Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

Malaria Detection Using Smear Images

Dr.Mahesh U.Nagaral[1] and Mohammadi Jamadar[2]

[1] BLDEA's V P Dr PGHCET Vijayapura, Karnataka, India
[2] BLDEA's V P Dr PGHCET Vijayapura, Karnataka, India
[1]maheshunagaral@gmail.com
[2]mohammadijamadar777@gmail.com

Abstract. Malaria is induced by parasites belonging to the Plasmodium genus, a blood disease spread by mosquitoes. The traditional method of diagnosing malaria involves examining the patient's stained blood cells under a micro- scope. The rapid and accurate detection of the illness in microscopic images of blood smears is essential for the treatment and control of malaria. This paper presents an approach to malaria detection utilizing a recurrent neural net- work(RNN) with long short-term memory(LSTM), renowned for its ability to effectively process sequential input. A skilled technician examines the slide while paying close attention to both visual and mental details. In this study, Plasmodium parasites in blood smear slides were identified and measured using a novel approach in image processing that produces a machine learning system that can gather, distinguish, and classify different types of infected cells.

Keywords: Malaria, Plasmodium, Edge Mask, RCB, Watershed Segmentation, Support Vector Machines.

1. Introduction

Malaria parasites are spread by Mosquito bites possess the capacity to be fatal. The two primary forms of blood smears used to diagnose malaria are thin and thick smears are commonly utilized. Finding parasites in a blood droplet requires a thick smear. Malaria in humans is triggered by five distinct Plasmodium spe- cies: Plasmodium knowlesi, Plasmodium knowlesi, Plasmodium malaria, Plas- modium vivax, and Plasmodium ovale. Among these, the plasmodium falciparum and plasmodium vivax are the most frequent types. The most serious symptom is because of plasmodium falciparum accounting for a majority of global malaria-related fatalities. In this study, Using LSTM, we suggested a method for detecting malaria. The proposed method encompasses various phases, including data collection, preprocessing, image patch extraction, LSTM model architecture design, data augmentation, training, model evaluation, testing, and post-processing. Each stage is carefully designed to ensure accurate detection and robust performance of the LSTM-based malaria detection system.

By automating the malaria detection process using LSTM, we envision that healthcare facilities in resource-constrained areas can benefit from faster and more reliable diagnosis, leading to timely treatment and improved patient out-

comes. The goal of the LSTM Method is to help develop an effective and automated method for the rapid and precise detection of malaria. Malaria parasites detected in our dataset include Plasmodium knowlesi, Plasmodium malaria, Plasmodium Ovale, and Plasmodium Vivax. The image processing approach involves several procedures, starting with the employment of an auto-generated segmenter, which completes the four critical stages: (1) Convert to grayscale (2) Initialize segmentation with Otsu‘s threshold. (3) Filter components by area (4) Form masked from input image and segmented image.

After successfully detecting the cells, the subsequent objective is to identify infections within those cells. To achieve this, we will utilize pixel intensity values and examine the intensity ranges where the infections are present. However, a challenge arises due to changes in illumination and color scales between different images. To address this, we will employ histogram matching to ensure a consistent color scale across all images. Once the segmentation of infected regions is complete, we proceed to classify the infected cells. To differentiate between infected and non-infected cells, machine learning techniques will be employed to fulfill this categorization challenge.

2. Related Work

WHO, Fact sheet: World Malaria Report [1] suggests Vitamin A has been linked to immunity and defense against infectious illnesses in research on animals. A comprehensive and methodical examination of the literature was performed, involving searches on PubMed and the Cochrane Library, to locate Randomized Controlled Trials (RCTs) pertaining to how vitamin A utilized to prevent disease and treatment of malaria throughout pregnancy and childhood. The results of Randomized Controlled Trails(RCTs) demonstrate that vitamin A has no advantage in preventing or treating malarial infection during pregnancy or infancy.

