



Malaria Prediction System using Deep Neural Network and Genetic Algorithm

¹Tawonga Mkandawire, ²Dr. Glorindal Selvam

¹ Master of Science in Computer Science student, ²Supervisor, HOD and Post graduate coordinator

^{1,2} School of Computer Science,

^{1,2} DMI ST Eugene University, Chimombo, Zambia

Abstract- Malaria is a global health issue that has affected most developing countries in Africa and several approaches have been used to predict malaria such as the golden standard (microscope) and several systems were developed using machine learning techniques. However, the performance metrics were not up to the optimal best. The gaps lead to development of a malaria detection system that works to automate the process of detection in blood smears by using malaria detection models such as deep neural network, which has a greater efficiency and accuracy thus used for training and testing. The genetic algorithm is used to enhance the accuracy in the neural network in order to improve the predictability of the infected and non-infected cells. The CNN classifier is used to transfer learning and classification this works in line with the Resnet-50 model which works best for deep networks thus good to work on the features to have a right prediction thereby having true positive and negative predictions.

Index Terms - Malaria, Deep neural network, CNN classifier, Genetic algorithm, Resnet- 50 model, RDT

I. INTRODUCTION

Malaria is a disease transmitted by mosquitoes and many cases come from the WHO Africa region. According to WHO between 2019 and 2020, malaria cases increased from 213 million to 228 million and deaths from 534 000 to 602 000 in the African region. This shows that 95% of cases and 96% of deaths globally were found in the region were 80% of all deaths are among children aged under 5 years. WHO further states that based on the need to harness innovation an investment is required to accelerate research and development, currently no tool that is available today will solve the problem of malaria in moderate and high burden settings. This will bring new diagnostics, vector control approaches that can help to speed the pace of progress against malaria and attain global targets [7].

The traditional approach of diagnosing malaria requires the use of symptom based analysis, microscopes and RDT [7]. Microscopy examination uses stained blood smear, which are examined by a microscope. The examination uses thin and thick blood smears that consists of a single layer of red blood cells (RBCs) while the other is 6 to 20 times thicker, allowing for a greater volume of blood to be examined. Malaria diagnosis through RDTs is detected on specific malaria antigens in the blood. This requires time and skill to have a better throughput that will lead to the right prediction. The lack of reliable resources and maintainability on microscopes affects the standard measuring thus mislead predictions and causes more deaths and server illness which costs a lot of money for governments [10]. However, advancements are there according to recent researches in artificial intelligence; this is addressing the gaps due to the development of malaria prediction systems that work to automate the process of detection in blood smears. were most focused on malaria detection using image processing and machine learning that brought an accuracy of ≥ 90 [2, 6], deep learning based approach for malaria detection in blood cell images using a small neural network [5]

This paper will present a malaria prediction system that uses the deep neural network and genetic algorithm to optimize and give the correct predictions by having cross validations, but also accommodation of huge datasets due to the usage of multi-layer

networks. It also has several sections such as literature survey in section II followed by materials and methods in section III and results and discussion for malaria prediction and conclusion in section IV and V respectively.

II. LITERATURE SURVEY

Research has been done to improve the automated prediction of malaria by using various approaches, several related work was analyzed to work on the proposed system simulation. The literature [1] focuses on an accuracy matrix which was highly based on the feature extraction of thin blood smears and classification of infected and non-infected cells using CNN. The model used a shallow network and smaller images. Some studies worked on detection by using morphological segmentation based on segmenting cells from the non- region of interest portion, machine learning helped to have a high score accuracy though the inconsistencies among the images was a challenge [2]. The usage of deep belief network, restricted boltzmann machine generic pattern recognition and deep learning for accuracy and cost-effective diagnosis of malaria in rural areas when experienced lab technicians are not available. Though it has a challenge in sensitivity [3]. Suman Kunwar and team proposed an approach to detect malaria by using Watershed segmentation, support vector machines automated malaria detection and quantification of malaria infection based on training with machine learning in order to have predictive value [5]. Recent studies show how deep learning models, random Forest, Decision Tree, Accurate and effective period protozoal infection detection based on input pictures and to scale back manual labor with a mobile application. The is need to work more on staining and lighting variations [6]

Micheal Olaolu and team proposed evaluation of dataset using Decision, it has the reduction of dimensionality and classification by using a tree based approaches. Genetic algorithm is used to work on feature extraction along with other algorithms in order to test the performance of decision tree [11]. Furthermore, Literature[9]ANN and SVM was proposed for detection .The experimental results show that ANN is more accurate than SVM, having 94.1% accuracy compared with SVM's 92.3% and reduction in time build the model was very clearly, 13.6 second and 5.8 times faster respectively after applying knowledge discover technique.

