IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

A Contrast of the Efficiency of Machine Learning Techniques for Identifying Signs of Malaria Employing Microscopically Images

[1]G N Beena Bethel, [2]B Pushpa, [3]Santhoshi M

[1]Professor, [2]Assistant professor, [3]UG Student [1,2,3]CSE Dept, GRIET

E-Mail: beenabethel@gmail.com, Vetukuripushpa1608@ grietcollege.com, santoshimummidi26@gmail.com

Abstract

Plasmodium parasites are a component of malaria, a blood-borne illness increase by mosquitoes. The conventional approach to diagnosing malaria is making a blood smear and by means of a microscope to look at the blood-stained smear in order to identify the pathogen genus Plasmodium. This approach is highly dependent on the knowledge of qualified specialists. Within the cover of this research, simple machine learning techniques are utilized over the conventional method that has certain problems regarding sympathy and specificity, in order to separate out the parasite from blood smears for malaria recognition. The suggested methodology uses patient photos to identify the presence of malaria with no the requirement for specialists or blood staining.

KeyWords: Plasmodium and malaria, blood smears, microscopy, machine learning, decision trees, random forest trees, Adaboost, and logistic regression

INTRODUCTION

According to a World Health Organization investigation, malaria might strike almost 33 million people [1]. Malaria is a blood-borne illness that is transmitted by the Anopheles mosquito species and is brought on by Plasmodium-infected red blood cells. A someone by malaria will exhibit a wide variety of clinical symptoms, from extremely mild to severe instances, and may even pass away. Malaria illness identification under a microscope is a laborious and time-consuming process. The extensive knowledge of a laboratory technician or microscopist is required for this conventional approach. In fact, a skilled malaria microscopes is crucial to the identification of parasites[2], [3]. It has been stated that of the 300–500 million belongings of sensitive malaria, 1-3 million are almost fatal, based on study done in [5], [8], and [13]. In areas where malaria is a severe problem, diagnosis is extremely challenging, and treatments are only provided based on symptoms.

In underdeveloped nations similar to Uganda [14], wherever barely half of the country health centers have microscopes and almost simply one-fourth of them have laboratory personnel skilled in diagnosing malaria, illness diagnosis is a significant challenge. Additionally, early and more accurate disease detection is crucial since it may facilitate the early administration of treatment to the diagnosed patient. Furthermore, false positives may increase needless financial burden and drug resistance, while false negatives may result in fatalities [21], [16]. Thus, the necessity to design an alternative diagnosing approach arises.

For diagnostics, computer vision techniques and image processing can be used. A new computer vision technique depending on the way of identifying the MP (Malarial Parasite) as of the light microscopy images was recently proposed by Khan and his hardworking team. This is a pixel-based method for segmenting tissue containing malaria parasites using the K-means clustering algorithm [22]. The machine learning algorithms in [4] had access to enough training data. Images in use by a conventional microscope are worn to recognize the parasites present in the blood smear. A small number of further researches focused even more on grouping the many species and phases of the parasite's life cycle. Image processing techniques are still used because we don't desire to totally replace human diagnosticians, just somewhat, so that a blood smear can at rest be used to make a final choice. By assisting in the implementation of the malaria diagnostic through a remote network connection and in the triage of their concentration, this procedure will increase the efficacy of lab personnel.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

In low quality blood smear images, the automated detection of malaria is the major emphasis of this work. This is achieved by restrict and categorizing the contaminated erythrocytes from the vigorous ones. Because standard algorithms are unable to process these poor-quality images, we turn to classical machine methods. As a result, our system is capable of detecting malaria without the need for human interaction, or at the very least, it can benefit technicians by reducing their workload and potentially improving the diagnosis accuracy.

