European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 4, 2020

Automated Diagnosis of Malarial Parasite in Red Blood Cells

[1]Mr K.P.K Devan, [1*]Dr G. S. Anandha Mala, [2*]Deepthi Salunkey. K, [2*]Grace Cynthia. R, [2*]Madhumitha. J
[1]Associate Professor, [1] Professor, Department of Computer Science and Engineering,*
Easwari Engineering College, Chennai - 89
[2]Students, Department of Computer Science and Engineering,
Easwari Engineering College, Chennai - 89.
deepthi.salunkey@gmail.com madhumithajayakumarj@gmailcom gracecynthiak@gmail.com

Abstract: The traditional system for detecting the infection has been the manual process of diagnosing the stained slides under a microscope. This manual process might consume more time for producing the results and the availability of medical experts is not always assured. Considering this as the primary concern we proposed a strategy which limits the human error while recognizing the presence of malarial parasite in the blood sample by using Image Processing. Hence by automating the diagnosis process, results can be acquired relatively quicker and more accuracy can be expected. The technologies and techniques to patently extract the required features and efficiently classify the infected samples are surveyed. This paper presents a survey of various approaches to automate the detection and classification of infected and uninfected cells.

Keywords: Malarial parasite, Image Processing, Parasite Detection Index, Classification, Segmentation

1. Introduction:

The female Anopheles mosquito is responsible for spreading the single-celled Plasmodium parasite. Once ingested into the host's blood, it multiplies among the red blood cells such as in the mosquito's intestine. It multiplies in the mosquito's gut and is passed on to any host that is fed on by the mosquito. Humans can also get Malaria by blood transfusion and an offspring from its mother during childbirth from the placenta. To detect Malaria the affected red blood cells are diagnosed first. Blood smear is analyzed under a microscope which is traditionally done by a medical expert. To automate the process, various Image Processing and Machine Learning methods were used. Various segmentation methods have been used to distinguish the presence of the parasite. The typical ring structure of the parasite that has held host in the red blood cells can be identified from the microscopic image. So, this image is segmented and classified to detect both infected and non-infected cells. The objective of this new system is to possibly increase the sensitivity, accuracy and the F-score of the previously existing systems to provide a more right diagnosis of the infectious disease.

2. Related Works:

[1.1] In this method the blood smear images are captured and a contrast enhancement technique known as partial contrast stretching technique is applied on the native malaria image. Then, the RGB color space is transformed into HSI color space after which an unsupervised segmentation technique that is the moving K-means clustering is applied to it. A 7x7 pixel median filter is applied to remove noise and smoothen out the image. The area in terms of pixels is obtained by applying the Seeded Region Growing Area Extraction algorithm. The holes in the segmented infected cell are filled based on morphological reconstruction algorithm. 100 malaria images have been analyzed so as to approve the segmentation procedure. During MKM Clustering some regions of the infected cell were not recognized due to similar intensity values. A segmentation accuracy of 99.49% was achieved from the results.

[1.2] In this paper, there are four convolutional layers that has been used to make up the OSICNN model. Normalization across channels are performed using the Cross Channel Normalization layer with the activation function ReLu. Max-pooling layer has been used to decrease the spatial size of the

convolutional layer's outputs. A fully connected layer and a SoftMax layer are also used to classify the extracted features and a Dropout layer with a dropout ratio of 0.5 preceding the fully connected layer to prevent over fitting problems. And as for the step increase method the values of alpha, beta, gamma and sigma are slowly increased with every iteration until they cause a decrease in the system's performance and those values are set to be optimal. Subsequent to accomplishing the ideal configuration, an augmented dataset is utilized for direct training for better preparing of OSICNN. Rotations by 90, 189, 270 degrees and reflections along X-axis and Y-axis were used as the augmentation operations. This model was run over 27,558 images of equally infected and uninfected cells acquired after segmenting 200 images sampled to 100x100 to suit the CNN.

