Plasmodium Infection Detection Using Machine Learning Algorithms

Dr. Shubhangi D C[1], Syeda Saima Batool[2]

[1]*Professor, Department of Studies in Computer Science and Engineering, VTU, CPGS, Kalaburagi, Karnataka, India*
[2]*Student, Department of Studies in Computer Science and Engineering, VTU, CPGS, Kalaburagi, Karnataka, India*

---***---

Abstract *- Malaria is parasitic infection spread via blood that is generated by Plasmodium parasites. It is spread via female anopheles mosquito bites. The most common and traditional method of detecting malaria is for trained technicians to visually examine blood smears under a microscope. This traditional technique is expensive and time intensive, and diagnosis is highly dependent on the examiner's expertise and understanding. Our primary objective is building prototype capable of detecting cells by pictures of numerous cells in thin blood smears on microscope slide and classifying them to be infected or uninfected. In addition, we use machine learning to classify the infected cell picture.*

Keywords: Plasmodium parasite, Wildlife, Microscope Slides, Machine Learning.

1.INTRODUCTION

Plasmodium infection detection with a microscope is a time-consuming and difficult task. This old approach necessitates the skills of a microscopist or a laboratory technician. Diagnosis is difficult in areas where malaria is prevalent, and therapies are focused solely on symptoms. For diagnosis, image recognition and computer vision techniques are used. A latest technique of computer vision based on the methodology used to differentiate Plasmodium parasites from light microscopy images was recently proposed. This is a pixel-based method for segmenting plasmodium parasite tissue that employs the K-means clustering algorithm. Photos taken with a regular microscope are used to identify the parasites found in blood smears. A few other studies went much further into clustering the various parasite species and phases of their life cycles. We don't want to fully eliminate human expert's diagnostic procedure, so image processing techniques are still used to some extent for final decision based on blood smears. By assisting lab technicians in triaging their focus and implementing the malaria diagnosis over a remote network link, this method would increase their efficacy.

2.RELATED WORK

In this paper, [1] G. Rollin, et al. took into account a network of 54,16,537 English Wikipedia entries retrieved in 2017 and created a reduced network of 11230 infectious disease articles and 195 world country articles using the newly advanced reduced Google matrix (REGOMAX) approach. The PageRank along with CheiRank algorithm were utilised for determining many influential disease such as TB, HIV/AIDS, and Malaria being top PageRank diseases. They used reduced Google matrix for determining susceptibility of global nations to particular illnesses, taking into consideration their impacts over past, comprising period of prehistoric Egyptian mummies. When acquired findings were compared to WHO data, it became apparent that Wikipedia network evaluation generates trustworthy findings including a precision of up to 80%.

In [2] W.M.I. Razza k, et al. deliberated state-of-art in architecture of deep learning and optimisation for medical image analysis as well as categorization. They reviewed problems of deep learning-based algorithm for medical imaging as well as an outstanding research question. Following accomplishment of deep learning including numerous real-world application, it's now giving intriguing clarifications for medical imaging with high precision, also seen as critical technique for future health-care application.

In this paper [3] Bias, S. D., et al. suggested an innovative, efficient, and low-complexity edge recognition technique, with a focus on the examination of microscopic thin blood smears infected with malaria. The method offers a simple, dynamic thresholding approach based on histogram analysis with the goal of collecting as much data about the blood cells as possible with the least amount of computing work, accompanied by morphological filter procedure to eliminate noise and artefacts.

In [4] G. Gopakumar, et al. employed a center stack of image obtained by bright field microscope to detect parasite, than a typical method of viewing a singular focused picture for parasite detection. Rather than using the traditional method of extracting specific characteristics, they used a Convolutional Neural Network, which could function straight on image without requirement of hand-engineered feature. They avoided the need for cell segmentation by working with picture patches at the presumed parasite location. They conducted tests and evaluated recognition accuracy when using a single focused picture against a focus stack of images.

In [5] Bashir, A., et al. proposed a method for detection of malaria that is reliant upon digital image processing. Images of thin blood smears are utilised, and parasite in cells is identified using an image processing approach. This strategy worked well for us in terms of providing enough motivation and direction for image processing techniques.

In [6] Z. Liang et al., uses a Convolution neural network approach for detecting parasite-infected red blood cells in tiny smear on routinely produced microscope slide. It is sparked by research into the physiological processes underpinning object recognition in feline visual cortex. The research work was very successful in classifying the individual cells but the research was based on individual RBCs rather than an entire blood sample.

