Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS 2018)
IEEE Xplore Compliant Part Number: CFP18K74-ART; ISBN:978-1-5386-2842-3

Segmentation and Classification of Acute Lymphoblastic Leukemia Cells Tooled with Digital Image Processing and ML Techniques

D. Umamaheswari[1]
Assistant Professor, Department of Computer Science, Vidyasagar College of Arts and Science, Udumalpet, Tamil Nadu, India
{ums1082@gmail.com}

Dr. S. Geetha[2]
Assistant Professor, Department of Computer Science, Government Arts College, Udumalpet, Tamil Nadu, India
{geet_shan@yahoo.com}

Abstract -

Medical science has been contributing its active role in fighting vigorously against the life ceasing diseases. The algorithm is proposed by investigating the existing segmentation algorithms in the field of leukemia research for the sake of supporting hematopathologists to recognize Acute Lymphocytic Leukemia (ALL) by analyzing the blood cell images. There are four levels of separating and classifying benign and malignant white blood cells (WBC). They are, preprocessing, segmentation, extraction of features and classification. As a preliminary task of image analysis, segmentation is done with morphological operators and Otsu's thresholding. Then utilization of nucleus features with supervised KNN classifier gains the classification accuracy of 95.96%, 95.92% of sensitivity and 96% of specificity.

Keywords-Acute Lymphoblastic Leukemia, Morphological Operators, Otsu's Thresholding, Segmentation, KNN Classifier, Blobs, WBC.

I. INTRODUCTION

The disorder leukemia is one amongst the most widespread childhood cancers, but this often affects the older community. This gets instigated in the lymphoid system of the body, consisting of bone marrow, blood, and lymphoid tissues. When DNA of WBCs distorted, sourcing irregularity in WBCs is produced and grow in count quickly, which generally called leukoblasts. This causes deficiency in the cells and stops the usual functioning of healthy blood cells. The reasons could not be made specific, but smoking, long-term constipation, ionizing radiation, certain chemicals, down-syndrome, and genetics are pointed broadly.

The malignant leukemic WBCs are identified by the characters like the abnormality in contour and size of nuclei, the improper organization of cells, high mitotic count, and loss of specialty and functional discrimination. This framework is developed to identify and classify cancerous and healthy blood cells by exploring their features. The algorithm consisting of preprocessing methods, Otsu's thresholding combined with morphological operators, GLCM and KNN classifier as the tools.

II. RELATED WORKS

Majority of research proposals used a fusion of diverse methods in the detection of leukemia through separating the nuclei of WBCs. The fundamental inspiration behind this work is expertise results produced by the morphological operators and Otsu's thresholding in separating the nuclei region of WBC [1]. The work of Raju Bhukya et al., is taken as the base work of this proposal. They used the shape and color information of cells to segment the nucleus region from cytoplasm [1], and then the biggest blob is separated based on the maximum area of contour. In our proposal, the maximum number of pixels within the segmented region is considered, and KNN classifier is used instead of SVM. Though the segmentation has been completed successfully, it was difficult to extract the features of all the WBCs found in the image [2]. Morteza Moradi Amin et al., [3] proposed a work for segmenting the nucleus of WBC using K-Means clustering method and Support Vector Machine as the classifier. Their work contributes the accuracy rate of 95.80%.

Therefore it is necessary to separate a single WBC from others. Since segmentation is a crucial task in further feature extraction step, cytoplasm, nucleus and any other part of blood cells are conscientiously segmented through Gram Schmidt orthogonalization and deformity methods [4], furthermore K-means clustering scheme [5][6].

Luis H. Vogado et al. also applied K-means clustering in the segmentation of nuclei with the Kappa index rate of 0.9306 on ALL-IDB 2 [7]. Ja-Won Gim et al. used GVF snakes and region merging schemes, which is providing a superior outcome in segmenting the nucleus of WBC [8]. In some blood samples, WBCs have impinged upon the others, they must be separated for detailed analysis. This is usually done by morphological erosion [9]. The holes in the separated nucleus are filled by morphological filling operation [10]. Cecilia and Lorenzo have used

histogram plots for implementing the calculations in setting up the threshold values [11]. Arindam Jati et al. also used histogram plot to fix different threshold values in finding the member pixels [12].

