Detection of Malaria Using Deep Learning Techniques

Md. Siam Islam
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
170104124@aust.edu

Md. Tanvir Hasan
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
170104115@aust.edu

Popin Saha
Department of Computer Science and Engineering
Ahsanullah University of Science and Technology
Dhaka, Bangladesh
170104132@aust.edu

I. Introduction

Malaria is a disease that originated on the African continent. Malaria is caused by the plasmodium falciparum virus, which was created by the virus plasmodium falciparum. The illness has spread all across the world thanks to a mosquito-borne virus. The virus can live in hot and mild temperatures, but not in extremely cold one. The illness was 40 million years old and had existed since the dawn of time. Malaria may infect individuals of all ages, from infants to adults. Starting with a fever and progressing to a coma and death. The illness targets the human body's blood cells directly, breaking white blood cells and halting the functioning of the human organs. Malaria can only be identified by collecting blood samples from humans and examining them under a microscope.

When an infected Anopheles mosquito bites a human, Plasmodium parasites in the form of sporozoites infect the victim. The sporozoites enter the human liver rapidly and multiply asexually in the liver cells for 7 to 10 days. The parasites will then develop merozoites, which will travel through the circulation and lodge in the capillaries of the lungs. Merozoites are subsequently transported into red blood cells and proliferate, causing the cell to rupture. This will go through a cycle of a mosquito carrier passing on the disease to another healthy individual.

If a person has malaria, he or she will be aware of the symptoms that his or her human body will send as a warning signal. The human body will begin to activate white blood cells in order to offer immunity to malarial cells. It can result in a high temperature, headache, nausea, vomiting, stomach discomfort, and even coma. despite the fact that there were several machine learning algorithms for predicting malaria. In the proposed study, we utilized a deep learning model to accurately predict malaria.

II. Related Works

Variability and artifacts are crucial for capturing microscopic pictures of malarial cells, according to Raghuveer et al. [1] They took Leishman blood smears for this experiment, as shown in the model. As a result, we were able to grasp the notion of Leishman blood smears and implement it in our own project.

Ratnaparkhe et [2] demonstrated the idea of image processing using OpenCV and the contour detection concept, which we utilized in our study to discover the blood cell's characteristics using contour detection. So, after we've discovered the characteristics, we'll start counting the amount of dots to see if the cell is malarial or not.

Zhaoui et al [3] wrote an introduction to the deep learning idea of convolutional neural networks and how they may be used to determine if a blood cell is contaminated or not. So, for our proposed work and other convolutional neural networks, we adopted the notion of creating a scratch convolutional neural network.

The backpropagation feedforward neural network idea was first presented by Ross et al [4]. This project has a faster learning rate than a simple convolutional neural network. As a result, comprehending the neural network idea employed in that model.

Gopalakrishna et al [5] pioneered the notion of creating an artificial microscopic slide of the malaria-causing plasmodium falciparum cell. In order to comprehend the morphology of the malarial cell, the idea of cell structure and cell characteristics is explored.

III. Project Objective(s)

The research offers a classification model for images that may be used to identify malaria-infected cells. We use deep

learning methods to identify parasite-infected red blood cells in tiny smears on conventional microscope slides. The most often utilized technique nowadays is microscopically analyzing thin blood smears and visually looking for infectious cells. A physician manually counts the parasite red blood cells, sometimes as many as 5,000 cells (per WHO standards) [6]. Malaria might be more effectively prevented, managed, and treated if more accurate and effective symptomatic methods were available. To determine the proximity of malaria-infected cells, we used image processing techniques. And we utilize deep learning to classify the malaria stage, i.e. parasitized or uninfected. Tasks of our system is given below:

1) Collect the required dataset.
2) Split the images according to train, test and validation using sklearn and save those images on the targeted folder in drive.
3) Image normalization.
4) Convert all images to be 224 x 224 and use 32 as batch size to process this number of images at a time.
5) Training data augmentation, i.e. Rotation, Width Shift, Height Shift, Zoom, Horizontal Flip.
6) Now insert the images into the model and run the model by using tensor flow and keras package.
7) Use Epochs = 20, 15 and 20 for DNN, CNN and Frozen CNN.
8) Classify malaria cells using DNN, CNN and Frozen CNN.
9) Compare all implemented models.

