Determination of Hemoglobin Percentage through K-Means Algorithm and Image Processing for White Blood Cell Diagnosis

[1*]Mrs. Siddamalla.Swarajyam, Dr. M. Narayanan[2], Dr. G. Jawaharlel Nehru[3], Dr. T. Poongothai[4]

[1*]Assistant Professor,Department of Computer Science And Engineering,St.Martin's Engineering College,Dhulapally,Secunderabad-500 100,Telangana, India

[2]Professor & HOD, Department of Computer Science And Engineering,St.Martin's Engineering College,Dhulapally,Secunderabad-500 100,Telangana, India

[3]Associate Professor,Department of Computer Science And Engineering,St.Martin's Engineering College,Dhulapally,Secunderabad-500 100,Telangana, India

[4]Professor & HOD[CSE-AIML], Department of Computer Science And Engineering,St.Martin's EngineeringCollege,Dhulapally,Secunderabad-500 100,Telangana, India

Abstract- White blood cell diagnosis is often carried out by physicians manually, using a microscope to examine blood smears for the presence of white blood cells. Due to the fact that it is a time-consuming, difficult, and prone to mistake procedure, an automated approach using a computerised system is ideal. The most critical steps of this automated procedure are the segmentation and categorization of white blood cells. This work proposes an automated segmentation approach for microscopic white blood cell pictures, with a particular emphasis on images obtained from fresh blood smears. It is recommended that the segmentation be carried out utilising an integration of numerous digital image processing techniques, which is implemented in the segmentation process. Sixty microscopic blood pictures were evaluated, and the suggested approach achieved a cytoplasm segmentation accuracy of 93 percent and a nucleus segmentation accuracy of 89.8 percent, respectively, on the images.

Keywords—white blood cell, blood smear,segmentation, image processing.

I. INTRODUCTION

Clusteringinimage Segmentation is described as the process of recognising groupings of things that are comparable to one another. Clustering methods may be divided into two categories: supervised clustering and unsupervised clustering. Segmentation is the process of dividing a digital picture into different segments depending on the pixels that make up the image in question. It is possible to get an image segmentation result in a collection of segments that will merge to create the whole picture.

Many different clustering techniques have been used for segmenting an image. A human body has three main kinds of cells: red blood cells (RBCs), white blood cells (WBCs), and blood platelets. The automatic identification of these three categories is accomplished by the use of WBC. It is this sort of approach that is used for the automated detection of illnesses. WBC can identify a variety of disorders, including lymphoblastic leukaemia, which is not visible to the naked eye. Some haematologists are required to manually count and categorise the cells in order to diagnose disorders that cannot be identified with high accuracy using a microscope. Normal human blood contains three types of cells: red blood cells, white blood cells (leukocytes), and blood platelets (platelet-

rich plasma). Red blood cells are, on the whole, straightforward and identical. White cells, on the other hand, contain a nucleus and cytoplasm, and there are various types of them. Neutrophils, eosinophils, basophils, monocytes, and lymphocytes are the five types of white cells that may be found in the body. The nucleus and cytoplasm have distinct textures, colours, sizes, and morphologies, which distinguish them from one another. Doctors use a microscope to examine human blood in the conventional method. Because this manual method is time-consuming, tiresome, and prone to mistake, it seems that an automated system would be beneficial and required. It is possible that the automated system will need four stages: acquisition, detection, feature extraction, and classification. Acquisition is the first step. In the second step, cell segmentation is employed to generate a large number of single cell pictures, which are then combined. Then, for each single cell picture, three sections are identified: the nucleus, the cytoplasm, and the background. In the third phase, colour, texture, and form feature vectors of the segmented cell and its nucleus are recovered from the segmented cell and its nucleus. Lastly, a classifier labels each white blood cell according to its extracted characteristics in the last phase of this process. The cell segmentation step is the most significant because the precision of the segmentation plays a critical role in the subsequent phases. Several studies have previously suggested strategies for separating white blood cells into different subsets. Theera-Umpon [1] developed a technique for white blood cell image segmentation that is based on Fuzzy C-Means (FCM) clustering and executed on a white blood cell picture.