Literature Review

Plasmodium is a parasite that feeds on red blood cells (RBCs) and is spread by mosquitoes. Malaria can range in intensity from moderate to quite dangerous, and it ultimately kills people. The potential presence of parasites and red blood cells in the blood smear is examined using neural networks. In [19], the Bayesian pixel classifier—which stain pixels—is worn to train the weighted KNN (K-Nearest Neighbors) method with the acquired data. In [15], an effort is completed to recognize multi-class parasites based on their lifecycle stage and kind. A histogram-based technique called basic thresholding is suggested in [9] to determine whether Plasmodium is present in blood smears. Smear preparation is crucial since variations in it could lead to changes in imaging conditions as well [7].

Morphological operators are worn to divide the overlapped RBCs [10], [11]. The irregularities are discovered by the analysis of blood cell pictures, wherever cells found in the image are labeled after the genuine image is binarized using a fuzzy measure approach [6]. Additionally, these labeled cells are divided into leukocytes, erythrocytes, and platelets employing a hierarchical neural network design that takes into account characteristics including color, size, and features. [20] describes an approach that is divided into four phases: edge detection, edge linking, clump splitting, and sponge detection. This approach uses adaptive histogram equalization as pre-processing. Relying on traditional supervised classification models, the color segmentation method in [20] divides the pixels into erythrocyte, parasite, and background categories. Various color models, such as the RGB (Red, Green, Blue), normalize RGB, HSV (Hue Saturation Value), and YCbCr models, been used to evaluate supervised categorization performance such Support Vector Machines, K-Nearest Neighbor, and Naive Bayes, respectively.

In the past few years, a lot of novel techniques for diagnosing malaria have been developed. These techniques include the fast antigen, fluorescence microscopy detection method, and PCR (Polymerase Chain Reaction) technique, that can recognize specific nucleic acid sequences [17]. Despite this, the most popular and widely applied method is the light microscopy diagnosis method [18]. The tumor is diagnosed using X-ray image categorization, and edges are found utilizing sparse banded filter matrices. Microscopy can measure parasitemia, distinguish among various species, and look at the many asexual phases of the parasite [18] and [20]. However, this method requires a skilled technician, and it takes time. The accurateness of the analysis depends on the microscopist's expertise and skill level as well as how much time they spend memorizing every slide. [20].

Proposed Model

The image dataset included in this work was taken at a 1000x magnification utilizing an oil immersion objective lens on 133 different people [4]. The photos that were difficult to identify as parasites, out-of-focus, and of low quality were removed. In conclusion, the collection includes 2703 blood smear pictures featuring bounding boxes containing 50,255 malaria parasites. Next, with the bounding box, every picture is divided into overlapping patches that are labeled as 0 or 1. Twenty-seven blood smear photos, or 75% of the labeled data, are utilized in the training dataset. The enduring 25% of the data, or 676 blood smear images, was utilized for testing purposes. Of these, 16312 patches have been identified as parasites.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

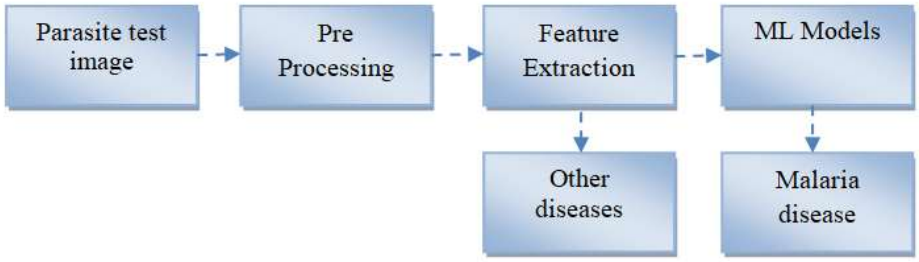


Figure 1: An approach that is suggested to use machine learning techniques to identify malaria in microscopic pictures

Figure 1 demonstrates the way the patient's test image is feature-engineered and pre-processed prior being fed into the machine learning system. Next, the existence or missing malaria in the imagepatcha is classified using a binary classification technique.

