International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Malaria Detection Using Machine Learning and Image Processing

MONISHA H R[1] , SHOBHA CHANDRA K [2], NIKHITHA A [3], NISHANTH M S[4],PAVITHRA S K [5]
[1]monisharajshekar1@gmail.com[2]ksc@mcehassan.ac.in, [3]nikhithanichi@gmail.com,
[4]nishanthgowda6992@gmail.com, [5]pavithrask53@gmail.com
Department of computer science and Engineering.
Malnad College of Engineering, Hassan

ABSTRACT - Malaria is a life-threatening disease caused by Plasmodium parasites and continues to be a significant global health concern, particularly in regions with limited access to healthcare facilities. Prompt and accurate diagnosis of malaria is crucial for effective treatment and control of the disease. So we proposed a novel approach for malaria detection using image processing techniques and machine learning algorithms. Machine learning algorithms are employed to train a classifier using the extracted features. Various supervised learning techniques, such as support vector machines (SVM), random forests, or convolutional neural networks (CNN), are explored for malaria classification. The classifier is trained on a large dataset of label led malaria-positive and malaria-negative images, ensuring the model's ability to distinguish between infected and uninfected cells accurately.

I. INTRODUCTION

Malaria is a mosquito-borne life threatening disease caused by Plasmodium parasite. Globally, an estimated 3.2 billion people are at high risk. According to the report, there were 212 million new cases of malaria worldwide in 2015 (range 148–304 million. The staining process somewhat colorizes the RBCs but highlight Plasmodium, WBCs and platelets. The detection of Plasmodium requires detection of the stained objects. However, we need to analyze stained objects further to determine if they are parasites or not to prevent false diagnosis.

Malaria parasite (MP) in blood sample can be identified by using image segmentation and feature extraction using minimum distance classifier . Based on Image Acquisition, Image Pre-processing, Image Smoothing, Thresholding and Dilation image segmentation is done. Feature extraction uses two phases in architectural model: Training Phase and Recognition Phase which helps to recognize the MP. In this work, we focus on automated detection and quantification of malaria detection, strategy to determine infected image using machine learning and discuss to improve the predictive value for detection of infected cells.

II. Literature Review

A. Ayesha Hoor Chaudhary, Javeria Ikhlaq, Muhammad Aksam Iftikhar, and Maham Alvi. Blood cell counting and segmentation using image processing techniques. Applications of Intelligent Technologies in Healthcare, pages 87–98, 2019

The results show that the accurate count of a patient's blood cells is vital for successful diagnosis of a plethora of diseases. Current systems deployed in Pakistan either rely on heavy and expensive machinery or is sometimes conducted manually. The propose the use of digital image processing techniques to build a cheaper alternative, that rely on digital images of blood smears, which are economical to produce, and are in fact a costless feature built-in to most existing lab microscopes. In this work, morphological image processing is deployed to segment the image and to differentiate and extract the blood cells from the plasma.[1]

B. S Afkhami and H Rashidi Heram-Abadi. Detection of malarial parasite in blood. 2020.

presents that there are many systems which describe the computerized methods of image analysis that commonly involves three main phases. In the first phase of preprocessing, luminance of the image is corrected and transformed to a constant color space. At the second step, a histogram-based image segmentation process is used which helps in avoiding maximum artifacts and over stained objects [2].

C. Ahmedelmubarak Bashir, Zeinab A Mustafa, Islah Abdelhameid, and Rimaz Ibrahem. Detection of malaria parasites using digital image processing. In 2019 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE).

proposes a parasite detection technique which is based on digital image processing. Images of thin blood smear are used and with the help of image processing approach the parasite in the cells are identified [3].

D. Lu´ıs Rosado, Jos´e M Correia Da Costa, Dirk Elias, and Jaime S Cardoso. Automated detection of malaria *parasites on thick blood smears via mobile devices. Procedia Computer Science, 90:138–144, 2019.*

This paper presents an image processing and analysis methodology using super vised classification to assess the presence of P.falciparum trophozoites and white blood cells in Giemsa stained thick blood smears. The main differential factor is the usage of microscopic images exclusively acquired with low cost and accessible tools such as smartphones, using a dataset of 194 images manually annotated by an experienced parasilogist[4].

E. Han Sang Park and T Matthew. 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.

This paper presents an automated analysis method for detection and staging of red blood cells infected by the malaria parasite Plasmodium falciparum at trophozoite or schizont stage. Unlike previous efforts in this area, this study uses quantitative phase images of unstained cells. Erythrocytes are automatically segmented using thresholds of optical phase and refocused to enable quantitative comparison of phase images. Refocused images are analyzed to extract 23 morphological descriptors based on the phase information[5].

