

DETECTION OF MALARIA USING IMAGE CELLS

Sumit Patil¹, Piyush², Rahul Kumar³, Akash Raju⁴, Susila M⁵.

^{1,2,3,4}Student, Department of Electronics and Communication Engineering,
SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India.

⁵Professor, Department of Electronics and Communication Engineering,
SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamil Nadu, India.

ABSTRACT - Malaria is a deadly disease that prompts numerous deaths every year. As of now, most malaria analysis are performed physically, which is time consuming which leads to delay in treatment. The key work of this paper is to address the improvement of computer-assisted detection of malaria and for easy detection and proper diagnosis. Furthermore, utilizing a Deep Learning approach dependent on microscopic pictures of fringe blood smears. Convolutional Neural Networks, a class of Deep Learning models guarantee exceptionally adaptable and better outcomes throughout feature extraction and characterization. In this paper convolutional neural network model such as VGGNet and Machine Learning model like Support Vector Machine were compared. Results appeared that all these profound convolution neural systems accomplished accuracy of over 95%, which is higher than SVM because the Deep Learning techniques have the advantage of consequently learning from the previous data or information.

Key Words: Deep Learning, Convolutional Neural Network, Machine Learning, Feature Extraction, VGGNet, Support Vector Machine.

1. INTRODUCTION

Malaria is a destructive, irresistible mosquito-borne ailment brought about by Plasmodium parasites. These parasites are transmitted by the bites of contaminated female Anopheles mosquitoes. The most generally utilized technique so far is analyzing thin blood smears under the microscope, and visually scanning for tainted cells [15]. The patients' blood is spread on a glass slide and stained with differentiating agents to more readily distinguish contaminated parasites in their red platelets. The analytic precision vigorously relies upon human mastery and can be badly affected by the onlooker variability and the risk forced by large scale scope analysis in infection endemic regions [14]. Alternative procedures are restricted in their performance and are less viable in disease-endemic regions [9]. About a large portion of the total population is at risk from malaria and there are more than 200 million malaria cases and around 400,000 passing's per year due to this disease [1]. This gives every one of us the more inspiration to make malaria recognition and treatment at early which should be simple and viable.

Automatic image recognition technologies based on Machine Learning and Deep Learning have been applied to both thick and thin malaria blood smears for microscopic diagnosis [21]. In this paper machine learning model SVM was used. SVM is an area explicit classifier and supervised learning algorithm utilized for characterization and regression analysis [12]. It develops a hyperplane for each class in the multi-dimensional space [13]. Convolutional Neural Networks have demonstrated to be extremely compelling in a wide range of PC vision tasks [3]. Convolution layers take spatial progressive examples from the data, which are likewise interpretation invariant. So, they can learn various parts of pictures [14].

2. MOTIVATION

It is really certain that malaria is pervasive over the globe particularly in tropical areas. The inspiration for this project is anyway based on the nature and casualty of this malady [14]. Initially if an infected mosquito bites you, the Red Blood Cells (RBC) are used as hosts and are destroyed afterwards [15]. Commonly, the main side effects of malaria are like flu or a virus. Hence, a postponement in the right treatment can prompt complications and even death. Subsequently early and effective testing and identification of malaria can spare lives [8].

The World Health Organization (WHO) has released several crucial facts on malaria that about a large portion of the total population is at risk from malaria and there are more than 200 million malaria cases and around 400,000 passing's per year due to this disease [2]. This gives every one of us the more motivation to make malaria recognition and treatment at early which should be fast, simple and viable.

3. SCOPE OF STUDY

DL models such as, CNN have demonstrated to be extremely viable in a wide assortment of computer vision assignments [5]. Quickly, the key layers in a CNN model incorporate convolution and pooling layers. Convolution layers take in spatial progressive examples from the information, which are likewise interpretation invariant [4]. In this manner, they can learn various parts of pictures [14]. Pooling layers help with down-sampling and dimension reduction. We will utilize open-source tools and frameworks which incorporate Python and TensorFlow to build our models [7].

