Malarial Parasite Classification using Recurrent Neural Network

Muhammad Imran Razzak *razzakmu@ngha.med.sa*
Health Informatics, CPHHI
King Saud bi Abdulaziz University for Health Sciences
Riyadh, 11426, Saudi Arabia

Abstract

Malaria parasite detection relies mainly on the manual examination of Giemsa-stained blood microscopic slides whereas it is very long, tedious, and prone to error. Automatic malarial parasite analysis and classification has opened a new area for the early malaria detection that showed potential to overcome the drawbacks of manual strategies. This paper presented a method for automatic detection of falciparum and vivax plasmodium. Although, malaria cell segmentation and morphological analysis is a challenging problem due to both the complex cell nature uncertainty in microscopic videos. To improve the performance of malaria parasite segmentation and classification, segmented the RBC and used RNN for classification into its type. Segmented RBCs are classified into normal RBC and infected cell. RNN identify the infected cells into further types.

Keywords: Malaria Detection, Segmentation, RBC Classification, Blood Cell Analysis.

1. INTRODUCTION

Half of world population is at risk of life threatening infectious disease Malaria. It is a potentially fatal parasitic disease of both human and animals and causes 219, 216 million infection cases of malaria and killed 6.6, 6.5 million people in 2010 and 2011 respectively. The major population that is at high risk mainly includes pregnant women and children who especially under five years [1]. Since 2000, malaria mortality rates are reduced to more than 25% through increased prevention and control measure [2].

Malaria disease is transmitted via a bite from the infected female mosquito which introduce the organisms form her saliva. Five species of plasmodium can infect and be transmitted by human whereas the majority of deaths are caused by falciparum and vivax plasmodium. Species diagnose is necessary for proper treatment in case of malaria. Some rules for species identification are shown in Table 1. However it does not guarantee that every parasite will exhibit the morphological characteristics. Typically, the malarial infection can be diagnosed by microscopic examination of blood using blood films or with antigen based rapid diagnostic. Even after fifty years of malaria eradication program, it still continues to increase. Whereas the control tools are getting less effective due to drug and insecticide resistance that are developed in mosquitos. This alarming situation has led the researchers to develop a rapid, accurate and affordable diagnostic method for early malaria parasite detection [1]. Early detection of malaria is vital in order to ensure prompt and effective treatment. Patient suffering from malaria disease should be diagnosed at early stage and should be given an effective and affordable treatment within 24 hours [0]. Microscopic blood image analysis is a standard technique used for blood cell analysis for diagnosing, however in remote area, a delay in obtaining results may lead to incorrect initial treatment due to unavailability of early diagnose system.

Malaria parasite diagnoses is a manual counting process that use microscopic examination of Giemsa-stained thick and thin blood smears. Manual counting methods are very long, tedious and prone to technician's ability to conduct the process correctly that requires training and skills,

i.e. a trained expert takes about 15 min to evaluate and count 100 cells and blood sample of millions of patients is performed every year [4]. Manual blood cell counting is not a reliable screening method when it is performed by non-expert due to lake of training expertise as it requires special training and considerable expertise [5-8]. Machine aided automatic analysis of microscopic blood cell is a powerful diagnostic tool that improves accuracy, saves time and reduces the required manpower as well as minimizes human errors. Automatic malarial diagnose has larger interest especially for clinics and laboratories; however, blood cell segmentation and morphological analysis is a challenging problem due to both the complex cell nature and uncertainty in microscopic videos. Comparison of microscopy diagnosis and rapid diagnosis tests is shown in table 2.

Recognition and inspection of blood can assist hematologists in analyzing the blood sample and diagnosing diseases like Aids, malaria and blood cancer. Blood samples study has been widely researched in the field of immunology, infectious diseases, transplantation, hematological malignancy and vaccine development. Various numbers of algorithms and techniques have been published that are related to blood cell segmentation and classification. This paper presents a survey of algorithms particularly focusing on blood cell morphology and cell counting for automatic screening of several diseases.

An advance in genomic technologies has opened a new realm for early detection of diseases that shows potential to overcome the drawbacks of manual detection technologies. Automatic instrument for blood cell counting such as flow cytometry and automatic counting machines can examine cell quantity but not qualitatively. Yet 21% of blood samples analysis still requires microscopic expert review [9-10]. Camera based methods provides quantitative as well as qualitative blood evaluation whereas not extensive research has focused in this area [11]. Various works are going on analyzing the microscopic images to point out the infection presence in human blood. Thus, several image processing methods have been presented in literature for the automatic segmentation and classification of blood cell. This paper explores present malarial parasite identification and classification using recurrent neural network. The organization of the rest of the paper is as follow: Section II briefly reviews the malarial parasite recognition approaches discussed earlier and Section III presents methodology for the segmentation and classification of malarial infected cells whereas the conclusion and future directions are presented in section IV.

