Deep Learning Based Approach For Malaria Detection in Blood Cell Images

Amogh Manoj Joshi[1], Ananta Kumar Das[2], Subhasish Dhal[3]
[1]Department of Electronics and Telecommunication Engineering
Vivekanand Education Society's Institute of Technology, Mumbai
[2]Department of Software Engineering
International Institute of Information Technology, Bangalore, India
[3]Department of Computer Science and Engineering
Indian Institute of Information Technology Guwahati, India

Abstract—Malaria, a life-threatening disease, develops due to the bite of female Anopheles mosquito. It spreads the plasmodium parasites in human blood, killing hundreds of millions of people every year. Modern scientific advancements play a pivotal role to combat the disease, along with biomedical research by the medical experts to possibly eradicate this disease from all parts of the world. With the significant development in deep learning research, faultless identification of medical imaging has become an important factor in medical diagnosis and decision-making. To this end, we present a deep learning based approach using a convolutional neural network for detecting malaria from microscopic cell images using image classification. The proposed CNN model implemented using 5-fold cross validation approach outperforms all the existing methods in terms of accuracy and other evaluation metrics, thus achieving the best results till date in malaria detection using deep learning.

Index Terms—Deep Learning, Malaria Detection, Image Classification

I. INTRODUCTION

Malaria is a life-threatening disease, which is caused by the bite of anopheles mosquitoes. It spreads plasmodium parasites in the human liver, affects the red blood cells and develops into life-threatening symptoms. According to the estimation of World Health Organization (WHO), in the year 2018, more than two hundred million cases of malaria have been reported and more than 90% of the cases are in the African region [1].To prevent, control and eliminate malaria, more than 3 billion dollars have been spent [1]. Computer-aided diagnostic (CADx) tools have the potential to lessen the burden on healthcare staff by aiding in disease interpretation and diagnostic process using some machine learning (ML) algorithms applied to microscopic blood cell images. [2]. To overcome challenges of developing hand-engineered features that capture even the minute variations in the underlying data, deep learning (DL) [3], also known as deep hierarchical learning, is widely used with significant success owing to its high accuracy. This paper, not being the subject's first study paper, several study papers have been released before, bearing witness to automated diagnosis of malaria in blood smear images [5] [6] [7] [9]. Early detection of malaria is crucial to ensure appropriate diagnostic process and increase the chances of the patient being cured [4]. Considering the severity of malaria by the amount of deaths caused by this illness, accepting possible negligible mistakes caused while execution by an automated method is justifiable. Deep learning techniques have advanced over the years and have proven to be much better than traditional methods as they ease the feature extraction process. Hence, in this paper we further utilize deep learning as a method and propose an approach using a Convolutional Neural Network (CNN) for detecting malaria from microscopic cell images using image classification. In this paper, we investigate two training approaches and compare both based on performance and select the better approach. We further evaluate our proposed model on testing set using metrics like F1 Score, AUC Score, Specificity and Sensitivity. The important contributions are:

1) Present a Convolutional Neural Network based approach for detection of malaria in cell images which outperforms the existing methods based on accuracy performance.
2) Perform a comprehensive evaluation of the proposed method using other performance metrics like F1 Score, AUC Score, Specificity and Sensitivity

The paper is designed as follows. Section I gives the introductory part and the importance of developing a deep learning approach for malaria detection. Section II gives an overview of the existing methods similar to our approach. Section III includes the proposed methodology for malaria disease detection. Section IV discusses the experimental setup while Section V discusses the results obtained using our proposed approach. Finally, section VI concludes the paper.

II. RELATED WORKS

In recent years, a variety of approaches have been proposed for automated detection of malaria parasites. The most accurate and widely used technique for diagnosing malaria is examining microscopic thick and thin

