IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

An Application of Machine Learning to Detect Blood Diseases

Shubham Janardhan Kadu
PG Scholar
Post Graduate Department of Digital Electronics
Prof Ram Meghe College of Engineering and Management
Badnera, Amravati

Dr. P. S. Choudhary
Associate Professor
Post Graduate Department of Digital Electronics
Prof Ram Meghe College of Engineering and Management
Badnera, Amravati

Abstract:

The human blood is a health indicator; it delivers necessary substances such as oxygen and substance that provides nourishment is necessary. Hence, segmentation of blood cells and identification of blood type is very important. The human blood consists of the RBCs, WBCs, Platelets and Plasma. Presently, lab technicians' tests blood groups manually and they use a device called Hemocytometer and microscope to count blood cells. But this method is extremely time consuming, monotonous and leads to the inaccurate results due to human errors.
To overcome the problems regarding time, accuracy and cost, a method is proposed based on processing of images acquired from laboratory. The image processing techniques such as Segmentation, Morphological operations and Circular Hough Transform will test with some samples.

Fast and reliable desktop application platform for blood image analysis and malaria diagnosis from Giemsa-stained thin blood film images. The application is based on novel Annular Ring Ratio Method which will test and validate in IMAGE PROCESSING TOOL. The method detects the blood components such as the Red Blood Cells (RBCs), White Blood Cells (WBCs), and identifies the parasites in the infected RBCs. The application also recognizes the different life stages of the parasites and calculates the parasitemia which is a measure of the extent of infection.

Keyword: RBC, WBC, Malaria Detection, Machine Learning, Feature Extraction, Python

Introduction:

Malaria is a deadly disease and the recent survey by the World Health Organization (WHO) has estimated that malaria causes over 200 million cases of fever annually [1]. The diagnosis of the disease requires powerful and expensive tools unavailable for the poorest countries of the world, where often the disease is endemic. Microscopic malaria diagnosis is, by far, considered to be the most effective diagnostic method, but it is highly time-consuming and labor intensive. The accuracy of the system solely depends on the expertise of the microscopic. Other techniques widely involved in Malaria diagnosis are Rapid Diagnostic Tests (RDTs) and Polymerase Chain Reaction (PCR) tests [3]. However, the accuracy of these tests depends on the extent of infection with sensitivity directly proportional to the level of infection. Various automated malaria related diagnostic studies are described in [3]-[9]. Recognizing the potential of mobile technology and internet to revolutionize the access to information throughout the developing countries like India and Africa as well as developed nations, the work reported in this paper exposes a reliable automated Android based diagnostic platform, without expert intervention for the effective treatment and eradication of the deadly disease, which can be deployed in all the Android based mobile phones and tablets.

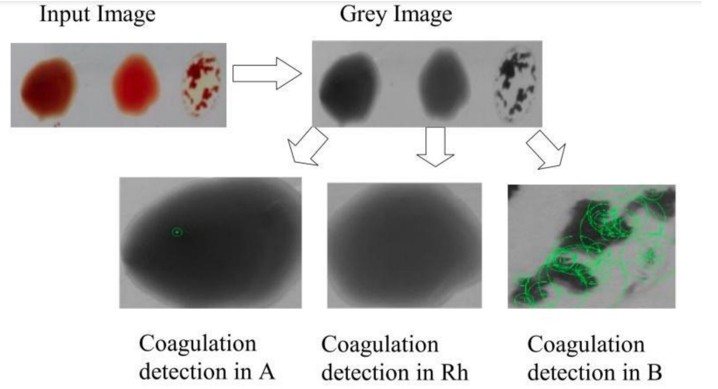


Fig 1. How blood analysis work.

