IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Survey On Image processing Based Automatic Detection Of Malaria

Miss. Kalyani Sakhare
Digital Electronics Department
Sipna College Of Engineering And Technology
Amravati,India

Prof.Dr. A. S. Joshi
Digital Electronics Department
Sipna College Of Engineering And Technology
Amravati,India

***Abstract*— Malaria is the deadliest disease in the earth and big hectic work for the health department. The traditional way of diagnosing malaria is by schematic examining blood smears of human beings for parasite-infected red blood cells under the microscope by lab or qualified technicians. This process is inefficient and the diagnosis depends on the experience and well knowledgeable person needed for the examination. Deep Learning algorithms have been applied to malaria blood smears for diagnosis before. Design propose system a new and highly robust deep learning model based on a convolutional neural network (CNN) which automatically classifies and predicts infected cells in thin blood smears on standard microscope slides. Testing on a small dataset of images gathered from a different source achieves similar performance, suggesting the model may generalize to different imaging conditions. The system achieves higher recall than existing non-deep approaches, and its accuracy, recall and precision of the highest performing CNN approach.**

Keywords—CNN,Deep Learn,Maleria etc

I Introduction

Malaria is a life-threatening disease. It's typically transmitted through the bite of an infected Anopheles mosquito. Infected mosquitoes carry the Plasmodium parasite. When this mosquito bites you, the parasite is released into your bloodstream. Once the parasites are inside your body, they travel to the liver, where they mature. After several days, the mature parasites enter the bloodstream and begin to infect red blood cells.Within 48 to 72 hours, the parasites inside the red blood cells multiply, causing the infected cells to burst open.The parasites continue to infect red blood cells, resulting in symptoms that occur in cycles that last two to three days at a time.Malaria is typically found in tropical and subtropical climates where the parasites can live. The World Health Organization (WHO) Trusted Source states that, in 2016, there were an estimated 216 million cases of malaria in 91 countries.What causes malaria? Malaria can occur if a mosquito infected with the Plasmodium parasite bites you. There are four kinds of malaria parasites that can infect humans: Plasmodium vivax, P. ovale, P. malariae, and P. falciparum.P. falciparum causes a more severe form of the disease and those who contract this form of malaria have a higher risk of death. An infected mother can also pass the disease to her baby at birth. This is known as congenital malaria.Malaria is transmitted by blood, so it can also be transmitted through an organ transplant, a transfusion, and use of shared needles or syringes.The blood smear masses are typically dense and complex with the fusion of multiple cells. When a blood smear gets infected with malaria, its cell texture changes with time and it becomes tedious with a difficult job to differentiate between a healthy and unhealthy sample. The analysis of blood smear images from multi-views can act as a key support to detect the infection with minimum time, effort and cost. Through this project our effort is to detect wheter the given smear sample image contains maleria pathogens or its clear providing a realtime unique solution on increasing demand for inspections and shortage of pathologists with a wide diversity of weather and illuminance has resulted in making this a big health threat socially and economically.

II LITERATURE SURVEY

Malaria being one of the most fatal disease has been at the focal point of some major studies in the recent past. Some of them are described briefly here. The paper [3] uses a CNN model to detect parasite infected red blood cells in thin smears on standard microscopic slides prepared using routine methods. It is inspired by experiments on the underlying physiological mechanisms in the visual cortex of felines for recognizing objects. Another study [4] presents that there are many systems which describe the computerized methods of image analysis that commonly involves three main phases. In the first phase of preprocessing, luminance of the image is corrected and transformed to a constant color space. At the second step, a histogram-based image segmentation process is used which helps in avoiding maximum artifacts and over stained objects. Later, a back propagation neural network was used for classifying objects. A more accurate method of counting blood cells using Python OpenCV is explored in [5]. Recently, deep learning models have greatly exceeded performance of human beings in identification of complex

images [8,9]. The latest trend going on in artificial intelligence is deep learning that has increased the performance in several medical areas. It is multilayer neural network classifier which is trained by back propagation and many added layers are also used for the classification and detection process of medical imaging [10-14]. Deep learning requires huge number of training set to train the program or machine. A convolutional neural network (CNN) is a type of deep learning algorithm that takes an input image, applies weights and biases to many layers of network and classifies the images [15].

Deep Learning Technique

malaria detection is definitely an intensive manual process which can perhaps be automated using deep learning With regular manual diagnosis of blood smears, it is an intensive manual process requiring proper expertise in classifying and counting the parasitized and uninfected cells. Typically this may not scale well and might cause problems if we do not have the right expertise in specific regions around the world. Some advancements have been made in leveraging state-of-the-art (SOTA) image processing and analysis techniques to extract hand-engineered features and build machine learning based classification models. However these models are not scalable with more data being available for training and given the fact that hand-engineered features take a lot of time.