For a given input image, our model will predict whether the output will be parasitized or uninfected. Input and probable output of an image is shown in Fig. 1.

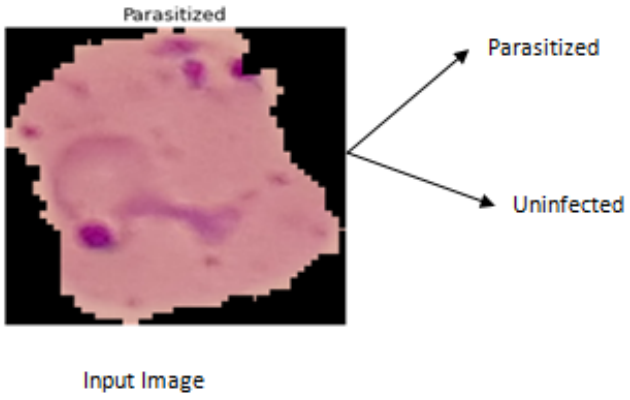

Fig. 1. Prediction of an input image

IV. Mehodology

The proposed framework relies on image classification. The details of the proposed method were depicted in Fig. 2. The whole workflow is divided into three phases to classify Malaria Cell Disease, including data acquisition, data processing, and data classification using the proposed machine learning classifiers.

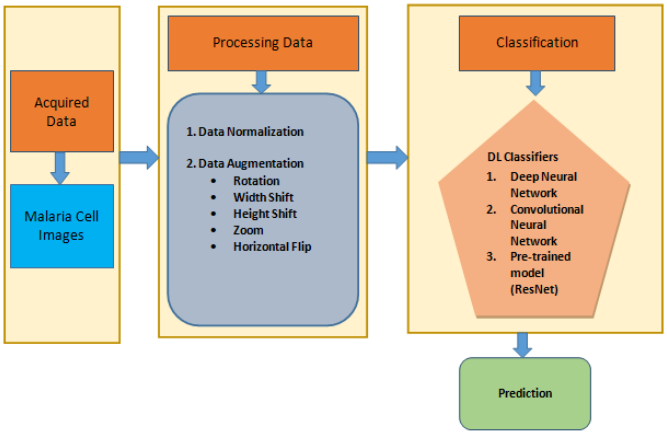

Fig. 2. Diagram of proposed methodology

A. Data Preprocessing

Once all of the pictures have been collected, the process of training, validation, and testing may begin. We utilized a 20% testing and validation set, with the remaining set serving as the training set. We normalized images, used batch size of 32 and target image size of (224,224) and set class mode as binary using ImageDataGenerator library of tensorflow. We also used augmented training data to train CNN, Frozen CNN model and some of the data augmentation properties are: Rotation, Width Shift, Height Shift, Zoom, Horizontal Flip. We used ResNet50 architecture as pre-trained CNN model to extract features of images and then forward the extracted images to the Dense Layer to classify images using Frozen CNN model.

We implemented various deep learning models, such as Deep Neural Network, Convolutional Neural Network, Frozen Convolutional Neural Network to classify malaria cells. The following subsections represent the implemented models.

B. Deep Neural Network

Artificial neural networks (ANNs) are networks of artificial neurons based on biological neurons. Every ANN includes at least three layers: an input layer that accepts input, a hidden layer that trains on the dataset given to the input layer, and an output layer that outputs a result based on the application. Deep learning has received widespread praise for producing outcomes that have never been seen in any previous machine learning technique. Deep Neural Networks (a kind of ANN) are widely utilized for classification tasks and have outperformed conventional machine learning classifiers. Fig. 3 shows a summary of our implemented DNNs.