Blood grouping tells us what type of blood a person has. Everyone may have different blood groups. These differences in human blood groups are because of the presence or absence antigens and antibodies on the ML ace of blood cells. Individuals have different combinations of antigens and antibodies and therefore have different blood groups. According to ABO and Rh blood grouping systems there are 8 different blood groups: A Rh+, A Rh-, B Rh+, B Rh-, AB Rh+, AB Rh-, O Rh+ and O Rh- [1-2]. The complete blood count (CBC) evaluates the health of person and detects the disorders like anemia, infection and leukemia. CBC is very important in medical diagnosis [3]. RBCs, WBCs, platelets, plasma these are constituents of human blood. The complete blood count involves counting of these four types of cells. The count of these cells determines the ability of an organism to resist a particular infection and capability of the body system [4]. The normal count of these cells is different for men, women, and children, etc.

Literature Survey:

Chima S. Eke, Emmanuel Jammeh et.al. presented Early Detection of Alzheimer's Disease with Blood Plasma Proteins using Support Vector Machines. The successful development of amyloid-based biomarkers and tests for Alzheimer's disease (AD) represented an important milestone in AD diagnosis. objective in this study is to developed a method to identify potential blood-based non-amyloid biomarkers for early AD detection [1]. Sneha Raina, Abha Khandelwal et.al. focused on Blood Cells Detection Using Faster-RCNN. They applied Faster R-CNN to the Blood Cell Count and Detection dataset to detect RBCs, WBCs, and Platelets [2]. Hasina khatoon Dr.Dipti Verma reviewed Identification of Diabetes Disease from Human Blood Using Machine Learning Techniques. They compared various machine learning and neural network-based approaches that are applied on publicly available datasets [3]. Fahad Kamal Alsheref, Wael Hassan Gomaa published a paper on Blood Diseases Detection using Classical Machine Learning Algorithms. They concluded that novel benchmark data set that contains 668 records. The data set is collected and verified by expert physicians from highly trusted sources. Several classical machine learning algorithms are tested and achieved promising results [4]. Christy Evangeline N, Annalatha M proposed the use of Computer aided system for Human Blood Cell Identification, Classification and Counting. They described work aims to generate a preliminary framework of Automatic Analysis of Human Blood smear for identification, classification and counting of blood cells by using various Image processing [5]. Vasundhara Acharya, Preetham Kumar presented Identification and Red Blood Cell Classification using Computer Aided System to Diagnose Blood Disorders. According to the research they provided accuracy of 98 percent correct identification of type of Anemia when the obtained results were compared with evaluation done by medical experts [6]. M. B. Menhaj et.al. Introduced A fuzzy based classifier for diagnosis of acute lymphoblastic leukemia using blood smear image processing. They used morphology method for the diagnosis of acute lymphoblastic leukemia [7]. S.D. Bias et.al. proposed the Novel Fuzzy Logic Inspired Edge Detection Technique for Analysis of Malaria Infected Microscopic Thin Blood Images. They mainly focus on the analysis of malaria infected microscopic thin blood smears. The algorithm proposes a simple, dynamic thresholding technique that is computed via histogram analysis, designed to capture as much information about the blood cells with minimal computational effort, which is followed by a morphological filtering process to remove noise and artifacts [8]. Hanung Adi Nugroho et.al. Introduced Plasmodium detection methods in thick blood smear images for diagnosing Malaria. In this paper, a comparison of some image processing methods will be reviewed to determine the best method applied in this field [9]. Krishna Kumar Jha et.al. Presented Detection of Abnormal Blood Cells on the Basis of Nucleus Shape and Counting of WBC. they utilized framework to segment the blood cells, to identify different type of normal and abnormal WBC and count number of WBC [10]. Pranati Rakshit et.al. Introduced a paper on Detection of presence of Parasites in Human RBC In Case of Diagnosing Malaria Using Image Processing. they used several pre-processing activities; area of the infested corpuscle is calculated and Sobel Edge detection method is used to find the boundary of the corpuscles [11]. Vincenzo Piuri et.al. presented research on the acute lymphoblastic leukemia image database for image processing. they proposed new public dataset of blood samples, specifically designed for the evaluation and the comparison of algorithms for segmentation and classification [12]. Fabio Scotti published a paper on Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images. The proposed system firstly individuates in the blood image the leucocytes from the others blood cells, then it selects the lymphocyte cells (the ones interested by acute leukemia), it evaluates morphological indexes from those cells and finally it classifies the presence of the leukemia [13]. David J. Foran et.al. proposed the computer-Assisted Discrimination Among Malignant Lymphomas and Leukemia Using Immunophenotyping, Intelligent Image Repositories, and Tele microscopy. they developed a distributed, clinical decision support prototype for distinguishing among hematologic malignancies. The system consists of two major components, a distributed tele microscopy system and an intelligent image repository [14]. P.G. Bhende proposed Analysis of Blood Cells using Image Processing they developed the differential blood counter (DBC) system [15].

