.14	0.04	7
Red_blood_cell	0.89	0.63	0.74	508
ring	0.02	0.10	0.04	21
Schizont	0.02	0.20	0.04	5
trophozoite	0.62	0.31	0.41	139
accuracy			0.53	686
Macro avg	0.27	0.26	0.21	686
Weighted avg	0.79	0.53	0.63	686

This is a system with high accuracy but poor recall, indicating it returns very few results, yet when compared to the training labels, the majority of its projected labels are right. The model's total accuracy is 53 percent. But. The frequency-based probability estimate of Nave-Bayes is 0 if no instances of a class label and a specific attribute value occur together. Naive Bayes assumes that all predictors (or traits) are independent, which is rarely the case in real life.Hence SVM algorithm can be used, in general it, do not suffer from overfitting and perform well when there is a strong indication of class separation. SVM is effective when the total sample size is less than the number of dimensions, and it is memory efficient.

.

Classification Report of SVM

Table 7 shows the classification report of SVM. The precision of red blood cells is 74%. The recall of red blood cells is 100%, then it tells us the model has detected all samples of red blood cells as positive. F1 score is 85%.

Table 7: Classification Report of SVM

	Precision	Recall	F1-Score	Support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
Red_blood_cell	0.74	1.00	0.85	508
ring	0.00	0.00	0.00	21
Schizont	0.00	0.00	0.00	5
trophozoite	1.00	0.01	0.01	139
accuracy			0.74	686
Macro avg	0.29	0.17	0.14	686
Weighted avg	0.75	0.74	0.63	686

For trophozoites, the precision score is 100% Which means that all the positive samples are classified as positive and none of the positive samples is classified incorrectly. Recall and f1 score is 1% which is very low because of false negatives. The overall precision of the model is high i.e.75%. The recall metric evaluates the model's ability to identify Positive samples. The greater the recall, the more positive samples are found. However, if the number of features exceeds the number of data points, the approach is prone to over-fitting. We are now attempting to ensemble models with high accuracy in order to test for improved performance.

Classification Report of Ensemble model

Table 8: Classification Report of Ensemble model.

	Precision	Recall	F1-score	support
gametocyte	0.00	0.00	0.00	6
leukocyte	0.00	0.00	0.00	7
red_blood_cell	0.82	0.98	0.89	508
ring	0.00	0.00	0.00	21
schizont	0.00	0.00	0.00	5
trophozoite	0.84	0.42	0.56	139
accuracy			0.81	686
Macro avg	0.28	0.23	0.24	686
Weighted avg	0.78	0.81	0.77	686

Table 8 shows the classification report of the ensemble model. The precision of trophozoites is more i.e. 84% and recall and f1-score of red blood cells are high i.e. 98% and 89% when compared to other stages. Recall gives a measure of how accurately our ensemble model is able to identify the relevant data Our ensemble model has an 81 percent recall rate. The recall of the model measures its ability to recognize positive samples. The more positive samples identified, the larger the recall. When compared to random forest, the overall accuracy of the ensemble model is 81 percent.

I. Confusion Matrix:

A confusion matrix is a table used to illustrate the performance of a classification model on a set of test data for which the true values are known. Seaborn is employed in this approach to make confusion matrices more intelligible and visually appealing.