blood smear images [8]. In this section, we present an overview of some of the recent works related to malaria detection from blood smear images. Pattanaik et. al [11] proposed a Computer Aided Diagnosis (CAD) scheme for identifying the presence of malaria parasites in thick blood smears. They performed 10-fold cross validation and achieved an accuracy of 89.10%, sensitivity of 93.90% and specificity of 83.10%. Olugboja et. al [6] presented a comprehensive analysis of different machine learning techniques like Linear Support Vector Machine(SVM), Fine Gaussian SVM, Subspace K-Nearest Neighbors(KNN), Cosine KNN and Boosted KNN for automatic detection of malaria parasite in stained blood smears. Among these, Subspace KNN achieved the best accuracy of 86.3% whereas Fine Gaussian SVM achieved the best True Positive Rate (TPR) of 99.8%. Bibin et. al [5] proposed a novel Deep Belief Network (DBN) approach for malaria detection from blood smear images. The proposed DBN is trained by extracting the features from the images and initializing the visible variables of DBN. This method achieved an F1 score of 89.66%, sensitivity of 97.60% and specificity of 95.92%. Yang et. al [8] developed a deep learning based malaria parasite detection method that can run on smartphones. Their CNN achieved an overall accuracy of 93.46%. Chowdhury et. al [9] used a CNN approach to perform blood cell count on blood smear images. The CNN was also trained to detect malaria pathogens in the blood smears if present. They achieved a mean average precision (mAP) of 95%. Kalkan et. al [7] used a deep learning approach using a custom CNN to detect malaria from cell images. Using 5-cross validation technique, they achieved a training accuracy of 97% and testing accuracy of 95%. Nayak et. al [10] performed a comparative evaluation of pretrained CNN models like DenseNet121, VGG16, Alexnet, ResNet50, FastAI and ResNet101 for malaria detection in blood cell images. Among these CNN models, ResNet50 outperformed others and achieved the highest accuracy of 97.5%. To summarize, a lot of deep learning methods have been proposed for detecting malaria from cell images. Some of them used large pretrained CNN models for achieving a high classification accuracy, whereas some used custom CNNs for reducing the computational time. In medical image classification, large wrongly classified data has catastrophic results and disrupts the idea of proposing a medical diagnosis aid. So along with accuracy, other metrics like F1 score, AUC score, Sensitivity, Specificity are vital in comprehensively evaluating any proposed method. Many of the above-mentioned proposed methods aren't evaluated using these metrics. So,in this study, we further evaluate our proposed method using these metrics thereby proving the robustness of our model.

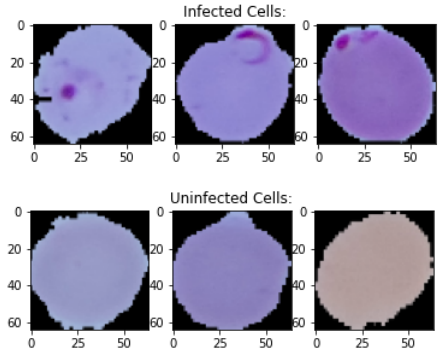

Fig. 1. Cell Images from Dataset

III. PROPOSED WORK

In this section, we brief over the dataset used in this study and the CNN model proposed for malaria detection. We also describe the training approaches followed and compare their performance using evaluation metrics.

A. DATA ACQUISITION

NIH Gov's Official Malaria dataset is used in this research work. The dataset contains a total of 27558 images of Malaria infected and Non-infected cells. The images are RGB images varying in sizes from 76 x 68 to 152 x 141. For standardized research purpose, all images were resized to 64x64x3 dimension. As the number of images were sufficient for training a CNN model and enough to prevent overfitting, no data augmentation was done on the images. The results obtained outperformed all the existing methods, so there was no need of applying data augmentation to improve the results. Fig. 1 shows a glimpse of resized images belonging to both the classes from the dataset. Dataset was first split into a Training set and Testing set in the ratio 70:30. To observe how the trained model performs on unseen images, we should have a good amount of images in the test set which is why we allotted 30% of the dataset for testing purpose. The Training set was further divided into Training and Validation set in the ratio 90:10.
So the final Image wise split was:

- Train: 17361
- Val: 1929
- Test: 8268

B. METHODOLOGY

Convolutional Neural Networks, a part of deep learning have proved to be of immense use in image recognition, identification and categorization [14] [15]. CNNs outperform traditional deep learning methods in terms of accuracy and efficiency in disease recognition studies [16] [17]. Convolutional Neural Network consists of several types of layers, each performing a particular operation.

