MALARIA DETECTION USING DEEP LEARNING

A.ROJARAMANI
Assistant Professor
Department of CSE
Teegala Krishna Reddy Engineering College, Hyderabad

K.SATHWIKA
Department of CSE
sathwikasai695@gmail.com
Teegala Krishna Reddy Engineering College, Hyderabad

S.NAVYALAXMI
Department of CSE
navyalaxmi23@gmail.com
Teegala Krishna Reddy Engineering College,Hyderabad

K.PREMCHAND
Department of CSE
premchandkusa217@gmail.com
Teegala Krishna Reddy Engineering College,Hyderabad

Abstract - Malaria is one of the deadliest diseases across the globe. This is caused by the bite of female Anopheles mosquito that transmits the Plasmodium parasites. Traditional malaria detection techniques require experts to test blood cells under a microscope. The shortage of skilled technicians and the unavailability of required equipment and infrastructure result in false diagnoses leading to an increase in morality rate. In existing system detection using Machine learning techniques like Support Vector Machine (SVM) which is tedious and requires hand-engineered features extraction to train data, and the results where not up to the mark. In our proposed system we used deep neural networks to detect the malaria virus in human blood cells. The proposed method shows a system with end-to-end automated models using a deep neural network that performs both feature extraction and classification using blood smear cell images. Models are evaluated based on accuracy, precision, recall and F1-score. Data preprocessing techniques like Data segmentation and Normalization are applied to maximize the model performance and a five layer convolutional network to perform best in class feature extraction.

Index Terms – Deep Neural Network, CNN, Image Data Generator.

I INTRODUCTION

Malaria is a mosquito-borne disease caused by a plasmodium parasite transmitted by the bite of infected mosquitoes. Worldwide, an estimated 300–500 million people contract malaria each year, resulting in 1.5–2.7 million deaths annually. According to the World Health Organization (WHO), approximately 219 million cases were diagnosed with malaria resulting in 435,000 deaths globally in 2017. Malaria is a deadly disease which is more frequently found in rural areas where medical diagnosis and health care options are not easily accessible. Worldwide accepted light microscopy technique is used by the practitioners for the diagnosis of malaria. The conventional light microscopy for malaria diagnosis uses thick and thin stained blood smears for diagnoses. Microscopy is well adapted in areas which are highly prone to malaria. The major drawback of this approach is its dependence on skilled technicians, of which there is a critical shortage. A nationwide study in Ghana found 1.72 microscopes per 100,000 population, but only .85 trained laboratory staff per 100,000 population. This results in a delay in an inaccurate diagnosis. Most of the time, diagnosis is often made based on clinical signs and symptoms alone, which is error-prone and this inaccurate diagnosis leads to higher mortality, drug resistance, and the economic burden of buying unnecessary drugs. Early and accurate malaria diagnosis and prompt treatment can cure a patient and save many lives by preventing severe malaria cases. All around the world, there are millions of people who are still lacking access to malaria prevention and treatment. Therefore, there is a need for a reliable alternative which will help to provide access to high-quality diagnosis routinely that is currently unavailable. This project focuses on designing an accurate malaria diagnosis model that can be implemented without any dependencies on skilled technicians and testing the model accuracy to get high-quality results. Automated image analysis software could remove the most serious limitation of the worldwide accepted microscopy method in general, dependency on human experts for diagnostic accuracy of the results.

II DNN

The neural network needs to learn all the time to solve tasks in a more qualified manner or even to use various methods to provide a better result. When it gets new information in the system, it learns how to act accordingly to a new situation. Learning becomes deeper when tasks you solve get harder. Deep neural network represents the type of machine learning when the system uses many layers of nodes to derive high-level functions from input information. It means transforming the data into a more creative and abstract component. In order to understand the result of deep learning better, let's imagine a picture of an average man. Although you have never seen this picture and his face and body before, you will always identify that it is a human and differentiate it from other creatures. This is an example of how the deep neural network works. Creative and analytical components of information are analyzed and grouped to ensure that the object is identified correctly. These components are not brought to the system directly, thus the ML system has to modify and derive them.