Confusion matrix of Logistic Regression

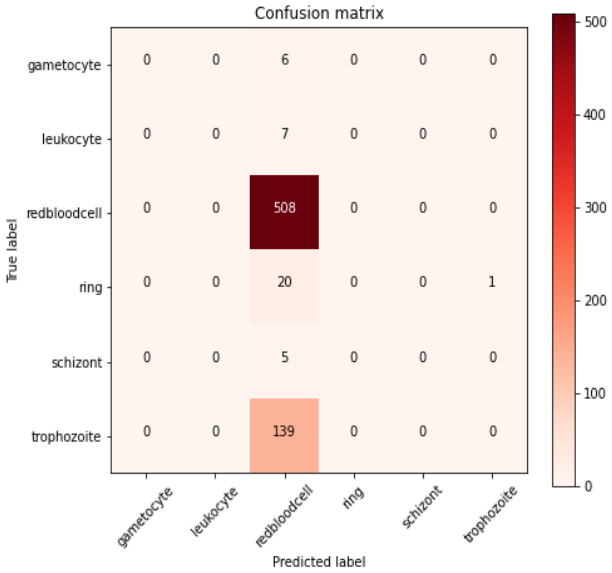

Figure 11. Confusion Matrix of Logistic Regression

If we look closely at the above diagram, you'll see that all of the highlighted cells in both rows and columns have the same label. Figure 11 represents the confusion matrix of the Logistic Regression model.For the red blood cell row, whatever is shown in a highlighted cell is truly positive for red blood cells and the values in other cells in the same row are false negatives for red blood cells. The other cells in that row are false negatives for the label. Red blood cells, for example, have 508 in the cell marked in brown. This means that out of the labels identified as red blood cells by our model 508, only 6 false negatives of gametocyte, 7 false negatives of leukocyte, 20 false negatives of ring, 5 false negatives of schizont, and 139 false negatives of trophozoite are true red blood cells, indicating that either gametocyte, leukocyte, ring, schizont, or trophozoite has been falsely identified as red blood cell.

Confusion matrix of LDA:

In the above image, If we look closely, we can see that all of the highlighted cells in both rows and columns have the same label. The remaining cells in that row are false negatives for the label. The confusion matrix of the LDA model is shown in Figure 12. For the red blood cell row, whatever is presented in a highlighted cell is a true positive for red blood cells, whereas the values in other cells in the same row are false negatives for red blood cells.