II. RELATEDWORKS

According to the author, the procedure is effective and produces better segmentation results than human segmentation would provide. By combining clustering with FCM, Chinwaraphat and colleagues [2] were able to enhance segmentation. In their modification of the conventional FCM clustering, the authors were able to increase the efficiency of the extraction area of the nucleus and the cytoplasm compared to the standard FCM. In addition to the segmentation methods described above, Dorini et al. [3] suggested an approach that is based on morphological procedures for the separation of the nucleus and cytoplasm of white blood cells from the blood picture. The scale-space features of the toggle operator were investigated by the author in order to increase the segmentation accuracy. This approach was able to be applied to a large number of photos and generated excellent results for a range of cell appearances and image quality. In addition, Sadeghian et al. [4] established a framework for performing white blood cell segmentation utilising an integration concept of digital image processing procedures, which they published in 2011. It comprises of two stages: segmentation of the nucleus based on morphological analysis and segmentation for cytoplasm separation based on the intensity of the colours used in each segment. Other color-based segmentation algorithms, such as those presented by P.S. Hiremath and colleagues [5] and F.D. Ratnasari [6,] have been offered. There have also been other studies that have used colour segmentation based on other colour spaces as a starting point. Using the HSI colour space, J. Duan and L. Yu [7] devised a white blood cell segmentation approach that, according to the authors, is more accurate and produces superior performance. In their paper [8, K. Jiang et al. introduced a white blood cell segmentation algorithm that operates well in HSV space, which is more suited than RGB space owing to the minimal correlation between cells.

III. PROPOSED SYSTEM

This study investigated the development of a segmentation approach based on numerous digital image processing techniques. The approach was created to be compatible with the data that was collected from freshblood smears, which were blood pictures. Fresh blood smears are acquired shortly after the preparation of blood smears from human blood. Because there is an obvious difference in the colour and shape of the cells between fresh blood smear and preserved blood smears, it is preferable to utilise fresh blood smears. Furthermore, the fresh blood smear is the kind of blood smear that is utilised in hospitals to diagnose and treat patients. Figure 1 is a sample of a fresh blood smear picture taken from a blood sample.

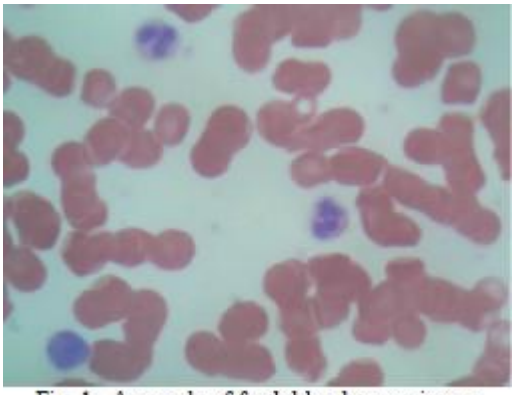

Fig. 1. A sample of fresh blood smear image

One of the most widely used algorithms is the K-Means clustering.In K-Means, data that belongs to one cluster could not be included in anothercluster.Kischosenlargeenoughbutsignificantlysmallerthanthenumberofsamples. TheKfarthestflowsdeterminethefirstpartition. K-Meansisperformed1000iterationsonallsamples as shown in Fig.2.

Eachclusteristhenrepresentedusingasubsetofsamples,oneortwoinouralgorithm. K-Means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem by Angela et al (2002). The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids should be placed in a cunning way because different location causes different results. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set andassociate it to the nearest centroid. When no point is pending, the first step iscompletedandanearlygroupageisdone.Atthispointtheyneedtore-calculate k new centroids as b-arycenters of the clusters resulting from thepreviousstep.Aftertheyhavethesek-newcentroids,anewbindinghastobe done between the same data set points and the nearest new centroid. A loophasbeengenerated.Asaresultofthislooponemaynoticethatthek-centroids change their location step by step until no more changes are done. Inother words centroids do not move any more. Finally, this algorithm aims at minimizing an objective function ,in this case as quared error function.

Algorithm for K-Means

1. Randomlyselectk-pointsastheinitialcentroids.
2. Formk-clustersbyassigningallpointstotheclosestcentroid
3. The centroid of each cluster is recomputed, until the centroidsremainunchanged
4. Distancefunctionisgivenby d(i,j)=sqrt((a(i,j)^2)-(b(i,j)^2)) whered(i,j)isaEuclideandistance,a(i,j)and b(i,j)arethepixelvalues.