In image processing, feature extraction leads to analyze the segmented images for getting a conclusion. Specifically Color, texture, shape and fractal dimensions help in knowing images. Shape features are appropriate to apply with morphological methods of segmentation of the Region of Interest (ROI) [13]. In leukemic WBC the nucleus is massive in size, but knowing this is not adequate for prediction. If the whole image of the blood sample is considered as a case of study, Hausdorff dimension feature [14] suites well in assessing the differential count. Analyzed images are being classified by means of employing Machine Learning (ML) algorithms. Nipon-Theera Umpon and Sompong Dhompongsa achieved 77% of classification accuracy with Artificial Neural Networks and Bayes' classifier [15].

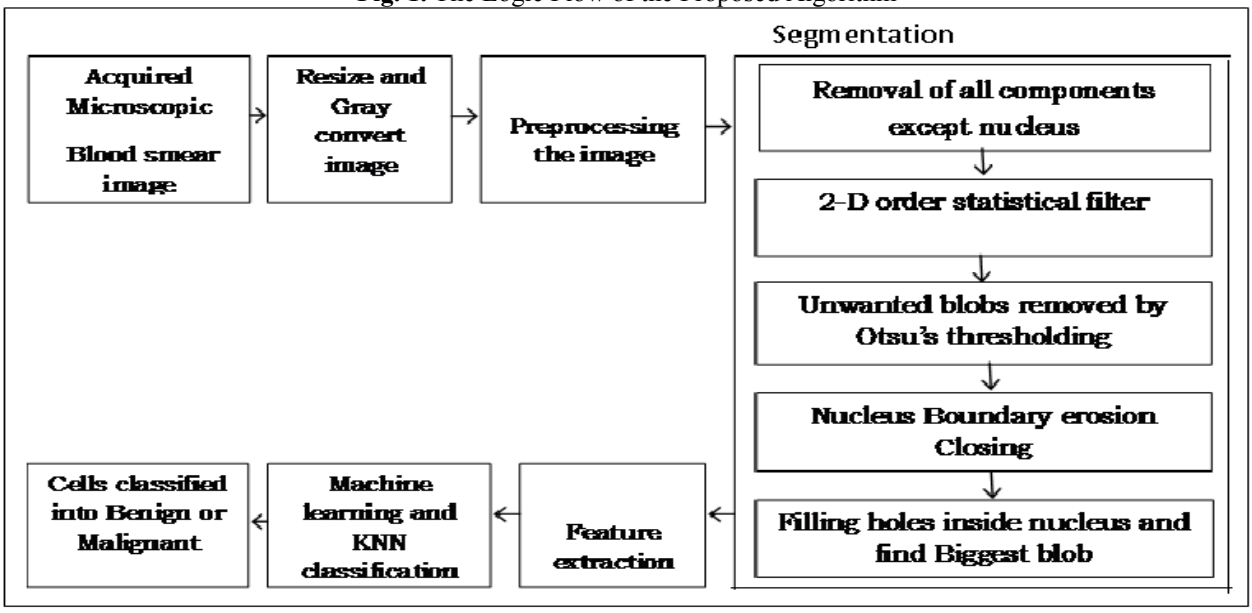
III. OVERVIEW OF PROPOSED METHODOLOGY

Nucleus segmentation is necessary for spotting out leukemia affected WBCs, this has been implemented in various stages shown in Figure 1. In image preprocessing stage the acquired color image of smeared blood cell is resized to make up with 256×256 pixels, this is done to avoid handling images of varying sizes and formats. The resized image is gray converted to decrease the computational complications and give supplementary credence to the Region of Interest (ROI) by conserving the brightness in accordance with the weights of red, green and blue components. The brightness of pixels recognized in terms of intensity, increasing intensity enhances the brightness of the image. The deprived contrast of the WBC region is stretched with histogram equalization through normalization of each pixel by the total value of pixels.

Adding up the brightened and contrast stretched images brightens all part of the blood image except the nucleus region of WBCs, say I1. This resultant image I1 is subtracted from the contrast stretched image, which highlights the borders of all components including the nucleus region, say I2. These two outcomes (I1, I2) are again added to remove all other components and to extract the nucleus region separately, results in I3. Then 2-Dimensional order statistical minimum filter applied to eliminate unnecessary noise. The image pixels more than a particular value are thresholded for binarization using 2D-Otsu's thresholding.

The foremost purpose of using morphological erosion is increasing and introducing the gapping between different blood components in the image. This is achieved with a structuring element H, which shrinks an image by eliminating away a coat of pixels from both inner and outer borders of regions. Whilst subtracting the eroded image from original grayscale image, boundaries of each region can be found. Morphological closing tends to smooth contours of boundaries, eliminate minute holes and filling gaps in contours. Through this, the nucleus is made as a whole single region.