Existing System

In existing system have shown detection using Machine learning techniques like Support Vector Machine (SVM) which is tedious and requires hand-engineered features extraction to train data, and the results where not up to the mark.

Disadvantages:

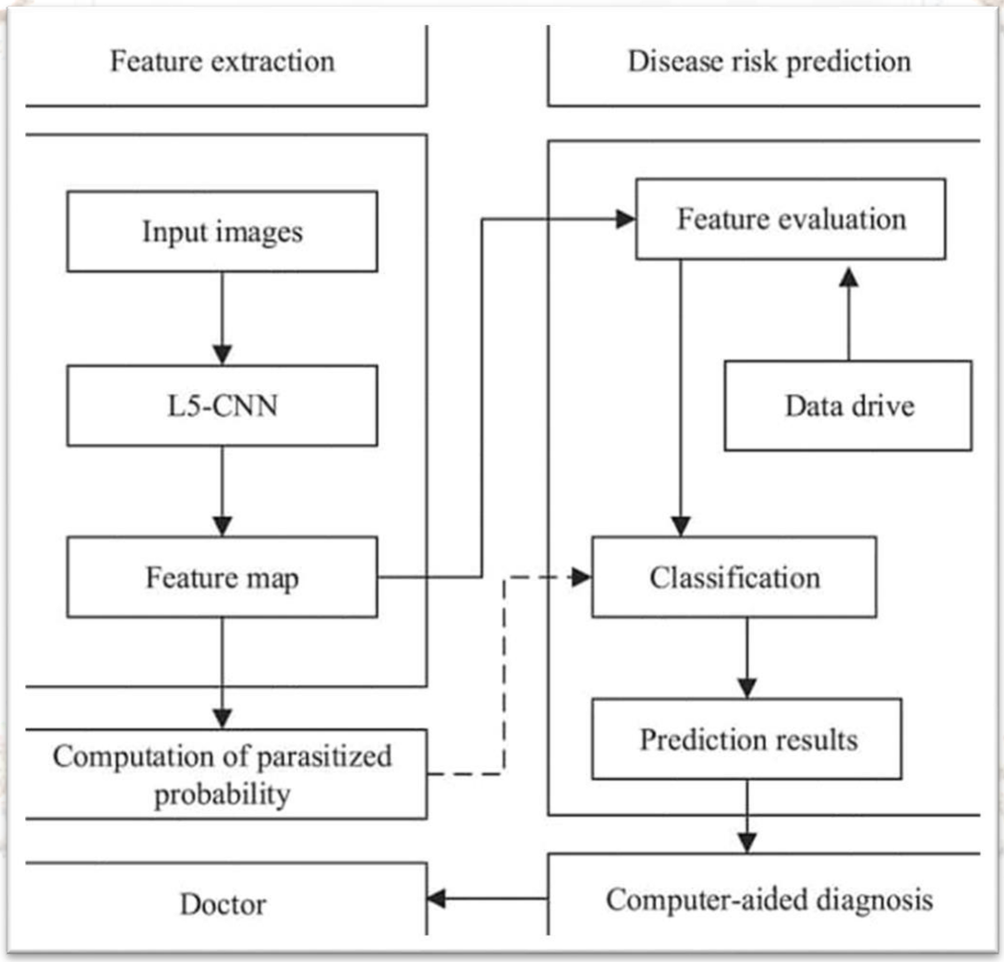
1. Low accuracy
2. Low feature consideration
3. Limited ability to explore

Proposed System

The proposed method shows a system with end-to-end automated models using a deep neural network that performs both feature extraction and classification using blood smear cell images. Models are evaluated based on accuracy, precision, recall and F1-score. Data preprocessing techniques like Data segmentation and Normalization are applied to maximize the model performance. In our project five layer deep Convolutional Neural Networks are used.