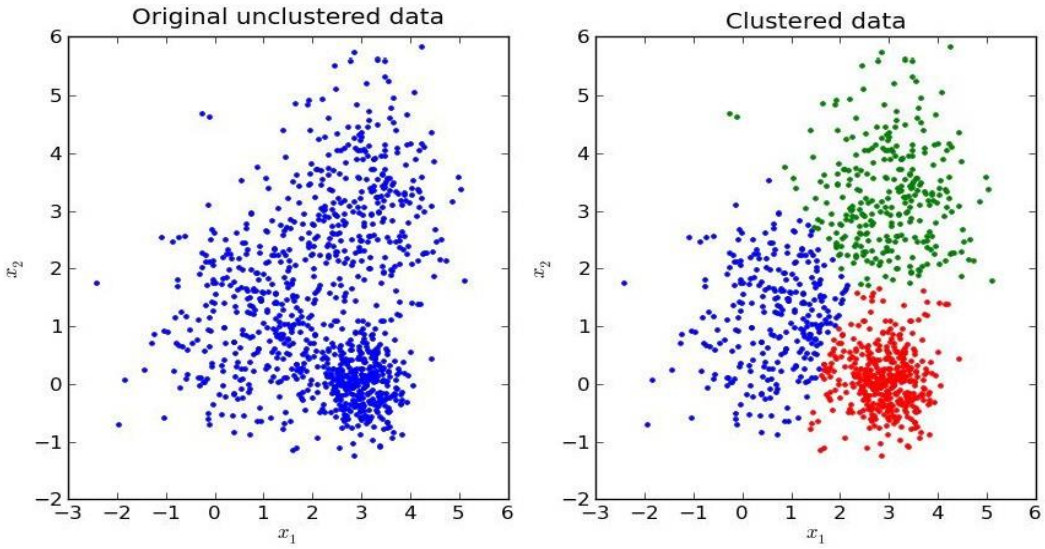

Fig.2 Clustering result from the moving K-Meansalgorithm

IV. RESULTS AND DISCUSSION

The experiments were performed to 60 bloodimages with size of 600x450 pixels producing 76images of white blood cells. The average time ofsegmentation is 1.22 seconds. Some sample imagesof white blood cells and their segmentation results are shown from figures 3 to 7. The results of cell segmentationarea are marked with white lines and the results of thesegmentation nucleus area are marked with yellowlines.