A 1024 x 768 image with a label of 0 or 1 was divided keen on approximately 475 overlapped patches. The size of every image patch is 50 by 50 pixels. We present the task of malaria detection as a model of binary classification using this dataset of annotated picture patches. For this sorting task, the raw format pixel data in image patches will not be immediately usable. Rather, we utilize a demonstration that is unaffected by translation, rotation, and continuous amount offsets. The primary concern in the Plasmodium recognition problem is the form of substance in the input patch. If the photos are gathered in varying sizes, we must scale them. We also need a representation that is invariant to rotation, translation, and intensity.

A crucial first step in creating an automated malaria diagnosis system is the engineering of features. In order to recognize additional associated problems with the same platform in the future, we first have to determine an illustration of the data that yields satisfactory results for identifying the plasmodium. After that, we need to have a specific illustration of the shapes found in the images that included blood smear, excluding objects like leukocytes or various hemoparasites. In general, color information is additionally very helpful, but when using blood films that are discolored by the field's stain, it is useless. Consequently, for this assignment, statistical descriptions of the geometries are used. For the purpose of feature extraction, the color patches often need to be converted to grayscale patches. This uses two different kinds of features: one that is generated from connected machinery and another that is generated by computing the moment's patches and threshold at various levels.

The identification of malaria can be approached as a classification problem using the labeled image patches; that is, identifying either 0 (another disease) or 1 (malaria sickness). To identify malaria, we employ a number of machine learning techniques, as well as AdaBoost, Random Forest, Decision Trees, and KNN. At 0.965 accuracy, Random Forest executes remarkably well in identifying malaria cases. Instead of evaluating every individual's performance at the level of their complete image, the current work bases its evaluation on the existence of parasites at the patch level. If present is at least one positive patch in the sample of photos, the person is considered infected. Given that the photos used in our research came from malaria-infected individuals, it is not feasible to provide sensitivity and specificity data for individual patients.

Technicians can quickly make the selection by using this system as an assisting method. As a result, the specialist attention is narrowed to only focus on the things in the microscope images that have the highest likelihood of carrying Plasmodium. A different, more sensitive threshold is worn for this motivation. We utilize different categorization levels to attain distinct false positives and negatives.

Python2 is used in conjunction by Sci-Kit Learn [26] and OpenCV2 [27] to construct this system. The CPU system used for this experiment had an i7 processor setup with 32 GB of RAM.

Results & Analysis

The present study uses a picture dataset that was sourced from [4]. 75% of the label information, or 2027 blood smear images with 37550 patches identified as parasites, is utilized in the training dataset. The remaining 25

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

percent of the data, or 676 blood smear images with 16312 patches marked as parasites, will be utilized for testing.