Most researches have worked on machine learning [8] and smaller networks [4,1], though a recent studies focused on deep learning[13] the was a need to work on optimizing the results based on tuning thus prompting a further study on how a hybrid deep neural network-genetic algorithm can be used to predict malaria without overlooking several performance metrics in order to have a robust and effective system

III. MATERIALS AND METHODS

The proposed system is an automated deep neural network model to diagnose and predict malaria in line with a genetic algorithm from microscopic blood smear images, thus detecting the infected cells with the accuracy of 98.5%. This system uses a deep neural network model called Resnet to pretrain for image classification, this is ideal as it is 50 Layers deep. The system has three blocks were the first works on having the data validated and preprocessed while the middle will have the feature extraction then lastly have a comparison and work on the prediction based on probability. The overall system design works to be reliable due to the delicate operation of the medical sector, this lead to a work flow which ensures that that error rate has to minimal by working on the parameters which affect the status of the prediction.

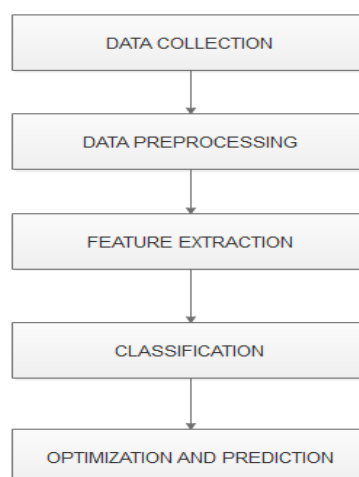
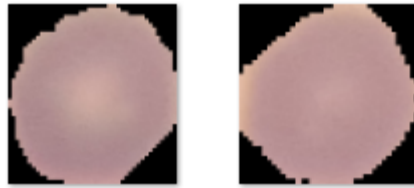


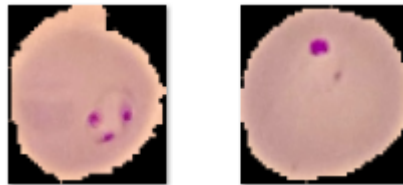
Figure 1: Block diagram for malaria prediction

A. DATA COLLECTION AND PREPROCESSING

Data is collected from a dataset of microscopic blood smears which is available at: <https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria>, it has 27,558 images, which consists of two classes infected and uninfected cells.



Non-infected malaria blood smear samples



Infected malaria blood smear samples

TABLE 1. DESCRIPTION OF DATASET

Descriptions	Infected	Uninfected
Number of images	13,779	13,779
Format	PNG	PNG
Position of Image	V and H	V and H
Color Model of Images	RGB	RGB

After the data is available preprocessing is done, this involves the resizing of image and normalizing the image using a preprocess input method. Splitting of data into training and testing data with the ratio of 75 and 25 % respectively, but also checking how robust the network can be by performing cross validation. Preprocessing is required as it helps to remove noise and enhance the quality of images through data augmentation.

B. FEATURE EXTRACTION

This is done at the middle layer and it works as an automatic feature extractor that extracts hidden and important features. Extracted features are passed to a fully connected neural network which performs classification images by maximizing the probability scores. The features which give more dominant difference between normal cells and infected cells are identified as feature set, this focuses on geometrical and textural features.

C. CLASSIFICATION

The malaria parasite classification used ResNet-50, which uses a deep network This is used to extract information of image classes digitally by using a supervised concept which classify images by labelling. The extracted test features are pass the image features to a classifier and get the known labels thus decides the status of RBC and identification of infected stages.

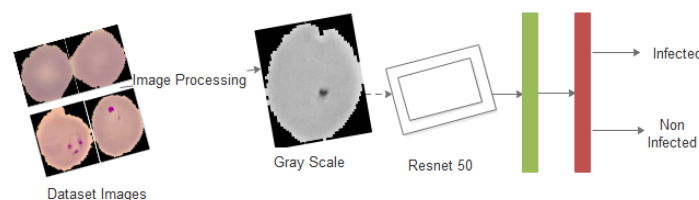


Figure 2: Classification model diagram

D. OPTIMIZATION AND PREDICTION

The genetic algorithm will work on comparisons thus having a probability score, this will work on the optimization of parameters to enhance the prediction in order to have a throughput that is reliable. This is based on the production of reasonable quality solutions with a sample of randomly generated outcomes by having a fitness criterion which works to optimize the deep neural network if it gets to the local optimization based thus focusing on a global approach.

