

Malaria Parasite Disease Classification using Deep Learning Neural Networks

Kancharagunta Kishan Babu¹
Department of CSE - AIML & IoT
VNR Vignana Jyothi Institute of
Engineering and Technology,
Hyderabad, India
kishan.kancharagunta@gmail.com

Musku Sai Ritish Reddy²
Department of CSE - AIML & IoT
VNR Vignana Jyothi Institute of
Engineering and Technology,
Hyderabad, India
sritish.10@gmail.com

Santhosh Reddy Bhavanam³
Department of CSE - AIML & IoT
VNR Vignana Jyothi Institute of
Engineering and Technology,
Hyderabad, India
bhavanamsanthoshreddy@gmail.com

Akhil Chimma⁴
Department of CSE - AIML & IoT
VNR Vignana Jyothi Institute of
Engineering and Technology,
Hyderabad, India
akhilchimma@gmail.com

Parthav Nuthalapati⁵
Department of CSE - AIML & IoT
VNR Vignana Jyothi Institute of
Engineering and Technology,
Hyderabad, India
parthavnuthalapati2019@gmail.com

Srikar Kovvuri⁶
Adlai E. Stevenson High School
Lincolnshire Illinois
srikarkovvuri5@gmail.com

Abstract— Malaria is a global health threat caused by the Plasmodium parasite. The early identification and classification of the parasite are vital for prompt treatment and improved patient outcomes. However, traditional detection methods depend heavily on expert knowledge for accurate diagnosis, making them challenging to implement. This paper aims to conduct a comprehensive analysis of 6 existing deep learning models including AlexNet, GoogleNet, VGG19, MobileNet, ResNet, and RCCNet utilizing the malaria parasite dataset. Additionally, we perform a comparative evaluation of these models by implementing various Data Augmentation techniques and improving accuracy by 3% reaching maximum accuracy up to 96.73 %. The assessment is centered around key factors such as training duration, precision, weighted f1 score, specificity, and sensitivity.

Keywords—Malaria parasite disease classification, Convolutional neural networks, Deep learning, Data augmentation.

I. INTRODUCTION

Malaria is a disease usually found in tropical countries such as sub-Saharan Africa, Southeast Asia spread to humans by some kinds of mosquitoes. In 2021 alone, it is estimated that about half the world's population ran a risk of being affected by Malaria. Of the 247 million cases recorded that year, the number of deaths stood at 619,000[1]. Malaria is known to be transmitted through bites and blood transfusion. Traditional methods depend completely on medical professionals to pass their judgement before providing the required medication.

Detecting life-threatening diseases like Malaria will be crucial in saving lives. Regular diagnosis allows for immediate treatment initiation, improving the chances of a full recovery. It can also reduce the disease from advancing to more severe stages, effecting health and getting complications. However, early detection of such diseases can be critical in saving lives of people and prevent the disease to progress in advance stages. By identifying the diseases early, we can allocate the required medical resources and take preventive measures against it. This will not only save lives but also reduce medical bills. Overall, detecting the diseases early can be a crucial component of healthcare systems by improving the outcomes of public health.

Image classification tasks have been performed extremely well by using deep learning methods rather than traditional methods. Over the past few years, medical imaging has

shown prominent results with the use of deep learning methods, including the detection of breast cancer [2], diabetic retinopathy [3], and skin cancer [4]. Deep learning methods can learn features using its complex and deep architecture which helps in identifying even complex data. These studies give significant evidence that deep learning methods hold potential for accurate classification on medical images, enabling regular treatment and better patient outcomes. These studies also highlight the importance of using advanced techniques to detect and classify the medical diseases in its early stages, as this can lead to better treatment outcomes.

This paper presents a comprehensive analysis of existing deep learning models including AlexNet, VGG19, GoogLeNet, MobileNet, ResNet-50 and RCCNet, on the dataset of malaria parasite images. Additionally, we have observed the effect of data augmentation techniques including shear, zoom, horizontal and vertical flipping with improving metrics in the models under identical conditions. The results of this study demonstrate the potential of deep learning for malaria parasite classification and highlight the importance of optimizing data augmentation techniques for improved performance in resource-limited settings.