4. CONVOLUTIONAL NEURAL NETWORK

In neural systems, ConvNets or CNNs is one of the fundamental classes to do pictures acknowledgment, pictures characterizations [11]. Object detection faces and so forth., are a portion of the territories where CNNs are generally utilized. Convolutional Neural Networks have demonstrated to be extremely compelling in a wide range of PC vision tasks [11]. Convolution layers take spatial progressive examples from the data, which are likewise, interpretation invariant. So, they can learn various parts of pictures [16]. Malaria identification utilizing deep learning models like CNNs could be compelling, modest and adaptable particularly with the appearance of move learning and pre-prepared models which work very well even with requirements like less information [14].

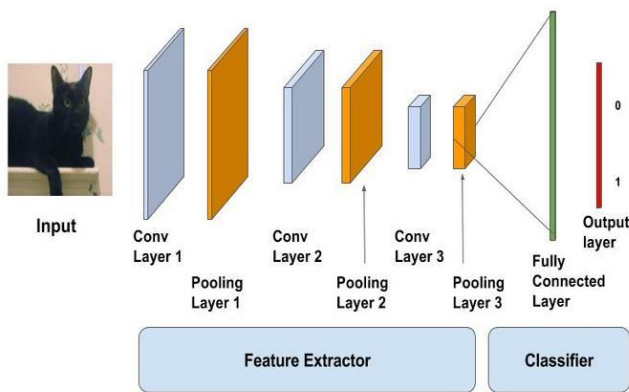


Figure 1 CNN Architecture [11]

5. SUPPORT VECTOR MACHINE

SVM is an area explicit classifier and supervised learning algorithm utilized for characterization and regression analysis [12]. It develops a hyperplane for each class in the multi-dimensional space. The possibility of SVM is basic: The calculation makes a line or a hyperplane which isolates the information into classes [13]. As indicated by the SVM calculation, we discover the focuses nearest to the line from both the classes. These focuses are called support vectors. Presently, we register the separation between the line and the support vectors [12]. This separation is known as the margin. The hyperplane for which the margin is most extreme is the ideal hyperplane [12].

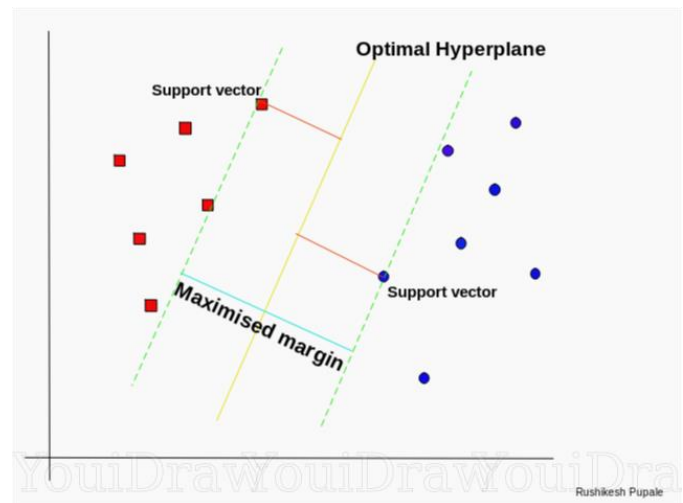


Figure 2 SVM Architecture [12]

6. DEEP TRANSFER LEARNING

Transfer learning is much the same as people have an inherent ability to move information across tasks, it empowers us to use information from recently learned assignments and apply them to latest, related tasks even with regards to Machine Learning or Deep Learning [6]. CNN's assist us with computerized and adaptable feature engineering [10]. Computerized malaria recognition utilizing DL models like CNNs could be exceptionally successful, low-cost and adaptable particularly with the appearance of transfer learning and pre-trained models which work very well even with limitations like less information [20].