3.PROPOSED SYSTEM

To detect plasmodium infected cells, we used image processing techniques. We classify the stage of malaria using machine learning technology, whether it is Healthy, the most dangerous stage of malaria, or Malaria infected. We apply an self-generated segmenter which executes four phases: First is, change to grayscale.

grayscale = ((0.3×R)+(0.59×G)+(0.11×B))

Secondly, Initialize segmentation with Otsu's threshold.

$$\sigma_{\omega}^{2}(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t)$$

where $\omega_1(t), \omega_2(t)$ are possibility related with two classes split with threshold t, whose significance is within the limit from 0 to 255 . Then sort components according to their area of origin. And by masked by input image as well as segment image.

Detecting edges: We utilised a variety of edge detection methods, including Canny, Gaussian Laplacian, Prewitt, Robert, Sobel, and Zero-cross operators, with LOG providing the best results.

edges = cv2.Canny(img, 100, 200)

If cell becomes infected during image processing phase, we perform check to determine if malarial parasite is falciparum or another kind. Machine learning algorithms such as Random forest, Adaboost are used, with finest algorithm having maximum precision score being chosen for classifying.

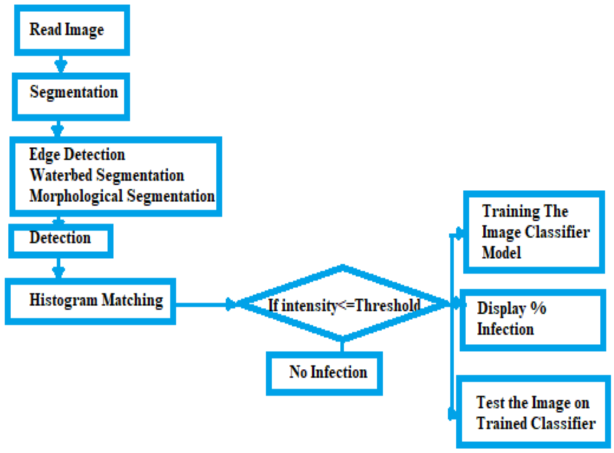


Figure 1: System Architecture

4.METHODOLOGY

The detection of Plasmodium parasites in single cell slides is currently done entirely by hand. This method could be improved by taking a picture of blood smear and then using suggested model for determining if cells are infected or not. Projected model employs image processing system for enhancing current techniques and reduce duration required for malaria parasite detection in blood test. Information was gathered by hand from the CDC's Division of Parasitic Infection and Malaria. The main goal is creating machine that can differentiate presence of the Plasmodium parasite in a small blood sample also quantify portion of RBC within contaminated sample. The first step is splintering contamination, which requires first segmenting cells. Procedures built on Edge Detection, Watershed Segmentation also Morphological Segmentation are used in segmentation techniques. The pathogens are segmented after the cells have been partitioned. It is accomplished by utilizing threshold intensity pixel value for infection. If pixel value falls within threshold's limit, infection is detected.