Bloland PB [2] suggest Resistance of parasites to cheap, first line malaria drugs is a significant obstacle to reducing the almost a million deaths due to malaria each year worldwide. Overdiagnosis rates at the community and peripheral levels, where self-treatment is common, are startlingly high when compared to the number of parasitologically proven cases of malaria. If more expensive drug treatments are introduced, this degree of misdiagnosis of malaria cannot be condoned or sustained.

D Ghate, C. Jhadav, and N. U. Rani et al. [3] suggest malaria comes under among the dangerous diseases in many countries. It is the primary reason for most of the causalities across the world. Therefore, to facilitate the early detection/diagnosis of malaria to lesser the mortality rate, an automated computational method is required with a high accuracy rate. different computer-assisted techniques has been outlined as follows: (i) acquisition of image dataset, (ii) preprocessing, (iii) segmentation of RBC, and (iv) extraction and selection of features, and (v) employ- ing blood smear pictures for classification in parasite identification in malaria. As

a result, it significantly enhances the current computational techniques for making an early diagnosis of malaria.

In this study, the viability of automated red blood cell abnormality detection is evaluated by Ahirwar A., Pattnaik S., and Acharya B et al. [4] and describes a method for recognizing and malaria parasite classification in light microscope images of blood samples. The primary method for diagnosis is a visual microscopic examination of blood smears stained with Giemsa. Using the image classification technique, malaria parasites found in thin blood smear can be positively recognized. Differentiating between parasitized and nonparasitic blood pictures is possible with a back propagation feed forward neural network-based classifier.

They devised an image classification algorithm proposed by Amit Kumar, Prof.Arun Choudhary, Prof.P.U.Tembhare, Prof.C.R.Pote et al. [5] In thin blood smears, malaria parasites can be seen. In this method, the segmentation of blood cells is possible by using Otsu thresholding to either the green or blue channel of an RGB blood picture.This process involve two distinct phases: initially, the overall count of parasites is computed, followed by the calculation of the total count of RBC cells in the subsequent phase. These counts are then utilized to calculate the percentage of parasitemia. The majority of ailments are caused by blood. An image segmentation method is used to analyse the blood cell image. Numerous researchers have contributed to this field by utilizing a different image processing methods to analyze blood cell images.

Extensive collections of digital micrographs acquired from high-throughput screening tests must be reliably analyzed using sophisticated image processing techniques. Ritter, H. et al., Nattkemper, T.W. Schubert, W. Hermann, et al. [6] The method is composed of model-based fitting methods, morphological opera- tors, and Artificial Neural Networks (ANN). Due to the constrained efficacy of individual methods, we support a hybrid framework that combines a multi-layer perceptron (MLP) network and the Hough transform. Our research shows that this amalgamation of techniques enhances performance, leading to heightened sensitivity and positive predictive value in obtaining cell body positions.

3. Proposed Method

• Convolutional Neural Network (CNNs) are commonly utilized for image recognition tasks, object detection, and natural language processing. CNNs are powerful models, but they have certain limitations. It's crucial to remember that CNNs have been quite effective in a variety of applications and still an important tool in the deep learning space.

In figure 1 a typical CNN comprises several layers: convolutional layer, pooling layer, a flattening layer, and a fully connected layer. The convolutional layer is the main component of the CNN network which operates on the principal of reducing computational complexity the kernal technique is utilized in this layer to extract various feartures from input data. The pooling layer, which comes after, is made to reduce connections between layers and consequently, the size of the underlying image.

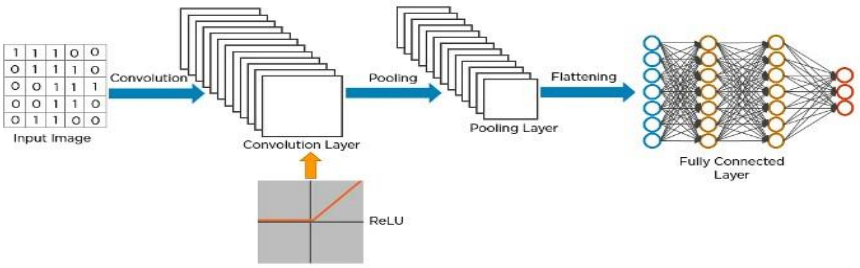

Fig 1: CNN Model.