Summary:

Sr. No.	Research Paper Name	Author	Year	Observation	Scope to Work
1.	Blood Cells Detection Using Faster - RCNN	Sneha Raina, Abha Khandelwal, Saloni Gupta	Oct 2-4, 2020	She applied Faster R-CNN to the Blood Cell Count and Detection dataset to detect RBCs, WBCs, and Platelets.	In identifying bounding boxes around the blood smear slide images, the accuracy around the clusters of red blood cells was depressed. It can be improved in various ways which include but are not limited to increasing the number of epochs, experimenting with the layers in convolutional networks etc. However, there is a conundrum to that method. By increasing the number of epochs, we can encounter over-fitting which will reduce limit accuracy.
2.	Early Detection of Alzheimer's Disease with Blood Plasma Proteins using Support Vector Machines	Chima S. Eke, Stephen Pearson	2020, IEEE	Existing ML models performed poorly in comparison at this stage of the disease suggesting that the underlying protein panels may not be suitable for early disease detection. Our results demonstrate the feasibility of early detection of AD using non amyloid based biomarkers.	The main objective is to develop a ML-based method (support vector machines (SVM) in particular – see later) to identify blood biomarkers of early AD based on non-amyloid proteins with the potential to identify the disease prior to accumulation of amyloid-beta in the brains.
3.	Identification of Diabetes Disease from Human Blood	Dr. Dipti Verma,	2020, IEEE	She compared various machine	Diabetes is portrayed as a social event of metabolic issue applying immense

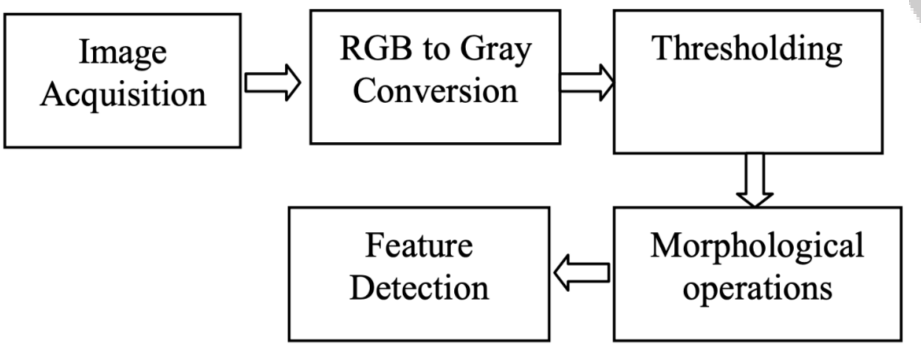
	Using Machine Learning Techniques	Ankit Arora		learning and neural network-based approaches that are applied on publicly available datasets.	load on human wellbeing around the world. Wide research in all pieces of diabetes has provoked the time of gigantic proportions of data. The purpose of the current examination is to coordinate a productive review of the employments of AI, and mechanical assemblies in the field of diabetes research with reverence to a. Forecast and Diagnosis, b. Diabetic Complications, c. Hereditary Background and Environment, and e. Human services and Management with the primary grouping having all the reserves of being the most conspicuous.
4	Blood Diseases Detection using Classical Machine Learning Algorithms	FahadKamal Alsheref , Wael Hassan Gomaa2	International Journal of Advanced Computer Science and Applications, 7, 2019 (IJACSA)	He concludes Machine learning becomes an essential technique for modeling the human process in many disciplines, especially in the medical field, because of the high availability of data. One of the essential disease detectors is the blood analysis; as it contains many parameters with different values that indicates definite proof for the existence of the disease.	The future work will focus on testing the proposed data set using different deep learning algorithms to compare between classical and deep learning approaches in this research area.
5	Computer aided system for Human Blood Cell Identification, Classification and Counting	Christy Evangeline N	2018 IEEE	This proposed works aims at both Qualitative and Quantitative Analysis simultaneously there by giving Complete Blood Count and the Differential WBC information at the same time. Thus, this work can effectively be used as an alternative to conventional microscopic studies and the Flow cytometry Methods.	Future work can be extended towards development of automated cell counters using Artificial Neural Networks that includes segmenting WBCs into nuclei and cytoplasm, feature extraction from each region, classifying WBCs into five categories, and accurate disease diagnosis.