5.IMPLEMENTATION

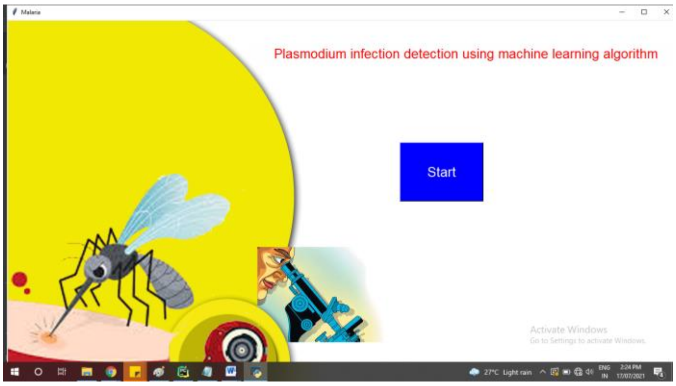


Figure 2: Home Screen Used to start the system

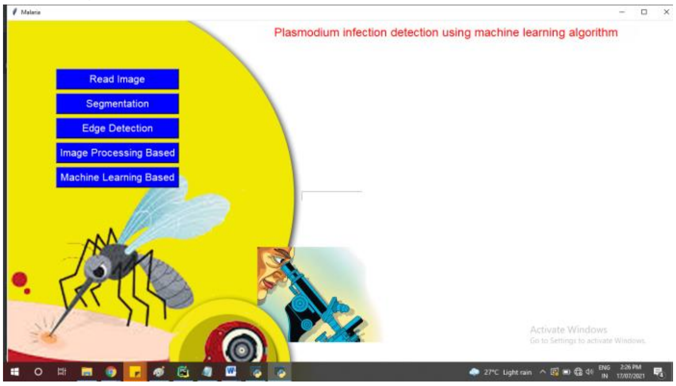


Figure 3: Menu Used for detecting Malaria

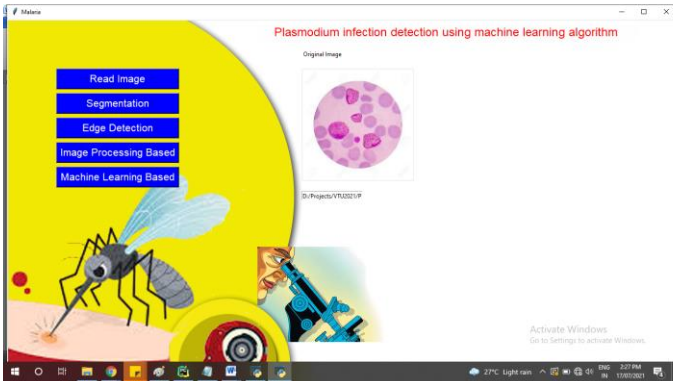


Figure 4: Read Image Used to select the image

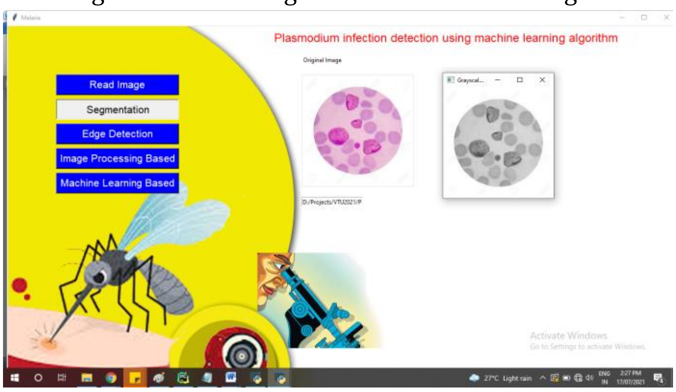


Figure 5: Segmentation Image Segmentation is procedure of classifying a image at the pixel level

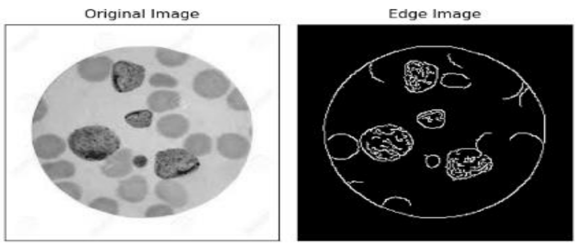
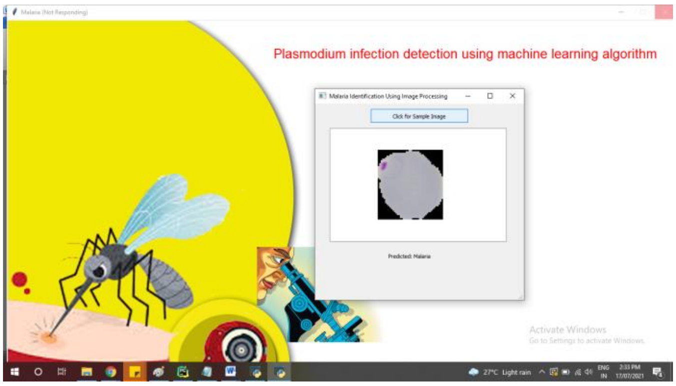


Figure 6: Edge Detection Edge detection is a technique for determining the boundaries of objects within photographs.



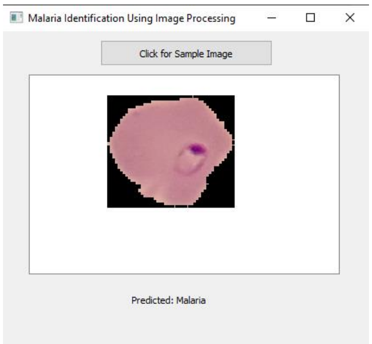


Figure 7: Predicted Malaria

```
              precision    recall  f1-score   support

 Parasitized       0.90      0.91      0.90      2765
  Uninfected       0.90      0.89      0.90      2747

    accuracy                           0.90      5512
   macro avg       0.90      0.90      0.90      5512
weighted avg       0.90      0.90      0.90      5512

Adaboost Accuracy: 0.8998548621190131
```

