International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2020 IJSRCSEIT | Volume 6 | Issue 3 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT206398

Malaria Detection Using Image Processing

Spandana E M, Kavya Hegde, Sneha Ummanabad
Department of Computer Science, Srinivas Institute of Technology, Mangalore, Karnataka, India

ABSTRACT

Malaria is one of a serious infectious disease in the world caused by a peripheral blood parasite of the genus Plasmodium. Traditional microscopy method, for malaria diagnosis is old method and has been occasionally proved inefficient since it is time consuming and results are difficult to reproduce. As it constitutes a serious global health problem, the evaluation process is of high importance. In this work, an accurate, easy, rapid and affordable model of malaria diagnosis using stained thin blood smear images was developed. The method makes use of the intensity features of Plasmodium parasites and also erythrocyte or red blood cell images. Images of infected and non-infected red blood cells were taken, pre-processed, and then relevant features were extracted from them and eventually diagnosis was made based on the features extracted from the images. A set of features based on intensity have been proposed, and the performance of these features on the red blood cell samples from the database have been evaluated using an artificial neural network (ANN) classifier. The results have shown that these features could be extremely successful if used for malaria parasite detection.

Keywords : Malaria, erythrocyte, blood smears, Parasite, Digital Image Processing, grayscale image.

I. INTRODUCTION

Malaria is life threatening disease caused by plasmodium parasites and the severity of Malaria varies based on the species of plasmodium. According to the World Health Organization(WHO) statistics, in 2000, it was estimated that there were 262 million cases of malaria globally, leading to 839.000 deaths across the world. By the year 2015, it was estimated that the number of malaria cases had decreased to 214 million, and the number of deaths decreased to 438.000.

This is due to the fact that, the environmental conditions are suitable for mosquitoes, in addition to the poor socio-economic conditions which make access to health care and disease prevention resources difficult. There are various techniques to diagnose malaria, of which manual microscopy is considered to be the standard.

a b c d e

The manual approach of diagnosis is time consuming and may also lead to inconsistency. And, it also demands trained, experienced technicians or pathologists. This approach if digitized once, it will reduce the time taken for screening the disease. This study helps to investigate the use and application of digital image processing for detecting malaria parasites using microscopic color images. Therefore, an efficient method is proposed for parasite detection based on the intensity and texture features. Parasite detection is the elemental function of this semi-automated diagnosis.

The rest of this paper is organized and as follows. Section II presents related work, Section III describes the model evaluation, Section IV presents results and and finally Section V concludes the paper.

II. RELATED WORK

A. Automatic Malaria Blood Smear Classification

Machine learning technique has been used to detect parasitemia in images of Giemsa-stained blood smears. This early study compared the correlation between automatic and manual parasite detection but did not classify infected and uninfected red blood cells. In 2009, a study based on support vector machine (SVM) to classify pre-processed blood cell images to detect infected erythrocytes was conducted. The algorithm worked well on a small dataset of 650 malaria images in both specificity and sensitivity. Unfortunately, the model performance decreases when it is applied to blood images in the infection stage.

B. Convolutional Neural Network

A convolutional neural network (CNN) is a specific deep learning architecture which is suitable for image recognition. A CNN model processes input data by its multiple layers and is characterized by four key features, they are: local connections, shared weights, pooling, and the use of many layers. Its use is then extended to handwriting recognition and later to natural image recognition. The performance of CNN models for image classification obtained another boost by the introduction of ImageNet by Alex Krizhevsky, thus also known as AlexNet, in 2012. AlexNet is a breakthrough application for CNNs useful in multi-categorical classification.

C. Challenges for CNN Applications

The idea of CNNs is to apply smaller convolutional filters in combination with deep network architecture to capture the distinguishable features in the images as much as possible. However, a more complex CNN architecture will inevitably increase the demand for powerful computing resources. The training efficiency can be effectively improved by technologies such as GPU and cluster computing, where in they are unable to ensure the performance of classification. Other factors such as data pre-processing and also the size of training dataset strongly affect the classification accuracy. Since the CNN classification accuracy depends on the amount of training data, small data sets such as those used in earlier approaches are not so large enough for training a deep model with its many parameters. As a compromise, the method of transfer learning has been introduced where a pre-trained network model is used to extract features that a conventional classifier, such as SVM, can use for a fine-tuned classification. Transfer learning can be used as an alternative to deep learning. Time for training the model is saved at the cost of performance. In this study, we will implement deep learning by both training a newly configured CNN model and applying transfer learning in order to evaluate the use for malaria blood smear classification.