III. Methodology

Our data consists of diverse datasets, including a wide range of images. Developing an algorithm that works well across such heterogeneous datasets is a challenging task. There are two main approaches we can consider: developing a robust algorithm that can handle the heterogeneity or transforming the data into a more homogeneous state.

A. Data Collection

The dataset includes a sufficient number of positive and negative samples to enable effective training and evaluation of machine learning models for malaria detection. Multiple image variations, such as different magnifications and lighting conditions, were included in the dataset to account for real-world variability. Publicly available datasets, such as the Malaria Dataset from the National Institutes of Health (NIH), were also included to augment the dataset and increase its size. The dataset was carefully curated to represent various malaria parasite species and morphological variations, ensuring the models' ability to generalize across different strains.

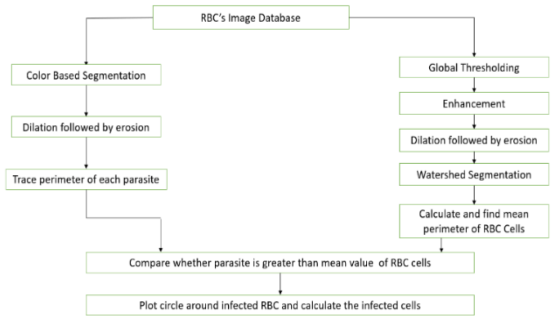

Figure 1. Methodology of Image Processing system for detection and quantification of plasmodium parasites.

B. Color Based Segmentation

Color based segmentation was employed to extract relevant regions of interest from blood smear images, separating them from the background and other non-relevant components. By applying color thresholding techniques, the image was segmented into distinct color regions, enabling the isolation of potential malaria parasites from the surrounding elements.

C. Global Thresholding

Global thresholding was employed to segment malaria-infected regions from blood smear images by selecting a single threshold value for the entire image.The grayscale intensity values of the blood smear images were used for global thresholding, allowing for the separation of infected areas from the background and healthy cells.

D. Watershed Segmentation

It is employed to separate individual cells and malaria parasites in blood smear images based on the local intensity gradients and spatial information. Machine learning models were trained to leverage the results of watershed segmentation for accurate identification and classification of malaria parasites. Watershed segmentation enabled the precise localization and separation of overlapping or clustered cells, allowing for accurate identification of malaria-infected cells.

E. Enhancement

Image enhancement is the process of adjusting digital image so that the result is more suitable than the original one and will be helpful for further processing.

F. Erosion and Dilation

Erosion and dilation processes were applied to blood smear images in malaria detection to refine the boundaries of detected regions and remove noise or artifacts. Erosion was used to erode the boundaries of segmented regions, helping to eliminate small false-positive regions and improve the accuracy of malaria detection. Dilation, on the other hand, was used to expand the boundaries of segmented regions, ensuring better connectivity and completeness of malaria-infected areas.

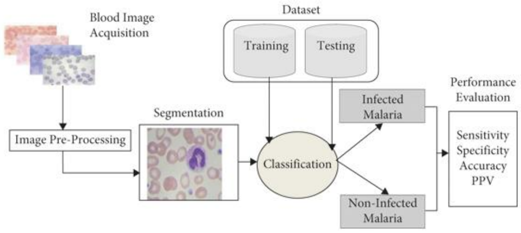

Figure 2 Methods for automated diagnosis system for malaria.

IV. Implementation

Preprocessing: Obtain a dataset of microscopic images of blood smears, including both infected and uninfected red blood cells(RBCs) with malaria parasites. Annotate the images to indicate the presence or absence of parasites. The input images are preprocessed to enhance features and remove noise. Common techniques include resizing, noise reduction, contrast enhancement, and color normalization. Segment the RBCs and parasites within the images using appropriate segmentation techniques such as thresholding, region-based segmentation, or edge-based segmentation. Refine the segmentation results using morphological operations or other post-processing techniques.

Feature Extraction: Relevant features are extracted from the preprocessed images. These features capture important characteristics of malaria-infected cells. Ensure that the extracted features capture meaningful information that distinguishes infected and uninfected RBCs. Feature extraction plays a vital role in malaria detectioning image processing and machine learning. It involves transforming the raw image data into a set of meaningful and representative features that capture relevant information for distinguishing between infected and uninfected red blood cells(RBCs) with malaria parasites. Here are some common feature extraction techniques used in this context: Shape-based Features, Texture-based Features, Intensity-based Features, Statistical Features.

