



Fast and Robust Method for Nuclei segmentation in Cytological Images of Breast FNAC Samples

Mr. Aditya Pise¹, Mr. Rushi Longadge²

¹Department of Computer Science and Engineering, G. H. Raison Academy of Engineering and Technology, Nagpur, India
adityapise1@gmail.com

²Department of Computer Science and Engineering, G. H. Raison Academy of Engineering and Technology, Nagpur, India
rushi.longadge@gmail.com

Abstract: This Paper presents two automatic segmentation methods to locate nuclei in cytological images. The analysis includes the watershed algorithm, active contouring. One can also find here a description of image enhancement, Hough transform for circle detection and automatic nuclei localization mechanism used in this approach. Preliminary experimental results collected present the quality of the methods in the analysed issue. Here we mentioned the common errors and possible future problems and points out regions that need further research.

Keywords: Hough Transform; Active contour; Watershed Algorithm

I. INTRODUCTION

In the last few years we have been observing a self-motivated growth in the number of research works conducted in the area of breast cancer diagnosis. Many institutes and industries [4] are working on this issue because of the fact that breast cancer is becoming the most common form of cancer disease of today's female population. The attention covers not only curing the external effects of the disease [24] [3] but also its fast detection in an early phase. Thus, the construction of a fully automatic cancer diagnosis system supporting a human expert has become a challenging task. These days many camera-based automatic breast cancer analysis systems have to face the problem of cells and their nuclei separation from the rest of the image content [11].

This process is very important because the nucleus of the cell is the position where breast cancer malignancy can be seen. Thus, much attention in the building the expert supporting diagnosis system has to be paid to the segmentation stage. The main obscurity of the segmentation process is due to the incompleteness and ambiguity of the information contained in the image. The flaw of the data acquisition process in the form of noise, chromatic distortion and deformity of cytological material caused by its preparation moreover increases the problem complexity. The nature of image acquisition and the method of scene illumination also affect the image luminance and sharpness. Many segmentation methods have been proposed [2] but, unluckily, each of them introduces several additional problems and usually works in practice under given assumptions and/or

needs the end-user's interaction/co-operation [11]. Since nowadays many cytological projects assume full automation and real-time operation with a high degree of effectiveness, a method free of drawbacks of the already known approaches has to be constructed. In this paper a group of modified versions of cytological image segmentation methods adopted for fine needle biopsy images are presented, that is, the watershed algorithm, active contours. One can also find here a description of de-noising and contrast enhancement techniques, pre-segmentation and fully automatic nuclei localization mechanism used in our approach.

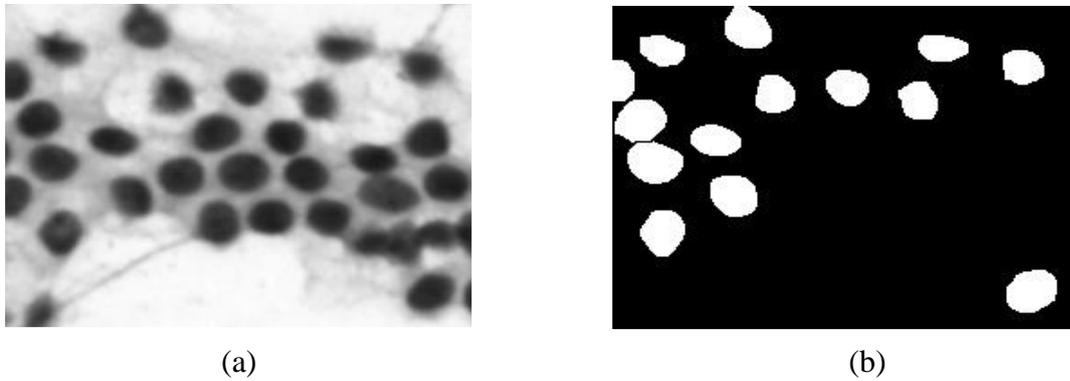


Fig. 1 Exemplary fragments: (a) cytological image, (b) appropriate segmentation mask.

II. PROBLEM FORMULATION

The formulation of the segmentation process is very hard because it is a poorly accustomed problem, and in many situations the segmentation process is domain specific. What we have on input is cytological images obtained using the Fine Needle Aspiration Cytology (FNAC) technique and imaged with camera mounted on top of an microscope. The image itself is coded using the RGB colorspace and is not subject to any kind of enhancement. The size of the image equals 1024×768 pixels. What we expect on output is a binary segmentation mask which will permit more robust morph metric parameters estimation in our future work. Additionally, the proposed segmentation algorithm should be insensitive to colors of contrasting pigments used for the preparation of the cytological material. Original image and its appropriate segmentation mask shown in Fig1.