[1.3] ResNet34 training networks were used for large recognition by reformulated formulae. This architecture calculates from layer 1 to layer 152 having minimum complexity. This pre- trained model consumes less time and obtains more features. CNN for the image reorganization, uses Convolution Nets. Procedure for ResNet34 Architecture needs residual blocks trained to achieve deeper networks. Initially calculate activations from layer 1 to the end of layers. This reduces the error in network and assumes weight and bias for computing layers of predictions. The mathematical form of ResNet can be evaluated. ResNet calculations on network layers to make connections among layers and enable cross layer connectivity are made. For the model selection, the Convolutional layer involves a group of convolutional kernels that associates with small image regions. This implementation additionally performs categorization with respect to various filters, padding and direction. Pooling layer downsampled the features pointing to a specific local region. Activation functions like maxout, ReLu, tanh, sigmoid helps us to work on complex patterns and provide decision function. Image is classified using a CNN at three different stages. Initially at first stage data is collected and performed feature generation, at second stage data will be pre-processed and feature selection will be performed, at third stage a supervised model has been selected and applied to tune on parameters and finally analyse the prediction data.

[1.4] It is highly important to automate the process of evaluation. For recognizing WBC and potential parasites present on microscopic slides, auxiliary and new threshold selection techniques are utilized. Image highlights based on shading, surface and geometry of the cells and parasites are generated. The keywords used here are Malaria, Parasite, Neural Network, Erythrocytes. The images used for processing were obtained from the Public Health Image Library. Oil immersion views of giemsa recolored blood films were caught utilizing a binocular microscope mounted with a digital camera. This is done to avoid multiple problems such as being specific to a certain species and having high costs per test while retaining the pros of a traditional microscope. Among the tried calculations, Susan edge detection technique gave great limitation of edges yet shaped a thick outskirt making cell separation troublesome. The highest performance was produced by the BFF neural network which was trained with the back propagation algorithm.

[1.5] The contrastive divergence method has been used to stack Restricted Boltzmann machines to pre-train a trained model of DBN. The visible variables of the DBN have been initialized by extracting the features from the images. The feature they have used is a concatenation of color (histogram and color coherence vector) and texture (LBP features, Haralick features and gray level run length matrix feature). There are 4 hidden layers that are each, independently trained as an RBM. After the pre-training process, the conditions of the hidden nodes from the prepared RBM are taken care of as the contributions to the following layer of the RBM. Similarly, a series of RBMs are trained and they are stacked to construct a DBN. The final layer of variables is represented with the newly formed DBM that represent the desired output values by performing back propagation. There are 484 visible layers and the output layer has two nodes with four hidden layers containing 600 hidden nodes in each layers. The DBN was applied on 4100 peripheral blood smear images and resulted in an F-score of 89.66%, a sensitivity of 97.60%, and specificity of 95.92%.

[1.6] The primary deep learning method that can distinguish malaria parasites in thick blood smear images and can run on cell phones has been developed. The two processing steps applied in this framework are an intensity-based Iterative Global Minimum Screening (IGMS), which plays out a quick screening of a thick smear image to detect the parasite candidates and a customized Convolutional Neural Network (CNN) that arranges each applicant has either parasite or background. At first, the white blood cells are detected by Otsu's method. Then Iterative Global Minimum Screening is done. After which a CNN model consisting of 7 convolutional layers, 3 Max-pooling layers, 3 fully connected layers and 1 softMax layers is used for classification of the positive and negative cases. The system was trained upon 150 patients over 1818 images. The parameters were calculated with an accuracy of 93.46%, a specificity of 94.33% and a sensitivity of 92.59%.

[1.7] This paper introduces the blood image processing so as to assess the parasitaemia of the blood. This paper intends to distinguish the red blood cells that are infected by malaria parasites utilizing statistical base approach. The keywords used are digital image processing, pattern recognition, shape analysis, invariant moments, and malarial blood images. Further assessment of the size and state of the nuclei of the parasite is additionally thought of. The point of this paper is to introduce a model to recognize the parasite utilizing digital image of stained malarial blood from a microscope so as to to assess the number of parasitaemia of the blood that is tally number of parasites per number of red blood cells. The image returned is statistically examined and contrasted to produce a statistic database.

[1.8] The images that are to be worked with are obtained in gray scale and are thresholded or masked to obtain the portion of the image required. After the portion of the RBC image is selected, the parameters like mean, standard deviation and coefficient of variation are calculated. The number of pixels that were counted during pre-processing were taken to find the area and the border pixels to find the perimeter after which circularity is found using a formula. These are the six parameters that are used as an input to the ANN: area, perimeter, circularity, average, standard deviation and variance. They were obtained using holographic images of 24 healthy RBC's and 24 infected ones. Binary bipolar outputs are obtained with -1 that stands for Malaria negative and 1 standing for Malaria positive. A Feed-forward neural network that was trained utilizing back propagation was created by MatLab. An SVM classifier was used to detect the infected cells and the outputs are obtained.