II. PROPOSED SYSTEM

Deep learning methods have demonstrated significant promise in effectively and precisely identifying malaria parasites within microscopic images. This investigation introduces a novel approach for categorizing malaria parasites, leveraging finely tuned deep learning models.

AlexNet [5] was introduced by Krizhevsky in 2012 which marked a pivotal role in deep learning and computer vision. It was the first deep convolutional neural network to drastically reduce the error rate in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) by developing an architecture comprising a total of eight layers, including five convolutional layers followed by three fully connected (FC) layers, comprising roughly sixty million trainable parameters. The key innovation in AlexNet was the introduction of ReLU activation function and dropout regularization.

The winner of the ILSVR 2014 was GoogLeNet [6], by introducing the concept of inception models by utilizing different sizes of parallel convolutional filters. This model is notably deep, with 22 layers, yet computationally efficient

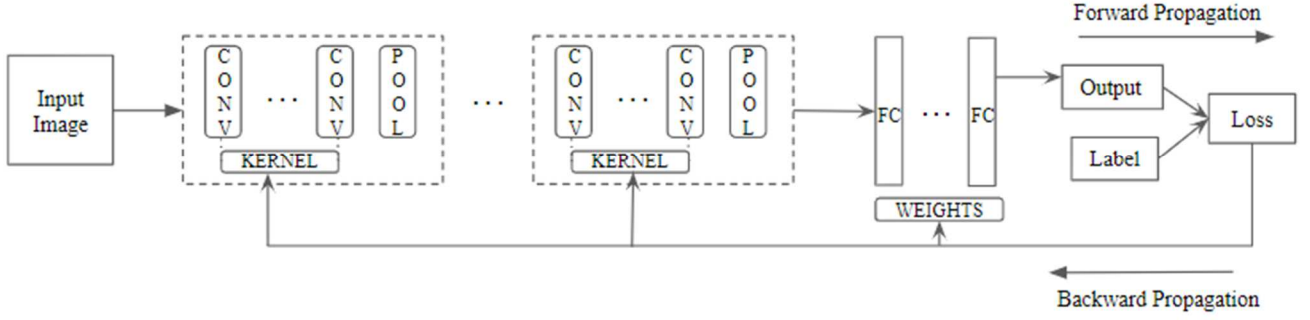


Fig. 1. Training phase of a CNN architecture

due to its architectural design unlike VGG-19 [7], the 1st runner-up of ILSVR 2014, having 19 layers including 16 convolutional layers and 3 FC layers. VGG-19 is computationally intensive due to its depth and trainable parameters. However, it is known for its remarkable ability to capture intricate spatial features in images, making it a popular choice for image classification task.

Vanishing gradients problems arises with the increase of depth in the neural networks. To overcome the challenge of training very deep neural network ResNet [8], a deep learning architecture renowned for its use of residual connections was introduced. ResNet contains 152 layers with skip/shortcut connection in residual networks. Although ResNet may have varying depth, the most common network is ResNet - 50, a ResNet architecture having a total of 50 layers in total convolutional and fully connected layer.

Although, deep networks help with effectiveness and accuracy these models are not suitable for deployment in servers and devices as they occupy lot of memory and are computational expensive. To address the above problem MobileNet [8] was introduced in 2017 which allows faster inference times and reduced memory footprint, making it ideal in resource constrained environments. Its architecture is characterized by depth wise separable convolutions, which operate both on spatial and depth dimensions simultaneously. The architecture has an overall of 30 layers including convolutional, depth wise and pointwise layers. This has made MobileNet a popular choice for deploying deep learning models with real time tasks.