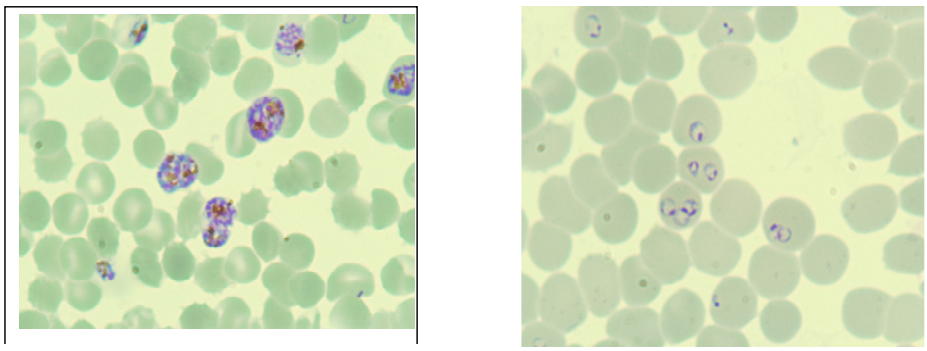

FIGURE 1: Malaria Effected Blood Images.

2. Related Work

Blood smear image consists of white blood cell, red blood cells, platelets as well as other possible contents such as malarial parasite, Aids virus etc. scattered across the background. The contents of blood cell images are very complex. The number or ratio of different blood cells in blood samples indicates different diseases, i.e. the white blood cells can be classified into more than 20 different classes according to their maturity [2]. Differential counting of blood cell can provide

valuable information for the accurate diagnosis of disease i.e. leukemia, cancer and other blood related diseases. Such differential counting involves classifying the blood cells with respect to color, size and shape. Blood cell counting is time consuming job for biologist, e.g. expert requires 15 min to count and evaluate a slide. Moreover, it is painstaking and subjective job which also requires training and skills. Due to these complexities and hematologist restrictions, manual counting accuracy decreases and thus machine based blood classification system is essential for biologist in diagnosing disease.

	P. Falciparum	P. Vivax	P. Ovale	P. Malariae
Size	Not enlarged	Enlarged	Enlarged	Enlarged
Shape	Round crescent gametocyte	Round or Oval	Round or Oval amoeboid	Round
Dots	Large red spots	Small red dots	Small red dots	Few tiny dots

TABLE 1: Some Rules for Species Identification.

An automated malarial diagnosis system can be designed by understanding the diagnostic expertise (hematologist knowledge) and representing it by specifically tailored image processing and pattern recognition algorithms. Image processing based malarial diagnostic methods has been widely studied in order to provide early and accurate detection of malaria parasite. An ideal diagnostic method would be accurate, non-invasive, and inexpensive. In order to perform malarial diagnosis, the system must be capable of differentiating between malaria infected cells, and healthy blood component. The key tasks for malarial parasite classification involve segmenting the malaria parasite infected cells from the complicated background.

There are many different methods for each phase that can be utilized to build machine based malarial parasite recognition system. We have discussed these approaches one by one in details. Most microscopes provide uniform or relatively uniform illumination images whereas several illumination and contrast enhancement technique have been applied in literature. One way to deal with uneven illumination is the predefined illumination correction but some time we don't have reference image [12]. Suradkar used the local histogram equalization for contrast enhancement of parasite and RBC [13] whereas Zou et al. and Sio et al. used adaptive histogram equalization [19] for image enhancement [14, 15]. Ruberto et al. used paraboloid modal for illumination correction [16]. Sriram et al. used diagonal modal for illumination modeling [17]. In diagonal modal, an image of unknown illumination is transformed to the known illuminant space by multiplying pixel values with a diagonal matrix. Das et al. performed gray world assumption for correcting illumination [18]. Nasir applied partial contrast stretching (PCS) to improve the image quality and contrast of malaria image. Diaz used adaptive local low-pass filter to correct luminance differences on luminance channel [20-21]. The filter was designed for a window size which contained the largest image feature, i.e. a typical erythrocyte size. Filter was selectively applied on higher luminance levels that represent the background pixels. Mehrjou used adaptive histogram shaping function for contrast enhancement [22]. The image is divided to several tiles and histogram shaping is applied to these tiles separately followed by bilinear interpolation to eliminate artificially induced boundaries.