Advantages:

1. Increased feature consideration.
2. Efficient computation.
3. Time taken to run is less.
4. Accuracy is increased drastically.



System Architecture

III Algorithms

1. CNN- Main learning algorithm.
2. ReduceLROnPlateau- Learning rate controller on abnormal statistics.
3. EarlyStopping- Optimizer to stop learning when maximum accuracy is achieved.
4. ImageDataGenerator- Extracts machine readable data frim images.

CONVOLUTIONAL NEURAL NETWORKS

Convolutional neural networks (CNN) are mostly used for image recognition, and rarely for audio recognition. It is mostly applied to images because there is no need to check all the pixels one by one. CNN checks an image by blocks, starting from the left upper corner and moving further pixel by pixel up to a successful completion. Then the result of every verification is passed through a convolutional layer, where data elements have connections while others don’t. Based on this data, the system can produce the result of the verifications and can conclude what is in the picture.

REDUCE LR ON PLATEAU

Reduce learning rate when a metric has stopped improving. Models often benefit from reducing the learning rate by a factor of 2-10 once learning stagnates. This scheduler reads a metrics quantity and if no improvement is seen for a ‘patience’ number of epochs, the learning rate is reduced.

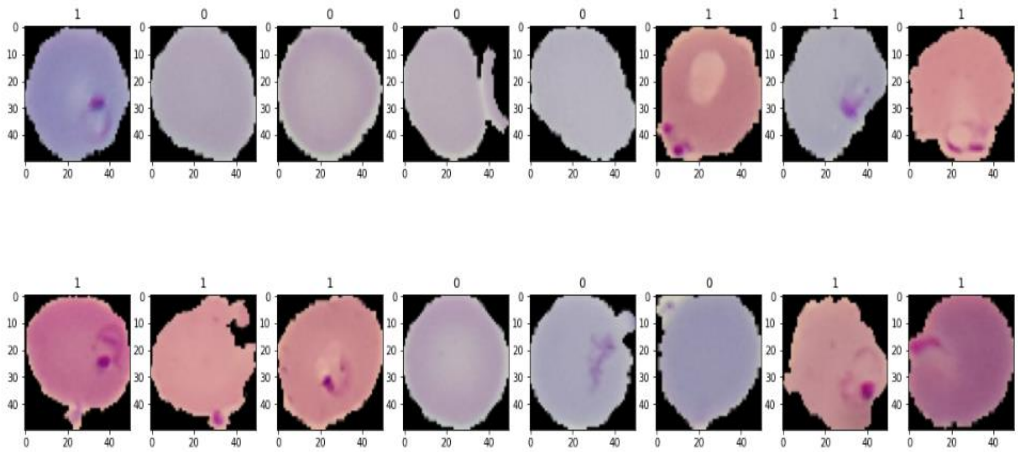
EARLY STOPPING

Early stopping is a form of regularization used to avoid overfitting when training a learner with an iterative method, such as gradient descent. Such methods update the learner so as to make it better fit the training data with each iteration. Up to a point, this improves the learner's performance on data outside of the training set. Past that point, however, improving the learner's fit to the training data comes at the expense of increased generalization error. Early stopping rules provide guidance as to how many iterations can be run before the learner begins to over-fit. Early stopping rules have been employed in many different machine learning methods, with varying amounts of theoretical foundation.