Research Methodology:

The system is divided into three significant phases: Blood group detection, RBC counting, WBC counting. Figure 1 shows general block diagram of digital image processing. From control panel of system, we have to select an option to perform required test i.e., Blood group detection, RBC counting, WBC counting. After this one window will appear from which we have to select an input image. All the images are stored by the name of patients. Then image undergoes some image processing techniques as mentioned below and we get the result in GUI edit window. Figure 4 shows general block diagram of In image acquisition of blood group detection images of blood samples is obtained from the laboratory consisting of a color image composed of three samples of blood and reagent. Simply a glass slide with blood sample is placed on a white paper and photo is taken by using a phone camera of 12MP. The image is read from phone storage at the time of system evaluation. In preprocessing rbg to gray conversion is performed. Next step is to detect MLfeatures that works on gray images. In computer vision, speeded up robust features (ML) is a local feature detector and descriptor. It can be used for object recognition, image registration, classification or 3D reconstruction [6]. Here we have used MLto detect the coagulation formed in an image. Based on this detection of coagulation we are detecting blood groups. Figure shows the output images of ML feature detection.

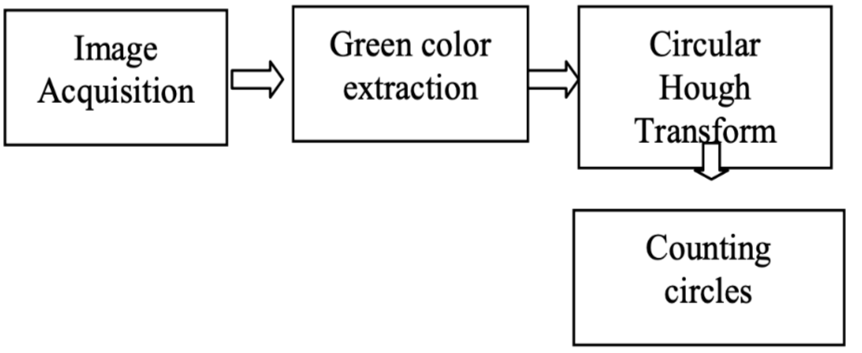
Phase-I: Blood Group Detection

This is the first phase of proposed system i.e., Blood Group detection. Figure shows flow of blood group detection. Block diagram for Blood Group Detection In image acquisition of blood group detection images of blood samples are obtained from the laboratory consisting of a color image composed of three samples of blood and reagent. Simply a glass slide with blood sample is placed on a white paper and photo is taken by using a phone camera of 12MP. The image is read from phone storage at the time of system evaluation. In preprocessing to gray conversion is performed. Next step is to detect ML features that works on gray images. In computer vision, speeded up robust features (ML) is a local feature detector and descriptor. It can be used for object recognition, image registration, classification or 3D reconstruction [6]. Here we have used ML to detect the coagulation formed in an image. Based on this detection of coagulation we are detecting blood groups. Figure shows the output images of ML feature detection.