In some cases, the illumination can be excessively uneven thus the use of histogram, adaptive Otsu etc may not work sometime. Sabino et al. performed non-supervised nucleus region detection before nucleus color segmentation using the G channel from RGB color coordinates [23]. For colored images segmentation into ROIs, supervised classification method that is based on RGB color space is used. Costa et al.minimizes the sensibility of cytoplasm to small color variations using cell-modeling and morphological filters [24], an alternative for noise reduction of transformation from color images to segmented ROIs.

Memeu et al. used both RGB and HIS for segmentation [425]. Green channel from RGB whereas hue and saturation channels are used for segmentation based on Otsu method and Zack algorithms for RBC and parasites respectively. The green channel gave good results for erythrocytes segmentation but it also added the parasite as part of the foreground. The hue

component resulted to a binary image whose foreground had noisy boundaries whereas the saturation component failed to produce erythrocytes as the objects.

Kaewkamnerd et al. segmented the background by using histogram on to HSV color format [26]. After background segmentation, the image is divided into small windows of 300 by 300 pixels for efficient processing. Finally, malaria parasites are identified based on their size. Sriram et al. segmented the background by using histogram on green component RGB color format [17].

Multilevel thresholding especially using Otsu method has been performed by many researchers for blood cell segmentation and it showed promising results. Savkare and Narote performed Global threshold and Otsu thresholdding on gray scale and green channel image [27]. Both images are added and median filter is applied to remove the unwanted points. Later on distance transform and watershed transform are applied to segment the cells. Kumar et al. used Otsu threshold on histogram of B component of RGB color space followed by the morphological operation [28]. Ahirwar et al. relied on two thresholds (one for erythrocytes, and one for parasites) for parasite segmentation [29]. The first threshold is selected to separate the erythrocytes from the background of the image. The second threshold is taking the first minimum after the principal mode of the histogram incorporating only the erythrocytes. Tek et al. modeled the stained and unstained pixel distributions with histograms and used the probability densities to determine whether a pixel on the input image is stained or not [12].

	Microscopy	**RDTs**
Requirement	Electricity Special Training Staining Chemical	None, Basic Training, None
Time	~60 Minutes	15-20 Minutes
Cost		
	Specification	
Detection Threshold	50 pal/ul	~100par/ul
Detection of all species	Yes	Some bands
Quantification	Yes	None
Specie Identification	Yes	None
Life Stage Identification	Yes	None

TABLE 2: Comparison of Microscopy Diagnosis and Rapid Diagnosis Tests [34].

3. Malarial Parasite Detection: Methodology

An automated malarial diagnosis system can be designed by understanding the diagnostic expertise (hematologist knowledge) and representing it by specifically tailored image processing and pattern recognition algorithms. Image processing based malarial diagnostic methods has been widely studied in order to provide early and accurate detection of malaria parasite. An ideal diagnostic method would be accurate, non-invasive, and inexpensive. In order to perform malarial diagnosis, the system must be capable of differentiating between malaria infected cells, and healthy blood component. The key tasks for malarial parasite classification involve segmenting the malaria parasite infected cells from the complicated background. There are three major steps of analyzing malarial parasite infected images.