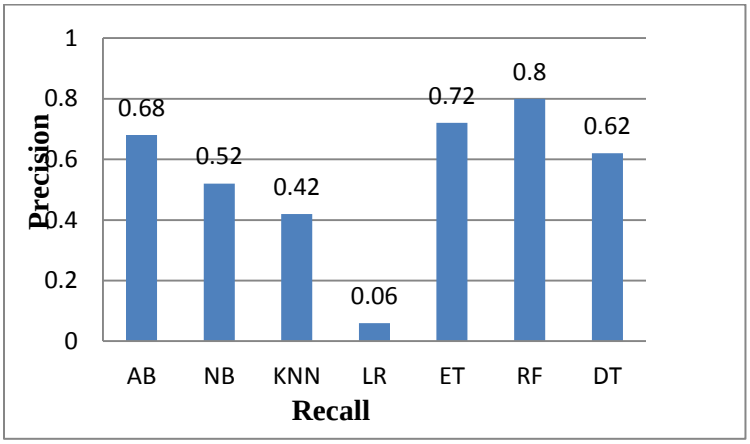
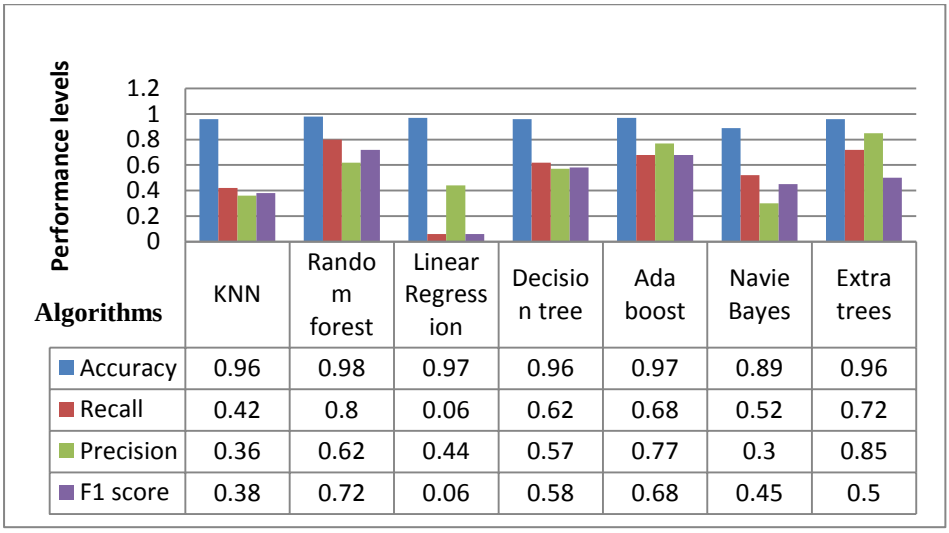


Fig .2. Precision-Recall Curve

The graph of the Precision-Recall (PR) Curve in Figure 3 illustrates the potential trade-offs among raising sensitivity and lowering the false alarm rate. The associated 20% recall would result in higher than any approach employed for assessing thin blood smears utilizing the identical number of fields of vision if the 90% precision was employed as the detection threshold. According to Fig. 3, if we were to use this recognition system for fully automatic malaria diagnosis, we might achieve a false alarm rate of less than 1/10 and a recall value of about 1/5 of what an experienced technician might achieve, meaning that the automated diagnosis system would perform nearly five times as well as a human technician.



	KNN	Random forest	Linear Regression	Decision tree	Ada boost	Navie Bayes	Extra trees
Accuracy	0.96	0.98	0.97	0.96	0.97	0.89	0.96
Recall	0.42	0.8	0.06	0.62	0.68	0.52	0.72
Precision	0.36	0.62	0.44	0.57	0.77	0.3	0.85
F1 score	0.38	0.72	0.06	0.58	0.68	0.45	0.5

Figure.3. Comparison of the efficacy of the suggested method for detecting malaria with traditional machine learning methods

As shown the figure 3 the results of the various machine learning methods. The precision, accuracy, Recall, and F-Score statistics are tabulated. Random Forest fared the best out of the seven customary machine learning methods we tested, closely surpassed by Ada Boost.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

Conclusion

It optional a shallow machine learning algorithm-based technique for detect malaria parasites. Health professionals can find great value in this way of identifying the malaria parasite in nations with limited resources and a shortage of skilled laboratory experts. In the current study, we separated the image into patches and performed an analysis according to the occurrence or disappearance of the malarial parasite. We include employed a number of traditional machine learning methods, including AdaBoost, Decision Tree, KNN, Random Forest, etc., to achieve this. Our model's accuracy helps laboratory professionals make decisions by highlighting areas of images that are vulnerable to Plasmodium infection. Data collection is as well aided by automated malaria diagnosis. Additionally, our feature extraction and classification methodology might be broadly applicable to additional diagnostic tests for instance those for tuberculosis, infestations of worms, or hemoparasites.