Deep Learning models, or to be more specific, Convolutional Neural Networks (CNNs) have proven to be really effective in a wide variety of computer vision tasks. Briefly, The key layers in a CNN model include convolution and pooling layers as depicted in the following figure 1

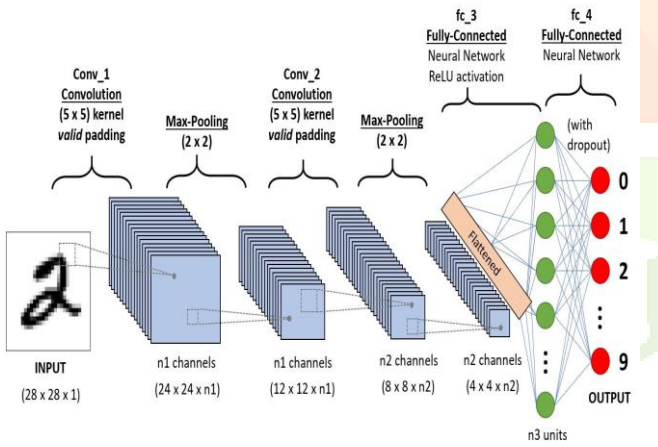

Fig 1. CNN Architecture

Convolution layers learn spatial hierarchical patterns from the data, which are also translation invariant. Thus they are able to learn different aspects of images. For example, the first convolution layer will learn small and local patterns such as edges and corners, a second convolution layer will learn larger patterns based on the features from the first layers, and so on. This allows CNNs to automate feature engineering and learn effective features which generalize well on new data points. Pooling layers help with down sampling and dimension reduction.

Thus, CNNs help us with automated and scalable feature engineering. Also, plugging in dense layers at the end of our model enables us to perform tasks like image classification. Automated malaria detection using deep learning models like CNNs could be very effective, cheap and scalable especially with the advent of transfer learning and pre-trained models which work quite well even with constraints like less data.

The paper by Rajaraman et al. , 'Pre-trained convolutional neural networks as feature extractors toward improved parasite detection in thin blood smear images' leverages a total of six pre-trained models on the data mentioned in their paper to obtain an impressive accuracy of 95.9% in detecting malaria vs. non-infected samples. Our focus would be to try out some simple CNN models from scratch and a couple of pre-trained models using transfer learning to see the kind of results we get on the same dataset! We will be using open-source tools and frameworks which include Python and Tensor Flow to build our models.

Deep Learning Model Training Phase

In the model training phase, we will build several deep learning models and train them on our training data and compare their performance on the validation data. We will then save these models and use them later on again in the model evaluation phase.

Model 1 : CNN from Scratch

Our first malaria detection model will be building and training a basic convolutional neural network (CNN) from scratch. First let's define our model architecture.

Model 2 Deep Transfer Learning

Just like humans have an inherent capability of being able to transfer knowledge across tasks, transfer learning enables us to utilize knowledge from previously learned tasks and apply them to newer, related ones even in the context of machine learning or deep learning.

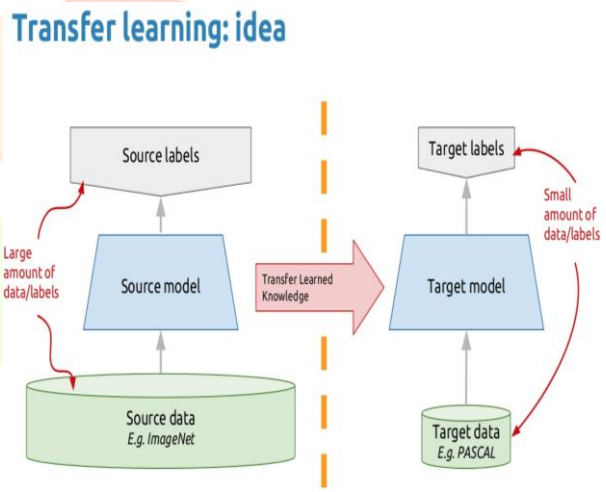

Fig 2 Transfer Learning Model

We will apply the two most popular strategies for deep transfer learning.

- Pre-trained Model as a Feature Extractor
- Pre-trained Model with Fine-tuning

We will be using the pre-trained VGG-19 deep learning model, developed by the Visual Geometry Group (VGG) at the University of Oxford, for our experiments. A pre-trained model like the VGG-19 is an already pre-trained model on a huge dataset (ImageNet) with a lot of diverse image categories. Considering this fact, the model should have learned a robust hierarchy of features, which are spatial, rotation, and translation invariant with regard to features learned by CNN models. Hence, the model, having learned a good representation of features for over a million images, can act as a good feature extractor for new images suitable for computer vision problems just like malaria detection! Let's briefly discuss the VGG-19 model architecture before unleashing the power of transfer learning on our problem