- Preprocessing
- Segmentation
- Classification

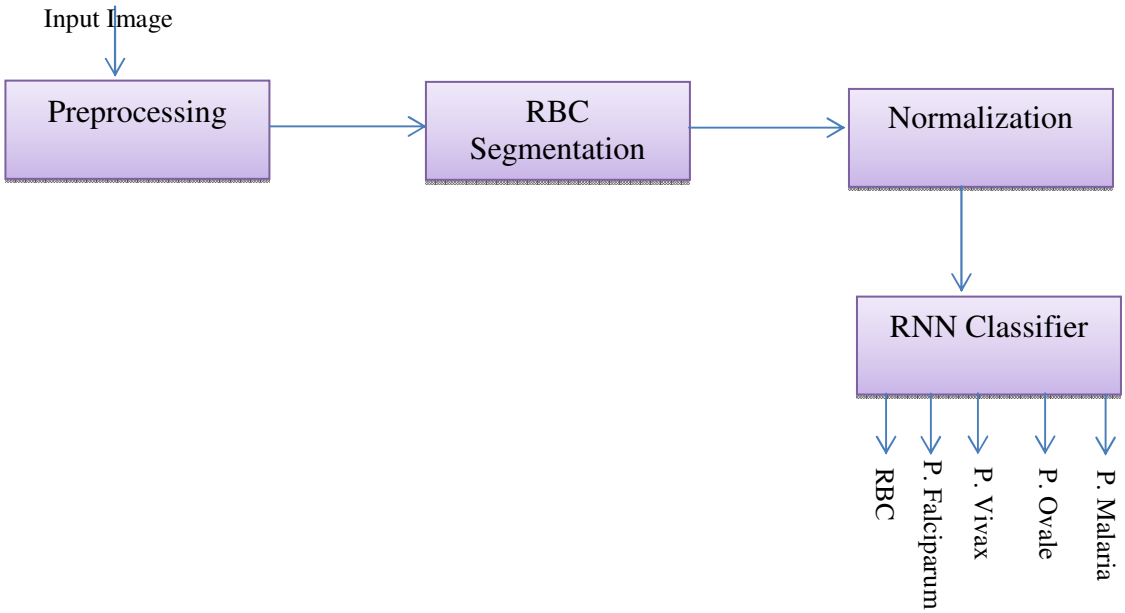

FIGURE 2: Malaria Parasite Classification [30].

2.1 Preprocessing

The aim of preprocessing is to obtain images with low noise, high contrast than original images for the further processing. Due to camera calibration and staining variability of blood smear, changes may occur in illumination and color distribution of blood images. This particular problem poses difficulties for classification of blood cells since it is hard to deal with proper segmentations of objects with quite similar colors. This process contains two operations image enhancement and noise reduction. We have used spatial filtering (median filter) for noise reduction. The median filter replaces pixel value with the median of its neighboring value.

To get the finest coefficients details of noise free image, a Forward Discrete Curvelet Transform (FDCT) is applied to the V channel as shown in Figure 3(b). It is a multi-dimensional transformation which can sense both the contours as well as curvy edges of the overlapping objects in the image. The FDCT has high directional sensitivity along with the capability to capture the singularities. Edge and singularity details are processed to extract the feature. After obtaining the highest detailed coefficients Inverse Discrete Curvelet Transform is applied to high frequency band to obtain the detailed image. This detailed image is now having the stronger edges than the original and would perform better in lending edge details to the segmentation step.

The next step is the adaptive equalization operation to spread out the intensity values along the total ranges of values in order to achieve better contrast. Adaptive histogram equalization differ from ordinary histogram equalization in respect that it computes several histogram of each corresponding to distinct section and use these histogram to redistribute the lightness value. After applying the adaptive histogram equalization, the background pixels have higher intensities than the cells.

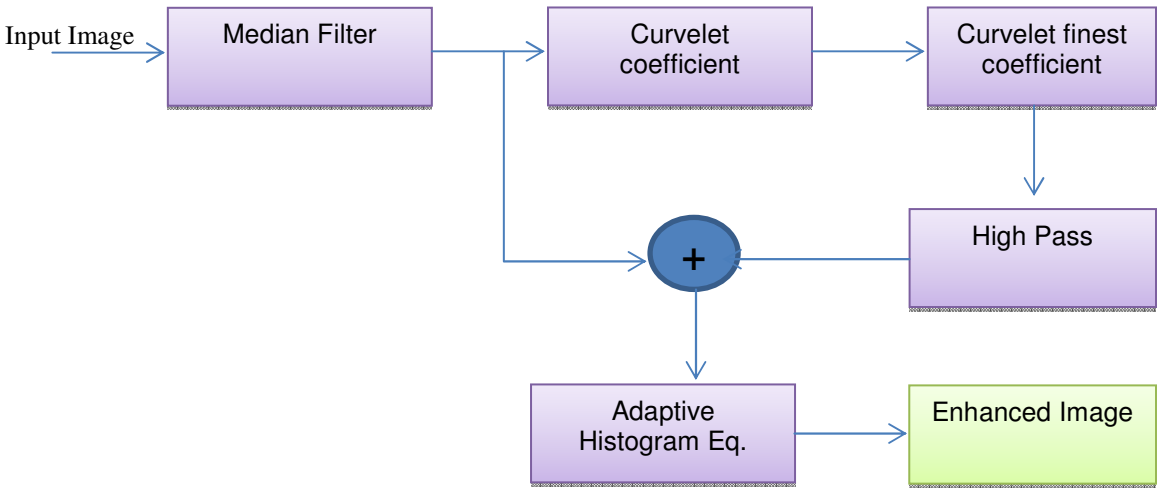

FIGURE 3: Preprocessing of Blood Cell Image.

2.2 Segmentation

In the analysis of automatic classification of malarial parasite procedures, the most important and difficult part is segmentation of malaria parasite infected blood cells from the background and other cells because the blood cells are often overlaid with each other and is the basis of quantitative analysis of its deformability and hence its filterability[12]. Cell shapes, light variation and noise are the other factors that make segmentation a difficult task. Accurate segmentation allows fruitful result in sub-sequent levels. Malarial parasite lies in erythrocytes thus we need to segment the erythrocyte form the blood images. We have used Rao's method for background segmentation. Rao's method extracts a rough foreground image using morphological rea top-hats. Two different threshold values are determined form these backgrounds and foreground that are used to produce the refined binary foreground mask.