Fig. 1. The Logic Flow of the Proposed Algorithm



The nucleus might have contained unfilled holes inside its region; they are filled by iteratively calling dilation, complementation, and intersections. The target of this process is filling up all the pixels of ROI with 1s, by captivating that all the non-boundary pixels have the value 0. The test image may have more than one WBC inside. Therefore it is obvious to single out the biggest nucleus blob for analysis. Areas of the blobs are calculated for this reason.

Gray Level Cooccurrence Matrix (GLCM) and some other feature extraction techniques are applied in finding out the texture features like homogeneity, contrast, correlation, energy and entropy, the statistical features such as root mean square and standard deviation and texture features for instance perimeter, filled area, circularity, equivalent diameter, kurtosis, skewness, smoothness and centroid. K-Nearest Neighbor supervised learning algorithm finds the 1 closest neighbor. The algorithm is trained with the observations of blood images. There is no industry standard in the ratio between training dataset and test dataset and it could be at the rate of 70:30 or 80:20.

Steps in algorithm

1. Acquire the microscopic blood smear image.
2. Resize the acquired RGB image into 256 × 256 matrix and Convert into the grayscale image.
3. Enrich the intensity of grayscale image by Linear Contrast Enhancement.
4. Enhance the contrast of the image using histogram equalization.
5. Brighten the image except for nucleus. Then extract the borders of all components.
6. Filter the image using 2-D order statistic minimum filtering.
7. Binarize the filtered image with Otsu's global thresholding and Complement to highlight ROI.
8. Erode boundary of the selected region as
   $I_{(Eroded\_Image)} = I \ominus H = (\bar{I} \oplus H^{*})'$.
9. Close the selected region and fill tiny holes in the region by closing and filling operations
   $I_{(Closed\_Image)} = I \bullet H = (I \oplus H) \ominus H$
10. Label the objects of the image to find the biggest blob.
11. Features extracted and Images classified using K-Nearest neighbor algorithm

IV. EXPERIMENTAL RESULTS

The proposed algorithm is implemented with Matlab 2015a and tested by inputting images from the ALL-IDB 2 benchmark dataset of 260 microscopic smeared blood cell images. Out of which 130 are malignant and 130 are benign images. Fig. 2 illustrates the outcome of the algorithm in diverse stages with non-cancerous blood cell image.