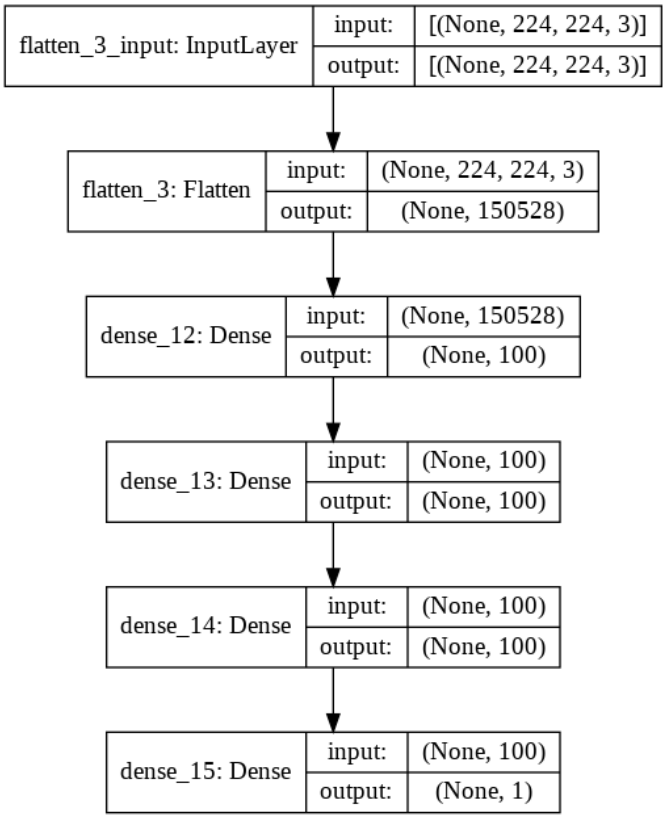

Fig. 3. Implemented Deep Neural Network model's summary

C. Convolutional Neural Network

The Basic Convolutional Neural Network is built from the ground up. As a result of utilizing Tensor Flow, the Python library frees Keras. Keras is a Python open-source neural network library. As a result, the Keras has an attribute called Conv2D. It is also Conv1D, however Conv2D is utilized in this project since the picture being used is two-dimensional. Then follows the max-pooling phase, which aims to maximize the cluster of neurons in the preceding layers. On CNN, max pooling is a down sampling technique. Max pooling is often used to converge any kind of matrix to the lowest feasible value. Take, for example, the 4X4 matrix in the matrix, which has four corner values. We determine the determinants of the four values using the max-pooling effect, and the result is a 2X2 matrix. There is another layer called Flatten in the Keras layers that is utilized to prepare to become a vector for the completely linked layers by utilizing the flatten command with regard to keras. So, after the keras characteristics have been completed, the sigmoid will be activated by computing the accuracy. After acquiring the keras's characteristics, the model is constructed using keras. Fig. 4 shows a summary of our implemented CNNs.

D. Frozen Convolutional Neural Network

The order of presentation is the same for both basic and frozen CNN, but the given attributes change as the frozen