III. IMAGE PREPARATION AND ENHANCEMENT

The raw image taken from camera does not carry important information, so it can be converted in to grayscale image to reduce processing complexity in stages that require only. An RGB color image can be converted to greyscale by removing blue and red chrominance components from the image defined in the YCbCr Explanation "Derived from" ITU-R BT.601". The luminosity component can be determined using:

$$Y = 0.299R + 0.587G + 0.114B.$$

Since the majority of images we deal with have low contrast, an enhancement technique is needed to improve their quality. In our approach we use simple histogram processing with a linear transform of the image levels of intensities [2].

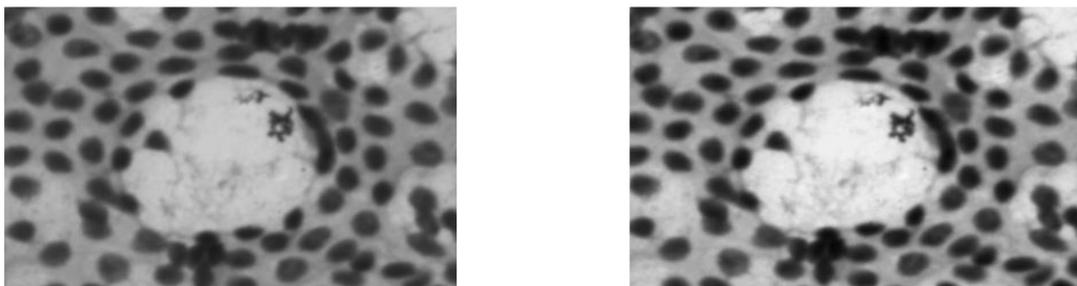


Fig.2 (a) raw image

(b) Contrast Enhancement

The contrast correction is conducted for each color channel separately resulting in an image being better defined for later stages of the presented hybrid segmentation methods. Fig.2a,b show the contrast enhancement due to histogram stretching.

IV. CIRCLE DETECTION BY HOUGH TRANSFORM

The Hough transform (Ballard, 1981; Duda and Hart, 1972; Zorski, 2000) can be easily adopted for the purpose of circle detection. Circular Hough Transform (CHT) [8] is used to detect circles. The edge detected from the canny edge detector forms the input to extract the circle using the Circular Hough Transform. In Circular Hough Transform, voting procedure is carried out in a parameter space. The local maxima in accumulator space obtained by voting procedure are used to compute the Hough Transform. Parameter space is defined by the parametric representation used to describe circles in the picture plane, which is given by equation (3).

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \tag{3}$$

An accumulator is an array used to detect the existence of the circle in the Circular Hough Transform. Dimension of the accumulator is equal to the unknown parameters of the circle. The equation of the circle in parametric form is given by Eq.3 implies that the accumulator space is three-dimensional (for three unknown parameters x_0 , y_0 and r).

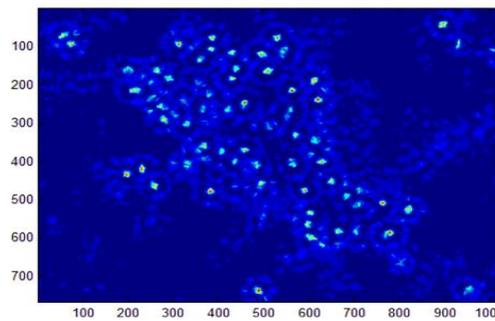


Fig.3 Accumulator (Hough Transform)

Fig.3 Accumulator (Hough Transform)

This equation (3) defines the locus of points (x, y) centered on an origin (x_0, y_0) with radius r . Each edge point in figure 3a defines a set of circles in the accumulator space [8]. These circles are defined by all possible values of the radius and they are centered on the coordinates of the edge point. Figure 3b shows three circles defined by three edge points labeled 1, 2 and 3.

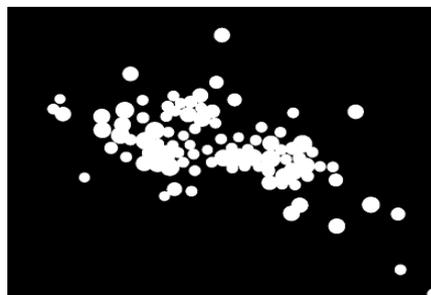


Fig.4 Selected circle by Hough transform after applying threshold on accumulator.