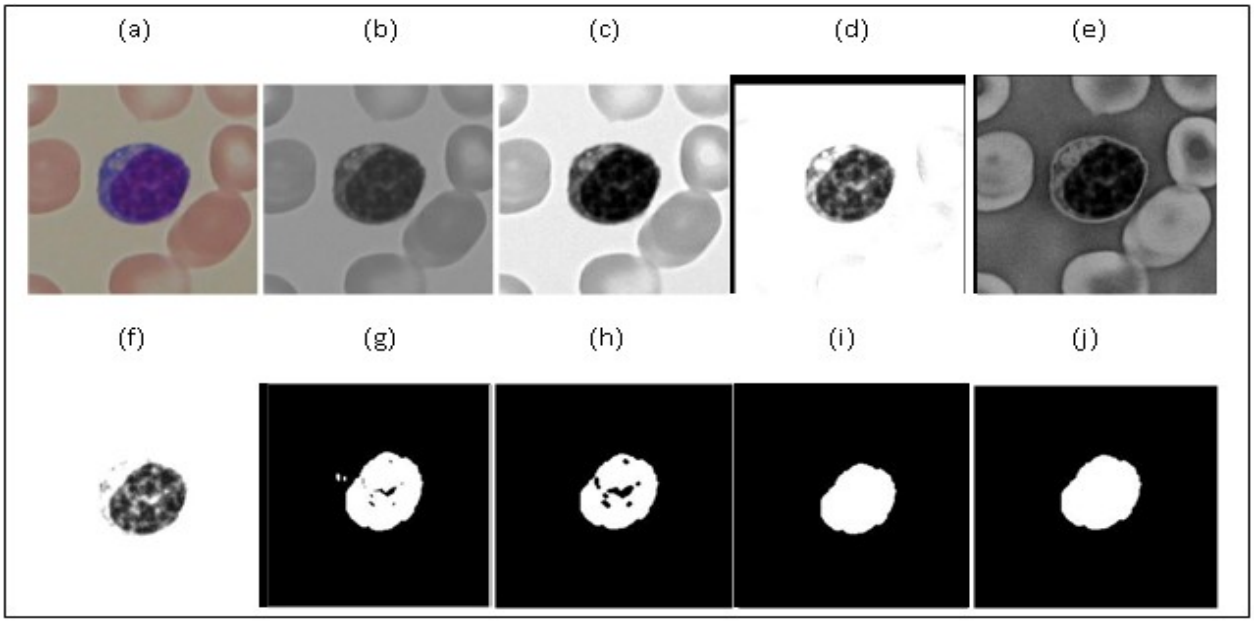


Fig. 2. Results of Proposed Algorithm with Benign Image: (a) Original Microscopic Blood cell Image, (b) Gray Scale Image, (c) Intensity Adjusted with LCE, (d) Brightened Image, (e) Borders of all components Highlighted, (f) All Blood components removed except Nucleus, (g) Thresholded with Otsu's, (h) Nucleus Eroded Image and closed, (i) Holes filled-in Nucleus Region, (j) The Biggest Blob is Found

Fig. 3 illustrates the same with cancerous blood cell image. This also points out the extraction of the biggest blob in case the image contains more than one WBC in it.

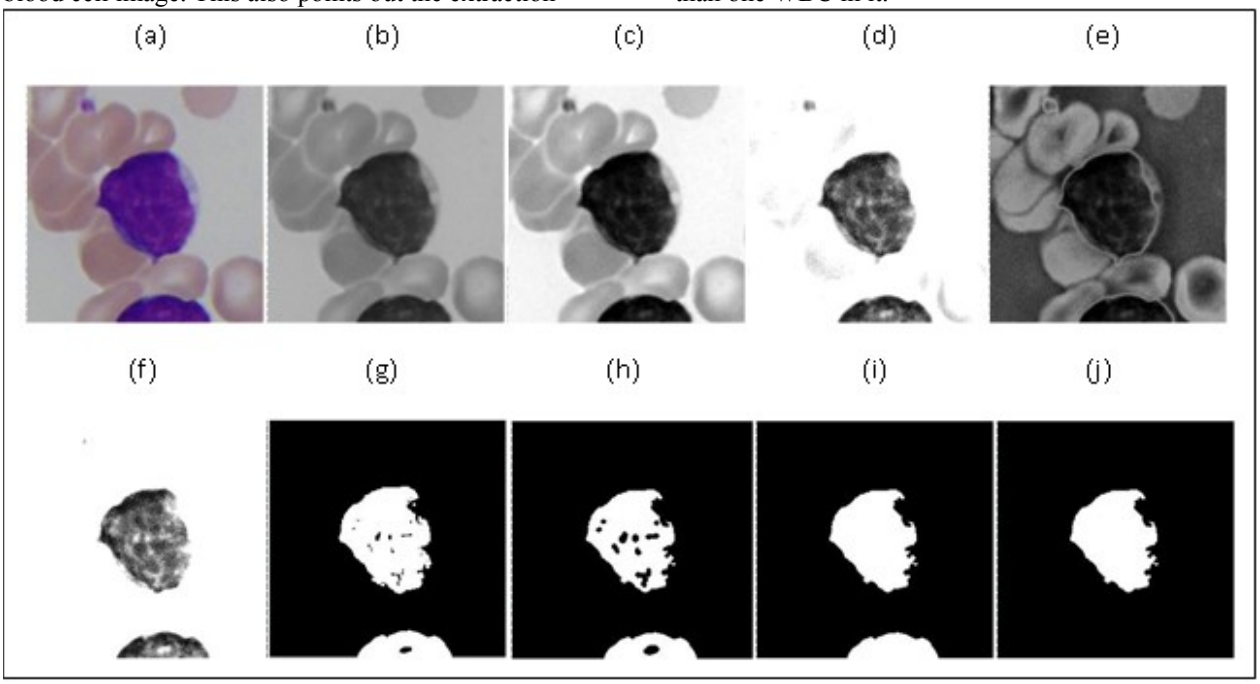


Fig. 3. Results of Proposed Algorithm with Malignant Image: (a) Original Microscopic Blood cell Image, (b) Gray Scale Image, (c) Intensity Adjusted with LCE, (d) Brightened Image, (e) Borders of all components Highlighted, (f) All Blood components removed except Nucleus, (g) Thresholded with Otsu's, (h) Nucleus Eroded Image and closed, (i) Holes filled-in Nucleus Region, (j) The Biggest Blob is Found

