Counting and Classification of White Blood Cell using Artificial Neural Network (ANN)

Shubham Manik[1]; Lalit Mohan Saini[2] and Nikhil Vadera[3]
[1,2,3]School of Bio-Medical Engineering; National Institute of Technology; Kurukshetra
E-mail: [1]shubhmanik1992@gmail.com; [2]lmsaini@nitkkr.ac.in; [3]nikhilnv123@gmail.com

***Abstract*—Quantitative microscopy has improved conventional diagnostic method through better comprehension of microscopic features from a clinical point of view. The extraction of the nucleus from the blood smear images of white blood cells (WBC) gives the profitable data to specialists for classification of various types of disease as the vast majority of the illnesses present in the body can be distinguished by investigating blood. It is extremely time-consuming and tedious to segment the nucleus manually and also classification which is done on the premise of the instruments which are utilized by specialists for segmentation and classification of white blood cells are not economical for every doctor or hospital; so proposed method is ideal which decreases the execution time of segmentation and classification. In this paper; a new framework is proposed to enhance detection and classification of Leukocytes i.e. Nucleus Enhancement by finding Intensity maxima and then classified on the basis of various features extracted from segmented images. Classification is done by Artificial Neural Network (ANN).**

Keywords*—White Blood Cells (WBCs); Leukocytes; Nucleus Enhancement; Artificial Neural Network (ANN)

I. INTRODUCTION

Blood is a connective tissue that gives a system of quick transport of supplements; waste items; respiratory gases in humans and different creatures [1]. Blood is generally fluid thicker than immaculate water due to various cells and proteins suspended in it. A normal individual has around 5 Litre of blood. About 55% of the blood; constitutes a fluid known as plasma [2].Plasma is comprised of proteins that assist in coagulation of blood and perform different functions like it carries various substances. Glucose and some other nutrients are present in plasma. Blood volume left contains blood cells:

- Erythrocytes; oxygen carrier.
- Leukocytes; fight against infection.
- Thrombocyteshelp in blood clotting.

White blood cells oppose diseases and parasites. White Blood Cell (WBCs) or known as Leukocytes which provides immunity against various infectious diseases and alien invaders. Due to the presence of nuclei in White blood cells; they can be distinguished easily from other blood cells. Five types of Leukocytes are Neutrophils; Lymphocytes; Monocytes; Eosinophil and Basophile.[3] Physical appearances and functionality help to distinguish. Leukocyte count is an important element of the CBC; as Leukocyte population in the blood helps in diagnosis of disease. The normal leukocyte count is usually between 4.0-11.0 x 10^3 per μL [4]. Leukocyte volume is only 1% of the total blood volume in a healthy adult; making them comparatively less numerous than the RBCs at 40% to 45%.

There are five sorts of WBCs in the circulating blood; including lymphocyte; monocyte; neutrophil; eosinophil and basophil. White blood cell or leukocytes are divided into:

- Granulocytic series.
- Agranulocytic series.

Granulocytic Series these cells contain granules in their cytoplasm. Because of the presence of amultilobatenucleus; these are also called as polymorph nuclear leukocytes (Polys or PMN). Neutrophil; Eosinophil; Basophil come under Granulocytic series. Agranulocytic series are without any granules in their cytoplasm. These are also called as mononuclear cells. Lymphocytes; Monocytes; Platelets come under Agranulocytic series. Polys and lymphocytes make up 75% to 90% of the total WBC count. [3],[4],[5] The total lifespan of leukocyte is 13 to 20 days. They are produced in 7 to 14 days and their life in the peripheral blood is just 6 hours.