•The reduction of dimensionality is the primary objective of the pooling layer and derive the prominent characteristics to facilitate the effective training of the model. After several convolutional and pooling layers, a flattening layer must be used before the fully connected layer, the output is flattened into a 1D vector. The fully connected layers learn global patterns and relationships between high-level features extracted by the convolutional layers. The final output layer of the CNN produces the network's predictions.

•The number of nodes in this layer will depend on the particular task. For instance, in image classification, the number of nodes and the number of classes are correlated. Long Short Term Memory (LSTM), a kind of recurrent neural network (RNN), is shown in Figure 2 and has the potential to identify temporal dependencies and pat- terns in sequential data. The creation of automated systems that accurately identify malaria-infected cells, LSTM algorithm can be used to analyze blood smear images.

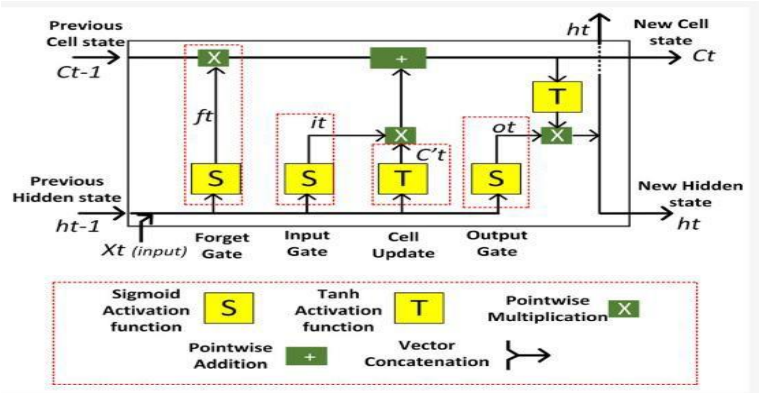

Fig 2: LSTM Model.

•The LSTM model incorporates memory cells and gates to control network infor- mation, allowing it to retain and utilize information effectively over extended dura- tions. This architecture demonstrates commendable predictive performance.

•It comprises memory blocks referred to as cells, with each cell possessing two states: the cell state and the hidden state. Within the LSTM network, these cells play a pivot- al role in crucial decision-making by retaining information about significant elements.

These basic components are known as gates, namely the forget gate, input gate, and output gate.

• LSTM model operates in three stages:

1] The system and forget gate work together to decide what information should be stored within the cell state. The calculation starts by using the sigmoid function to take into account both the input at the current time step, written as x_t, and the previous value of the hidden state, marked as h_{t-1}. The following formula is used to calculate within the forget gate.

$$f_t=\sigma(W_f \cdot [h_{(t-1)}, x_t]+b_f)$$

2] The next stage involves the network's computation changing the old cell state ($C_{(t1)}$) into the new cell state (C_t). The transformation of the previous cell state ($h_{(t1)}$) into the new cell state (h_t) is also a part of this continuing calculation. The cell state, which is determined by this mechanism, influences whether new information is stored in long-term memory. It is crucial to take the reference values from the cell update gate, input gate, and forget gate into account when calculating the new value for the cell state. Below are the formulas for this phase.

$$i_t=\sigma(W_i \cdot [h_{(t-1)}, x_t]+b_i) \quad (1)$$

$$C'_t=\sigma(W_C \cdot [h_{(t-1)}, x_t]+b_C) \quad (2)$$

$$C_t=(C_{(t-1)} \cdot f_t)+(i_t \cdot C'_t) \quad (3)$$

The final step involves determining the value of the hidden state (h(t)) after the cell status has been updated. The hidden state is meant to act as the network's memory at this stage, encompassing insights from past data and employed for predictive purpos- es. The output gate (ot) and the reference value from the recently updated cell state must both be included in the calculation in order to determine the value of the con-cealed state. The formula governing this procedure is indicated below.