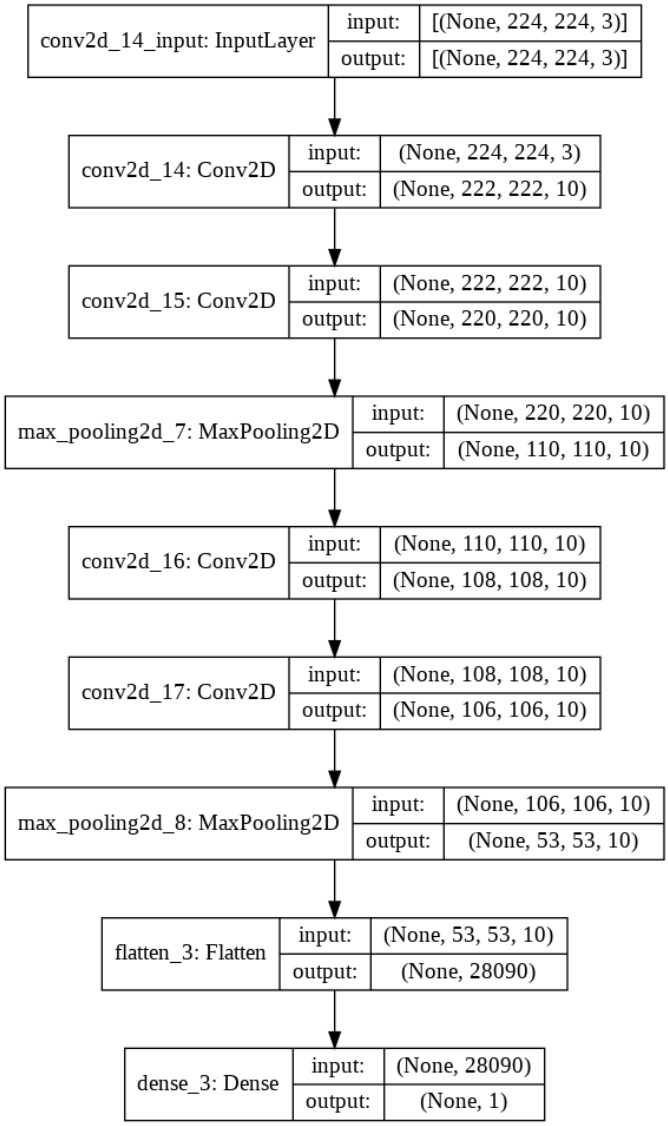

Fig. 4. Implemented Convolutional Neural Network model's summary

CNN takes place. The same thing happens in this CNN, but this time Keras is combined with ResNet. TensorHub contains a plethora of ResNet variants. The main difference between ResNet versions is that the higher the version, the more complex the model. ResNet50 is used in this project to make it run effectively on the images and model for training. We made no changes to the base model of the ResNet50 layers because we used a Frozen CNN model with all trainable parameters set to False. We used ResNet50 architecture as feature extractor for our classification problem. Then we fed the extracted features to the Dense Layer for classification. Fig. 5 shows a summary of our implemented Frozen CNNs.

V. Experiments

A. Dataset

The primary goal of this project is to create an effective deep learning model for predicting Malaria illness. Malaria

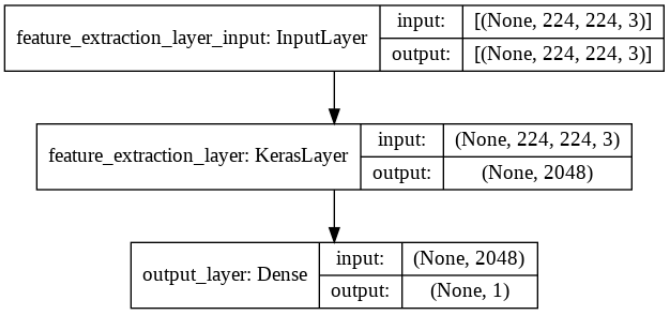

Fig. 5. Implemented Frozen Convolutional Neural Network model's summary

cell images dataset [7] consist a total of 24849 images where Uninfected and Parasitized cell images are of 11069, 13780 respectively. Data distribution of the acquired dataset is shown in Fig. 6.

We splitted our dataset into train, test and validation set where

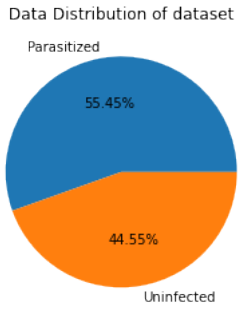

Fig. 6. Data Distribution of dataset

both test and validation set consists of 20% of the data and remaining portion is the training set.

B. Results

We implemented three deep learning models including DNN, CNN and Frozen CNN. Performance metrics are calculated for all of these models. Loss curves of CNN Model for each epochs on training and validation set is shown in Fig. 7. Evaluation metrics of CNN on test data is measured and shown in Fig. 8 and Fig. 9.

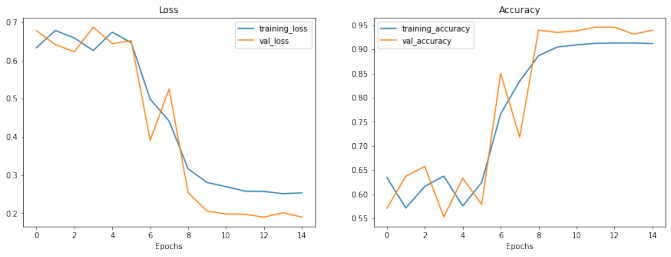

Fig. 7. Loss curves of CNN Model for each epochs on training and validation set

Loss curves of DNN Model for each epochs on training and validation set is shown in Fig. 10. Evaluation metrics of DNN on test data is measured and shown in Fig. 11 and Fig. 12.