[1.9] In this paper the authors have proposed a new image processing based framework which incorporates two algorithms. One is Haar wavelet for image transformation and the other one is K- Nearest algorithm for image classification, the main objective in the proposed system is to develop a malaria parasite detection system in which pathology admin will transfer the patient's scanned RGB report. To build up an expert system for patients after uploading image transformation, feature extraction and image classification. The feature extractions are done by uploading image and scale those images onto 256*256 pixels and transform the original image using Haar wavelet algorithm. It is used to compress the images and store those pictures for further classification. Image classification is done by using KNN algorithm by calculating Euclidean distance with the help of extracted features. In Euclidean distance systems will form clusters of multiple stages among these clusters suitable cluster will be considered as a final malarial stage. Then the K-Nearest Neighbour algorithm which is a method that does not use the estimation of parameters is used. Input consists of K-closest nearest sample in the feature. So, this system is interactive, hence is faster and more accurate than manual process. This system will help limit the human mistake while recognizing the presence of malaria parasites in the blood sample by using Image Processing and limit human blunder by automation.

[1.10] Image segmentation and feature extraction using minimum distance classifier was used to identify the parasites in the blood sample . Feature extraction uses two phases in architectural model: Training phase and Recognition phase which helps to recognize the Malaria parasite. In this work, they focus on automated detection and quantification of malaria detection, the strategy to determine infected images

using machine learning to improve the predictive value for detection of infected cells. The image is acquired that may contain impurities and noise. It is converted to gray scale and the zones are segmented by recognizing the similar properties. The image is thresholded by creating binary images for grey-level ones by converting all pixels below some threshold to zero and all pixels above to one. Then, the image was enhanced to make it more suitable for further processing based on intensity property. Erosion and Dilation are applied to remove a considerable amount of noise. After this, the images are segmented using watershed segmentation. It tends to separate touching objects so that overlapped RBCs will be separated and will be helpful for counting the RBCs. Using the CIE system they have specified any color in terms of its coordinates and have measured the sensitivities of three broad bands by suiting spectral colors to certain mixtures of three colored lights. After the segmentation the mean perimeter of the RBCs are found with the help of Matlab function region props helps to quantify properties of image regions. Then the parasite compares whether it is greater than the mean value of RBC cells. A circle is plotted around the infected RBC and they are calculated.