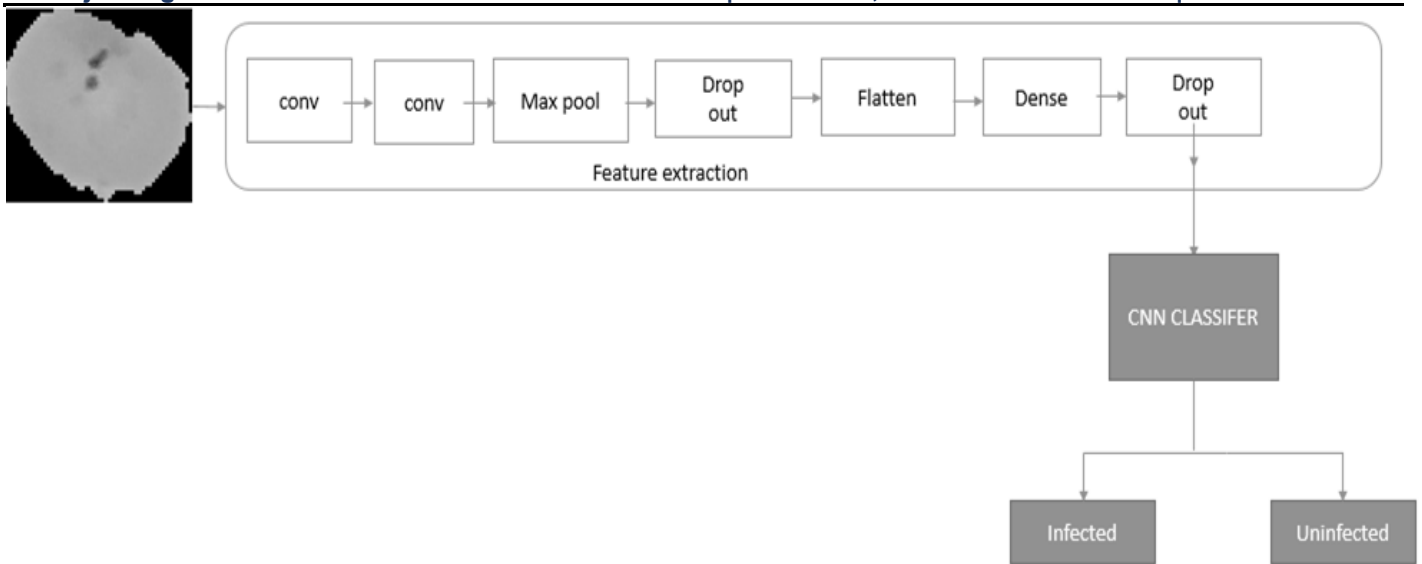


Figure 3: CNN ARCHITECTURE

IV. RESULTS AND DISCUSSION

The test of the simulation is focused on performance metrics that has several parameters, which will help in having an efficient and reliable system, the results were based on tests that were made in both categorized folders. The status was checked by considering the true values but a check was done on false predictions that were present either due to labeling or other factors of evaluation.

TABLE 2. EVALUATION METRICS

Parameter	Operation	Formula	Value in %
Accuracy	The <i>accuracy</i> is how close a measured value is to the actual (true) value.	Number of correct predictions / Total number of predictions OR $TP+TN/TP+TN+FP+FN$	98.6
Sensitivity (recall)	The percentage of the people with the disease that test positive	$TP/TP+FN$	99
Specifications	The number of people not having the disease that test negative	$TN/FP+TN$	97.9
Classification error	It is the number of cases which are incorrectly identified.	$FN+FP/TN+TP+FP+FN$	0.01

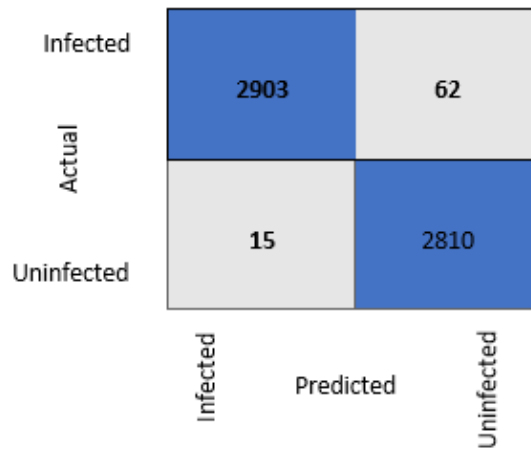


Figure 4: Confusion matrix

The predictions were 5790 with the diagonal and off diagonal elements that represent true predictions and false respectively. This is used to work on the evaluation metrics in table 2 were TN, TP stand for (True negative and positive) and FN, FP stand for (False negative and false positive).