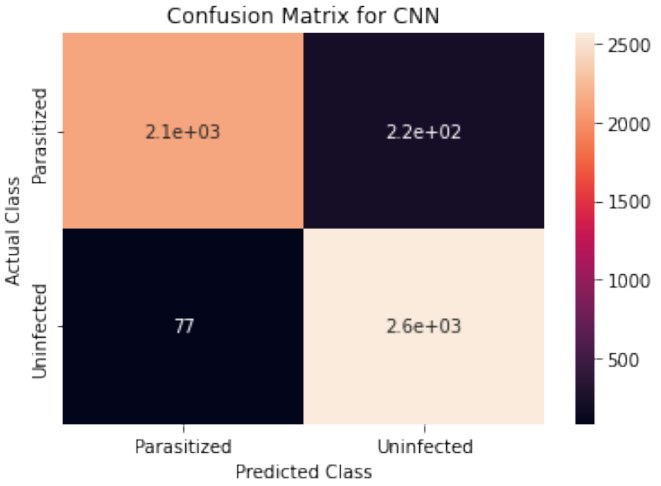

Fig. 8. Confusion Matrix for CNN

	cnn
Accuracy	94.07
Precision	90.67
Recall	96.51
F1-Score	93.50

Fig. 9. Performance metrics for CNN

Fig. 10. Loss curves of DNN Model for each epochs on training and validation set

Loss curves of Frozen CNN Model for each epochs on training and validation set is shown in Fig. 13. Evaluation metrics of Frozen CNN on test data is measured and shown in Fig. 14 and Fig. 15.

The findings of the classification of parasitized or uninfencted malaria cells are presented in Table. I. As indicated, many models have been evaluated to accomplish the classification task, and it is clear from the results that CNN outperforms other models. Our obtained accuracy on test set for CNN

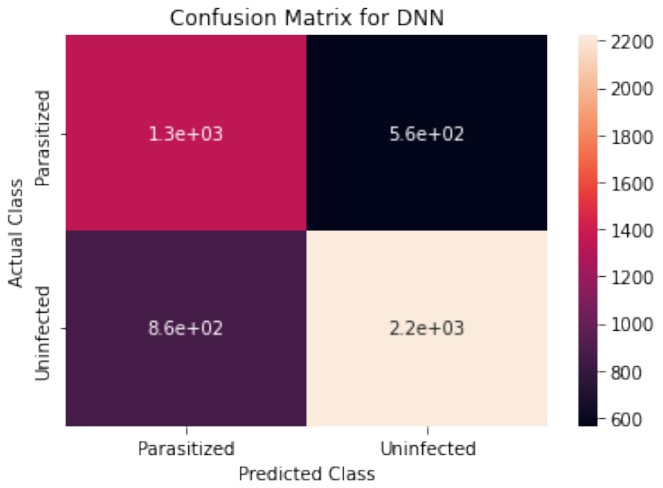

Fig. 11. Confusion Matrix for DNN

	DNN
Accuracy	71.45
Precision	70.48
Recall	60.92
F1-Score	65.35

Fig. 12. Performance metrics for DNN

is 94.07%. A comparison of performance metrics for these models is plotted in Fig. 16.

TABLE I
COMPARISON OF THE OUTCOMES ACHIEVED USING VARIOUS DEEP LEARNING MODELS

Method	*Accuracy*	*Precision*	*Recall*	*f1-Score*
DNN	71.45	70.48	60.92	65.35
CNN	94.07	90.67	96.51	93.50
Frozen CNN	92.49	87.44	96.92	91.94

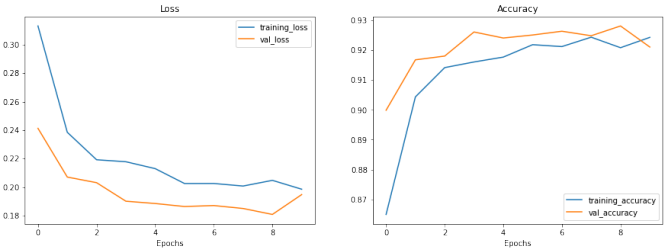

Fig. 13. Loss curves of Frozen CNN Model for each epochs on training and validation set