3. Functional Architecture:

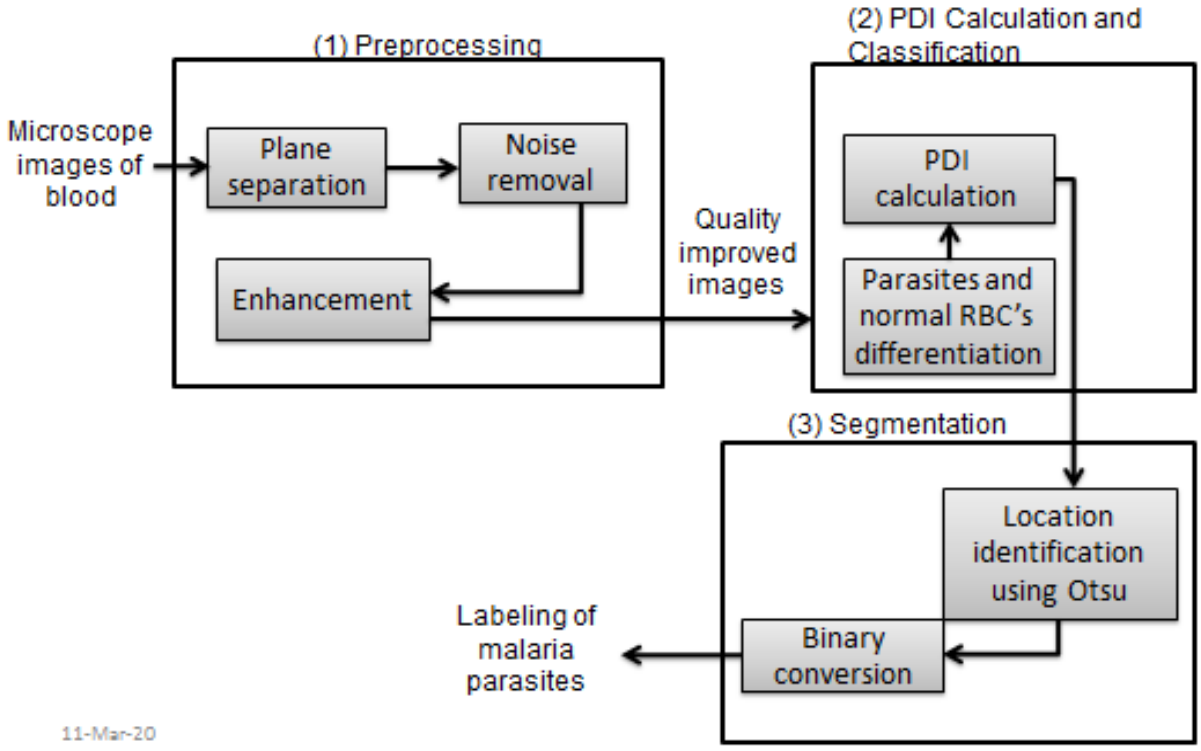


Figure 1.1. Functional Architecture diagram

4. Preprocessing

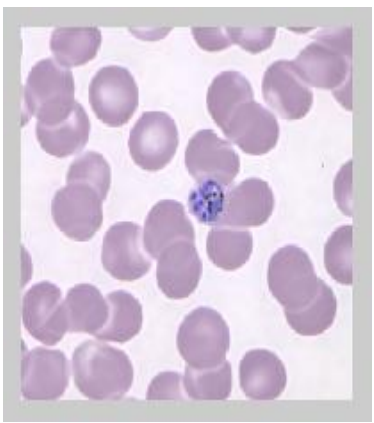


Figure 2.1

Figure 2.2

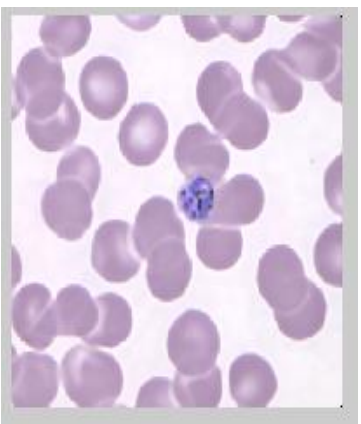


Figure 2.3

Figure 2.1. Original input image before pre-processing
Figure 2.2. Image after noise is removed by Median Filter application
Figure 2.3. Enhanced Image

The Geimsa stained microscopic images of blood samples are acquired. The predecessor to detecting our required features is to preprocess the image. The RGB images are divided into their individual components using plane separation method. The green plane is used for further processing as it is closest in resemblance to the RGB image itself. Median filter is applied to remove shot noise from the images acquired from the electron microscope (Figure 2.2). After the noise removal the images are enhanced to highlight it's features for further processes (Figure 2.3). This constitutes the first module.

5. PDI Calculation and Classification

The enhanced images are manipulated among the three different planes in order to bring out the parasite if present. The Green channel is masked with the blue channel to intensify the parasite's location. The resultant channel's intensity is doubled and is masked onto red channel. The resultant image is what we have termed to be the Cell Detection Index (Figure 3.1). An appropriate integer is chosen to multiply with the CDI such that the parasite's intensity does not exceed 255 to obtain the Parasite Detection Index (Figure 3.2). In the PDI if there are intensities greater than 127 present we will conclude that Malaria is present. The respective channels for masking are chosen in such a way for the sole purpose of making the Parasite Detection easier with a specific threshold. Based on the intensity of the PDI, the existence or absence of the parasite will be identified using an SVM classifier and the result will be shown. This PDI calculation and Classification constitutes the second module.