III. MODEL EVALAUTION

A. Data Source

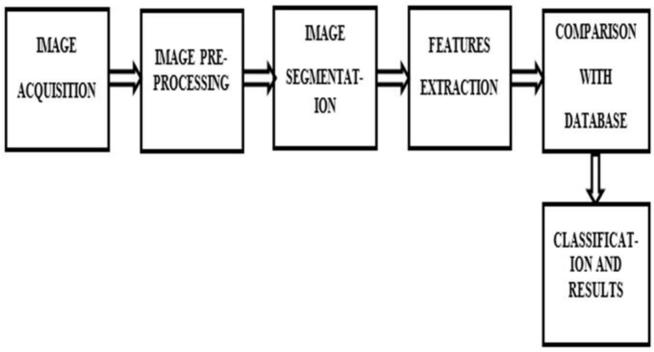
We used archived blood smear images acquired from dataset source called Kaggle and segmented the visual region of the erythrocyte cell from the original images. Our data set contains 26,187 erythrocyte images where the ratio of infected cells to uninfected cells is 1:1. All images are normalized to the median width and height for the training and classification experiments, at 50*50 pixels.

B. Data Preprocessing

All the erythrocyte images are read into MATLAB, resized, and serialized to generate the input to the MatConvnet toolbox. Before the data is passed to the CNN network for training, normalization is done in order to improve the brightness and contrast, and eigenvalue decomposition operation is applied on covariance matrix in order to whiten the entire dataset.

C. CNN Model Training

For training and evaluating the CNN model, a ten-fold cross-validation is implemented on the entire dataset, where 90% of the images are used for training, and 10% of the images are used for testing. In model training, 90% of the images are separated from the Figures training set for the actual training and the remaining 10% are used for back-propagation validation.



The performance evaluation criteria are the average of accuracy, sensitivity, specificity, precision, F1 score, and Matthews correlation coefficient. A pre-trained AlexNet based on the CIFAR-100 data set is applied as feature extractor for transfer learning. It is linked to a conventional SVM classifier to implement transfer learning as a comparison to the CNN model.

IV. RESULT

Table I presents the results of the two ten-fold cross-validations:

Table 1: Outcomes Of Cross-Validation

Measure	CNN model	Transfer Learning
Accuracy	97.37%	91.99%
Sensitivity	96.99%	89.00%
Specificity	97.75%	94.98%
Precision	97.73%	95.12%
F1 Score	97.36%	90.24%
Matthews correlation coefficient	94.75%	85.25%

The results show that the our new CNN model has a superior and best performance compared to the transfer learning model. The average classification accuracy of the CNN model is 97.37%, and the model sensitivity, specificity, and precision all reach upto 97% level. F1 score and Matthews correlation coefficient of the trained CNN model are both more than 6% which is larger than the transfer learning model. This shows that the trained CNN model is a much better representation of the training images than the transfer learning model, which depends on feature extraction from a pre-trained model that is trained on an entirely different image set.

V. CONCLUSIONS AND DISCUSSIONS

Based on the results of the above experiment, we conclude that our newly designed convolutional neural network model is a suitable solution for blood cell classification. Compared to transfer learning and other similar studies, the CNN model shows much better classification performance after training with more than 26,000 images. Its performance is affected by both architecture and the volume of training data. In future we expect that the deep learning method will significantly improve the working efficiency and accuracy of malaria diagnosis and other health-related applications, following previous studies on deep learning.

VI. REFERENCES

[1] C.J.L. Murray, L.C. Murray, S.S. Lim, K.G. Andrews, K.J. Foreman, D. Haring, N. Fullman, M. Naghavi, R. Lozano and A.D. Lopez, "Global malaria mortality between 1980 and 2010: a systematic analysis." The Lancet, 2012, 379(9814), pp. 413-431.

[2] S. Mali, S.P. Arguin, and P.M.Arguin, "Malaria surveillance—United States, 2010." MMWR Surveillance Summary, 2012, 61(2), pp. 413-431.

[3] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning." Nature. 2015 May 28;521(7553), pp. 436-44.

[4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions." IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1-9.

[5] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition." arXiv preprint, arXiv: 1512.03385, 2015 Dec.

[6] D.H. Hubel and T.N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex." The Journal of physiology. 1959 Oct 1;148(3), pp. 574-91.

[7] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," Neural networks, 2000 Jun 30, 13(4), pp.411-430.

Cite this article as :

Spandana E M, Kavya Hegde, Sneha Ummanabad, "Malaria Detection Using Image Processing", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 6 Issue 3, pp. 285-288, May-June 2020. Available at doi : https://doi.org/10.32628/CSEIT206398
Journal URL : http://ijsrcseit.com/CSEIT206398