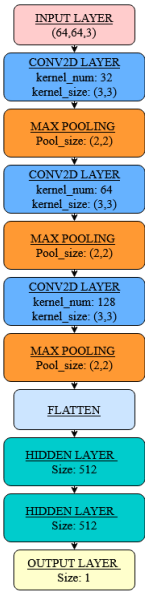

Fig. 2. Proposed Model Architecture

We use three main types of layers to build CNN architectures:

1) Convolution Layer
2) Pooling Layer
3) Fully Connected Layer

A diagrammatic representation of the proposed CNN model is shown in Fig. 2. It consists of three convolution layers each followed by a max pooling layer. The activation function used for convolution layers is ReLU which stands for Rectified Linear Unit. Activation functions are applied to feature maps to increase the non-linearity in a network [14]. ReLU activation eliminates all the negative values from an activation map by replacing them with zero. The output layer of the model contains 1 neuron and has sigmoid activation function. Sigmoid activation is used for binary classification and outputs a value between 0 and 1. The model is compiled using Adam optimizer with a learning rate of 0.01. The loss function used is binary crossentropy. After the model is compiled, it is ready to be trained on the dataset. The flow chart in Fig. 3 describes the work flow followed in this study. During the training process, the model learns to map the inputs to the outputs from the training set. The model training process essentially involves finding a set of weights that prove to be the best possible fit at solving the required problem [18].

C. CONVENTIONAL APPROACH

We initially trained the model using conventional training approach of training the model for a fixed number of epochs. One epoch is when the entire training set is passed both forward and backward through the neural network once. The model was initially trained for

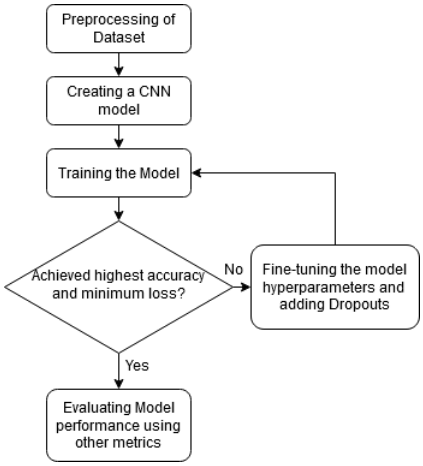

Fig. 3. Work Flow

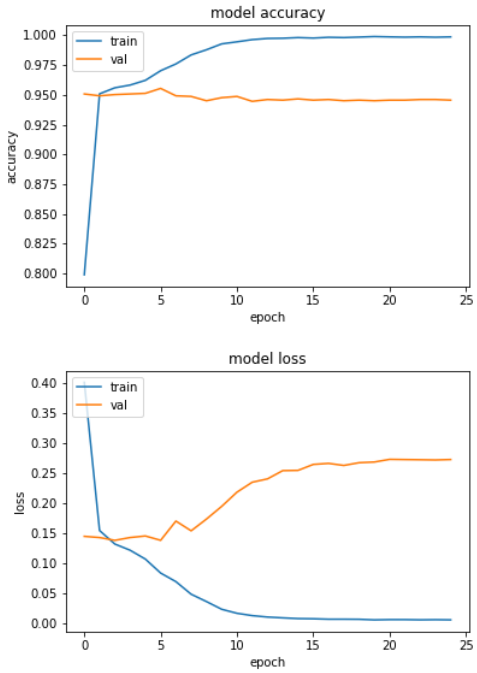

Fig. 4. Performance Graphs using Conventional Approach

25 epochs and achieved a training accuracy of 99.80%, but the validation accuracy stagnated at 94.56%. The accuracy and loss graphs of the model are shown in Fig. 4. The model performance was observed on the test set to see how well the model generalizes on unseen images. The performance was evaluated using metrics like Test Accuracy, F1 Score, AUC Score, Sensitivity and Specificity. The Table 1 depicts the performance on test set.

Although this is considered as an excellent result in deep learning experiments, the 5.2% difference in training accuracy and validation accuracy was a clear indication that the model wasn't generalizing that well

TABLE 1
Conventional Approach Performance on Test Set

Metrics	Performance
Training Accuracy	99.80%
Validation Accuracy	94.56%
Test Accuracy	94.48%
F1 Score	95.54%
AUC Score	95.48%
Sensitivity	95.90%
Specificity	95.12%

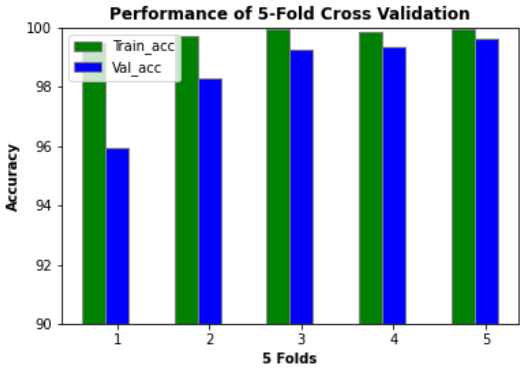

Fig. 5. Model Performance using K-Fold CV Approach

on the validation set as compared to the training set. So, we tried another training approach to check if that could yield better results. We tried the K-fold cross validation approach.