Though the image may contain more than WBCs in it, the proposed segmentation algorithm better segments a single nucleus, this is shown in Fig. 2(j) and Fig. 3(j). Subsequent to the successful segmentation and separation of the biggest nucleus region, 15 features are extracted and which are used by the KNN classifier in classifying the benign and malignant images. Performance of the KNN classifier along with the proposed segmentation algorithm is evaluated by confusion matrix consisting of True Positives, False Negatives, True Negatives, and False Positives. From this, the performance measures such as Accuracy, Sensitivity, Specificity, Error rate, false positive rate, and false negative rate are drawn and shown in Table 1.

Table 1. Evaluation of the KNN classifier using Confusion matrix

Measure	Accuracy	Sensitivity	Specificity	Error Rate	FPR	FNR
Result in %	95.96	95.92	96.00	4.04	4.00	4.08

The proposed algorithm gives 2.63% of the increase in classifying the nucleus of WBC when compared to the same work carried out by Raju Bhukya et al., [1] with the same dataset ALL-IDB2. Their work gives 93.33% of accuracy with the application of Support Vector Machines as the classifier. The following graph shows the enhancement in the accuracy rate of our work.

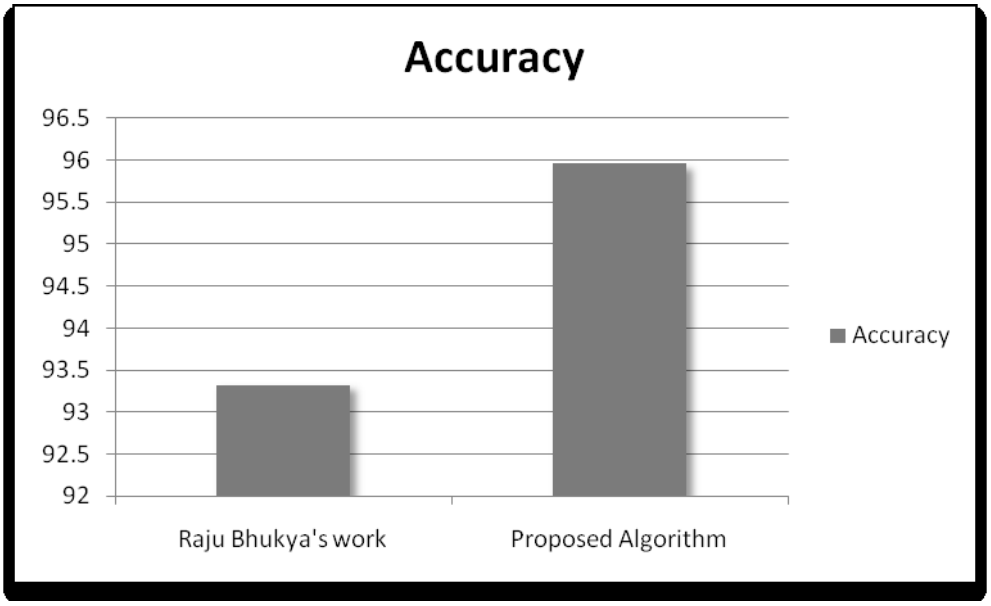


Fig. 4. The Efficiency of the Proposed Algorithm

V. CONCLUSION

The proposed algorithm proficiently segments the largest nucleus blob in the obtained blood cell image for the cause of supplementing hematologists in detecting leukemia. The KNN classifier classifies the given test images into benign or malignant with the accuracy of 95.96%, the sensitivity of 95.92%, and 96% of specificity. This is much superior to the algorithms developed for segmenting and classifying ALL cells in these days. The algorithm gets matured further with different sets of training data since the reduction in training reduces efficiency. In future work, the algorithm can be enriched to segment and classify the four types of leukemia. The ensemble of several machine learning algorithms can be brought together to get up-gradation in the accuracy.

REFERENCES

[1] Raju Bhukya, B.Prasanth, V. Sasank Vihari, Y.Ajay, "Detection of acute lymphoblastic using microscopic images of blood", International Journal of Advanced and Applied Sciences, 4(8), pp 74-78, 2017.

[2] Sonali C. Sonar, K.S. Bhagat, "An Efficient Technique for White Blood Cells Nuclei Automatic Segmentation", in International Journal of Scientific & Engineering Research, 2015, Volume 6(5).