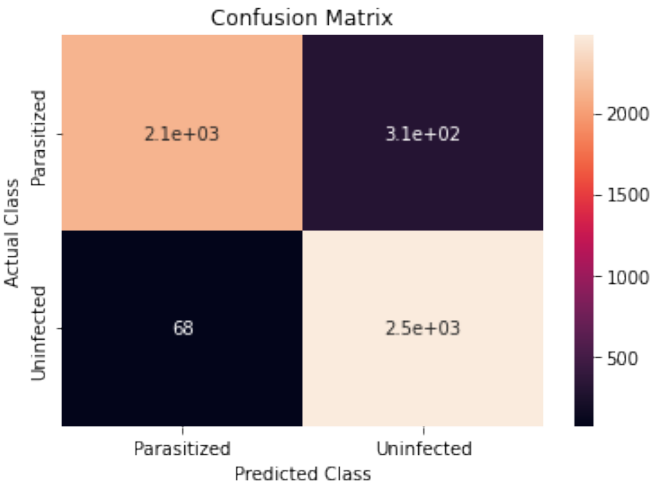

Fig. 14. Confusion Matrix for Frozen CNN

	Frozen CNN
Accuracy	92.49
Precision	87.44
Recall	96.92
F1-Score	91.94

Fig. 15. Performance metrics for Frozen CNN

VI. Conclusion and Future Directions

The conventional approach of delivering samples and evaluating cell development takes more time to identify malaria. As a result, a deep learning model has been developed in the proposed study to predict Malaria with a high accuracy rate and a short time span. Three deep learning models were built, with the most accurate model being chosen. In comparison to the other implemented models, we achieved a high accuracy rate using CNN. Work on disease diagnosis,

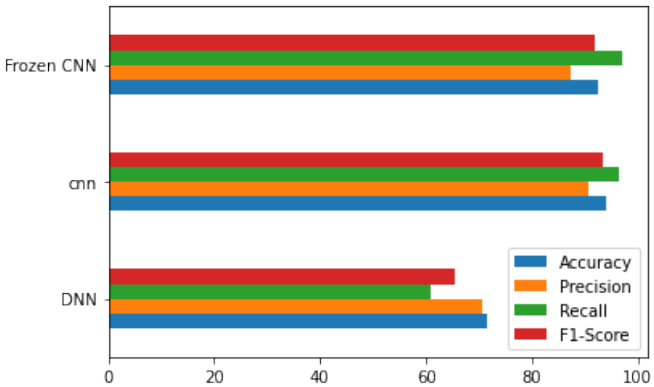

Fig. 16. Bar plot for performance metrics.

such as pneumonia and breast cancer, will be done in the future, as well as preparing for the detection of COVID19 smears in the human lungs.

REFERENCES

[1] V. V. Makkapati and R. M. Rao, "Segmentation of malaria parasites in peripheral blood smear images," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2009, pp. 1361–1364.

[2] K. Khatri, V. Ratnaparkhe, S. Agrawal, and A. Bhalchandra, "Image processing approach for malaria parasite identification," in *International Journal of Computer Applications, National Conference on Growth of Technologies in Electronics, Telecom and Computers-India's Perception*. Citeseer, 2013.

[3] G. Shekar, S. Revathy, and E. K. Goud, "Malaria detection using deep learning," in *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*. IEEE, 2020, pp. 746–750.

[4] N. E. Ross, C. J. Pritchard, D. M. Rubin, and A. G. Duse, "Automated image processing method for the diagnosis and classification of malaria on thin blood smears," *Medical and Biological Engineering and Computing*, vol. 44, no. 5, pp. 427–436, 2006.

[5] Y. Dong, Z. Jiang, H. Shen, W. D. Pan, L. A. Williams, V. V. Reddy, W. H. Benjamin, and A. W. Bryan, "Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells," in *2017 IEEE EMBS international conference on biomedical & health informatics (BHI)*. IEEE, 2017, pp. 101–104.

[6] "World Health Organization, Malaria," https://www.who int/news-room/fact-sheets/detail/malaria(2018), [Accessed: 29 September, 2021].

[7] S. Rajaraman, S. Jaeger, and S. K. Antani, "Performance evaluation of deep neural ensembles toward malaria parasite detection in thin-blood smear images," *PeerJ*, vol. 7, p. e6977, 2019.