Deep neural networks in the cancer domain were also considered such as RCCNet [9], a convolutional neural network which has produced significant results on the colon cancer classification task. The architecture has over 1,512,868 learnable parameters with seven learnable layers divided over three blocks. The RCCNet is a CNN model which is able to producer higher accuracy compared with other models.

The Fig. 1 shows the training phase of a CNN architecture. A CNN model contains an input image followed by convolutional layers followed by pooling layers and finally connected to a fully connected layer. The number of these layers are fixed by various CNN models, The CONV layer in the figure is a convolutional layer which is responsible for the convolution operation. The weights of these kernel are updated through backpropagation. The POOL layer in the figure is the pooling layer which performs the subsampling of the convoluted output by taking either the maximum,

minimum or average of the pixels present in the filter applied. By performing convolutional and pooling we are able to detect the features of the image. The number of CONV-POOL block are decided by the CNN model, latest CNN models have deeper blocks which help to detect better features and result in higher accuracy. The FC layers in the figure are the fully connected layers which are responsible for stretching the output of CONV-POOL block into a 1D numerical vector which then trains and predicts the output using neural network weights. Each output predicted is used to calculate the loss based on the label and is propagated back to the layers to update the weights in the FC and CONV layers.

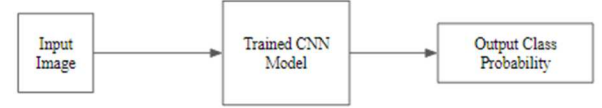


Fig. 2. Testing phase of a CNN architecture

During the testing phase, the trained CNN model takes the image as input, producing the probability of the class label to which the input belongs. The output with the maximum probability is assigned to the input image.

In this paper we conduct a comparative study of various models including AlexNet, GoogLeNet, VGG-19, ResNet-50, MobileNet and RCCNet under identical conditions. Their effectiveness is then ranked using metrics like accuracy, F1 score, sensitivity and specificity. To enhance model performance, data augmentation techniques are employed, including shear operations with a range of 0.2, zoom transformations within a range of 0.2, and performing both vertical and horizontal flips.

III. EXPERIMENTAL SETUP

This section provides the information regarding description of the dataset, training specifics and the criteria for evaluation.

A. Dataset Description

Publicly available dataset from the National Institute of Allergy and Infectious Diseases[7], consisting of microscopic images of malaria parasites, is used for comparison of the models. The dataset contains 27,558 images of two balanced

classes each patch with the dimension 142 X 148. The class 'Parasitized' consists of 13,779 images, and 'Uninfected' with 13,779 images. The dataset was split in 80-20 ratio, with the latter allocated for testing purposes.

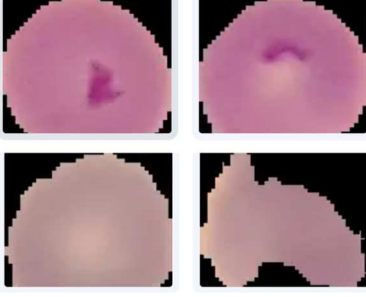


Fig. 3. Blood smears of parasite cells in top row and uninfected cells in bottom row.

B. Evaluation Criteria

Model performance is evaluated through a set of metrics, including the time taken for training also into account as an evaluation criterion.

The paper utilizes the metric accuracy which is defined as the ratio of correctly classified to the total classification made,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Specificity also known as True negative ratio, the ratio of true negatives to the total actual negatives,

$$Specificity = \frac{TN}{TN + FP}$$

Sensitivity or recall also called as True positive is the ratio of true positives to the total actual positives.

$$Sensitivity = \frac{TP}{TP + FN}$$

Since we are working with a balanced dataset, the values of sensitivity and specificity are the same. The F1-measure metric is a harmonic mean between recall and precision,

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

and finally, the training time which isn't typically a major factor but relevant in cases where model performs well in other evaluation metrics but requires a significant amount of training time.