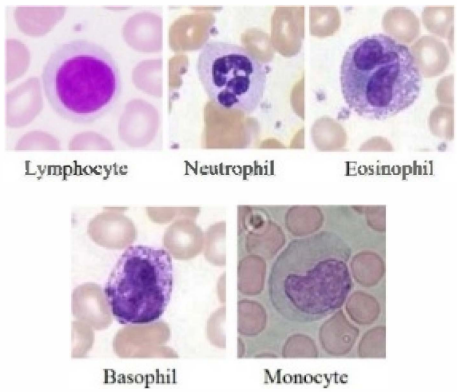

Fig. 1: Five Types of Leukocytes

II. LITERATURE REVIEW

Flow Cytometry: A programmed system that executes using theflow of leukocytes in a mono row through detectors. [6] These detectors measure the visual and electrical trace of the flow of liquid to anindividual count

of five types of Leukocytes. This process has advantages of breaking down vast volumes of human blood tests in the maximum proficient way and is turning out to be more sparing with advances in microfluidics stream cytometry.[7] On the other hand; it doesn't have any image capturing system of the cells being analyzed; in addition; examination concerning the morphology of cells and deviation is not possible. Additionally; the system can't recognize cells with unusual physiology in light of the fact that its method relies upon accepted ranges which are predefined.

The Manual Method of counting is a substitute method to count Leukocytes; though with much lower accuracy. The manual method of leukocyte counting can also be fulfilled on a hemocytometer utilizing a conventional magnifying instrument (microscope). For a blood test; the areas with single layer are physically examined for the total counting of leukocytes [8]; whereas for a hemocytometer; the region with grids is examined for the tally reason. [9] In spite of the fact that the manual system is more relentless and tedious; it gives researchers the adaptability to utilize a wider range of magnifying lens accessible of the conventional microscope for cautious visible examination of the samples. There is a tradeoff between the resolution of the image and the Field of Vision (FOV). For the most part; keeping in mind the end goal to see and total leukocyte counting under a standard microscope; magnification power of at least 10x (0.25Numerical Aperture) is utilized. [10] For an individual counting of leukocytes; an oil-immersed objective lens with around 100x amplification (1.4 Numerical Aperture) is utilized. At these high magnification levels; the FOV becomes less; that requires physical examine of the glass slide while the tally procedure.

Image processing Guclu Ongun *et al.* [11] proposed a new algorithm based upon snake segmentation. In this method; the center of mass of merged regions was used to find the initial position of the snake. After the initial position is known; snakes are put on the image and minimization procedure is done. The same process is repeated to find nucleus regions; but with changing energy constraints. For the classification purpose various shape based and texture based features were extracted. An SVM classifier yielded best results in most of the cases. Wang Shitong *et al.*[12] gave an advanced detection technique based on Fuzzy Cellular Neural Networks (FCNN). In this detection technique R; G; B images were separately compressed using the pyramidal method. Then mean gray value of the cytoplasm was taken as a threshold value and binary segmentation was done followed by erosion/dilation. Nuclei were located; detection of WBCs using correct window size. Mohammad Hamghalam *et al.*[13] used Giemsa-Stained images of aperipheral blood smear. Using Giemsa Stain; Otsu Thresholding can be performed efficiently. Afterthresholding; noise was removed from thebinary image. As neutrophils; which are abundant in number; have multiple lobules connected with the thin strand; different lobule may cause an error in counting so they were merged. To avoid the error of two near nuclei getting merge; the distance between nuclei is measured. Distance less than the diameter of nucleus got merged. The accuracy obtained was near 96.7%.

Jaebum Chung *et al.*[14] considered a method to increase Field of View by using the Fourier Ptychographic Microscopy. To obtain alarge field of view; there is a tradeoff with resolution. Fourier Ptychographic Microscopy is a novel method that can automatically stitch a series of lower resolution images in the frequency domain to get higher resolution wider Field of View image. Large regions can be captured using FPM. Lorenzo Putzu *et al.*[15] used the method of foreground and background separation. Foreground region of the image is a region of interest and else part is background. After background removal; grouped leukocytes were identified and then separated; to avoid error in counting. This separation is performed by measuring maximum concavity. Nucleus and cytoplasm selection were performed by thresholding green component of RGB color space using Otsu threshold.