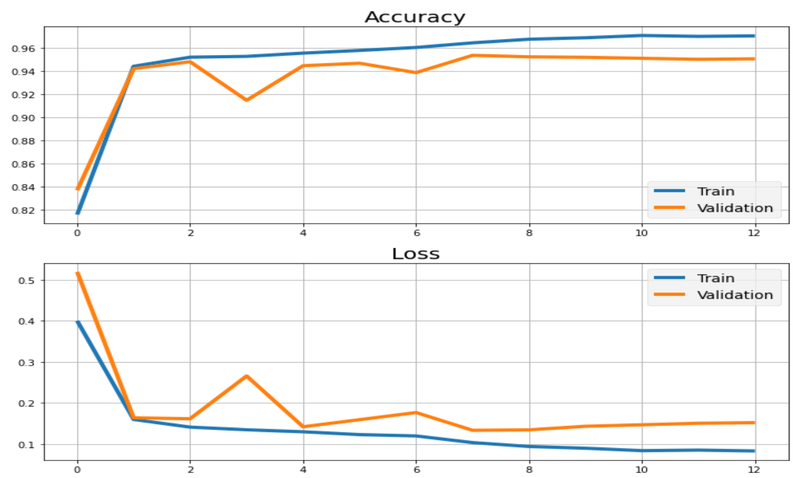
IMAGE DATA GENERATOR

When training a model, the Keras deep learning package allows you to employ data augmentation automatically. The ImageDataGenerator class is used to do this. First, the class must be constructed, and the kinds of data augmentation must be configured using parameters sent to the main approach. A variety of approaches, as well as image scaling methods, are supported. The dataset’s photos aren’t directly used. Instead, the model is only given enhanced pictures. So because augmentations are done at random, it is possible to produce and employ changed pictures as well as near imitations of the actual pictures during training. The validation, as well as the test data, can both be specified using a data generator. A second ImageDataGenerator instance is frequently used, which may have the same pixel scaling setup as the ImageDataGenerator example utilized for the learning data, but does not require data augmentation. It is because augmentation is used only to increase the training dataset artificially in terms of improving model performance on an unenhanced dataset.

IV RESULTS



From the above cells, 1 is Positive It is infected cell , 0 is Negative It is Uninfected cell, and it shows 99% Accuracy.



VI CONCLUSIONS

Here is our project that address problems with existing system and solves them effectively. In the end, we have achieved a model that can be integrated into any screening device and detect malaria infected cells.

VI REFERENCES

1. Agriculture Role on Indian Economy Madhusudhan L https://www.omicsonline.org/openaccess/ agriculture-role-on-indianeconomy-2151-6219-1000176.php?aid=62176
2. Priya, P., Muthaiah, U., Balamurugan, M. International Journal of Engineering Sciences Research Technology Predicting Yield of the Crop Using Machine Learning Algorithm.
3. Mishra, S., Mishra, D., Santra, G. H. (2016). Applications of machine learning techniques in agricultural crop production: a review paper.Indian J. Sci. Technol,9(38), 114.
4. Manjula, E., Djodiltachoumy, S. (2017). A Model for Prediction of Crop Yield.International Journal of Computational Intelligence and Informatics,6(4), 2349-6363.
5. Dahikar, S. S., Rode, S. V. (2014). Agricultural crop yield prediction using artificial neural network approach.International journal of innovative research in electrical, electronics,instrumentation and control engineering,2(1), 683-686.
6. Gonzlez Snchez, A., Frausto Sols, J., Ojeda Bustamante, W. (2014). Predictive ability of machine learning methods for massive crop yield prediction.
7. Mandic, D. P., Chambers, J. (2001). Recurrent neural networks for prediction: learning algorithms, architectures and stability. John Wiley Sons, Inc..
8. Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8),1735-1780.
9. Sak, H., Senior, A., Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In Fifteenth annual conference of the international speech communication association.
10. Liaw, A., Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.
11. Chen, T., Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acmsigkdd international conference on knowledge discovery and data mining (pp.785-794). ACM. 38
12. Cover, T. M., Hart, P. E. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27. 13. Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., Klein, M. (2002). Logistic regression. New York: Springer-Verlag.
14. Seber, G. A., Lee, A. J. (2012). Linear regression analysis (Vol. 329). John Wiley Sons.
15. urada, J. M. (1992). Introduction to artificial neural systems (Vol. 8). St. Paul: West publishing company.
16. http://www.indiawaterportal.org/met data/
17. District-wise, season-wise crop production statistics Vkhullar - https://data.gov.in/catalog/district-wise-season-wise-crop-productionstatistic