Traditional ML vs Transfer Learning

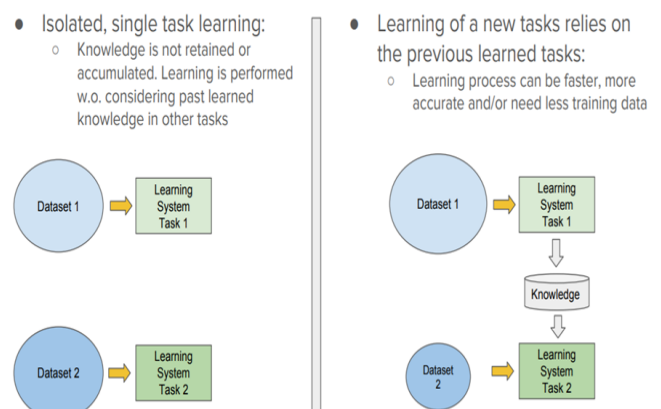


Figure 3 Transfer Learning [10]

7. PRE-TRAINED CNN MODEL (VGG-19)

The VGG-19 model is a 19-layer CNN model based on the ImageNet database, which is worked with the end goal of picture acknowledgment and characterization [19]. We will utilize the pre-prepared VGG-19 deep learning model, created by the Visual Geometry Group (VGG) for our investigations. A pre-prepared model like the VGG-19 is

trained on a dataset such as ImageNet with a ton of different picture classifications [14]. Consequently, the model, having taken in a decent portrayal of highlights for over a million pictures, can go about as a decent component extractor for new pictures appropriate for PC vision issues simply like malaria recognition [7].

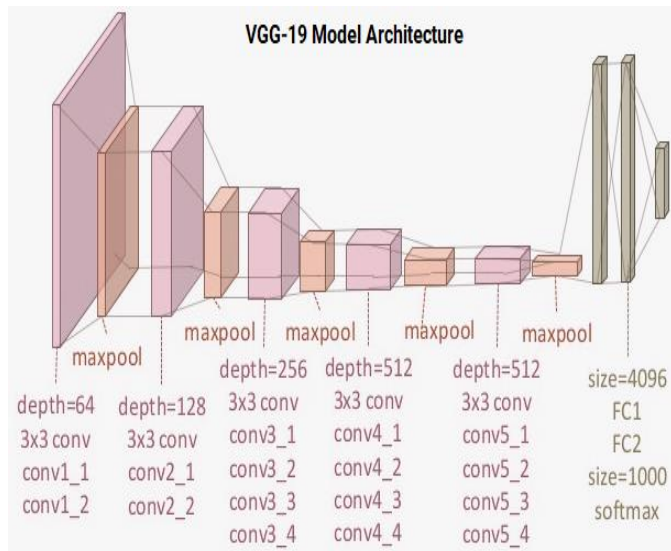


Figure 4 VGG-19 Architecture [14]

For building this model, we will use TensorFlow to stack up the VGG-19 model and freeze the convolution blocks so we can utilize it as a feature extractor [14]. Save the obtained result graphs for future assessment.

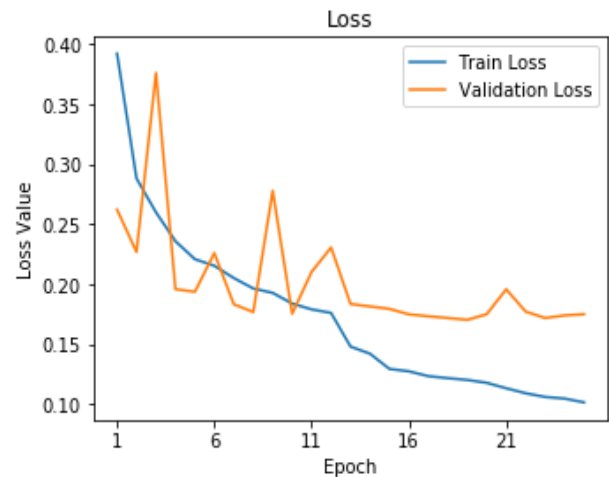
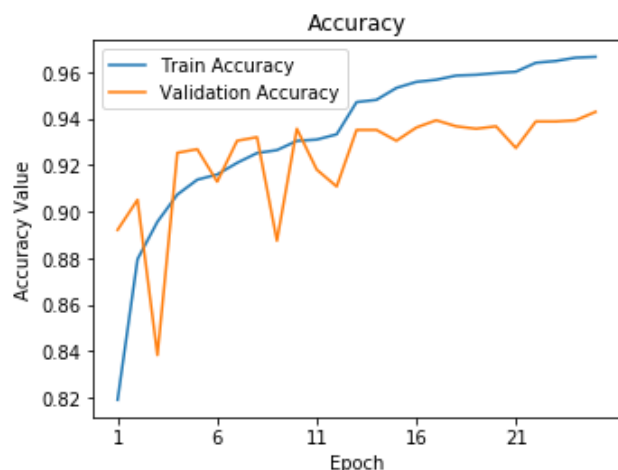