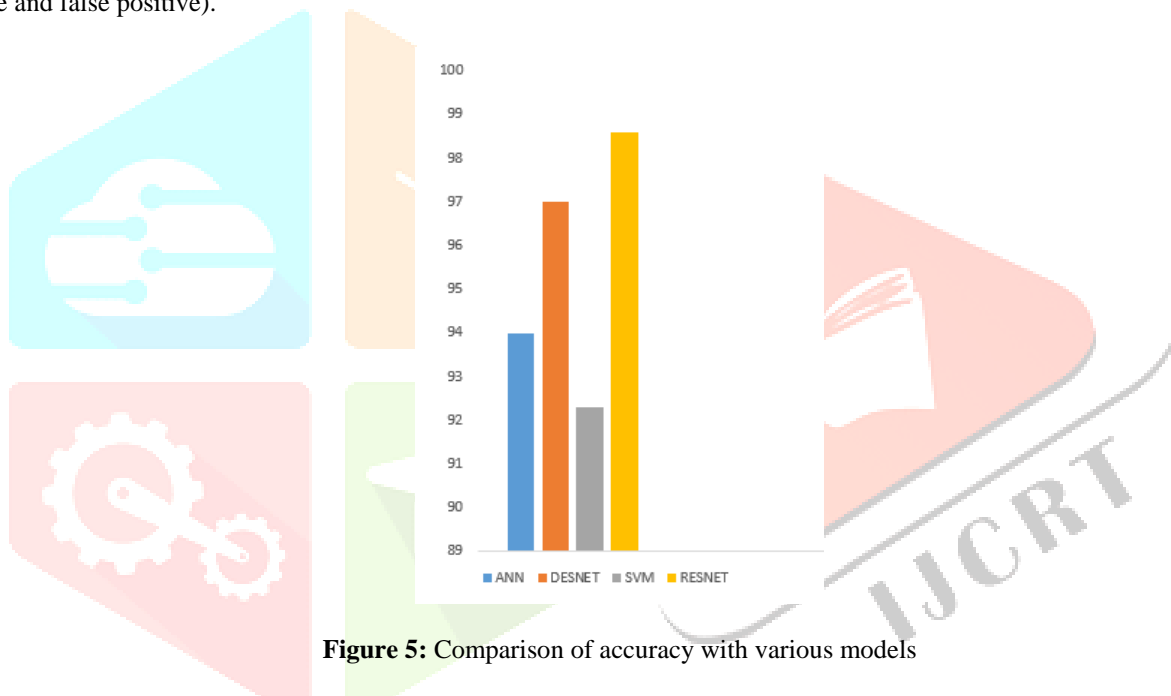


Figure 5: Comparison of accuracy with various models

This proposed system used Resnet50 which helped in achieving an accuracy of 98.6 than other models in various related work

MODEL DEPLOYMENT

The simulation was done in Matlab but a consideration on the processor was needed to ensure a faster operation. The overall simulation involves training the system based on the deep neural network used. At a later stage the user need to select an image which will be resized and then passed on to the deep neural network to work on the analyzing and identifying the infected and non-infected cells in line with the training and testing, the result will undergo optimization and a prediction will display.