At the end, a box counting algorithm is used to the segmented the blood cell images. The box counting algorithm counts the number of blood cells boxes having side length r needed to cover the surface of fractal objects and the number of boxes N, occupied by more than one pixel of the image. We have defined two procedures by two parameters for box counting. First one is the selection of r whereas the second is the range of r. The blood cell image has finite set of points and the upper limit is the image size while the lower is the pixel unit. Various researches propose using 2, 4, 8, 16, 2n pixels as box sizes to have a uniform spread of observation. The quadratic boxes cover the object, and the number of the boxes is recorded. The fractal dimension (FD) measures the dependence between the number of boxes N and the box side length r.

2.3 Classification of Malarial Parasite

We have used recurrent neural networks (RNN) for classification of the malarial parasite. RNNs are the computational modals that simulated structure and function of biological neural networks. Training is an important task in utilizing the neural network.

The BLSTM Network

We employed Bidirectional Long Short Term Memory (BLSTM) networks for the detection and classification of malarial parasite into four types. The BLSTM is a Recurrent Neural Network (RNN) approach to learn the sequential patterns. For maintaining the sequence, it is important to use previous computations at current point in time. In addition to that the future point should be predicted by keeping the current and previous calculations. It means RNN can also work in a bidirectional way. The bidirectional feature of RNN makes it vulnerable for learning patterns. The RNN is primarily meant to deal with sequences but it has a limit that not to deal with complex and

large sequences. As RNN iteration proceeds, the previous computation becomes faded with increase in time lag. But it can be the situation if we require the most previous computation at current point in time then it would be difficult to have that particular computation. The LSTM memory cell is depicted in Figure 6.

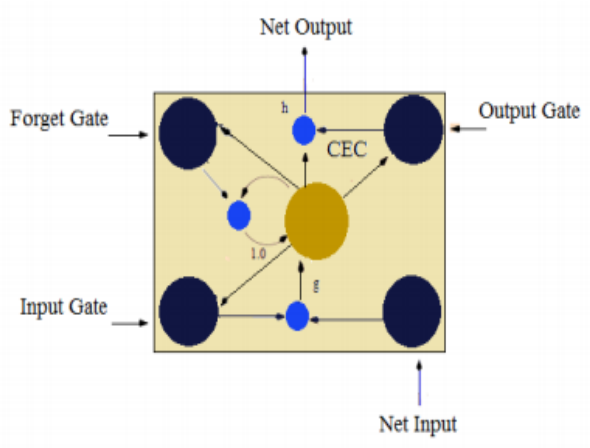

FIGURE 4: LSTM Memory Cell.

Training

We have segmented each Red blood cell using the bounding box in previous section. The RBC image is normalized to LxM size for RNN. We have five types of classes i.e. RBC, R.Falciparum, P.Vivax, P. Ovale and P. Malaria and the RNN is trained on these five types of cells. This normalized segmented RBC image with their corresponding ground truth (RBC, R.Falciparum, P.Vivax, P. Ovale and P. Malaria) is given to BLSTM classifier individually. The classifier then computes training and validation errors in series of epochs. In every epoch the classifier computes the forward pass in addition to backward pass. After every epoch the training and validation errors were computed and stored with the best results. When there is no significant change in training and validation error computation then training terminates. The training and validation were saved and used for testing phase. There are four parameters (like input image size, hidden layer neurons, learning rate and momentum) which we can tune them to produce variation in results. But in our experiments the momentum, learning rate and image size kept fixed. The only tuning parameter is number of hidden layer neurons. We reported the best results that have produced on 120 memory cells in hidden layer.

BPNN	RBC	P. Falciparum	P. Vivax	P. Ovale	P. Malariae
RNN	98.2	92.6	93.2	94.1	96.5

TABLE 3: Recognition Result.

Testing

We have five types of classes i.e. RBC, R.Falciparum, P.Vivax, P. Ovale and P. Malaria. The RNN is trained on these five types of cells. The normalized segmented red blood cell is forward to RNN for classification into one of five types (RBC in case of not infected and other four types in case of malarial infected cell).