Classification: Machine learning algorithms are employed to classify the extracted features into malaria-infected or uninfected classes. Machine Learning Algorithms are Support Vector Machines(SVM), Random Forest, Convolutional Neural Networks(CNN). Choose a suitable machine learning algorithm for malaria detection. Design the architecture of the machine learning model, taking into consideration the input features, the number of classes(infected and uninfected),and any specific requirements or constraints. Train the model using the training dataset, adjusting the model's parameters or hyperparameters to optimize its performance.

Model training and validation: Model training and validation are critical steps in developing a malaria detection system using image processing and machine learning. These steps involve training a machine learning model on a labelled dataset and evaluating its performance to ensure its effectiveness in detecting malaria-infected red blood cells(RBCs).The selected algorithm is trained on a labelled dataset, which includes images with known infection status. The dataset is split into training and validation sets. The model's performance is evaluated using metrics such as accuracy, precision, recall.

Testing and evaluation: Testing and evaluation are crucial steps in the development of a malaria detection system using image processing and machine learning. These steps involve assessing the performance of the trained model on an independent testing dataset to measure its accuracy, reliability, and generalization ability. Once the model is trained and validated, it is tested on unseen images to assess its performance on real-world data. The performance is evaluated using similar metrics as in the validation phase.

V. Results

Using the test results achieved from the application of pre-defined Machine learning algorithms and analyzing the factors that play a part in making the system perform better, we have implemented those Machine learning algorithms on the described dataset, which comparatively gave reliable results among all. More than 40 images have been tested and some of the tests images are listed below

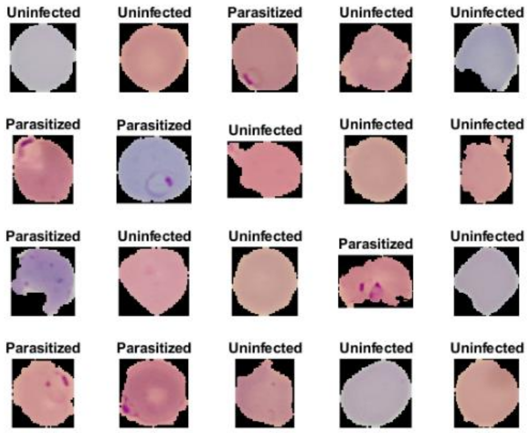

Figure 3. Tested Results

V. Conclusion

The integration of image processing and machine learning techniques has provided a great potential in the area of detection of malaria. By using Machine learning algorithms, such as convolutional neural networks (CNNs),Support Vector Machine(SVM), we can analyze and classify malaria-infected blood smear images. These algorithms can automatically learn and extract relevant features from the images, enabling accurate identification of malaria parasites. Image processing techniques play a crucial role in preprocessing and enhancing the quality of malaria images. Various methods, including image segmentation, noise reduction, and image enhancement, have been employed to improve the visibility of parasites and facilitate their detection. The combination of machine learning and image processing holds great promise for malaria detection.

References

[1] Ayesha Hoor Chaudhary, Javeria Ikhlaq, Muhammad Aksam Iftikhar, and Maham Alvi. Blood cell counting and segmentation using image processing techniques. Applications of Intelligent Technologies in Healthcare, pages 87–98, 2019

[2] S Afkhami and H Rashidi Heram-Abadi. Detection of malarial parasite in blood. 2020.

[3] Ahmedelmubarak Bashir, Zeinab A Mustafa, Islah Abdelhameid, and Rimaz Ibrahem. Detection of malaria parasites using digital image processing. In 2019 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE).

[4] Lu´ıs Rosado, Jos´e M Correia Da Costa, Dirk Elias, and Jaime S Cardoso. Automated detection of malaria parasites on thick blood smears via mobile devices. Procedia Computer Science, 90:138–144, 2019.

[5] Han Sang Park and T Matthew. 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.

[6] Usha Kumari, Madeha Muzafar Memon, Sanam Narejo, and Mashal Afzal. Malaria disease detection using machine learning. In 2nd International Conference on Computational Sciences and Technologies (INCCST 20), 2020.

[7] Negi Alok, Kumar Krishan, and Prachi Chauhan. Deep learning-based image classifier for malaria cell detection. Machine Learning for Healthcare Applications, pages 187–197, 2021