$$o_t=\sigma(W_o \cdot [h_{(t-1)}, x_t]+b_o$$

$$h_t=o_t \cdot \sigma(C_t)$$

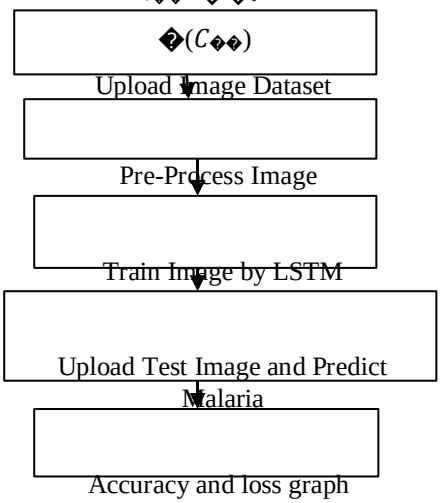

Fig 3: Flow chart for proposed system.

Figure 3 show flow chart for proposed system identify malaria upload the image of the dataset, pre-process that image, train those images, upload text images, de-

termine whether malaria is there or not, then give the accuracy and loss graph and next confusion matrix.

Here's a methodology for detection of malaria using LSTM (Long Short-Term Memory) method:

- Data Collection: Images obtained by microscopy may be utilized to compile a dataset of red blood cell in blood smears that are both malaria-infected and uninfected. Make sure the labels on the images are accurate.
- Data Pre-processing: Pre-process the images by resizing them to a consistent resolution and converting them to grayscale. Apply image enhancement techniques, such as contrast adjustment or histogram equalization, to improve image quality. Normalize the pixel values to a suitable range, typically between 0 and 1.
- Image Patch Extraction: Divide the preprocessed images into smaller patches or tiles. This step helps in capturing local information and reducing the computational complexity. Each patch should be of sufficient size to contain relevant cell structures.
- Patch Labeling: Label each patch based on whether it contains infected or uninfected cells. Use the labels from the original images to determine the patch labels. For instance, if a patch contains even a single infected cell, label it as infected.
- LSTM Model Architecture: Create an malaria detection using an LSTM-based deep learning model. Each patch will be considered as a sequence of image frames. The LSTM layers will capture the temporal dependencies within the patch sequence. Additional layers, such as feature extraction using convolutional layers and pooling layers for downsampling, can be incorporated into the model.
- Data Augmentation: Augment the dataset by applying various transformations to the patches, such as rotation, scaling, flipping, and adding noise. This step increases the diversity of the traning data and helps in improving the model's generalization ability.
- Training: Create training and validation sets from the dataset. Feed the training set, consisting of the labeled patches, into the LSTM model. Training the model with an appropriate optimizer, like Adam or RMSprop, and a suitable loss function, such as binary cross-entropy. Experiment with different hyperparameters, such as learning rate, batch size, and number of LSTM units, to raise the model's efficiency.
- Model Evaluation: Evaluate the trained model on the validation set to assess its performance. To accurately detect malaria-infected cells, compute assessment measures such as accuracy, precision, recall, F1 score, and AUC-ROC.
- Testing: Apply the trained model to an independent test set to evaluate its generalization performance. Measure the performance indicators for the model on the test set and assess its effectiveness in malaria detection.
- Post-processing: Apply a threshold to the model's output probabilities to classify each patch as infected or uninfected. Additionally, you can perform morphologi- cal operations or post-processing techniques to refine the predictions, such as removing small false positives or filling gaps in the predicted boundaries.

4. Results and Discussion

The process includes several stages, including data collection, preprocessing, image patch extraction, LSTM model architecture design, data augmentation, training, model evaluation, testing, and post-processing. Each stage is carefully designed to ensure accurate detection and robust performance of the LSTM-based malaria detection system. By automating the malaria detection process using LSTM, we envision that healthcare facilities in resource-constrained areas can benefit from faster and more reliable diagnosis, leading to timely treatment and improved patient outcomes. Our dataset (all images) comes from Center of Dis- eases Control and Prevention https://data.cdc.gov/. The dataset includes 722 cell images with parasitized and healthy RBCs. Moreover, the proposed method has the capacity to help healthcare professionals by reducing their workload and al- lowing them to focus on other critical tasks.