III. Proposed Framework

Each image acquired contains a large number of erythrocytes and a few leukocytes. The image is segmented between foreground and background using the proposed method. The foregroundis the region of interest i.e. Leukocytes. Leukocytes can be easily identified as they are nucleated cells in comparison with erythrocytes which anucleated and also the size of leukocyte is comparatively larger than an erythrocyte. Therefore; two factors are there to identify leukocytes for choosing aregion of interest (foreground region). After detecting and segmenting Leukocytes various features are extracted; on the basis; these features Leukocytes are classified using Artificial Neural Network (ANN) and then counting of leukocytes differentially and non-differentially is done. The various steps involved for proposed framework are:

1. Image acquisition.
2. Image segmentation.
3. Feature extraction.
4. Classification using ANN & Counting.

A. Image Acquisition

Normal/ abnormal Giemsa-stained blood slides are obtained from amedical laboratory. Further; multiple images per slide are captured using a Coslab light microscope with image capturing module and with are solution of 1280x1024 at 40x magnification power of the objective lens for better resolution. Better resolution is required to differentially detect leukocytes.

B. *Image Segmentation*

Image segmentation is divided into two parts:

1. Cell Segmentation.
2. Nucleus Segmentation.

1) *Cell segmentation*

1. RGB to grayscale.
2. Adaptive histogram equalization.
3. Binary conversion.
4. Hole filling.
5. Image opening.

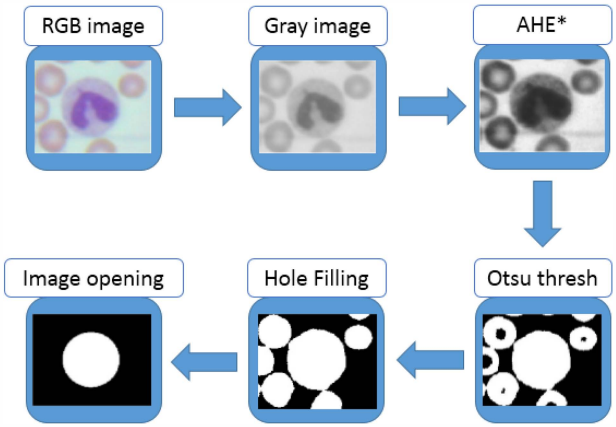

Fig. 2: Cell Segmentation Process

2) *Nucleus segmentation*

1. RGB to HSV color space.
2. Obtaining G segment of an RGB image.
3. Obtaining S segment of HSV image.
4. G-S segment.
5. Image closing.

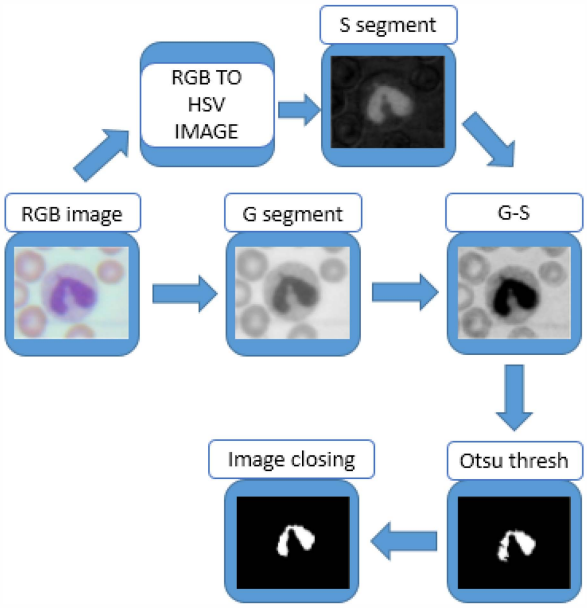

Fig. 3: Nucleus Segmentation Process

C. *Feature Extraction & Cell Identification*

After performing segmentation; morphological features are extracted namely the cell area; nucleus area; nucleus/cell area ratio; perimeter; eccentricity; major axis; minor axis and their ratio; the number of lobules in nucleus etc. Manually we can identify all the five types of leukocytes by using these features. Various other parameters that are based on texture i.e. contrast; correlation; energy; homogeneity etc. are also extracted to further help in classification of leukocytes. Here the Table gives the feature extracted values of 15 samples (Eosinophil; Lymphocyte; Neutrophil each having 5 samples).