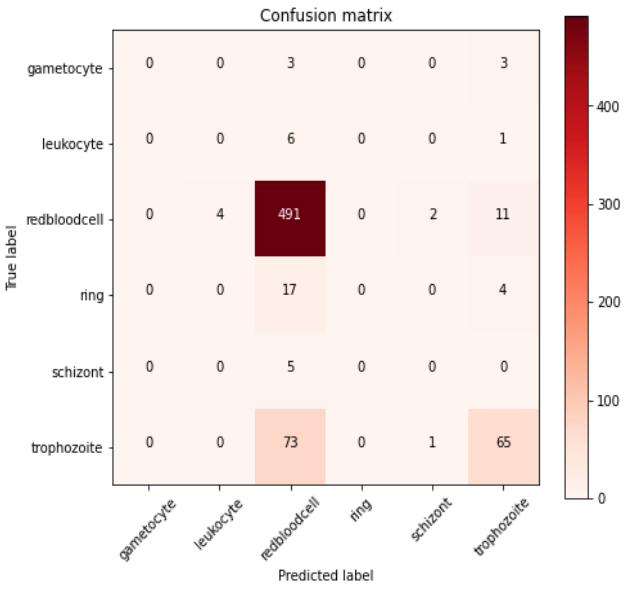

Figure 12 Confusion Matrix of LDA

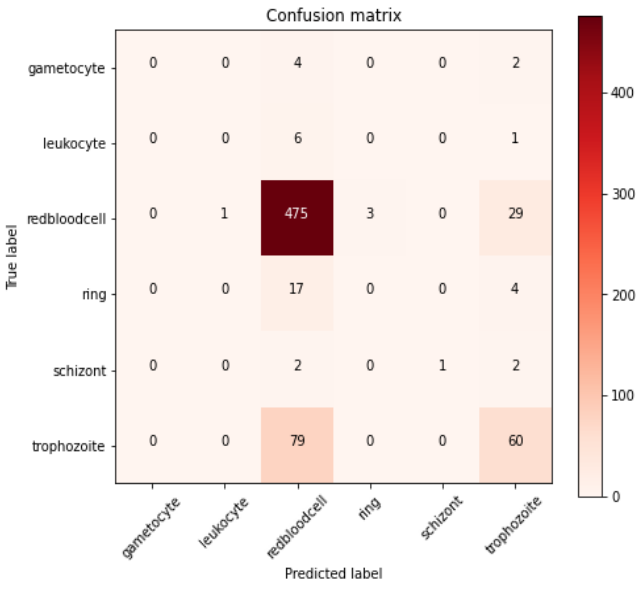

Figure 13 Confusion Matrix of KNN

For example, red blood cell has 491 in the cell highlighted in brown, Which means out of the labels which have been identified as red blood cell by our model 491 are truly red blood cells and we have 3 false negatives of gametocyte,6 false negatives of leukocyte,17 false negatives of the ring,5 false negatives of schizont and 73 false negatives of trophozoite. which means either gametocyte, leukocyte, ring, schizont or trophozoite are has been falsely identified as a red blood cells. For example, Trophozoite has 65 in the cell highlighted in pink. This means out of the labels which have been identified as trophozoite by our model 65 are truly trophozoite. and we have 3 false negatives of gametocyte,1 false-negative of leukocyte,11 false-negative of red blood cell and 4 false negatives of a ring, No false negatives of schizont which means either gametocyte, leukocyte, ring, red blood cell, schizont is has been falsely identified as a trophozoite.

Confusion Matrix of KNN

If we look closely at the figure above, we can see that all of the highlighted cells in rows and columns have the same label. Other cells in that row are labelled as false negatives. The confusion matrix of KNN is seen in Figure 13. Whatever is presented in a highlighted cell is actually positive for red blood cells, whereas the values in other cells in the same row are false negatives for red blood cells.

For example, a red blood cell has 475 in the cell highlighted in brown. This means out of the labels which have been identified as red blood cells by our model 475 are truly red blood cells and we have 4 false negatives of gametocyte,6 false negatives of leukocyte,17 false negatives of the ring,2 false negatives of schizont and 79 false negatives of trophozoite, which means either gametocyte, leukocyte, ring, schizont or trophozoite are has been falsely identified as red blood cells. In red blood cell row, 3 are falsely predicted as ring stage. In schizont row 1 true positive of schizont is identified. For in trophozoite column 60 samples are truly predicted as trophozoite and then the rest all are falsely identified as trophozoite by the KNN model

Confusion matrix of Decision Tree

If we look closely at the figure above, we can see that all of the emphasized cells in rows and columns have the same label. Other cells in that row are labelled as false negatives. Figure 14 depicts the Decision tree's confusion matrix. For the red blood cell row, whatever is reported in a highlighted cell is actually positive for red blood cells, whereas the values in other cells in the same row are false negatives. For example, red blood cell has 443 in the cell highlighted in brown, which means out of the labels which have been identified as red blood cells by our model 443 are truly red blood cells and we have 3 false negatives of gametocyte,6 false negatives of leukocyte,12 false negatives of the ring,3 false negatives of schizont and 57 false negatives of trophozoite. which

means either gametocyte, leukocyte, ring, schizont or trophozoite are has been falsely identified as a red blood cells. for the trophozoite row,72 samples are truly positive. So only two stages are correctly identified by the model accurately.