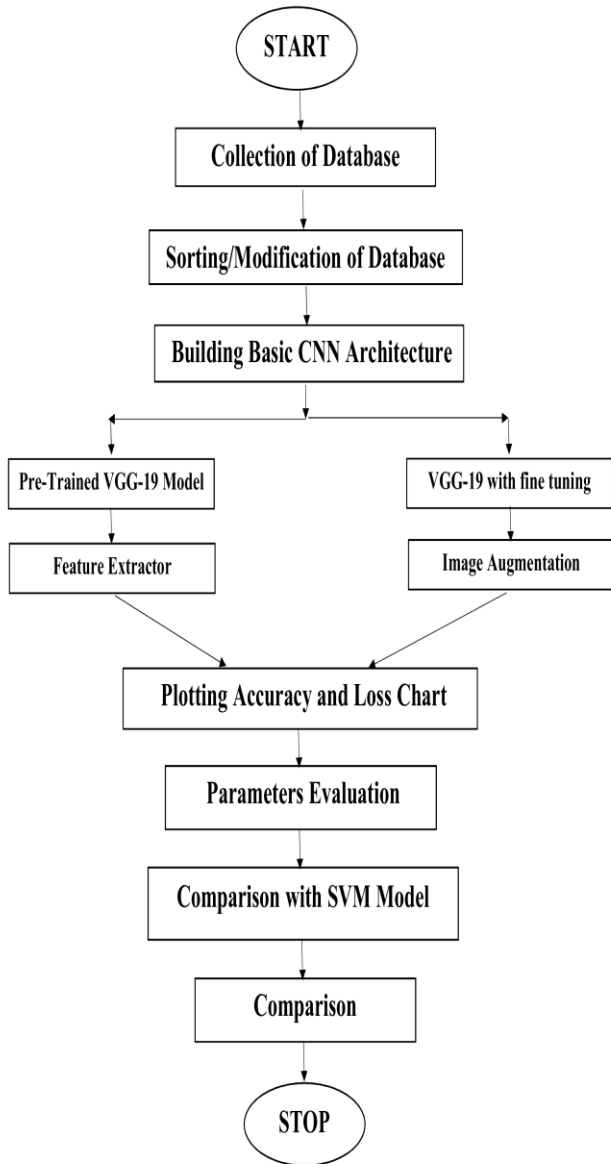
Figure 5 VGG-19 Output

8. MATERIALS AND METHODS

The dataset used in this research is taken from the official National Institute of Health (NIH) website. Dataset then was explored and extracted using Zip file opener. They have gathered and classified this dataset into healthy and infected blood smear pictures. The dataset was divided into train, validation and test datasets respectively in the ratio of 60:10:30. The fundamental CNN model design was made. A validation accuracy of 95.6% has, however, the model seemed to be overfitting with the training accuracy. At that point, the pre-prepared VGG-19 model as a feature extractor was taken. The model was not overfitting as much as our fundamental CNN model however the performance was not so much better and it was somewhat lesser than our fundamental CNN model.

The proposed model for this research paper is to fine tune the pre-trained VGG-19 model with image augmentation method. This appears to be the best model yet giving us a validation accuracy of almost 96.5% and based on the training accuracy, it doesn't seem like the model is overfitting as much as the first model. At last, the proposed model was evaluated based on the four parameters such as accuracy, f1 score, precision and recall, results were compared with the SVM model.