C. Training Details

The training of the models has been performed by using an Adam optimizer [10] with the initial learning rate 6×10^{-5} , $\beta_1 = 0.9$, $\beta_2 = 0.99$ and decay of 10^{-6} . The learning rate is reduced by a factor of $\sqrt[2]{0.1}$ when the validation loss has reached a plateau. Rectified Linear Unit (RELU) [11] was used as an activation function in the network. The models were trained for 100 epochs each with a batch size of 16. The binary cross entropy loss function was used in the network. Data augmentation techniques including shear operations with a range of 0.2, zoom transformations within a range of 0.2, and performing both vertical and horizontal flips.

IV. RESULTS AND DISCUSSIONS

We have conducted a comprehensive analysis of 6 existing deep learning models including AlexNet, GoogleNet, VGG19, MobileNet, ResNet, and RCCNet on the malaria parasite dataset with and without data augmentation. Table 1 contains the results obtained without Data Augmentation Techniques and table 2 presents the comparison with data augmentation with an increasing order of testing accuracy.

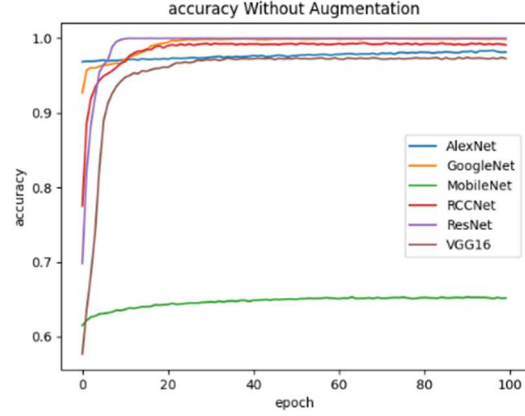


Fig. 4. Accuracy graph of all models without Data Augmentation

The insights gained from TABLE-1 is as follows:

1. A maximum of 96.08 testing accuracy is achieved by RCCNet closely followed by GoogLeNet (95.8) and VGG19 (94.77).
2. RCCNet also precedes other models in F1 score (0.9609) followed by GoogleNet (0.9582)
3. Highest Sensitivity and Specificity is achieved by GoogLeNet with values [0.9852, 0.9827].

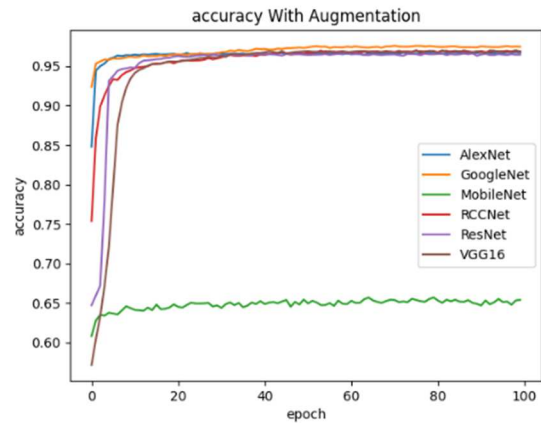


Fig. . Accuracy graph of all models with Data Augmentation

Table. 2. Results obtained without Data Augmentation Techniques

Model Name	Without Data Augmentation											
		Classification Accuracy			Loss		Weighted Average F1 Score		Specificity		Sensitivity	
	Training Time(s)	Train %	Test %	Overfit %	Train	Test	Train	Test	Train	Test	Train	Test
AlexNet	10770.653	98.15	48.11	50.04	0.0531	402.7	0.9815	0.48175	0.9815	0.48175	0.9815	0.4817
MobileNet	1587.7419	65.81	64.44	1.37	0.6161	0.6267	0.6518	0.6441	0.6518	0.6441	0.6518	0.6441
ResNet_50	2550.7592	100	93.28	6.72	1.21E-05	0.5363	1	0.9331	1	0.9331	1	0.9331
VGG 19	1217.1751	99.11	94.77	4.34	0.0274	0.2174	0.9911	0.9479	0.9911	0.9479	0.9911	0.9479
GoogLeNet	1652	99.83	95.8	4.03	0.0081	0.0938	0.9983	0.9582	0.9983	0.9852	0.9983	0.9827
RCCNet	2922.9697	97.28	96.08	1.2	0.0819	0.4974	0.9729	0.9609	0.9729	0.9609	0.9729	0.9609