Phase-II: Counting of RBC

This is the second phase of proposed system i.e., counting of RBC. shows flow of RBC counting. Figure: General block diagram of RBC counting to examine the RBCs and WBCs blood images are captured with the help of thin glass slides and digital microscope. In Digital Image Processing there are 3 color planes, R, G and B which contains color information in images. The green component is extracted because it contains maximum value [7]. 3.2.1 Image Segmentation by Circular Hough Transform The circular hough transform is then applied to green color image. This transform searches for the blood cells in the image and then detects them. The function “draw circle” draws circles around the detected cells. Even the overlapped circles are detected. Equation (1) is a standard equation of circle. $r^2 = (x - a)^2 + (y - b)^2$ (1) Here a and b are the coordinates for the center, and r is the radius of the circle. sin () x = a cos () y b r r (2) Equation (2) represents parametric equations of circle. Circular Hough transform detects edge points on each circle, and draws a circle with that point as origin and radius r. The circular Hough transform uses a three-dimensional array with the first two dimensions representing the coordinates of the array and increases every time when circle is drawn with the desired radii over every edge point. An accumulator keeps count of how many circles pass through coordinates of each point, and finds the highest count [8].

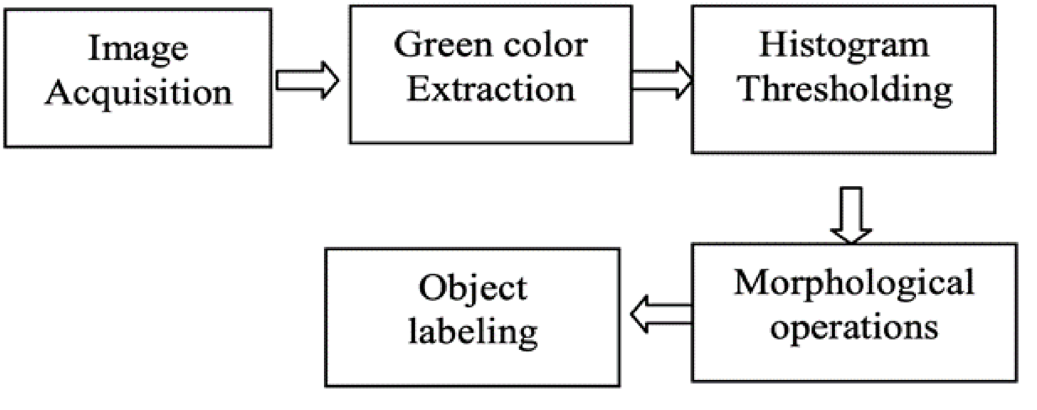


Objective

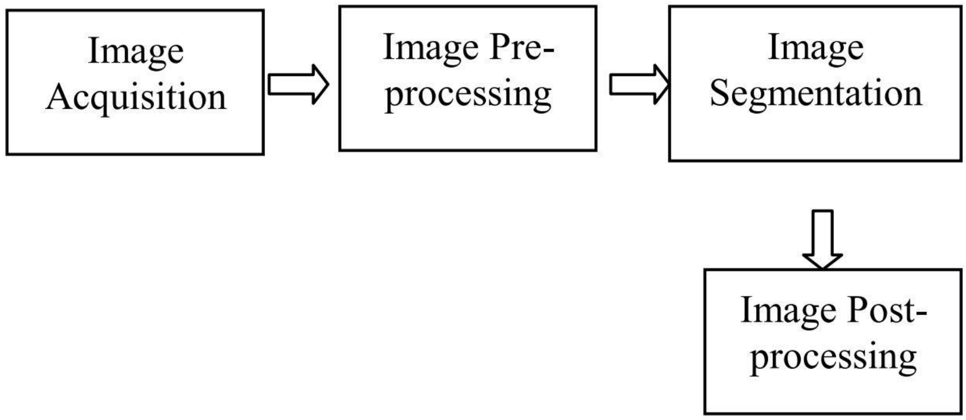
- IN this model we will detect count of RBC, WBC, Malaria, platelets, blood group using image processing method.
- The input of the application is an RGB Giemsa stained thin blood film image. In order to speed up the process, the image is converted to gray scale.
- A small threshold is applied to the peaks to remove the background clutter. The method detects the foreground cells and the location information (pixel coordinates) of the centroid of the cells stored which will be used at a later stage to identify the infected cells.