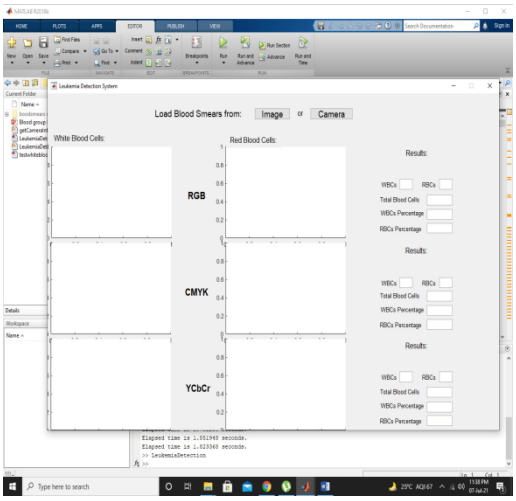

Fig.3 User interface

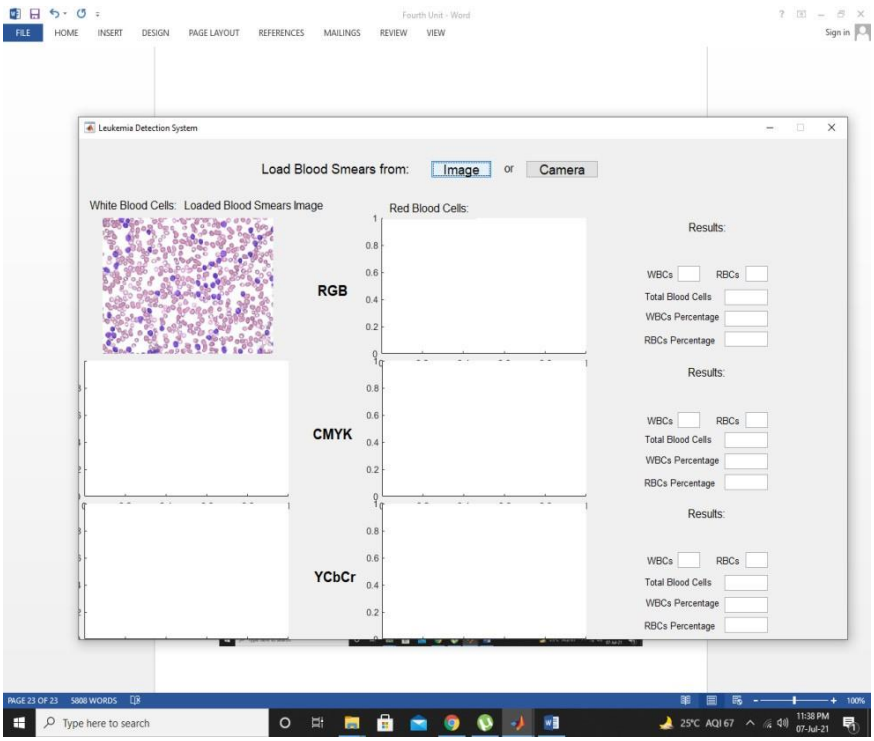

Fig.4 White Blood cells: loaded blood smears image

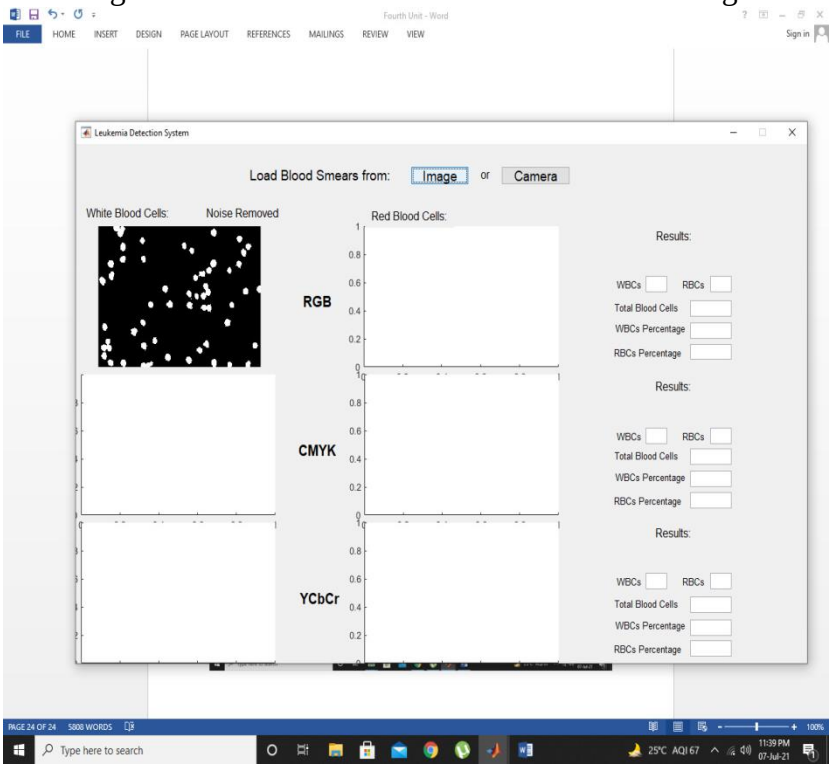

Fig.5 Noise removal

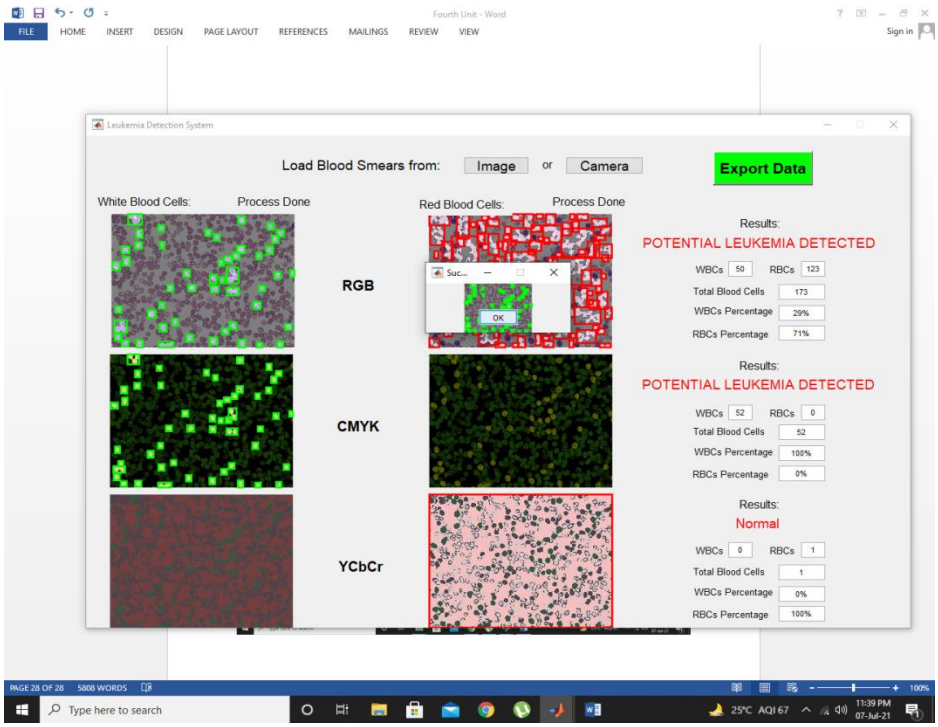

Fig.6 Extracted White blood cells

Fig.7 Enhanced Image

CONCLUSION

Utilizing integration of ideas in digital image processing, this research revealed a potential approach for segmenting white blood cells from fresh blood smear pictures using fresh blood smear images. Sixty tiny blood pictures were investigated in this experiment. In general, the presence of white blood cells included in the blood images may be determined using the suggested approach. The suggested technique achieved 92 percent accuracy for cytoplasm segmentation and 89 percent accuracy for nucleus segmentation when object-based accuracymeasurement was used to determine accuracy. This task will be completed in the next phase of the research.

REFERENCES

[1] N. Theera-Umpon, "Patch-Based White Blood Cell NucleusSegmentation using Fuzzy Clustering", *ECTI Transactions onElectrical Engineering, Electronics, and Communications,*Vol. 3, No. 1, pp. 15-19, 2005.

[2] S. Chinwaraphat, A. Sanpanich, C. Pintavirooj, M.Sangworasil, and P. Tosranon, "A Modified Fuzzy Clusteringfor White Blood Cell Segmentation", *The 3rd InternationalSymposium on Biomedical Engineering*, 2008.