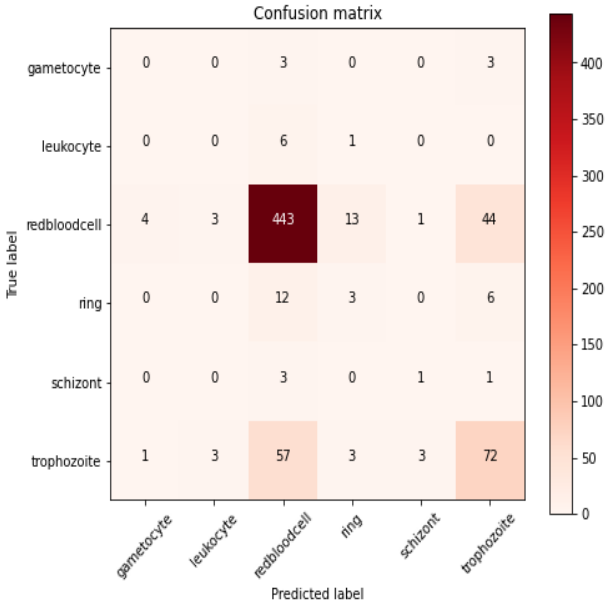

Figure 14 Confusion Matrix of Decision Tree

Confusion Matrix of Random Forest

If we look closely at the figure above, we can see that all of the emphasized cells in rows and columns have the same label. Other cells in that row are labelled as false negatives. Figure 15 depicts the Random forest's confusion matrix. For the red blood cell row, whatever is presented in a highlighted cell is a true positive for red blood cells, whereas the values in other cells in the same row are false negatives. For example, red blood cell has 485 in the cell highlighted in brown, which means that 485 of the labels identified as red blood cells by our model are truly red blood cells, and we have 3 false negatives of gametocyte, 5 false negatives of leukocyte, fifteen false negatives of the ring, three false negatives of schizont, and 57 false negatives of trophozoite. This indicates that a gametocyte, leukocyte, ring, schizont, or trophozoite was mistakenly labelled as red blood cells. There are 82 samples in the trophozoite row that are genuinely positive. As a result, the model properly identifies only two stages.

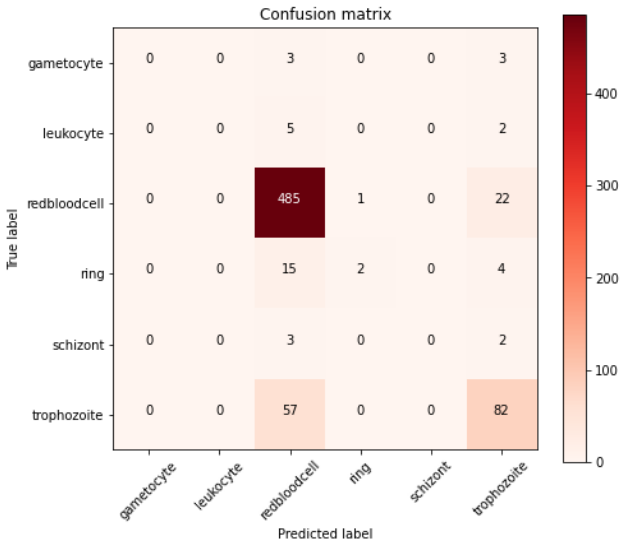

Figure 15 Confusion Matrix of Random Forest

Confusion Matrix of Naïve Bayes

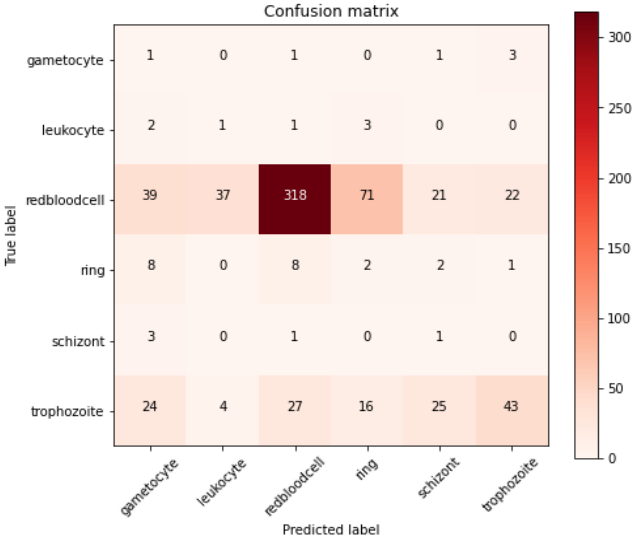

Figure 16 Confusion Matrix of Naïve Bayes

If we look closely at the figure above, we can see that all of the emphasized cells in rows and columns have the same label. Other cells in that row are labelled as false negatives. The confusion matrix of the Nave Bayes is shown in Figure 16. For the red blood cell row, whatever is reported in a highlighted cell is actually positive for red blood cells, whereas the values in other cells in the same row are false negatives. For example, red blood cell has 318 in the cell highlighted in brown, which means out of the labels which have been identified as red blood cells by our model 318 are truly red blood cells and we have 1 false negative of gametocyte,1 false negative of leukocyte,8 false negatives of the ring,1