Counting of WBC

This is the third phase of proposed system i.e., counting of WBC. Figure shows flow of WBC counting. Figure: General block diagram of WBC counting in thresholding if the image intensity is greater than some fixed constant each pixel in an image is replaced with a black pixel and if the image intensity is smaller than some fixed constant each pixel in an image is replaced with a white pixel [9]. (3) Thus, pixels labeled 0 corresponds to objects and pixels labeled 1 correspond to background. After thresholding we have taken the complement of that image. And further morphological operation like erosion and delusion are performed to smoothen the image and to fill the holes and gaps. Using Dilation and erosion secondary operations like closing and opening are performed. Closing operation is used to fill the holes and gaps. It is the process of dilation which is followed by erosion. Opening operation is used to smoothen the contours of cells and parasites. It is process in which erosion is followed by dilation.



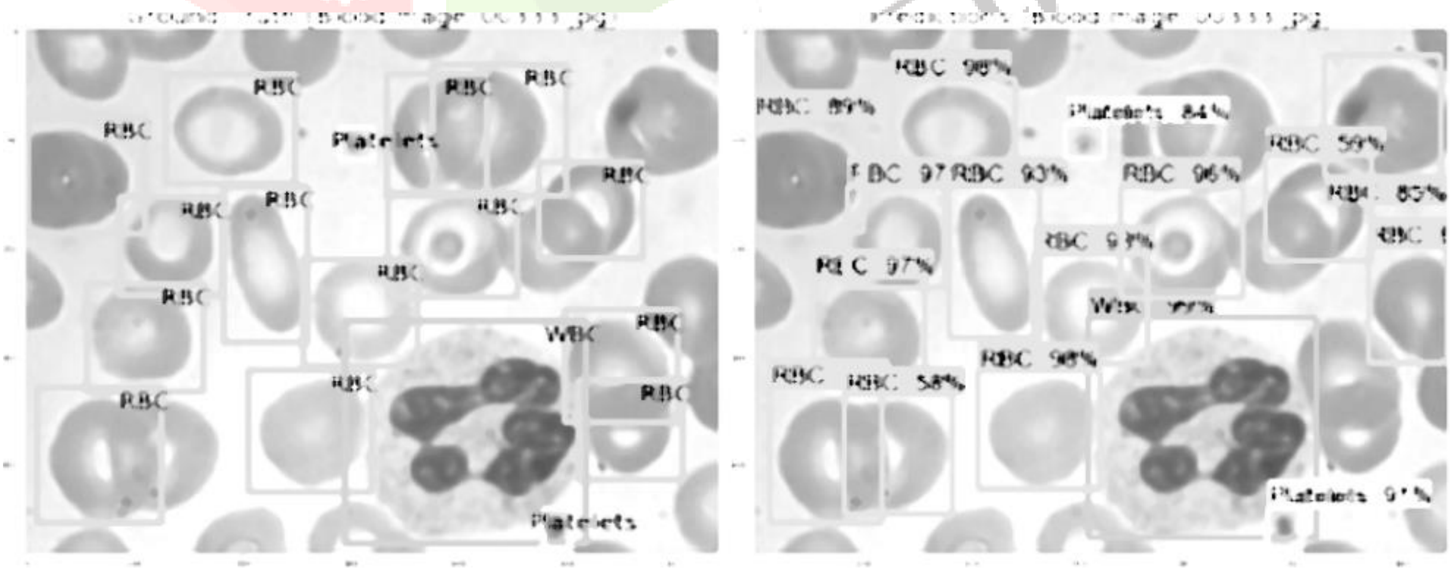
Block diagram

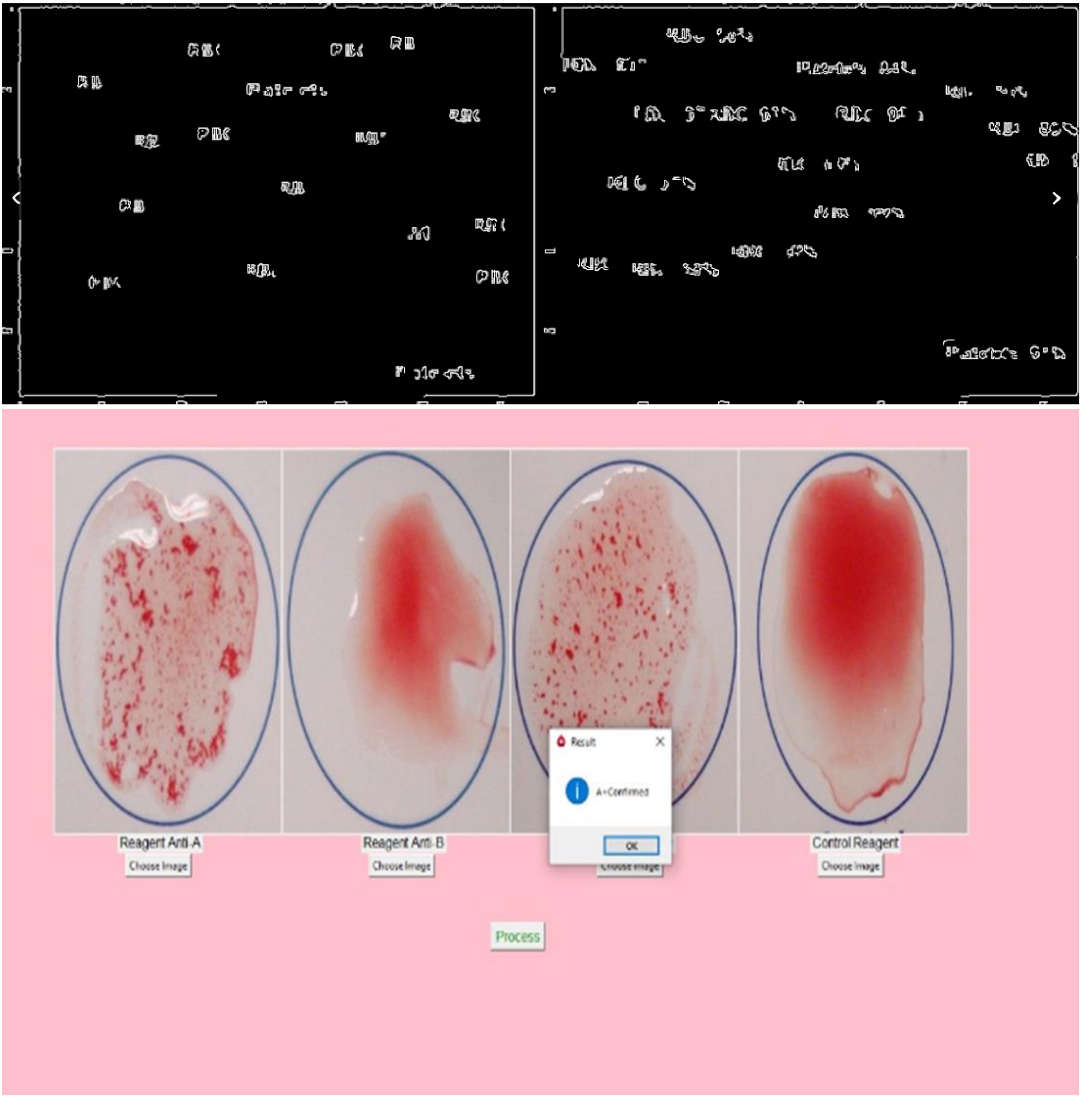


- IN this model we will detect cout of RBC, WBC, Maleria, platelets, blood group using image processing method.
- The input of the application is an RGB Giemsa stained thin blood film image. In order to speed up the process, the image is converted to gray scale.
- A small threshold is applied to the peaks to remove the background clutter. The method detects the foreground cells and the location information (pixel coordinates) of the centroid of the cells stored which will be used at a later stage to identify the infected cells.