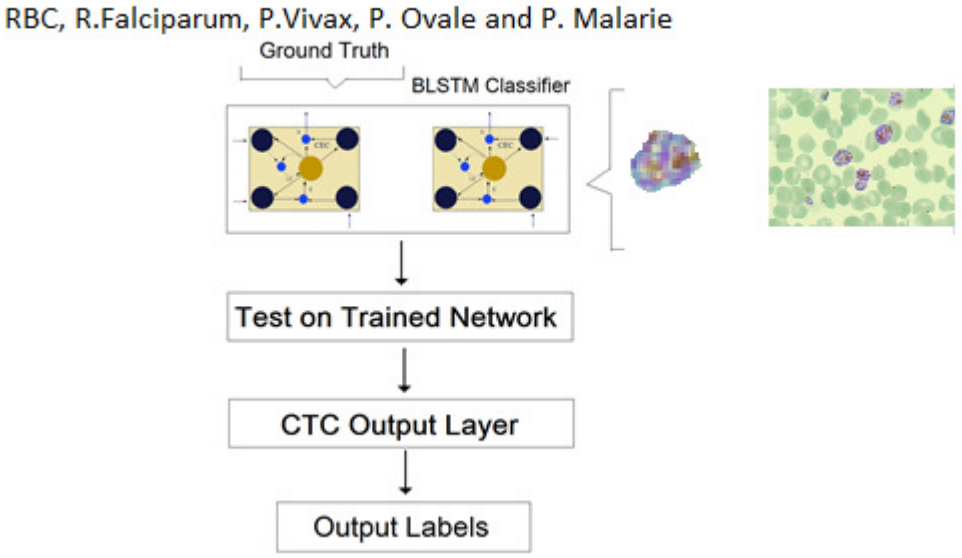

FIGURE 5: RNN for Malarial Parasite Classification.

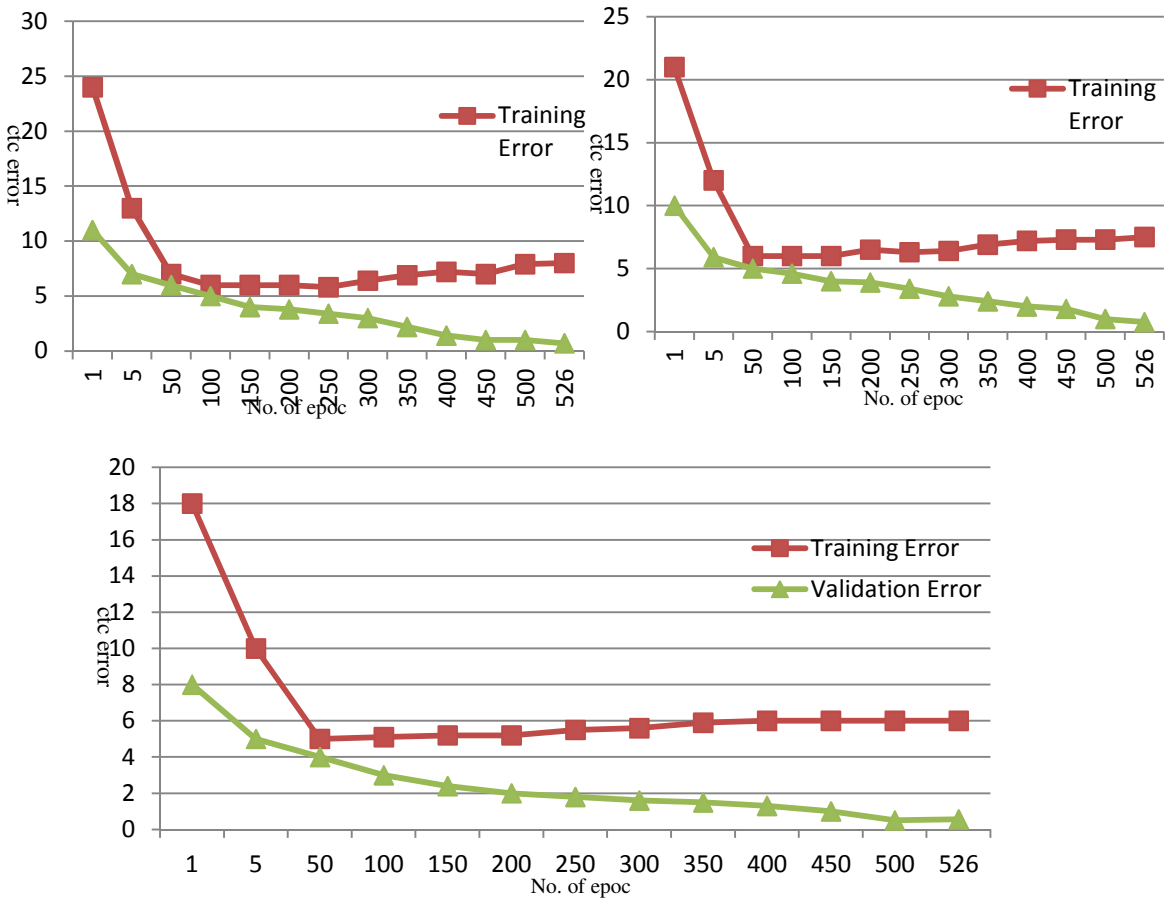

FIGURE 6: Learning sequences on various (a=60, b=80, c=120) LSTM memory cells.