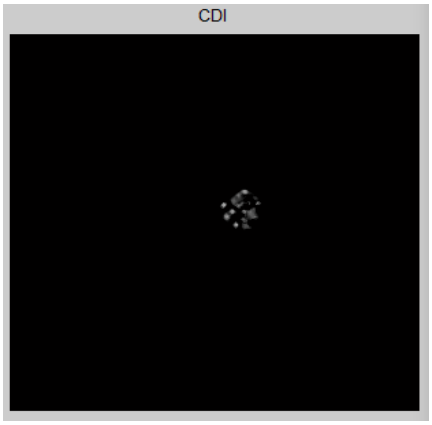


Figure 3.1

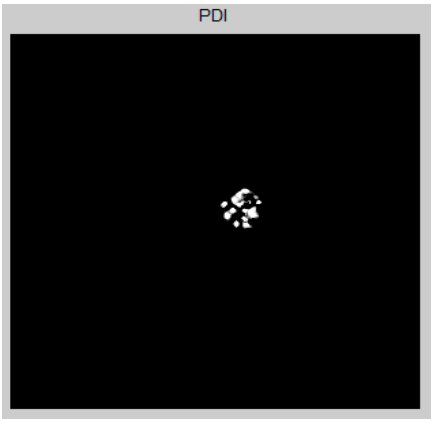


Figure 3.2

6. Segmentation

After the presence of the parasite is confirmed, the location of the parasite is also rounded up by segmenting the image. Otsu's Segmentation is used to locate the parasite.
Algorithm:

1. Compute the histogram and probabilities of each intensity level
2. Optimal threshold is found
3. Minimum intra-class variance in computed
4. Maximum inter-class variance is computed to separate the distinct classes

Since our images are distributed in a multi-modal fashion Otsu's is preferable. The microscopic images will have three classes. The background, cells other than parasite and the foreground that is our parasite. After the classes are plotted as a histogram the thresholds are found out and threshold variance between two classes is made maximum and the variance of intensity levels within a class are kept the lowest for a

vivid distinction between classes. To avoid losing the boundary region of the parasite to the uninfected region holes-filling is done. The dimensions of the bounding box that shows the extracted parasite is calculated and the output is produced.

The preprocessing and masking processes are done using functions of the Image Processing Toolkit in MatLab R2013a.

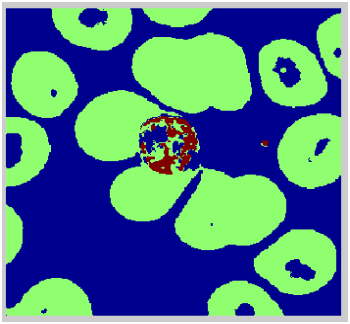


Figure 4.1 Segmented Image

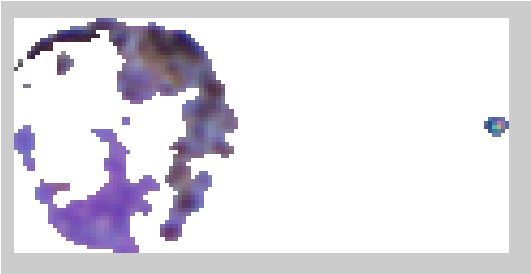


Figure 4.2 Parasite extracted Image

Implementation:

A report is generated based on the PDI of the input images. Our ROI (region of interest) must have an intensity that is greater than that of 127. A training dataset is created. Affected and unaffected samples are marked as Groups 1 and 2 for the sake of the dataset. The groups and their PDI intensities are tabulated for the reference of producing the results. Due to the CDI and PDI we can eliminate the doubt of white blood cells that are closely as dark as the parasite nuclei. The input required for this implementation is thin smear blood samples. It can be implemented in any medical unit that lacks medical experts for diagnosis.

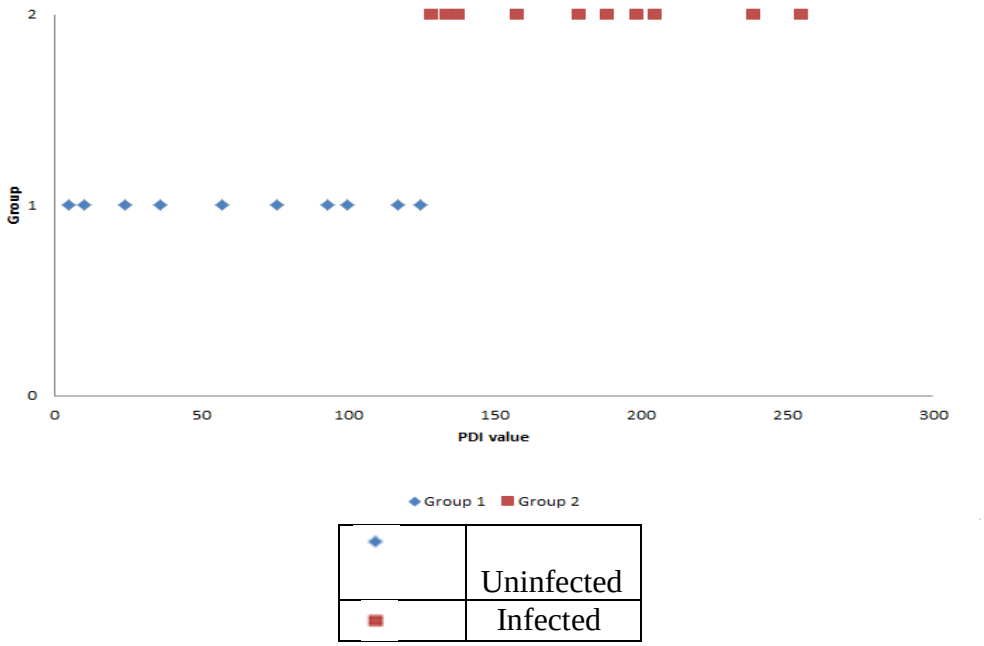
5. Results and Discussion:

A clear distinction between the positive and negative cases classified on the basis of the PDI value is represented graphically in Graph 6.1. The advantage of this system is that using the image processing functions to detect the parasite makes the process easier and faster. Accurate results can be produced at the cost of a smaller training dataset when compared to the other methods.

Technique used	Accuracy	Specificity	Sensitivity
CNN	93.46%	94.33%	92.59%

PDI calculation	94.30%	81.81%	100%
Restricted Boltzmann Machine	94.04% ()	95.92%	97.60%

Table 5.1. Comparison of parameters with alternate methods



Graph 6.1. Relationship between PDI value and Parasitaemia

6. References:

[1] A.S Abdul Nasir, M.Y Mashor, Z.Mohamed, "*Segmentation based approach for detection of malaria parasite using moving k-means clustering*" in **[2012]** IEEE EMBS International Conference on Biomedical Engineering and Sciences.

[2] Amit Doegar, Poonam Kashtriya, Varun Gupta, Vikas Kashtriya "*Identifying Malaria Infection in Red Blood Cells using Optimized Step-Increase Convolution Neural Network Model*" on July**[2019]** published in International Journal of Innovative Technology and Exploring Engineering (IJITEE).

[3] Andino Maseleno, G. Jose Moses, E.Laxmi Lydia, K.Shankar, N.Sharmili, "*Image Classification using Deep Neural Network for Malaria Disease Detection*" published in International Journal on Emerging Technologies in **[2019].**

[4] Bibhudendra Acharya, Neetu Ahirwar, Sapnojit Pattanaik "*Advanced Image Analysis Based System for Automatic Detection and Classification of Malarial Parasite in Blood Images*" published in International Journal of Information Technology and Knowledge Management.

[5] Dhanya Bibin, Madhu.S.Nair, P.Punitha published on May 18, **[2017]** "*Malaria Parasite Detection from Peripheral Blood Smear Using Deep Belief Networks*" in IEEE.

[6] Feng Yang, Hang Yu, Jian Yu, Kamolrat Silamut, Mahdieh Poostchi, Richard J Maude, Sameer Antani, Stefan Jaeger, Zhou Zhou published "*Deep Learning for Smartphone- based Malarial Parasite Detection in Thick Blood Smears*" in IEEE Journal of Biomedical and Health Informatics submitted on Dec 19**,[2018].**

[7] Guarav Bajpai, S.Raviraja, Sudhir Kumar Sharma "*Analysis of detecting the malarial parasite infected blood images using statistical based approach*" in **[2014].**

[8] Hassan Rashidi Heram - Abadi, Saeid Afkhami, "*Detection of Malarial Parasite using Two Classification Methods: Support Vector Machines and Artificial Neural Network*" accepted on 02 April**, [2017]** in IJOCIT.
[9] S.A. Ladhake, Kanchan N.Poharkar published "*Implementation of Malaria Parasite Detection System Using Image Processing*" in IJSRD on March **[2018].**
[10] Machana Shrestha, Rojesh Man Shikhrakar, Suman Kunwar "*Malaria Detection Using Image Processing and Machine Learning*" published by in July **[2012].**
[11] Raja, S. Kanaga Suba, and T. Jebarajan. "Reliable and secured data transmission in wireless body area networks (WBAN)." European Journal of Scientific Research 82, no. 2 (2012): 173-184.