D. K-FOLD CROSS VALIDATION APPROACH

K-fold Cross Validation gives a model with less bias as compared to other methods. In this method, we have a parameter K, which represents the number of folds the dataset is divided into. Every fold appears in the training set (k-1) times. The value of k used is generally 5 or 10. We used a 5-Fold cross validation approach and trained the model for 20 epochs in each fold. Using this approach, the model achieved a training accuracy of 99.95% and validation accuracy of 99.61%. A significant improvement was observed in validation accuracy as compared to earlier approach. At the same time, training accuracy also showed a rise of 0.15% as compared to previous approach. The bar plot in Fig. 5 plots the performance of model after each fold. The model performance was also evaluated on testing set and the results obtained outperformed all the existing methods proposed by researchers for malaria detection. Table 2 compares performance of both approaches. As 5-Fold cross validation approach yields us outstanding results, we select and propose this approach for our model. Fig. 6 sums up the training methodology followed in this study.

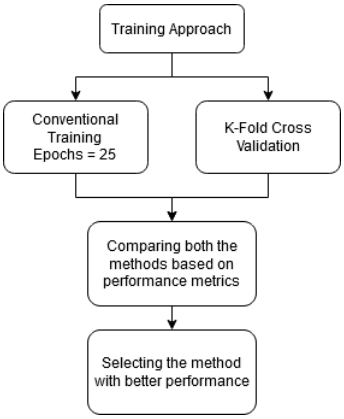

Fig. 6. Training Methodology

TABLE 2
Performance Comparison of Both Approaches

Metrics	Conventional Training	K-Fold CV
Training Accuracy	99.80%	99.95%
Validation Accuracy	94.56%	99.61%
Test Accuracy	94.48%	99.44%
F1 Score	95.54%	99.40%
AUC Score	95.48%	99.40%
Sensitivity	95.90%	99.92%
Specificity	95.12%	99.90%

IV. EXPERIMENTAL SETUP

In this study, the experiment environment was developed using Keras and Tensorflow libraries. The code was written and executed in ipython notebook (ipynb).

A. PARAMETERS

The training process is governed by model hyperparameters. They include variables which inturn determine how the model is trained. Table 3 describes the parameters used while training the model.

TABLE 11
Parameters Used

Parameters	Values
Training Method	5-Fold Cross Validation
Loss Function	Binary Crossentropy
Optimizer	ADAM
Batch Size	64
Epochs per Fold	20

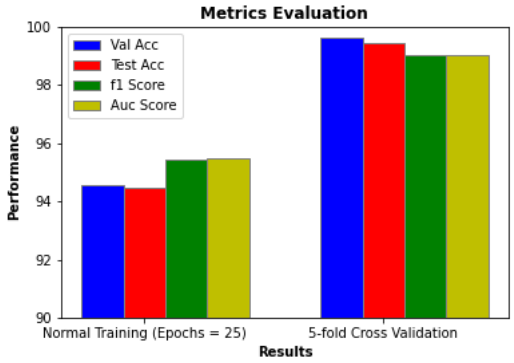

Fig. 7. Metrics Evaluation of both approaches

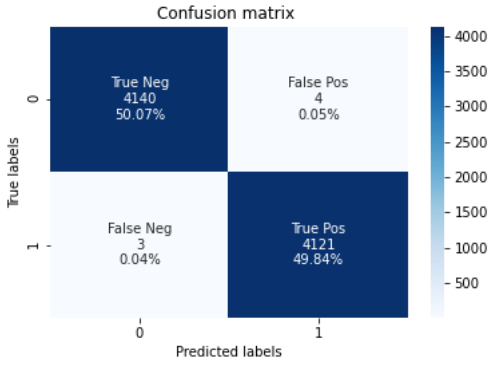

Fig. 8. Confusion Matrix for Test Set

B. METRICS

Evaluating a proposed method is an essential part of any research study. Metrics are used to evaluate the proposed model and to check its performance. Apart from accuracy, we evaluate our model using the following metrics:

1) F1 Score: F1-score is the harmonic mean of Recall and Precision, which are also performance metrics. The higher the F1-score, the better is the performance.