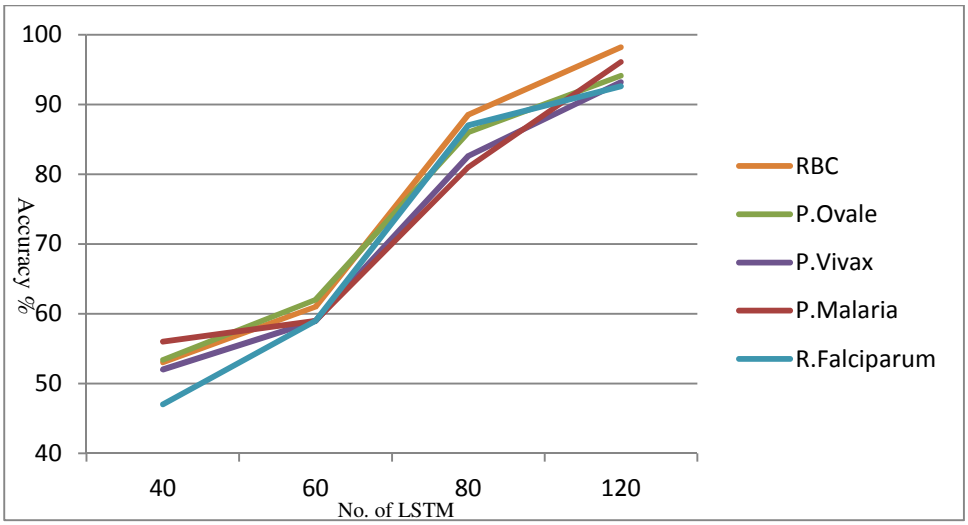

FIGURE 7: Accuracy and no of LSTM.

4. CONCLUSION

The paper presented a method for automatic detection of falciparum and vivax plasmodium. Although, malaria cell segmentation and morphological analysis is a challenging problem due to both the complex cell nature uncertainty in microscopic videos. To improve the performance of malaria parasite segmentation and classification, we have extracted the red blood cells and forward to Recurrent neural network for further classification into five types RBC, R.Falciparum, P.Vivax, P. Ovale and P. Malaria. RNN provided 98.2%, 92.6%, 93.2%, 94.1% and 96.1%, respectively for RBC, R.Falciparum, P.Vivax, P. Ovale and P. Malaria.

5. REFERENCES

[1] Yunda L., Ramirez, A.A., Millan, J., Automated Image Analysis Method for p-vivax Malaria Parasite Detection in Thick Film Blood Images, Revista S&T, 10(20), 9-25

[2] World Health Organization: World Malaria Report 2011. Geneva; 2011

[3] Chan LL, Laverty DJ, Smith T, Nejad P, Hei H, Gandhi R, Kuksin D, Qiu J., Accurate measurement of peripheral blood mononuclear cell concentration using image cytometry to eliminate RBC-induced counting error" , Journal of Immunological Method, Vol. 238, pp 25-32, 2013 Feb

[4] Chan, Y.-K., Tsai, M.-H., Huang, D. C., Zheng, Z.-H., Hung, K.-D., 2010. Leukocyte nucleus segmentation and nucleus lobe counting. BMC Bioinformatics 11, 558.

[5] M. M. Kettelhut, P. L. Chiodini, and H. Edwards, Moody A: External quality assessment schemes raise standards: evidence from the UKNEQAS parasitology subschemes. J. Clin. Pathol. 56, 927 (2003).

[6] R. E. Coleman, N. Maneechai, N. Rachaphaew, C. Kumpitak, R. Miller, V. Soyseng, K. Thimasarn, and J. Sattabongkot, Comparison of field and expert laboratory microscopy for active surveillance for asymptomatic plasmodium falciparum and plasmodium vivax in western thailand. Am. J. Trop Med. Hyg. 67, 141 (2002).

[7] Bates, V. Bekoe, and A. Asamoa-Adu, Improving the accuracy of malaria related laboratory tests in Ghana. Malar J. 3, 38 (2004).

[8] K. Mitiku, G. Mengistu, and B. Gelaw, The reliability of blood film examination for malaria at the peripheral health unit. Ethiopian J. of Health Dev. 17, 197 (2003).