false-negative of schizont and 27 false negatives of trophozoite. which means either gametocyte, leukocyte, ring, schizont or trophozoite are has been falsely identified as a red blood cells.71 samples of red blood cell are falsely predicted as ring ,21 samples of red blood cells are falsely predicted as schizont and 22 samples are falsely predicted as trophozoite.for the trophozoite row,43 samples are truly positive. So only two stages are correctly identified by the model accurately.

Confusion Matrix of SVM

In the above image, If we observe carefully, all the highlighted cells have the same label in rows and columns. Other cells in that row are false negatives for the label. Figure 17 shows the confusion matrix of the SVM. For the red blood cell row, whatever is shown in a highlighted cell is truly positive for red blood cells and the values in other cells in the same row are false negatives for red blood cells. For example, a red blood cell has 508 in the cell highlighted in brown, which means out of the labels which have been identified as red blood cells by our model 508 are truly red blood cells and we have 6 false negatives of gametocyte,7 false negatives of leukocyte,21 false negatives of the ring,5 false negatives of schizont and 138 false negatives of trophozoite. which means either gametocyte, leukocyte, ring, schizont or trophozoite are has been falsely identified as red blood cells.

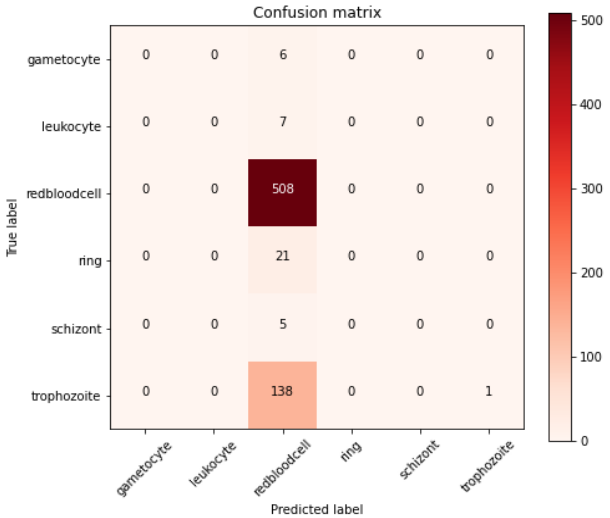

Figure 17 Confusion Matrix of SVM

for the trophozoite row, only 1 sample was truly positive. So only two stages are correctly identified by the model accurately.

Confusion Matrix of Ensemble model:

In the red blood cell row, whatever is presented in a highlighted cell is a true positive for red blood cells, but the values in other cells in the same row are false negatives for red blood cells. Figure 18 shows the confusion matrix of the ensemble model.

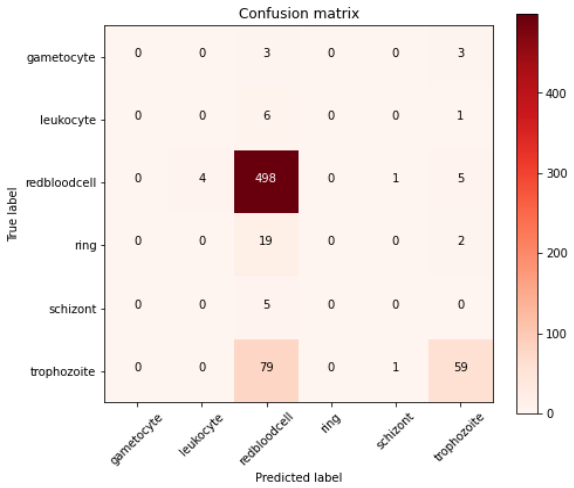

Figure 18 Confusion Matrix of the Ensemble model.