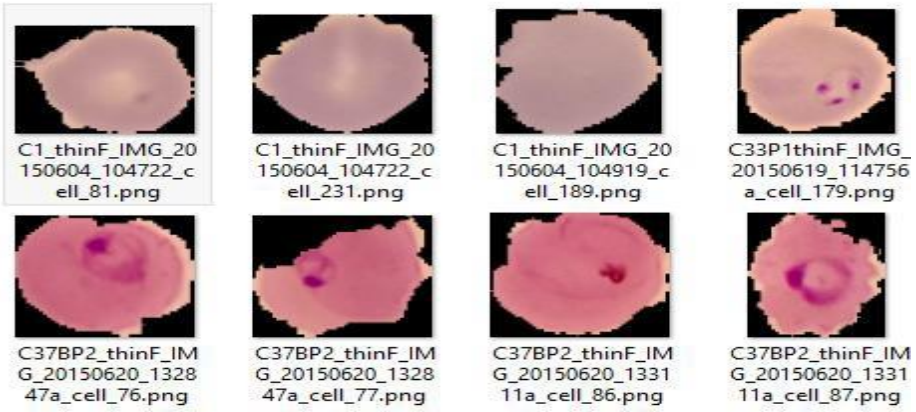

Fig 4: Sample dataset Of Malaria.

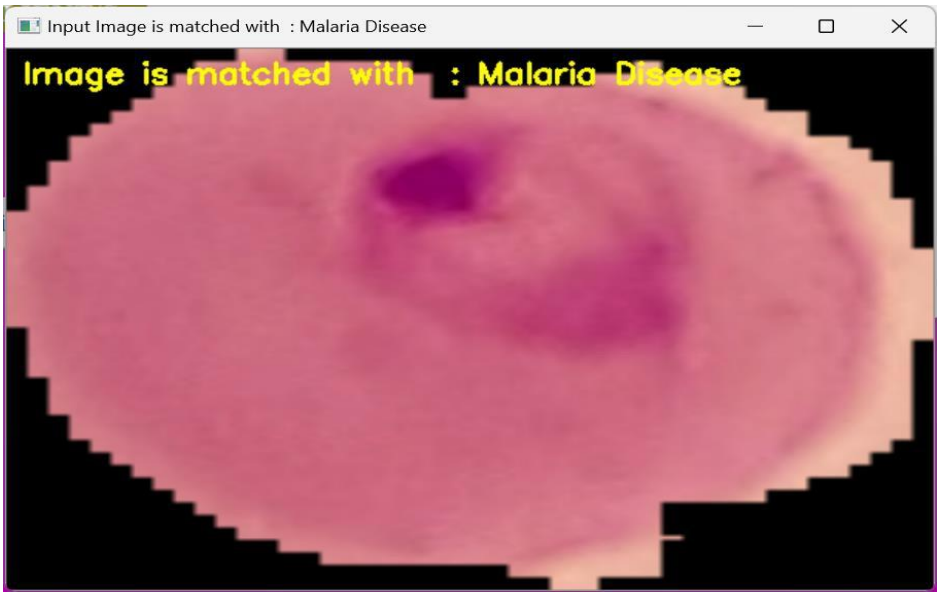

Fig 5: Malaria disease identified in given image.