[3] F. Sadeghian, Z. Seman, A.R. Ramli, B.H.A. Kahar, and M.I.Saripan," A Framework for White Blood Cell Segmentation inMicroscopic Blood Images Using Digital Image Processing",*Biological Procedures Online*, Vol. 11, No. 1, 2009.

[4] P.S. Hiremath, P. Bannigidad, and S. Geeta, "AutomatedIdentification and Classification of White Blood Cells(Leukocytes) in Digital Microscopic Images", *IJCA SpecialIssue on "Recent Trends in Image Processing and PatternRecognition"*, 2010.

[5] F.D. Ratnasari, "Identifikasi dan KlasifikasiJenisSel DarahPutihDenganPengolahan Citra Digital", *Thesis*, Departementof Physics, Faculty of Mathematics and Natural Science,Universitas Gadjah Mada, Yogyakarta, 2010.

[6] J. Duan and L. Yu, "A WBC Segmentation Method Based OnHSI Color Space", *Proceedings of IEEE IC-BNMT 2011*.

[7] L.B. Dorini, R. Minetto, and N.J. Leite, "SemiautomaticWhite Blood Cell Segmentation Based on MultiscaleAnalysis", *IEEE Journal of Biomedical and HealthInformatics*, Vol. 17, No. 1, 2013.

[8] K. Jiang, Q.M. Liao, S.Y. Dai, "A Novel White Blood CellSegmentation Scheme Using Scale-Space Filtering andWatershed Clustering", *Proceedings of the SecondInternational Conference on Machine Learning andCybernetics*, pp. 2820-2825, 2003.

[9] R.C. Gonzalez, and R.E. Woods, *Digital Image Processing*,Prentice Hall, New Jersey (2008).

[10] A. Kirillov, AForge.NET Framework, *http://www.aforgenet.com/framework/*, [accessed December 10, 2012].