2) AUC Score: AUC stands for Area under the ROC Curve. The ROC Curve is a graphical plot that illustrates the diagnostic ability of a binary classifier. ROC curve is created by plotting True Positive Rate(TPR) against False Positive Rate(FPR) at variable threshold points. The AUC Score is a diagnostic measure of how well the model distinguishes between the two classes.

3) Sensitivity: Sensitivity is a measure of the proportion of actual positive cases that got predicted as positive i.e (True Positives).

$$Sensitivity = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

4) Specificity: Specificity is a measure of the proportion of actual negative cases that got predicted as negative i.e (True Negatives).

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$

V. EXPERIMENTAL RESULTS

In this study, we proposed a CNN based approach for malaria detection from cell images. Further, we used 5-Fold Cross Validation method for training our CNN model. The bar plot shown in Fig. 7 compares both the approaches using evaluation metrics and justifies our method of selecting the K-fold cross validation approach over the conventional approach. Fig. 8 shows the confusion matrix plotted for the predictions on the testing set using K-fold approach. From the figure, it can be observed that from the total 4144 negative images, only 4 were classified wrongly as positive. Also, from the total 4124 positive images, only 3 were classified wrongly as negative. Sensitivity and specificity metrics are calculated from the confusion matrix using the formulae mentioned in the previous section. We achieved a sensitivity of 99.92% and a specificity of 99.90%. The performance on testing Set can also be visualised with a ROC curve shown in Fig 9. We got an ideal ROC curve with an area of 1. We further compare our model performance with existing methods proposed for malaria detection. Table 4 compares proposed CNN with existing methods based on training and validation accuracy. Olugboja et. al [6], Yang et. al [8] and Bibin et. al [5] didn't mention validation accuracy for their model in their respective proposals. Also, Chowdhury et. al [9] didn't propose the training accuracy achieved for his method. Such cases have been indicated by a blank in the tables. Table 5 shows a similar comparison based on test set performance. In this table, we perform a detailed comparison based on test accuracy, F1 Score, AUC Score, specificity and sensitivity. From Table 4 and Table 5, it can be observed that our proposed method outperforms all the existing methods in terms of performance on all the three sets i.e Training Set, Validation Set as well as Testing Set. The outstanding performance achieved using the proposed method makes it robust and fit for real time deployment.

VI. CONCLUSION

Malaria, being a life threatening disease, its early diagnosis can save a lot of lives. The accuracy of diagnosing malaria from blood smears relies on the efficiency of medical professionals and the quality of instruments used in the diagnostic process. This leads to a heavy strain on medical professionals in rural areas with less medical facilities. Deep learning methods with high accuracy in diagnosing the disease can alleviate this strain on healthcare system and make the diagnostic process easier and faster. The proposed method can prove to be an effective

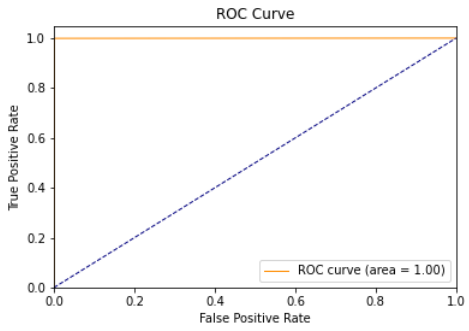

Fig. 9. ROC Curve for Test Set

TABLE IV
Performance Comparison based on Accuracy

Methods	Training Accuracy	Validation Accuracy
Olugboja et. al [6]	86.30%	-
Yang et. al [8]	93.40%	-
Bibin et. al [5]	96.35%	-
Kalkan et. al [7]	97.00%	95.00%
Chowdhury et. al [9]	-	97.07%
Proposed Method	99.95%	99.61%

TABLE 5
Performance Comparison based on Other Evaluation Metrics

Methods	Test Accuracy	F1	AUC	Specificity	Sensitivity
Yang et. al [8]	97.21%	80.81%	97.34%	98.39%	82.73%
Bibin et. al [5]	96.21%	89.66%	-	95.92%	97.60%
Pattanaik et. al [11]	89.10%	94.50%	-	83.10%	93.90%
Proposed Method	99.44%	99.40%	99.40%	99.90%	99.92%

medical diagnosis aid following its near negligible wrong predictions and high classification accuracy. Also, the fact that such a superior performance is achieved even after using a small neural network architecture, reduces the computational time, additionally giving our method an edge over existing methods that use heavy CNN models.