These circles are defined for a given radius value. Each edge point defines circle for the other value of the radius. These edge points map to a cone of votes in the accumulator space (g), which is also called as feature space.

V. THRESHOLD SELECTION

Thresholding the values in the accumulator by a given θ value we can obtain a very good presegmentation mechanism with a lower threshold strategy (see, e.g., Fig. 5). Since the threshold value strongly depends on the database and the feature image, the method can only be used as a presegmentation stage. A smaller value of the threshold causes fast removal of unimportant information from the background, and what we achieve is a mask which approximately defines the places where the objects we have to segment (nuclei in this case) are located and where the background is. Such a mask can constitute a basis for more sophisticated and detail-oriented algorithms.

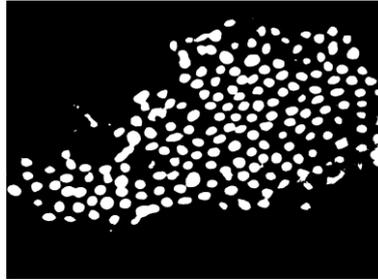


Fig.5. Nuclei probable image.

The results obtained at the presegmentation stage can lead to the estimation of an average background color. Such information can be used to model the nuclei as a color distance between the background and the objects, which fulfills the requirements of the lack of any color dependency in the imaged material. In our research we tried several distance metrics: Manhattan's, Chebyshev's, the absolute Hue value from the HSV colorspace, but the Euclidean one gives us visually the best results where B is the average background color estimated for the input image I .

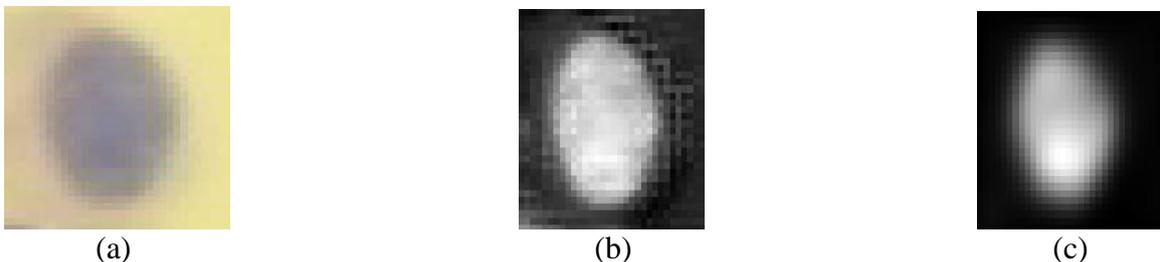


Fig. 6 Exemplary fragments: (a) cytological image, (b) Euclidean distance to the mean background color, (c) smoothed version of (b).

VI. PRE-SEGMENTATION METHODS

Though we have nuclei probable image but still it contains some non nuclei part. Following we introduce methods for removal of such objects.

VI.1 Removal of Small object

Object which are too small to be nuclei, for such object we are using well know method 'Opening' of the image; with this method we remove almost all the small object. Here we select 120 pixel areas as the threshold value and reject all the object which are smaller than selected threshold. We are clearly seen in fig..... The output image is much better than previous one.(See fig.7a,b)

VI.2 Removal of Connected component

Because of the overlapping of the nuclei or cytoplasm some time the large spots can appear on image which get segmented because it has the same intensity range that of nuclei hence to remove this type of large item we calculate the average size of the nuclei and then we remove the object larger than that. Because of these methods we not only enhance the images but also we decrease the process complexity (see fig.7 c,d).