References

1. W. H. O. (World H. Organization), *Global Health Observatory -Malaria*, (2011), Available: http://www.who.int/gho/malaria.
2. Mehrjou A, Abbasian T, Izadi M, Automatic malaria diagnosis system, *International Conference on Robotics and Mechatronics(ICRoM)*, (2013), 205-211.
3. McKenzie FE, Sirichaisinthop J, Miller RS, Gasser Jr RA, Wongsrichanalai C, Dependence of malaria detection and species diagnosis by microscopy on parasite density,*The American journal of tropical medicine and hygiene*, **69**(2003), 372-376.
4. Quinn JA, Andama A, Munabi I, Kiwanuka FN, Automated blood smear analysis for mobile malaria diagnosis, *Mobile Point-of-Care Monitors and Diagnostic Device Design*, (2014), 31-115.
5. Rafael ME, Taylor T, Magill A, Lim YW,Girosi F, Allan R, Reducing the burden of childhood malaria in Africa: the role of improved, *Nature*,(2006), 39-48.
6. Kim KS, Kim PK, Song JJ, Park YC, Analyzing blood cell image to distinguish its abnormalities (poster session), *ACM international conference on Multimedia*, (2000), 395-397.
7. Basic Malaria Microscopy: Tutor's guide, *WHO(World Health Organization*, (2010).
8. Roca- Feltrer A, Carneiro I, Armstrong Schellenberg JR, Estimates of the burden of malaria morbidity in Africa in children underthe age of 5 years, *Tropical medicine & international health*, **13**(2008), 771-783.
9. Anggraini D, Nugroho AS, Pratama C, Rozi IE, Pragesjvara V, Gunawan M, Automated status identification of microscopic images obtained from malaria thin blood smears using Bayes decision: a study case in Plasmodium falciparum, *International Conference on Advanced Computer Science and Information System (ICACSIS)*,(2011), 347-352.
10. Di Ruberto C, Dempster A, Khan S, Jarra B, Automatic thresholding of infected blood images using granulometry and regional extrema, *Pattern Recognition*,**3**(2000), 441-444.
11. Di Ruberto C, Dempster A, Khan S, Jarra B, Analysis of infected blood cell images using morphological operators, *Image and vision computing*, **20**(2002),133-146.
12. Díaz G, Gonzalez F, Romero E, Automatic clump splitting for cell quantification in microscopical images, *Iberoamerican Congress on Pattern Recognition*, (2007), 763-772.
13. Samba EM, The burden of malaria in Africa, *Africa health*,**19**(1997), 17.
14. Tumwebaze M, Evaluation Of The Capacity To Appropriately Diagnose And Treat Malaria At Rural Health Centers In Kabarole District, Western Uganda, *health policy and development*,**9**(2011),46-51.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 11, 2022

15. Tek FB, Dempster AG, Kale İ, Parasite detection and identification for automated thin blood film malaria diagnosis, *Computer vision and image understanding*,**11**(2010), 21-32.
16. Lee JH, Jang JW, Cho CH, Kim JY, Han ET, Yun SG, Lim CS, False-positive results for rapid diagnostic tests for malaria in patients with rheumatoid factor, *Journal of clinical microbiology*, **52**(2014), 3784-3787.
17. Haditsch M, Quality and reliability of current malaria diagnostic methods, *Travel medicine and infectious disease*,**2**(2004), 149-160.
18. Thung F, Suwardi IS, Blood parasite identification using feature based recognition, *International Conference on Electrical Engineering and Informatics*,(2011), 1-4.
19. Tek FB, Dempster AG, Kale I, Malaria Parasite Detection in Peripheral Blood Images, *British Machine Vision Conference*, (2006), 347-356.
20. Pammenter MD, Techniques for the diagnosis of malaria, *South African medical journal= Suid-Afrikaanse tydskrif vir geneeskunde*,**74**(1988), 55-57.
21. Agnihotri N, Agnihotri A,Wrong Sample Dispensing May Cause False Positive Malaria Test, *Journal of Clinical and Diagnostic Research*,9(2015), EG01-EG02.
22. Khan NA, Pervaz H, Latif AK, Musharraf A, Unsupervised identification of malaria parasites using computer vision, *International Joint Conference onComputer Science and Software Engineering*,(2014), 263-267.