[3] Morteza Moradi Amin, Saeed Kermani, Ardeshir Talebi, and Mostafa Ghelich Oghli, "Recognition of Acute Lymphoblastic Leukemia cells in Microscopic Images Using K-means Clustering and Support Vector Machines", Journal of Medical Signals and Sensors, 5(1): 49- 58, 2015.

[4] S.H. Rezatofighi, R.A. Zoroofi, R. Sharifian, H. Soltanian-Zadeh, "Segmentation of Nucleus and Cytoplasm of White Blood cells using Gram-Schmid Orthogonalization of Deformable Models", in 9th International Conference on Signal Processing, IEEE Explore, 2008.

[5] Jakkrich Laosai, Kosin Chamnongthai, "Acute Leukemia Classification by Using SVM and K-means clustering", in Proceedings of the International Electrical Engineering Congress, IEEE, 2014.

[6] Rajesh Kumar, Rajeev Srivastava, Subodh Srivastava, "Detection and Classification of cancer from microscopic Biopsy Images using Clinically significant and Biologically Interpretable Features", in Journal of Medical Engineering, Article ID 457906, Hindawi Publishing Corporation, 2015, Volume 2015.

[7] Luis H. S. Vogado, Rodrigo de M. S. Veras, Alan R. Andrade, Romuere R. V. e Silva, Flavio H. D. de Araujo, Fatima N. S. de Medeiros, "Unsupervised Leukemia cells segmentation based on multi-space color channels", in IEEE International Symposium on Multimedia, 2016.

[8] Ja-Won Gim, Junoh Park, Ji-Hyeon-Lee, Byoung Chul Ko, Jae-Yeal Nam, "A Novel Framework for White Blood cell Segmentation Based on Stepwise rules and morphological Features", in Proceedings of SPIE- IS&T 7877, Image Processing Machine Vision Applications, San Francisco, 2011, pp 1-6.

[9] Hayan T. Madhloom, Sameem Abdul Kareem, Hany Ariffin, "Computer-Aided Acute Leukemia Blast Cells Segmentation in Peripheral Blood Images", in Journal of Vibroengineering, 2015, 17(8), pp 4517-4532.

[10] Farnoosh Sadeghian, Zainina Seman, Abdul Rahman Badrul Hisham Abdul Kahar, M-Iqbal Saripan, "A Framework for White Blood Cell Segmentation in Microscopic Blood Images Using Digital Image Processing", in Biological Procedures Online, Springer, pp 196 – 206, New York, 2009.

[11] Cecilia Di Ruberto, Lorenzo Putzu, "Accurate blood cells segmentation", in 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems, 2014.

[12] Arindam Jati, Garima Singh, Rashmi Mukherjee, Madhumala Gosh, Amit Konar, Chandan Chakraborty, Atulya K. Nagar, "Automatic Leukocyte nucleus

Segmentation by intuitionistic Fuzzy Divergence Based thresholding", in Micron 58, 55-65, Elsevier Ltd 2014.

[13] Vincenzo Piuri, Fabio Scotti, "Morphological Classification of Blood Leukocytes by Microscopic Images", in IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2004.

[14] Sos Agaian, Monica Madhukumar, Anthony T. Chronopoulos, "Automated Screening System for Acute Myelogenous Leukemia Detection in Blood Microscopic Images", in IEEE Systems Journal, 8(3), 2014.

[15] Nipon Theera-Umpon, Sompong Dhompongsa, "Morphological Granulometric Features of Nucleus in Automatic bone Marrow White Blood Cell Classification", in IEEE transactions on Information Technology in Biomedicine, 11(3), pp 353–359, 2007.

AUTHORS BIOGRAPHY

D. Umamaheswari has completed her M.Sc., and M.Phil degrees in Computer Science and published four papers in international journals. She has qualified in State-level Eligibility Test (SLET). At present working in Vidyasagar College of Arts and Science, Udumalepet, Tamilnadu. She got 12 years of tetaching experience in various colleges. Currently pursuing Ph.D. (Part-time) in Computer Science, Government Arts College, Udumalpet, Tamilnadu, India. Her Area of interest is image processing.

Dr. S. Geetha has completed her M.Sc., M.Phil, Ph.D. degrees in Computer Science. She has published various papers in international journals of repute. Her area of specialization is soft computing. She has got 20 years of experience in PSG College of Technology, Coimbatore, Tamilnadu. Presently working at Government Arts College, Udumalpet, Tamilnadu,India.