TABLE 1: FEATURE EXTRACTED

Sr. No.	Cell Area	Nucleus Area	Nucleus/Cell Area	Contrast	Correlation	Energy	Homogeneity	Leukocyte
1	5022	1745	0.34	0.62	0.91	0.31	0.94	E
2	5121	1689	0.33	0.69	0.90	0.27	0.92	E
3	5693	1736	0.30	0.62	0.91	0.32	0.94	E
4	5054	1547	0.30	0.74	0.92	0.26	0.93	E
5	4980	1539	0.31	0.66	0.9	0.31	0.93	E
6	2449	1368	0.56	0.93	0.73	0.46	0.92	L
7	2418	1355	0.56	1.06	0.78	0.41	0.92	L
8	3106	1808	0.58	0.98	0.8	0.36	0.93	L
9	2493	1441	0.57	0.95	0.78	0.41	0.92	L
10	2605	1497	0.57	0.93	0.78	0.43	0.93	L
11	4672	2124	0.45	0.84	0.90	0.41	0.94	N
12	4604	1724	0.37	0.80	0.91	0.39	0.94	N
13	4222	1282	0.30	0.83	0.90	0.40	0.94	N
14	4221	1289	0.30	0.82	0.90	0.40	0.93	N
15	4365	1697	0.40	0.77	0.90	0.41	0.93	N

*(E = Eosinophil; L=Lymphocyte; N=Neutrophil)

D. *Classification using ANN & Counting*

A biological neural system in neuroscience is a progression of interconnected neurons whose initiation characterizes a conspicuous straight pathway. The neural network is a large network of neurons; in which a neuron interact with other neighboring neurons through an interface that consists of several axon terminals connected to dendrites via synapses. In the event that the aggregate of the input signals into one neuron surpasses a specific value; the neuron sends an Action Potential (AP) at the axon hillock and transmits this electrical signal along the axon.

Artificial neural networks are arange of models inspired by the biological neural network that is used to estimate a function that depends on a large number of inputs. Basically; ANN is inspired from the biological neural network that consists of a large network of interconnected neurons which exchange signals between each other. These connections are weighted according to the past experiences; thus; it is adaptive network and capable of learning.

Like in this problem; classification of leukocytes in three categories i.e. Eosinophil; Neutrophil; Lymphocyte; based on uniformity of morphological features (cell area; nucleus area; etc.) and textural features (energy; homogeneity; etc.). 120 example cases in which 7 items of data i.e. features and the correct classification in three categories (based upon manual identification) are arranged. The approach is to form two classes i.e. input class and target class. The input class is arranged as a set of 120 input vectors in a matrix. Then the target class is arranged as a set of 120 target vectors so that they indicate the classes to which input vectors are assigned. The basic approach of arranging target class is to classify by using static 1 or 0; 1 corresponding to target category and 0 for others.

To solve the purpose of classification Neural Network Pattern Recognition tool (nprtool) has been used. By using this tool a neural network is designed to classify inputs into aset of defined categories. 7 feature values have been given as an input according to which output is classified in three categories i.e. Eosinophil; Neutrophil; Lymphocyte There are three kinds of samples:

- *Training:* During training these samples are given to the network and the network are adjusted accordingly.
- *Validation:* These samples are used to halt training when generalization stops improving
- *Testing:* These samples are used to measure performance parameters of the network when training and validation are done.

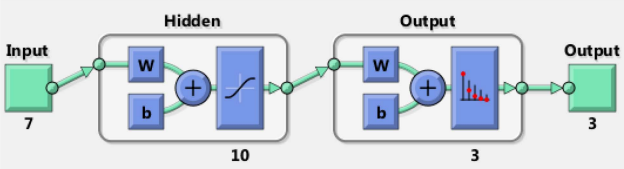

Fig. 4: Specification of ANN

3) Specification of Neural Network

Total no of samples ……90.
Samples for training…….....54(60%).
Samples for validation…….....18(20%).
Samples for testing ……..……18(20%).
Pattern of network ……..Two layer feed forward.
Hidden neurons ……Sigmoid.
Output neurons ……Softmax.
Training……Scaled Conjugate Gradient.
No of hidden neurons……..10.
Input……..……...7 input vectors.
Output Categories……..3(L; N; E).