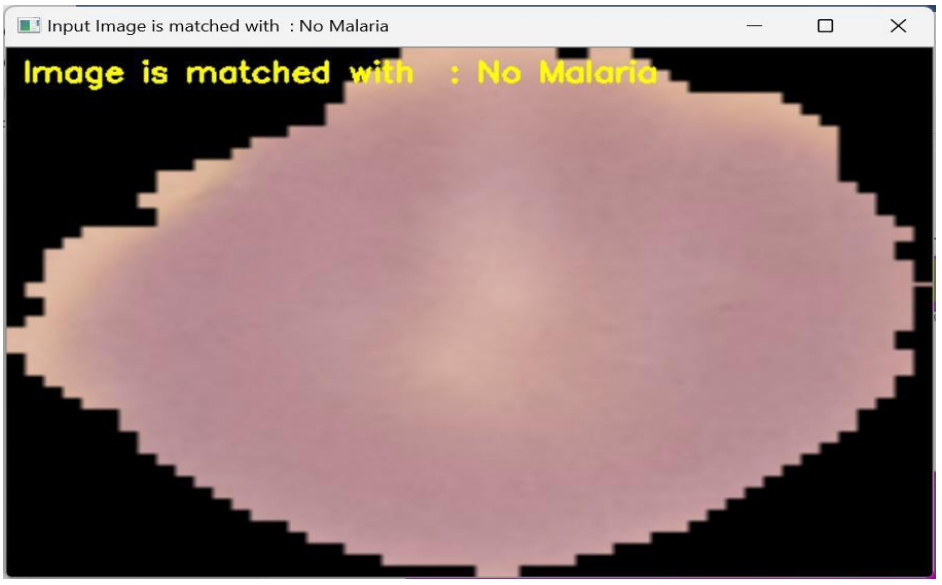

Fig 6: Malaria disease not identified in given image.

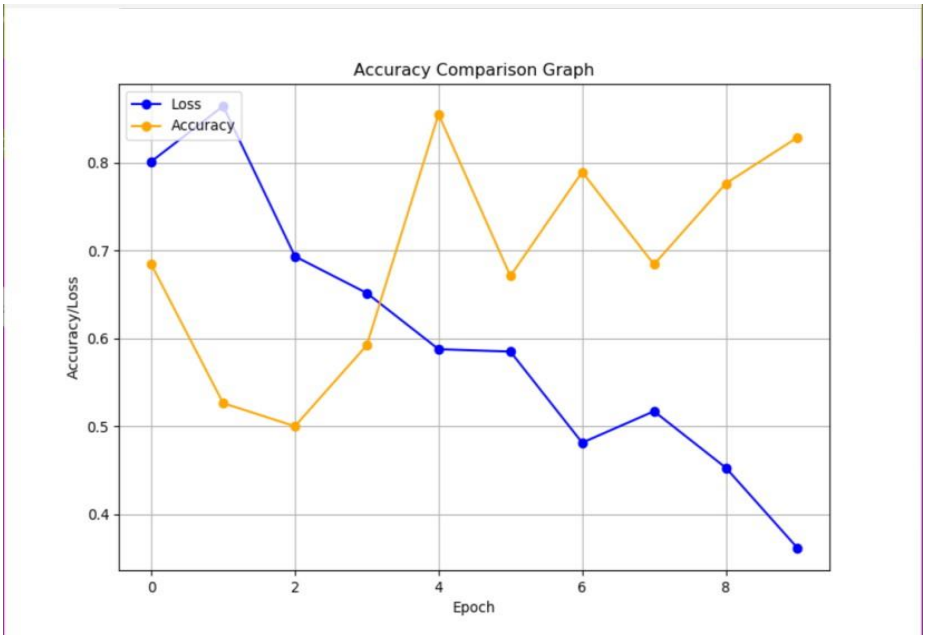

Fig 7: Accuracy and Loss comparison Graph.