From Table 2. We can deduce that,

1. An Increment of [0.93, 0.0092] in Accuracy and Weighted F1 score and a Decrement of 3.30 overfit for GoogLeNet when compared to the results obtained by the same model without using Data Augmentation techniques.
2. An Increment of [1.76, 0.0175] in Accuracy and F1 score and a Decrement of 4.01 in Overfit is observed for RCCNet compared to the previous results.

also, between same models implemented with and without Data Augmentation. Data augmentation reduced the overfitting percentage in all models. The sensitivity and specificity of the models were high, indicating that they were able to correctly classify both infected and uninfected cells. Our study found that GoogLeNet had the highest testing accuracy (96.73%), followed closely by RCCNet (96.53%). We can further improve the accuracy by increasing the data samples using generative adversarial networks (GANs).

Table. 2. Results obtained through Data Augmentation Techniques

Model Name	With Data Augmentation											
		Classification Accuracy			Loss		Weighted Average F1 Score		Specificity		Sensitivity	
	Training Time(s)	Train %	Test %	Overfit %	Train	Test	Train	Test	Train	Test	Train	Test
MobileNet	3011.579	65.42	63.47	1.95	0.654	0.6344	0.6542	0.6347	0.6542	0.634	0.654	0.6342
AlexNet	10318.08	96.68	70.66	26.02	0.097	21.017	0.9668	0.7057	0.9668	0.705	0.966	0.7057
ResNet50	7534.214	96.45	96.18	0.27	0.103	0.109	0.9645	0.9618	0.9645	0.961	0.964	0.9618
VGG19	3395.051	96.84	96.38	0.46	0.095	0.1714	0.9684	0.964	0.9684	0.964	0.968	0.964
RCCNet	2193.37	96.86	96.53	0.33	0.093	0.1051	0.9686	0.9654	0.9686	0.965	0.968	0.9654
GoogLeNet	2963.735	97.46	96.73	0.73	0.07	0.0938	0.9745	0.9674	0.9745	0.967	0.974	0.9674

V. CONCLUSION AND FUTURE SCOPE

This paper explores a comparative study of several CNN models for classifying malaria parasites. We evaluated the performance of all models on the publicly available malaria parasite classification dataset. This study is not only focused on comparison between different models but

ACKNOWLEDGMENT

We thank the Department of CSE - (AIML & IoT), Vallurupalli Nageswara Rao Vignana Jyothi Institute of Engineering and Technology for their generosity in providing the NVIDIA GPU k40 which we used to conduct experiments in this research.

REFERENCES

- [1] <https://www.who.int/news-room/fact-sheets/detail/malaria>
- [2] Bar, Yaniv, et al. "Deep learning with non-medical training used for chest pathology identification." *Medical Imaging 2015: Computer-Aided Diagnosis*. Vol. 9414. SPIE, 2015.
- [3] Ting, Daniel Shu Wei, et al. "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes." *Jama* 318.22 (2017): 2211-2223.
- [4] Afza, Farhat, et al. "Multiclass skin lesion classification using hybrid deep features selection and extreme learning machine." *Sensors* 22.3 (2022): 799.
- [5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105
- [6] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
- [7] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [9] Basha, S. S., Ghosh, S., Babu, K. K., Dubey, S. R., Pulabaigari, V., & Mukherjee, S. (2018, November). Recnet: An efficient convolutional neural network for histological routine colon cancer nuclei classification. In *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)* (pp. 1222-1227). IEEE.
- [10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [11] Agarap, A. F. (2018). Deep learning using rectified linear units (relu). *arXiv preprint arXiv:1803.08375*.