References

[1] World Health Organization (WHO), "World malaria report," https://www.who.int/malaria/publications/world-malaria-report-2018/en/, 2018.

[2] P. Mahdieh, S. Kamolrat, J. M. Richard, J. Stefan and T. George, "Image analysis and machine learning for detecting malaria," Translational Research, April 2018, vol. 194, pp. 36-55.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, 2015, vol. 521, pp. 436–444.

[4] S. Jürgen, "Deep Learning in neural networks: An overview," Neural Networks, January 2015, vol. 61, pp. 85–117.

[5] B. Dhanya, S. N. Madhu, P. Punitha, "Malaria parasite detection from peripheral blood smear images using deep belief Networks," IEEE Access, May 2017, vol. 5, pp. 9099–9108.

[6] O. Adedeji, W. Zenghui, "Malaria parasite detection using different machine learning classifier," International Conference on Machine Learning and Cybernetics (ICMLC), July 2017, pp. 246-250.

[7] C. K. Soner, K. S. Ozgur, "Deep Learning Based Classification of Malaria from Slide Images," Scientific Meeting on Electrical-Electronics and Biomedical Engineering and Computer Science (EBBT), June 2019, pp.1-4.

[8] F. Yang, M. Poostchi, H. Yu, Z. Zhou, K. Silamut, J.Yu, R. Maude, S. Jaeger and S. Antani, "Deep Learning for Smartphone-Based Malaria Parasite Detection in Thick Blood Smears," IEEE Journal of Biomedical and Health Informatics, May 2020, vol. 24, no. 5, pp. 1427 – 1438.

[9] A. Chowdhury, J. Roberson, A. Hukkoo, S. Bodapati and D. Cappelleri, "Automated Complete Blood Cell Count and Malaria Pathogen Detection Using Convolution Neural Network," IEEE Robotics and Automation Letters, January 2020, vol. 5, no. 2, pp. 1047 - 1054

[10] S. Nayak, S. Kumar and M. Jangid, "Malaria Detection Using Multiple Deep Learning Approaches," 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), September 2019.

[11] P. Pattanaik, M. Mittal and M. Khan, "Unsupervised Deep Learning Cad Scheme for The Detection Of Malaria In Blood Smear Microscopic Images," IEEE Access, May 2020, vol. 8, pp. 94936 – 94946.

[12] S. Militante, "Malaria Disease Recognition through Adaptive Deep Learning Models of Convolutional Neural Network," IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS), June 2019.

[13] A. Krizhevsky, I. Sutskever and G.E.Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, May 2017, vol. 60, no. 6.

[14] J. Redmon, S. Divvala, R. Girshick and A. farhadi, "You only look once: unified, real-time object detection," IEEE conference on computer vision and pattern recognition (CVPR), June 2016.

[15] S. Khan, N. Islam, Z. Jan, I. U. Din and J. J. P. C Rodrigues, "A Novel Deep Learning based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning," Pattern Recognition Letters, July 2019, vol. 125, no. 1, pp. 1-6.

[16] S. Fairuz, M.H. Habaebi and E. M. A. Elsheikh, "Finger Vein Identification Based on Transfer Learning of AlexNet," 7th International Conference on Computer and Communication Engineering (ICCCE), September 2018.

[17] J.S. Chima, A. Shah, K. Shah and R. Ramesh, "Malaria Cell Image Classification using Deep Learning," International Journal of Recent Technology and Engineering (IJRTE), March 2020, vol. 8, no. 6.

[18] M. Talo, U.B. Baloglu , O. Yildirim and U.R. Acharya, "Application of deep transfer learning for automated brain abnormality classification using MR images," Cognitive Systems Research, December 2018, vol. 54, pp. 176-188.

[19] A. Almisreb, N. Jamil and N. M. Din, "Utilizing AlexNet Deep Transfer Learning for Ear Recognition," Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), September 2018.

[20] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing and P. Xie, "Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans," IEEE TRANSACTIONS ON MEDICAL IMAGING, April 2020.