In the image, If we look closely, all the highlighted cells have the same label in rows and columns. Other cells in that row are false negatives for the label. A red blood cell, for example, has 498 in the cell highlighted in brown, which means that 498 of the labels identified as red blood cells by our model are truly red blood cells, and we have 3 false negatives of gametocyte, 6 false negatives of leukocyte, 19 false negatives of the ring, 5 false negatives of schizont, and 79 false negatives of trophozoite. This indicates that a gametocyte, leukocyte, ring, schizont, or trophozoite was mistakenly labelled as a red blood cell. Only 59 of the trophozoite samples were genuinely positive. As a result, the model properly identifies only two stages.

J.Comparison of models*:*

The performance measuring metrics like precision, recall,f1-score and accuracies of the seven ML algorithms and one ensemble algorithm, we just explored for our multi-stage dataset, can be summarized using the graph shown in figure 19.

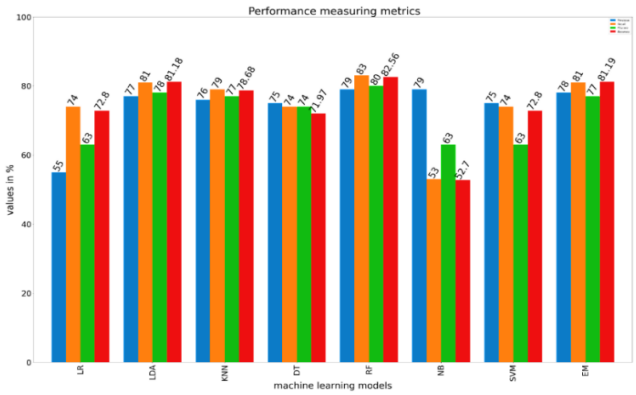

Figure 19 Performance metrics of different machine learning models.

Precision measures how well a model predicts a certain category.RF and NB shows a higher precision score of 79% followed by EM with 78% precision, LDA with 77% precision ,KNN with 76% precision, SVM and DT with 75% precision, and LR shows a low score of 55% precision. Hence RF and NB will predict more relevant stage results than irrelevant ones.

The recall score of RF is high with 83% followed by LDA and EM with 81% recall, KNN with 79% recall , SVM,LR, and DT with 74% recall and NB with low recall of 53%. Hence by this, we can conclude that RF will correctly identify the exact class.

The F1 score of RF is high with 80% followed by LDA with 78% f1-score,KNN and EM with 77% f1-score,DT with 74% f1-score,LR,NB and SVM with 63% f1-score.Hence we can say that RF can correctly predict the real class of multi stage malaria parasite.

Random Forest Classifier shows the best performance with 82.56% accuracy followed by LDA with 81.18% accuracy, KNN with 78.68% accuracy, SVM and Linear Regression with 72.80% accuracy, Decision Tree with 71.97% accuracy, Gaussian Naïve Bayes with 52.70% accuracy and ensemble model with 81.19% accuracy. Thus, Random Forest exhibits the best performance and Naïve Bayes the worst.

K. Snapshots of the web application

The web application is developed using a flask. Flask is a web framework that allows user to develop applications easily.

Figure 20 A user interface to choose the sample image

Figure 20 shows the user interface of the web application developed to identify the stage of malaria.

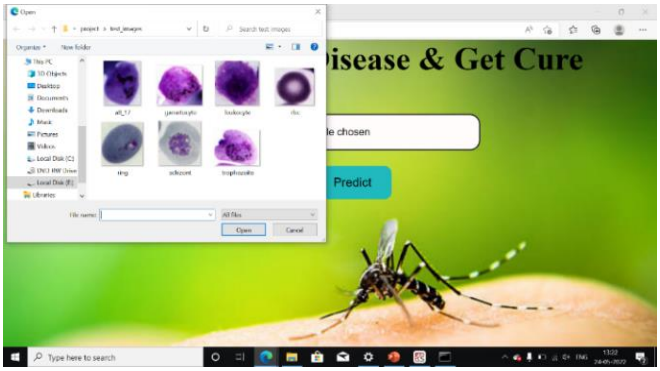

Figure 21. Select and upload the sample from the disk

Figure 21 shows the user interface of the web application developed where user can select the image sample and upload it to identify the stage of malaria.