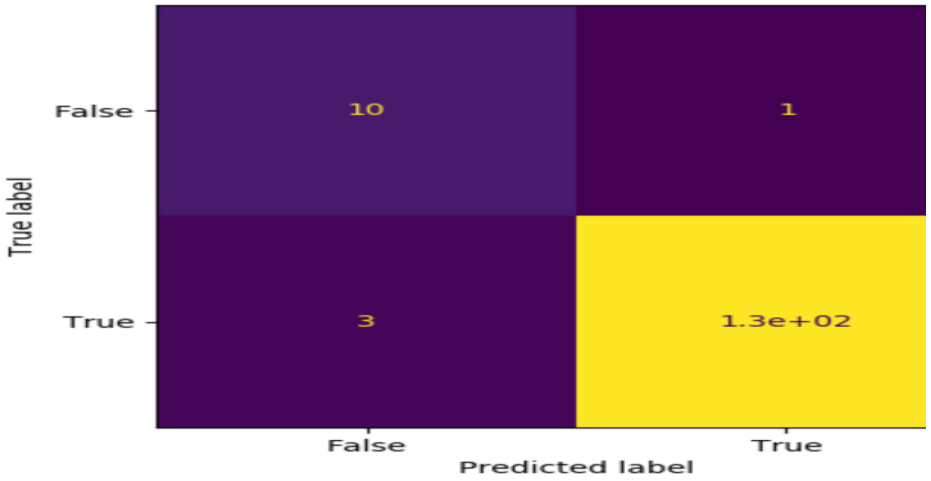

Fig 8: confusion matrix for testing data.

- Confusion matrix shows the performance of classifier.
- The Confusion Matrix created has four different quadrants:

  False Negative (Top-Left Quadrant) = 10

  True Positive (Top-Right Quadrant) = 1

True Negative(Bottom-Left Quadrant) = 3
False Positive (Bottom-Right Quadrant) =1.3e+02

- True means that the values were accurately predicted, False means that there was an error or wrong prediction.
- The CNN Model prediction accuracy is less than LSTM that is 68%
- Accuracy measures how often the model is correct.

Accuracy = (False Positive + True Negative) / Total Predictions.
= (1.3e+02+3)/144
= 133/144 = 0.92= 92%

5. Conclusion

Here LSTM outperforms both conventional machine learning and CNN models. Performance graph using LSTM shows that as increase in epochs there is increase in accuracy. The performance is demonstrated through a graph illustrating accuracy and loss. The dataset comprises two categories of images: one depicting malaria-infected samples and the other consisting of healthy images. In conclusion, the methodology for malaria detection using LSTM (Long Short-Term Memory) offers a systematic approach was developed to use deep learning algorithms to detect malaria-infected cells in microscopic images of blood smears. By following the outlined steps, we can build an effective model for malaria detection.
Future work, Researchers are actively working on addressing the limitations of CNN through advancements in architecture design, regularization techniques, and the integration of other models, such as recurrent neural networks (RNNs) and attention mechanisms.

References

1. WHO, "Fact sheet: World Malaria Report 2016," in World Health Organization, world Health Organization, 2016. [Online]. Available: http://www.who.int/malaria/media/world-malaria-report-2016/en. Accessed: Jan. 25, 2017
2. Bloland PB (2001) Drug resistance in Malaria, WHO/CDS/ CSR/DRS/ 2001.4. World Health Organization, Switzerland, 2001
3. D Ghate, C. Jadhav, and N. U. Rani, "AUTOMATIC DETECTION OF MALARIA PARASITE FROM BLOOD IMAGES,". [Online]. Available: http://ijact.org/volume4issue1/IJ0410050.pdf. Accessed: Jan. 25, 2017
4. Ahirwar A., Pattnaik S., Acharya B., Advancced Image Analysis Based System for Automatic Detection and Classification of Malarial Parasite in Blood Images. International Journal of Information Technology and Knowledge Management. Jan-June 2012, Volume 5, No 1, pp. 59-64
5. Amit Kumar, Prof. Arun Choudhary, Prof. P. U. Tembhare, Prof. C. R. Pote. Enhanced Identification of Malarial Infected Objects using Otsu Algorithm from Thin Smear Digital Images. International Journal of Latest Research in Science and Technology Vol.1, Issue 2: Page No. 159-163, July-August (2012).
6. Nattkemper, T. W., Schubert, W., Hermann, T., & Ritter, H. (2004). A Hybrid System for Cell Detection in Digital Micrographs. In B. Tilg (Ed.), Biomedical Engineering, Proc. (Vol. 417). Innsbruck, Austria: ACTA Press.