Figure 8: ML Analysis Using Random Forest and Adaboost

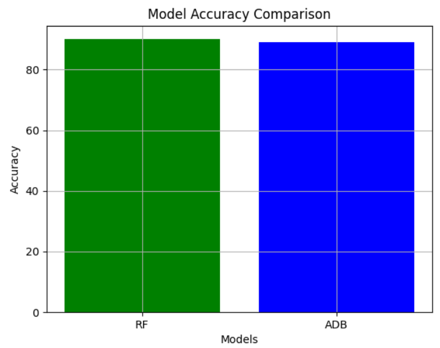


Figure 9: Model Accuracy

Model	Accuracy
Random Forest	90%
Adaboost	89%

Table 1: Model Accuracy Table

6.CONCLUSION

Malaria is deadly but preventive disease. In majority of times, a pathologist uses a microscope to detect the disease. With the aid of machine learning algorithms and image processing techniques, there has been an increasing trend in recent years for utilization of computerized system for detection of malaria disease. In proposed system we have used Adaboost and Random forest machine learning algorithms. This system helps the physician in diagnosing the plasmodium parasite. The obtained research outcomes revealed that applied method produces precise reliable results also makes fast decisions. We got the accuracy of the models i.e. Random Forest 90% and Adaboost 89% using this. In future, we can implement for some more machine learning classification algorithms and image processing techniques that can give better accuracies in detecting plasmodium parasite.

REFERENCES

[1] G. Rollin, J. Lages and D. L. Shepelyansky, “World Influence of Infectious Diseases from Wikipedia Network Analysis,” in IEEE Access, vol. 7, pp. 26073-26087, 2019

[2] W.M.I. Razzak, S. Naz, A. Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and the Future”, Lecture Notes in Computational Vision and Biomechanics book series (LNCVB, volume 26), 323-350, 2018.

[3] Bias, S. D., Reni, S. K., & Kale. “A novel fuzzy logic inspired edge detection technique for analysis of malaria infected microscopic thin blood images”, 2017

[4] G. Gopakumar, M. Swetha, S. Siva, and G R K S Subrahmanyam. 2016. Automatic detection of Malaria infected RBCs from a focus stack of bright field microscope slide images. In Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing.

[5] Bashir, A., Mustafa, Z. A., Abdelhameid, I., & Ibrahem, R. (2017). Detection of malaria parasites using digital image processing. 2017 International Conference on Computing and Electronics Engineering(ICCCCEE) doi:10.1109/iccccee.2017.7867644.

[6] Z. Liang et al., "CNN-based image analysis for malaria diagnosis," 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, 2016, pp. 493-496, doi: 10.1109/BIBM.2016.7822567

[7] H.S. Park, M.T. Rinehart, K.A. Walzer, J.T.A. Chi and A. Wax, Automated Detection of P.falciparum Using Machine Learning Algorithms with Quantitative Phase Images of Unstained Cells. PLoS ONE 11(9).

[8] S. S. Savkare, S. P. Narote,”Automatic Detection of Malaria Parasites for Estimating Parasitemia”, International Journal of Computer Science and Security (IJCSS), Volume - 5 Issue - 3, pp: 310 - 315, 2011.

[9] K. Torres, et al. ”Automated microscopy for routine malaria diagnosis: a field comparison on Giemsa-stained blood films in Peru”. Malaria journal, 17(1), 339. 2018.

[10] M. Kanojia, N. Gandhi, L. Armstrong and P. Pednekar, ”Automatic Identification of Malaria Using Image Processing and Artificial Neural Network”, Advances in Intelligent Systems and Computing, 846-857, 2018.

[11] Amer, G. M. H., & Abushaala, A. M. (2015). Edge detection methods. 2015 2nd World Symposium on Web Applications and Networking

[12] Chourasiya, S. (2014). Automatic Red Blood Cell Counting using Watershed Segmentation.

[13] Bhusare, M.P., & Bhosale, S.D. (2019). Automatic Blood Cell Counting using Morphological Image Processing operations.

[14] Mohamad Sharif, Johan & Miswan, M. & Ngadi, Md & Abdul Jamil, Muhammad Mahadi. (2012). Red blood cell segmentation using masking and watershed algorithm: A preliminary study 2012 International Conference on Biomedical Engineering,ICoBE 2012 10.1109/ICoBE.2012.6179016.

[15] Shapira, D., Avidan, S., & Hel-Or, Y. (2013). Multiple histogram matching. 2013 IEEE International Conference on Image Processing, doi:10.1109/icip.2013.6738468.

[16] Sowmya V, Mohan N, & Soman K. P, Edge detection using sparse banded filter matrices, Procedia Computer Science, 58(2015), 10-17.

[17] Ross NE, Pritchard CJ, Rubin DM, Duse AG, Automated image processing method for the diagnosis and classification of malaria on thin blood smears, Medical and Biological Engineering and Computing,44(2006),427-436.

[18] Dobson A, Lafferty KD, Kuris AM, Hechinger RF, Jetz W, Homage to Linnaeus: how many parasites? How many hosts?, National Academy of Sciences,105(2008), 11482-11489.

[19] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J,Scikit-learn: Machine learning in Python, Journal of machine learning research,12(2011), 2825-2830.

[20] Bradski G, Kaehler A, Learning OpenCV: Computer vision with the OpenCV library, O'Reilly Media, Inc.,(2008).