Figure 22 Prediction of selected sample-gametocyte

Figure 22 shows the interface of the web application developed where user can view the result of the prediction of the image sample uploaded.

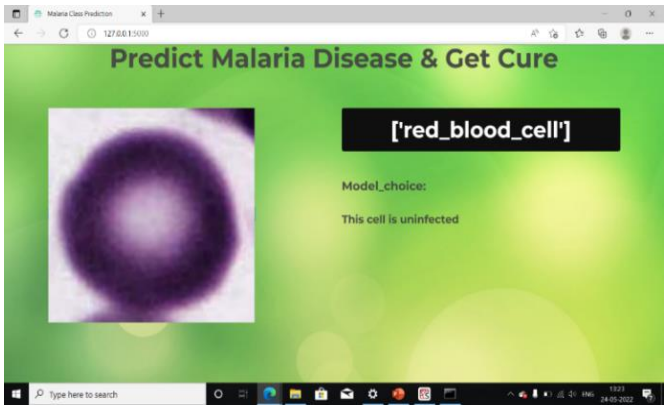

Figure 23 Prediction of selected sample-red blood cell

Figure 23 shows the interface of the web application developed where user can view the result of the prediction of the image sample uploaded.

Figure 24 Prediction of selected sample-schizont

Figure 24 shows the interface of the web application developed where users can view the result of the prediction of the image sample uploaded.

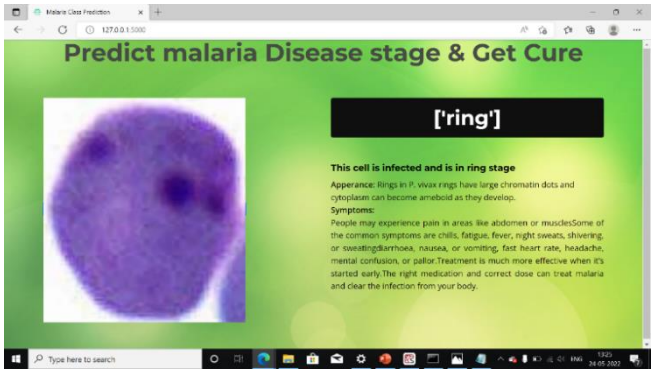

Figure 25 Prediction of selected sample-Ring

Figure 25 shows the interface of the web application developed where users view the result of the prediction of the image sample uploaded.

V.CONCLUSION AND FUTURE WORK

The proposed method for microscopic image analysis of multi-stage malaria parasites could be extremely useful in the creation of a low-cost, automated malaria diagnostic solution. Images are classified by using Random Forest, Decision tree, LDA, NB, Linear regression, SVM, and KNN in this proposed work. Ensemble of machine learning models has also been implemented. However, all the Machine learning algorithms perform poorly as indicated by the accuracies. The maximum score is just 82.56 percent, although deep learning systems outperform them with higher accuracies, which can be considered in future research. This enables a quick, efficient and easy way to manage malaria by facilitating accurate diagnosis and hence the correct treatment at the right time. In areas with limited resources, this can dramatically enhance efficiency and eliminate the need for specialist pathologists.

REFERENCES

[1]. Li, Sen, et al. "Multi-stage malaria parasite recognition by deep learning." *GigaScience 10.6* (2021): *giab040.*

[2]. Abubakar, Aliyu, Mohammed Ajuji, and Ibrahim Usman Yahya. "DeepFMD: Computational Analysis for Malaria Detection in Blood-Smear Images Using Deep-Learning Features." *Applied System Innovation 4, no. 4* (2021): 82.