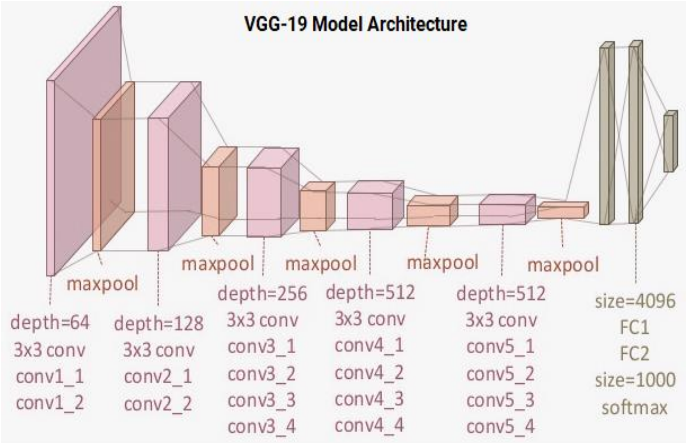

Fig 3 VGG-19 Model

Understanding the VGG-19 model

The VGG-19 model is a 19-layer (convolution and fully connected) deep learning network built on the ImageNet database, which is built for the purpose of image recognition and classification. This model was built by Karen Simonyan and Andrew Zisserman and is mentioned in their paper titled 'Very Deep Convolutional Networks for Large-Scale Image Recognition'. Fig. VGG-19 Model its clearly see that we have a total of 16 convolution layers using 3 x 3convolution filters along with max pooling layers for downsampling and a total of two fully connected hidden layers of 4096 units in each layer followed by a dense layer of 1000 units, where each unit represents one of the image categories in the ImageNet database. We do not need the last three layers since we will be using our own fully connected dense layers to predict malaria. We are more concerned with the first five blocks, so that we can leverage the VGG model as an effective feature extractor.For one of the models, we will use it as a simple feature extractor by freezing all the five convolution blocks to make sure their weights don't get updated after each epoch. For the last model, we will apply fine-tuning to the VGG model, where we will unfreeze the last two blocks (Block 4 and Block 5) so that their weights get updated in each iteration (per batch of data) as we train our own model.

Model 2: Pre-trained Model as a Feature Extractor

For building this model, we will leverage TensorFlow to load up the VGG-19 model, and freeze the convolution blocks so that we can use it as an image feature extractor. We will plugin our own dense layers at the end for performing the classification task. Thus it is quite evident from the preceding output that we have a lot of layers in our model and we will be using the frozen layers of the VGG-19 model as feature extractors only. We will n train our model using similar configurations and callbacks which we used in our previous model.

Model 3: Fine-tuned Pre-trained Model with Image Augmentation

In our final model, we will fine-tune the weights of the layers present in the last two blocks of our pre-trained VGG-19 model. Besides that, we will also introduce the concept of image augmentation. The idea behind image augmentation is exactly as the name sounds. We load in existing images from our training dataset and apply some image transformation operations to them, such as rotation, shearing, translation, zooming, and so on, to produce new, altered versions of existing images. Due to these random transformations, we don't get the same images each time. We will leverage an excellent utility called ImageDataGeneratorkeras that can help us build image augmentors.completing our model training phase and we will test the performance of our models on the actual test dataset.

Deep Learning Model Performance Evaluation Phase

We will evaluate the three different models that we just built in the training phase by making predictions with them on the data from our test dataset, because just validation is not enough! We will try built a nifty utility module called model_evaluation_utils, which we will be using to evaluate the performance of our deep learning models with relevant classification metrics. The first step here will be to obviously scale our test data. The next step involves loading up our saved deep learning models and making predictions on the test data. The final step is to leverage our model_evaluation_utils module and check the performance of each model with relevant classification metrics.Kaggle dataset is used in this analysis. It consists of cells segmented from thin blood mark slide images of malaria infected patients. The dataset contains 32,353 segmented cell images, out of which 80% are the training images and rest are used for testing the algorithm. The samples which are infected contain plasmodium and samples which are not infected contain no plasmodium but can also contain some types of staining impurities. Various pre-processing methods are used to achieve faster convergence which will be addressed in detail later. Figure 4 and Figure 5 show samples from dataset containing segmented red blood cells that are parasitized and uninfected. Figure 4 show that the images have different color distributions resulting from different stains during data processing and Figure 5 shows various types of parasite in red blood cells.

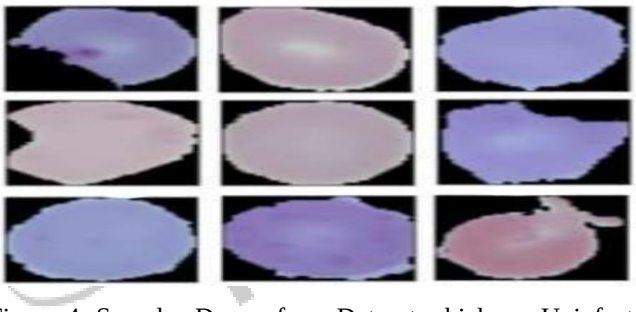

Figure 4: Samples Drawn from Dataset which are Uninfected Red Blood Cells.