9. FLOW CHART



10. PROPOSED MODEL- FINE-TUNING OF VGG-19 WITH IMAGE AUGMENTATION

The VGG-19 model is a 19-layer deep learning system based on the ImageNet database, which is worked with the goal of image recognition and classification. To build the model, TensorFlow library was used to stack up the VGG-19 model. In this model, the weights of the layers present in the last two blocks of the pre-prepared VGG-19 model were fine-tuned [18].

A processing technique, image augmentation was used [19]. The existing pictures from the training dataset were loaded and some picture change activities were applied to them, for example, rotation, shearing, translation, zooming, etc. to create new, adjusted variants of existing images [19]. Similar pictures each time were not obtained because of these arbitrary changes. The performance of the model is checked with relevant classification metrics and the accuracy obtained was 96% which was similar to the f1 score for the VGG-19 model which is comparably higher than the machine learning model SVM which is 93%.

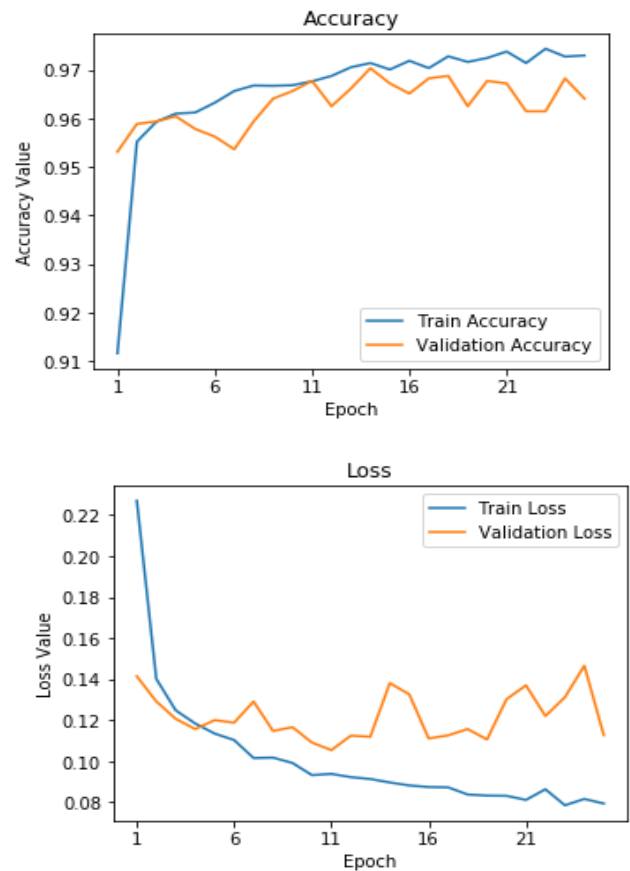


Figure 6 Proposed VGG-19 model output

11. PARAMETERS EVALUATION

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative

Figure 7 Confusion Matrix [17]

$$\text{Precision} = \frac{\text{True Positive}}{\text{Actual Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{Predicted Results}} \quad \text{or} \quad \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{Total}}$$

Figure 8 Accuracy, Recall and Precision [17]

$$F1 = 2 \times \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 9 F1 Score Calculation [17]

12. Results and Discussion

The performance of the model is checked with relevant classification metrics and the accuracy obtained was 96% [Table 1] which was similar to the f1 score for the VGG-19 model which is comparably higher than the machine learning model SVM which is 93% [Table 1].

Model	Accuracy	Recall	Precision
VGG-19	96.00	96.00	96.10
SVM	93.06	96.13	90.58

Table 1. Accuracy, Precision, Recall for the VGG-19 and SVM Model [22].

13. CONCLUSIONS

In this article, the fascinating true clinical imaging contextual investigation of malaria detection was handled. A simple model was built to construct open-source procedures utilizing learning which gives us the best accuracy and precision in detecting malaria. The results indicate that the deep learning model of VGG-19 [Fig.1] is better in performance parameters such as accuracy, precision, recall and f-score compared to the machine learning model of SVM [Fig.1]. But, deep learning model based on CNN consume much higher time for training the model than SVM. The detection of malaria isn't a simple system and the accessibility of the correct workforce over the globe is additionally a genuine concern. Based on our project, we propose that our newly designed convolutional neural network model is a suitable solution for blood smear classification. Its performance is affected by both the architecture and the volume of training data. We expect that

deep learning solutions will significantly improve the working efficiency and accuracy of malaria diagnosis and other health-related applications.