[3]. Sultani, Waqas, Wajahat Nawaz, Syed Javed, Muhammad Sohail Danish, Asma Saadia, and Mohsen Ali. "Towards Low-Cost and Efficient Malaria Detection." arXiv preprint arXiv:2111.13656 (2021).

[4]. Dutta, Prakhar, and J. B. Jeeva. "Automation of Malarial Cell Count and Stage Classification Using Morphological Operations and Variable Optimization Using Hyper-Parameter Tuning." *In Advances in Automation, Signal Processing, Instrumentation, and Control, pp. 2997-3003. Springer, Singapore,* 2021.

[5]. Masud, Mehedi & Alhumyani, Hesham & Alshamrani, Sultan & Cheikhrouhou, Omar & Saleh, Ibrahim & Muhammad, Ghulam & Hossain, M. Shamim & Mohammad, Shorf. (2020). “Leveraging Deep Learning Techniques

for Malaria Parasite Detection Using Mobile Application. Wireless Communications and Mobile Computing. 2020." 1-15. 10.1155/2020/8895429.

[6]. C V, Aravinda & Meng, Lin & Reddy K R, Udaya & Prabhu, Amar. (2021). "Microscopic peripheral malarial parasite detection and classification in blood smears using Gabor Filters and Machine learning algorithms". 10.21203/rs.3.rs265728/v1.

[7]. Rahman, Aimon, Hasib Zunair, M. Sohel Rahman, Jesia Quader Yuki, Sabyasachi Biswas, Md Ashraful Alam, Nabila Binte Alam, and M. R. C. Mahdy. "Improving malaria parasite detection from red blood cell using deep convolutional neural networks." *arXiv preprint arXiv:1907.10418* (2019*).*

[8]. Md. Khayrul Bashar "Improved Classification of Malaria Parasite Stages with Support Vector Machine Using Combined Color and Texture Features " *IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT)* 2019

[9]. Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. "Image analysis and machine learning for detecting malaria". *Translational Research,2018*

[10]. Park, Han Sang, Matthew T. Rinehart, Katelyn A. Walzer, Jen-Tsan Ashley Chi, and Adam Wax. "Automated detection of P. falciparum using machine learning algorithms with quantitative phase images of unstained cells." *PloS one 11, no. 9 (2016): e0163045.*

[11]. Vinayak K. Bairagi1 and Kshipra C. Charpe "Comparison of Texture Features Used for Classification of Life Stages of Malaria Parasite " *Hindawi Publishing Corporation International Journal of Biomedical Imaging Volume* 2016,

[12]. Akshay Nanoti, Sparsh Jain, Chetan Gupta, Garima Vyas "Detection of malaria parasite species and life cycle stages using microscopic images of thin blood smear".*International Conference on Inventive Computation Technologies (ICICT) August* 2016.

[13]. Dekel, E., Rivkin, A., Heidenreich, M., Nadav, Y., Ofir-Birin, Y., Porat, Z., & Regev-Rudzki, N. "Identification and classification of the malaria parasite blood developmental stages, using imaging flow cytometry Methods".2017

[14]. https://lhncbc.nlm.nih.gov/LHCresearch/LHCprojects/imageprocessing/ malaria-screeer.html

[15]. https://www.who.int/news-room/factsheets/detail/malaria

[16]. https://towardsdatascience.com/face-recognition-how-lbph-works

[17]. https://www.pyimagesearch.com/2015/12/07/local-binary-patterns-withpython-opencv/

[18]. T.Ojala.M.Pietikainen and D.Harwood,"A comparative study of texture measures with classification based on feature distributions"Pattern Recognition vol 29,1996.

Cite this Article

Miss. Spoorthi B, Dr. Aravinda C V, "Image Analysis for Detecting Malaria Cell Using Otsu Thresholding and Machine Learning Models", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 3, pp. 453-470, May-June 2022. Available at doi : https://doi.org/10.32628/CSEIT2283111

Journal URL : https://ijsrcseit.com/CSEIT2283111