[9] T. Chen, YongZhang, C. Wang, ZhenshenQu, FeiWang, and Tanveer Syeda-Mahmood, Complex local phase based subjective surfaces (CLAPSS) and its application to DIC red blood cell image segmentation. Neurocomputing 99, 98 (2013).

[10] M.I.Razzak, B. AlHaqbani, "Automatic Detection of Malarial Parasite Using Microscopic Blood Images" Journal of Medical Imaging and Health Informatics Vol. 5, 1–8, 2015

[11] S. H. Rezatofighti and H. Soltanian-Zadeh, Automatic Recognition of Five types of White blood cell in peripheral blood. Computerized Medical Imaging and Graphics, Volume 35, Issue 4, Pages 333–343, June 2011.

[12] F. B. Tek, A. G. Dempster, and I. Kale, Malaria parasite detection in peripheral blood images. Proceeding of British Machine Vision Conference (2006).

[13] T. P. Suradkar, Detection of malarial parasite in blood using image processing. International Journal of Engineering and Innovative Technology (IJEIT) 2, (2013).

[14] M. I. Khan, B. Acharya, B. K. Singh, and J. Soni, Content based image retrieval approaches for detection of malarial parasite in blood images. International Journal of Biometrics and Bioinformatics (IJBB) 5 (2011).

[15] W. S. Selena Sio, W. Sun, Saravana Kumar, W. Z. Bin, S. S. Tan, S. H. Ong, H. Kikuchi, Y. Oshima, and K. S. W. Tan, Malaria Count: An image analysis-based program for the accurate determination of parasitemia. Journal of Microbiological Methods 68, 11 (2007).

[16] C. D. Ruberto, A. Dempster, and S. B. Khan Jarra, Analysis of infected blood cell images using morphological operators. Image and Computer Vision 20 (2002).

[17] R. Sriram, M. Chandar, and K. Srinivas, Computer aided malarial diagnosis for JSB stained white light images using neural networks. International Journal of Advanced Research in Computer Science and Software Engineering 3 (2013).

[18] D. K. Das, M. Ghosh, M. Pal, A. K. Maiti, and C. Chakraborty, Machine learning approach for automated screening of malaria parasite using light microscopic images. Micron 45, 97 (2013).

[19] Z. Karel, Contrast limited adaptive histogram equalization. Graphics Gems IV 474–485 (1994), (code: 479–484).

[20] G. Díaz, F. A. González, and E. Romero, A semi-automatic method for quantification and classification of erythrocytes infected with malaria parasites in microscopic images. Journal of Biomedical Informatics 42, 296 (2009).

[21] Gloria Díaz, Fabio Gonzalez, and Eduardo Romero, Infected cell identification in thin blood images based on color pixel classification: Comparison and analysis. Lecture Notes in Computer Science 4756, 812 (2007).

[22] A. Mehrjou and T. Abbasian, Automatic Malaria Diagnosis System, First International Conference on RSI/ISM.

[23] D. M. U. Sabino, L. da Fontoura Costa, E. Gil Rizzatti, and M. Antonio Zago, A texture approach to leukocyte recognition. Real-Time Imaging 10, 205 (2004).

[24] F. Costa L da and R. Marcondes, Jr., Shape analysis and classification: theory and practice. Boca Raton, FL: CRC Press (2001).

[25] D. M. Memeu, K. A. Kaduki, and A. C. K. Mjomba, Njogu Samson Muriuki,Lucy Gitonga, Detection of plasmodium parasites from images of thin blood smears. Open Journal of Clinical Diagnostics 3, 183 (2013).

[26] S. Kaewkamnerd and C. Uthaipibull, Apichart Intarapanich, Montri Pannarut, Sastra Chaotheing, Sissades Tongsima, An automatic device for detection and classification of malaria parasite species in thick blood film. BMC Bioinformatics 13, S18 (Suppl 17) (2012).

[27] S. S. Savkare and S. P. Narote, Automatic detection of malaria parasites for estimating parasitemia. International Journal of Computer Science and Security (IJCSS), 5 (2011).

[28] A. Kumar, A. Choudhary, P. U. Tembhare, and 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 ISSN (Online):2278-5299, 1, 159 (2012).

[29] N. Ahirwar, S. Pattnaik1, and B. Acharya, Advanced image analysis based system for automatic detection and classification of malarial parasite in blood mages. International Journal of Information Technology and Knowledge Management 5, 59 (2012).

[30] Muhammad Imran Razzak, “Automatic Detection and Classification of Malarial Parasite”, International Journal of Biometrics and Bioinformatics (IJBB) Volume-9 Issue-1, 2015