IV. Experimental Results

The designed system has been evaluated on image set which has been acquired by a digital microscope equipped with a digital camera at 40 x magnification. The resolution of the images is 1280 x 1024. The image dataset consist of 70 microscopic images of ablood smear. There are 90 leukocyte samples that comprises of Lymphocyte; Neutrophils and Eosinophil. Only three types of leukocytes out of five are under the region of interest because these three constitutes of about 97% of leukocytes. The classification process is performed using Neural Network Pattern Recognition tool in MATLAB. Using this method overall accuracy is 98.9%. Differential accuracies are Eosinophil 100%;

Lymphocyte 96.7%; Neutrophil 100%.

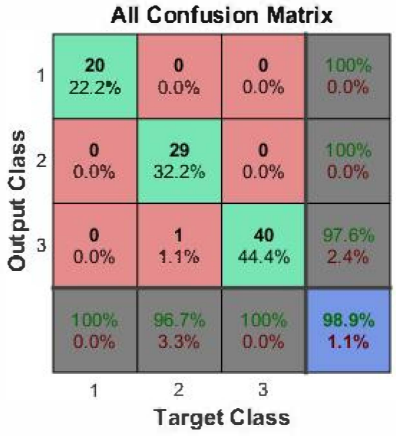

Fig. 5: Confusion Diagram

V. Conclusion

For the analysis of blood cell physically it is extremely time-consuming and tedious to segment the nucleus and on the other side; the instruments which are utilized by specialists for segmentation and classification of blood cells are not economical for every doctor or hospital. Analysis of blood cell in microscopic images of blood smear using digital image processing technique is better in terms of accuracy and efficiency considering time and cost in comparison with present techniques of blood cell investigation. MATLAB is one of the widely used software for this purpose. Nowadays; research work in this field is evolving and in order to get more accuracy in the result; various image processing techniques are implemented.

References

[1] http://www.merriam-webster.com/dictionary/blood

[2] The Franklin Institute Inc. "Blood–The Human Heart". Retrieved 19 March 2009.

[3] LaFleur-Brooks; M. (2008). *Exploring Medical Language: A Student-Directed Approach* (7th ed.). St. Louis; Missouri; US: Mosby Elsevier. p. 398.

[4] "Vital and Health Statistics Series 11; No. 247 (03/2005)

[5] Handin; Robert I.; Samuel E. Lux; Thomas P. Stossel (2003).*Blood: Principles and Practice of Hematology* (2nd ed.). Philadelphia: Lippincott Williams and Wilkins. p. 471. ISBN 9780781719933. Retrieved 2013-06-18.

[6] Ghaznavi F; Evans A; Madabhushi A; Feldman M. Digital imaging in pathology: whole-slide imaging and beyond. Annual Review of Pathology: Mechanisms of Disease. 2013; 8:331–59.

[7] Guclu Ongun; Ugur Halici; Kemal Leblebicioglu; Volkan Atalay "An Automated Differential Blood Count System".

[8] Wang Shitong and Wang Min "A New Detection Algorithm Based on Fuzzy Cellular Neural Networks for White Blood Cell Detection".

[9] Mohammad Hamghalam; Ahmad Ayatollahi "Automatic Counting of Leukocytes in Giemsa-Stained Images of Peripheral Blood Smear.

[10] Jaebum Chung; Xiaoze Ou; Rajan P. Kulkarni2; Changhuei Yang "Counting White Blood Cells from a Blood Smear using Fourier Ptychographic Microscopy".

[11] Der-Chen Huang; Kun-Ding Hung; Yung-Kuan Chan "A computer-assisted method for leukocyte nucleus segmentation and recognition in blood smear images".

[12] Byoung Chul Ko; Ja-Won Gim; Jae-Yeal Nam "Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake".

[13] Ningning Guo; Libo Zeng; QiongshuiWu "A method based on amultispectral imaging technique for White Blood Cell Segmentation".

[14] Anjali Gautam; Harvindra Bhadauria "Classification of White Blood Cells Based on Morphological Features".

[15] Wei-Liang Tai1; Rouh-Mei Hu Han C.W. Hsiao; Rong-Ming Chen and Jeffrey J.P. Tsai "Blood Cell Image Classification Based on Hierarchical SVM".