Digital image processing encompasses a broad range of hardware, software and theoretical underpinnings. To identify and discovering the blood cells from each other, the segmentation and edge detection techniques are helpful. Before examine the structure of different blood cell, the images can be recorded with the help of glass slides and images get captured using microscopes. The images get converted color images to gray level images. Classifying the image by gray-level pixels may reduce and simplify some image processing operations such as edge detection, edge smoothing, feature extraction, image processing and image registration. [4]

Result





Conclusion:

By using digital image processing, analysis of blood cell image is more accurate as well as this method is efficient in terms of cost and time consuming compared to existing techniques of blood cell analysis. MATLAB software use for this analysis. Day by day research work is increasing in this field and various image processing techniques are implemented in order to get more accurate result. For medical diagnosis and blood cell counting use of image processing techniques is useful and better than existing techniques provided that standardization of blood smear is done properly to obtain blood cell image.

REFERENCES

[1] Chima S. Eke et al., “Early Detection of Alzheimer’s Disease with Blood Plasma Proteins using Support Vector Machines”, 2020 IEEE Journal of Biomedical and Health Informatics.
[2] Sneha Raina et al., “Blood Cells Detection Using Faster-RCNN”, 2020 IEEE International on Computing, Power and Communication Technologies (GUCON). ©2020 IEEE
[3] Hasina khatoon, Dr.DiptiVerma, and Ankit Arora, " Identification of Diabetes Disease from Human Blood Using Machine Learning Techniques," 978-1-7281-7213-2/20/2020 IEEE
[4] Fahad Kamal Alsheref, and Wael Hassan Gomaa, “Blood Diseases Detection using Classical Machine Learning Algorithms,” Vol. 10, No. 7, 2019
[5] Christy Evangeline N and Annalatha M, " Computer aided system for Human Blood Cell Identification, Classification and Counting," 978-1-5386-4472-0/18/2018 IEEE
[6] Vasundhara Acharya, and Preetham Kumar, “Identification and Red Blood Cell Classification using Computer Aided System to Diagnose Blood Disorders”. 978-1-5090-6367-3/17/2017 IEEE
[7] M. A. Khosrosereshki, and M. B. Menhaj, "A fuzzy based classifier for diagnosis of acute lymphoblastic leukemia using blood smear image processing", 978-1-5090-4008-7/17/2017 IEEE
[8] S.D. Bias and S.Kareem Reni, I. Kale, "A Novel Fuzzy Logic Inspired Edge Detection Technique for Analysis of Malaria Infected Microscopic Thin Blood Images," 978-1-5386-1030-5/17/2017 IEEE
[9] Chyntia Raras Ajeng, Hanung Adi Nugroho,and Igi Ardiyanto. “Plasmodium detection methods in thick blood smear images for diagnosing Malaria”, 2012. 978-1-5090-1567-2/16/2016 IEEE
[10] KRISHNA KUMAR JHA, BIPLAB KANTI DAS, HIMADRI SEKHAR DUTTA, "Detection of Abnormal Blood Cells on the Basis of Nucleus Shape and Counting of WBC ", IEEE Trans. on Medical Imaging, vol. 31, no. 2, pp. 207-216, Feb. 2019.
[11] Pranati Rakshit, Kriti Bhowmik, "Detection of presence of Parasites in Human RBC In Case of Diagnosing Malaria Using Image Processing", IEEE/ACM Trans. on Computational Biology and Bioinformatics, vol. 11, no. 6, pp. 1131-1145, Nov./Dec. 2019.
[12] Ruggero Donida Labati, Vincenzo Piuri, Fabio Scotti, "Gene expression profiling predicts clinical outcome of breast cancer", Nature, vol. 415, no. 6871, pp. 530-536, Jan. 2021.
[13] Fabio Scotti. "Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images", 2017. [Online]. Available: https://www.fullstackpython.com/why-use-python.html. [Accessed: 12- Sep- 2017].
[14] David J. Foran, Dorin Comaniciu, Peter Meer, and Lauri A. Goodell, "Computer-Assisted Discrimination Among Malignant Lymphomas and Leukemia Using Immunophenotyping, Intelligent Image Repositories, and Telemicroscopy", Computing in Science & Engineering, vol. 9, no. 3, 2017.
[15] Mrs.P.G.Bhende, "Analysis of Blood Cells using Image Processing", Computing in Science & Engineering, vol. 13, no. 2, pp. 13- 21, 2018.