ACKNOWLEDGEMENT

The authors would like to express thanks to Dr. M Susila, Professor, Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, for her assistance throughout this research work.

REFERENCES

- [1] World Health Organization, Fact Sheet: World Malaria Report 2016. (13 December 2016).
- [2] World Health Organization, Malaria, <https://www.who.int/news/-room/fact/-sheets/detail/malaria> (19 November 2018).
- [3] Carlos Atico Ariza – “Malaria Hero: A web app for faster malaria diagnosis” (Nov 6, 2018).
- [4] Rajaraman et al. – “Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images” (2018). PeerJ 6: e4568; DOI 10.7717/peerj.4568.
- [5] K. He, X. Zhang, S. Ren, and J. Sun – “Deep Residual Learning for Image Recognition” (2015), (<https://arxiv.org/abs/1512.03385>)
- [6] A. Rosebrock – “Deep Learning and Medical Image Analysis”, pyimagesearch blog, 2017.
- [7] A. Rosebrock – “Deep Learning for Computer Vision with Python”, pyimagesearch blog, 2017.
- [8] Nadia Jmour, Sehla Zayen, Afef Abdelkrim – “Convolutional neural networks for image classification”, International Conference on Advanced Systems and Electric Technologies (IC_ASET 2018).
- [9] Muhammad Imran Razzak, Saeeda Naz, Ahmad Zaib – “Deep Learning for Medical Image Processing: Overview, Challenges and the Future Classification in BioApps”, Springer (2018).
- [10] Dipanjan (DJ) Sarkar – “A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning”, towardsdatascience (2018)
- [11] Sai Kumar Basaveswara – “CNN Architectures, a Deep-dive”, towardsdatascience (2019)
- [12] Rushikesh Pupale – “Support Vector Machines (SVM) – An Overview”, towardsdatascience (2018)
- [13] Oscar Contreras Carrasco – “Support Vector Machines for Classification”, towardsdatascience (2019)
- [14] Dipanjan (DJ) Sarkar – “Detecting Malaria with Deep Learning AI for Social Good — A Healthcare Case Study”, towardsdatascience (2018)

- [15] Gracelyn Shi - "Detecting malaria using deep learning". towardsdatascience (2019)
- [16] Vikas Gupta - "Image Classification using Convolutional Neural Networks in Keras", Learn OpenCV (2017)
- [17] Koo Ping Shung - "Accuracy, Precision, Recall or F1?" towardsdatascience (2018)
- [18] Nima Tajbakhsh, Jae Y. Shin, Suryakanth R. Gurudu, R. Todd Hurst, Christopher B. Kendall, Michael B. Gotway, Jianming Liang - "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging (2016).
- [19] Agnieszka Mikołajczyk, Michał Grochowski - "Data augmentation for improving deep learning in image classification problem", International Interdisciplinary PhD Workshop (IIPhDW 2018).
- [20] Michael White and Patrick Marais - "Supervised learning and image processing for efficient malaria detection" CEUR,2019.
- [21] Yuhang Dong¹, Zhuocheng Jiang¹, Hongda Shen¹, W. David Pan, Lance A. Williams, Vishnu V. B. Reddy, William H. Benjamin, Allen W. Bryan - Evaluations of Deep Convolutional Neural Networks for Automatic Identification of Malaria Infected Cells, IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) 2017
- [22] Zhaohui Liang, Andrew Powell, Ilker Ersoy, Mahdieh Poostchi, Kamolrat Silamut, Kannappan Palaniappan, Peng Guo, Md Amir Hossain, Antani Sameer, Richard James Maude, Jimmy Xiangji Huang, Stefan Jaeger, George Thoma - CNN-Based Image Analysis for Malaria Diagnosis" 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).