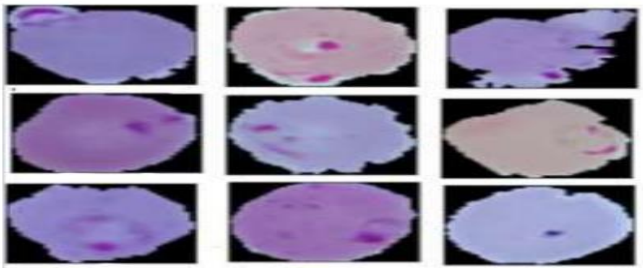

Figure 5: Samples Taken from Dataset which are Red Blood Cells Infected with Malaria.

CONCLUSION

In this survey paper studied Malaria afflicts hundreds of thousands of people around the world and is among the leading causes of death in developing countries [5]. The disease is primarily diagnosed through the technique of microscopic analysis, which necessitates trained laboratories and operational facilities that are often unavailable. Seeking to alleviate this predicament, we developed a compact and efficient convolutional neural network architecture, Our model attained high performance metrics . We hope that our model

will be used to further the development of efficient algorithms for disease detection.

REFERENCES

[1] Mohammed HA, Abdelrahman IAM. Detection and classification of malaria in thin blood slide images. In: International conference on communication, control, computing and electronics engineering (ICCCCEE). IEEE; 2017:1–5.

[2]Tek FB, Dempster AG, Kale I. A colour normalization method for giemsa-stained blood cell images. In: 14th signal processing and communications applications. IEEE; 2006:1–4.

[3] Di Rubeto C, Dempster A, Khan S, Jarra B. Segmentation of blood images using morphological operators. In: 15th international conference on pattern recognition, Vol. 3. IEEE; 2000:397–400.

[4]. Halim S, Bretschneider TR, Li Y, Preiser PR, Kuss C. Estimating malaria parasitaemia from blood smear images. In: 9th international conference on control, automation, robotics and vision. IEEE; 2006:1–6.

[5] Anggraini D, Nugroho AS, Pratama C, Rozi IE, Pragesjvara V, Gunawan M. Automated status identification of microscopic images obtained from malaria thin blood smears using Bayes decision: a study case in Plasmodium falciparum. In: Internationalconference on advanced computer science and information system. IEEE; 2011:347–52.

[6]. Kareem S, Morling RC, Kale I. A novel method to count the red blood cells in thin blood films. In: International symposium on circuits and systems (ISCAS). IEEE; 2011:1021–4.

[7] Kareem S, Kale I, Morling RC. Automated malaria parasite detection in thin blood films: a hybrid illumination and color constancy insensitive, morphological approach. In: Asia Pacific conference on circuits and systems (APCCAS). IEEE; 2012:240–

[8]. Nasir AA, Mashor M, Mohamed Z. Segmentation based approachfor detection of malaria parasites using moving k-means clustering. In: EMBS conference on biomedical engineering and sciences (IECBES). IEEE; 2012:653–8.

[9] Malihi L, Ansari-Asl K, Behbahani A. Malaria parasite detection in giemsa-stained blood cell images. In: 8th Iranian conference on machine vision and image processing (MVIP). IEEE 2013:360–5.

[10]. Berge H, Taylor D, Krishnan S, Douglas TS. Improved red blood cell counting in thin blood smears. In: International symposium on biomedical imaging: from nano to macro. IEEE; 2011:204–7.

[11]. Di Ruberto C, Dempster A, Khan S, Jarra B. Automatic thresholding of infected blood images using granulometry and regional extrema. In: International conference on pattern recognition, Vol. 3. IEEE; 2000:441–4.

[12]. Kareem S, Kale I, Morling RC. Automated P. falciparum detection system for post-treatment malaria diagnosis using modified annular ring ratio method. In: International conference on computer modeling and simulation. IEEE; 2012:432–6.

[13]. Litjens, Geert, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I. Sánchez. "A survey on deep learning in medical image analysis." Medical image analysis 42 (2017): 60-88.

[14]. Liang, Zhaohui, Andrew Powell, Ilker Ersoy, Mahdieh Poostchi, Kamolrat Silamut, Kannappan Palaniappan, Peng Guo et al. "CNN-based image analysis for malaria diagnosis." In 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 493-496. IEEE, 2016.

[15] Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. "Imagenet: A large-scale hierarchical image database." In 2009 IEEE conference on computer vision and pattern recognition, pp. 248-255. Ieee, 2009.