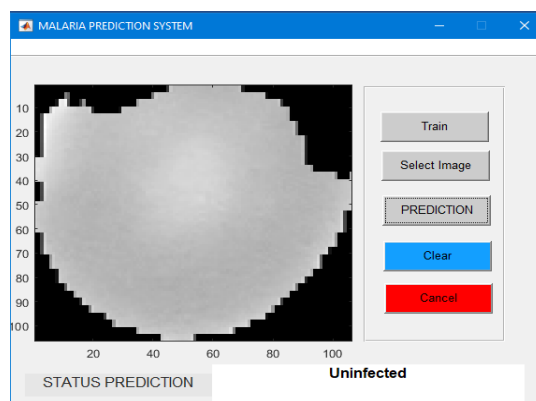


Figure 6: Malaria negative cell

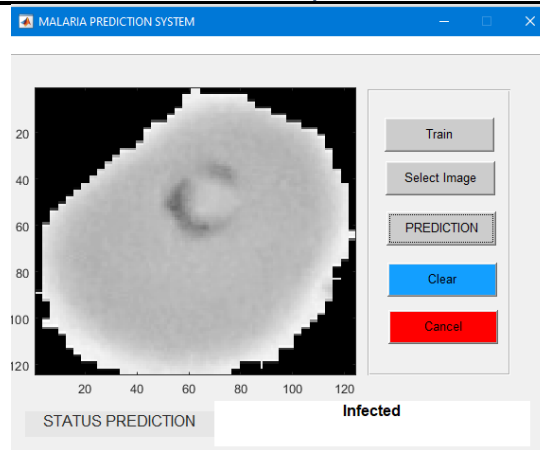


Figure 7: Malaria positive cell

V. CONCLUSION

This paper presents a study on how deep learning works to predict malaria by using different techniques such as normalization, data augmentation, feature extraction and classification. Kaggle dataset consisting of 27558 red blood cell images are compared during classification through the Resnet50 model. From the proposed model, an accuracy of 98.5% and sensitivity of 95 is achieved which outperforms most models. The is performance increase on the network due to pre-trained models thus using transfer learning technique on images and having more extracted features to classify the test images. The optimization method tunes the deep neural network by improved optimal prediction accuracy. This hybrid approach of genetic algorithm optimized neural network increases performance and stability than other techniques, which requires complex data and engineering.

VI. ACKNOWLEDGMENT

I sincerely thank Dr Glorindal, HOD In Charge School of Computer Science, PG coordinator and my internal supervisor for her support and contributions towards this research, which will help in improving the health sector in Malawi. Lastly, I appreciate the valuable inputs given by the review team thus contributing to the improvements at every stage of the research.

REFERENCES

- [1] Chakradeo, K., Delves M.J., Titarenko, S.,(2021). Malaria Parasite Detection Using Deep Learning Methods,e349644405 <https://DOI: 10.5281/zenodo.4569849>
- [2]Motwani, K., Kanojiya, A., Gomes, C., & Yadav, A. (2021). Malaria Detection using Image Processing and Machine Learning. International Journal of Engineering Research & Technology (IJERT), 9(3), <https://DOI : 10.17577/IJERTCONV9IS03011>
- [3] Bibin, D., Nair, M. S., & Punitha, P. (2017). Malaria Parasite Detection From Peripheral Blood Smear Images Using Deep Belief Networks. *9100-9108*, <https://doi: 10.1109/ACCESS.2017.2705642>
- [4] Joshi, M. A., Das, A. K., & Dhal, S.(2020) Deep Learning Based Approach For Malaria Detection in Blood Cell Images. IEEE region 10 conference (TENCON),241-246, <https://doi:10.1109/TENCON50793.2020.9293753>
- [5] Kunwar, S., Shrestha,M., & Shikhrakar, R.M. (2018) Malaria Detection Using Image Processing and Machine Learning, Corpus ID: 4855305
- [6] Ravalji, R., Shah, N., & Nai, M.(2020)Malaria Disease Detection using Machine Learning. International Journal of Engineering Research & Technology (IJERT), 7(12), 1180-1188
- [7] World Health Organization (2021), “World malaria report, Available at: <https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2021>
- [8] Iradukunda, O., Che, H., Uwineza, J., Bayingana, J. Y., Imam, M.S.B., & Niyonzima, I. (2019). Malaria Disease Prediction Based on Machine Learning. <https://doi:10.1109/ICSIDP47821.2019.9173011>
- [9] Alkrimi, J. A., Toma, Sh. A., Mohammed, R. S. , & George, L. E. (2020). Using Knowledge Discovery to Enhance Classification Techniques for Detect Malaria-Infected Red Blood Cells. <https://DOI:10.26438/ijrsmc>
- [10] World Health Organization (2015), Malaria microscopy quality assurance manual. 2.

- [11] Arowolo, M. O., Adebisi, M. O., & Adebisi, A. A. (2021)Enhanced Dimensionality Reduction Methods for Classifying Malaria Vector Dataset using Decision Tree <https://doi.org/10.17576/jsm-2021-5009-07>
- [12] CDC, Malaria Diagnostic Tests, (2020) Available at: [Global Health, Division of Parasitic Diseases and Malaria](#)
- [13] Shekar,G.,Revathy,S., & Goud, E. K.(2020). Malaria Detection using Deep Learning. International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number: CFP20J32-ART