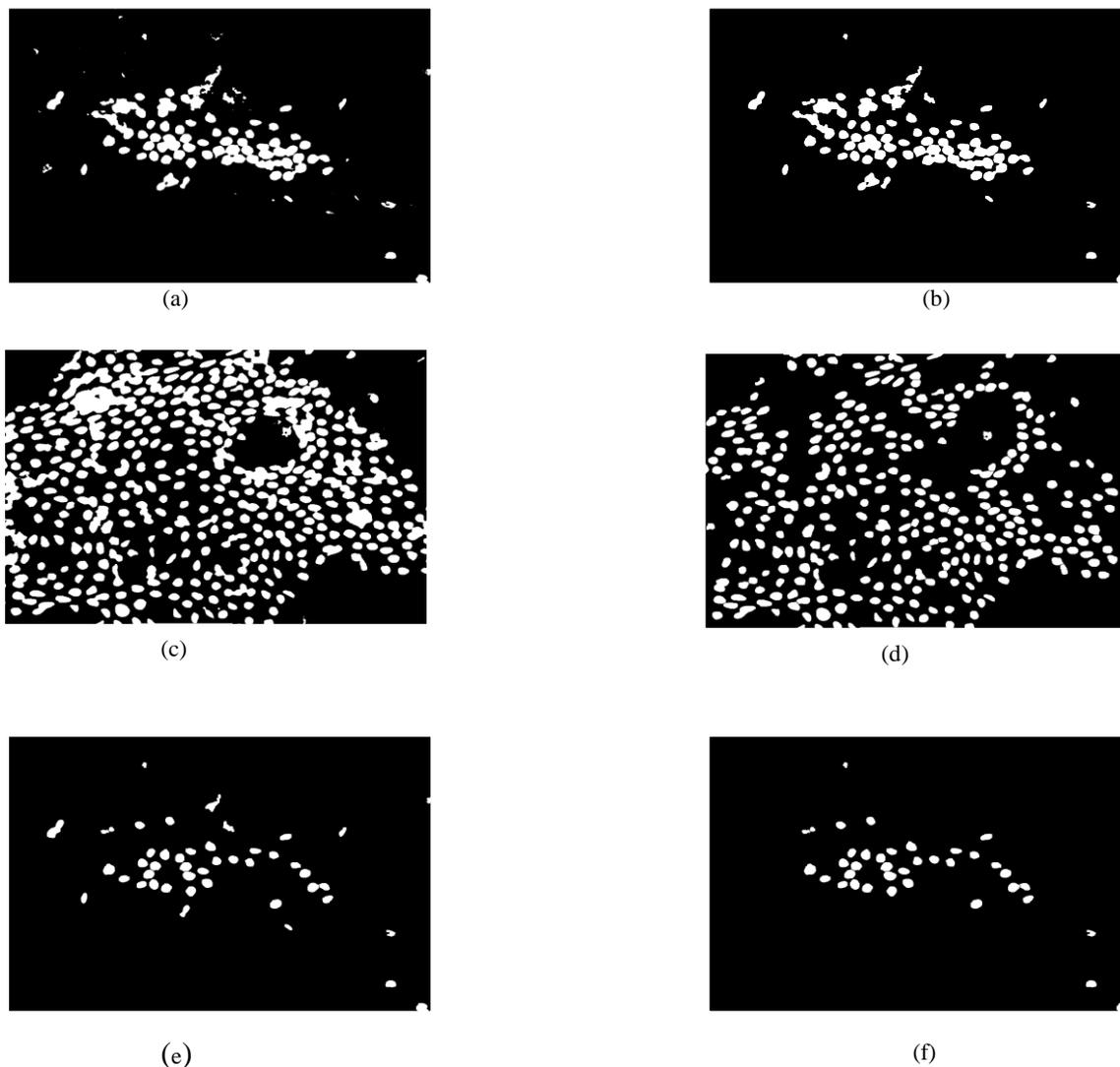


Fig.7 a, b, c, d, e, f (Resulting images of removal of irrelevant images)

The fig. b is resulting image of removal of small element, fig. d is resulting images of removal of large (connected) object and fig. e is the image after removal of the eccentric objects.

VII.SEGMENTATION

Now we ready to apply segmentation algorithm on images which are output of the previous methods. We are Applying watershed and active contour and showing their output and error.

VII.1 Watersheds

The watershed segmentation algorithm is inspired by natural observations, such as a rainy day in the mountains [4]. A given image can be defined as a terrain on which nuclei correspond to valleys .The terrain is flooded by rainwater and arising puddles start to turn into basins. When the water from one basin begins to pour away to another, a separating watershed is created. The flooding operation has to be stopped when the water level reaches a given threshold θ [3]. The threshold should preferably be placed somewhere in the middle between the background and a nucleus localization point. In our approach the nuclei are flooded to the half of the altitude between the nucleus localization point and the average height of the background in the local

neighbourhood. Since the images we have to deal with are spot illuminated during the imaging operation, this mechanism protects basins from being over flooded and, in consequence, nuclei from being under segmented [3].

VII.2 Active contours

The active contouring technique [3] can be considered as a more advanced region growing method [4]. The algorithm groups neighbouring pixels when a given homogeneity and similarity criteria are met. All joined pixels create a segment whose boundary spreads in all directions until another segment is met or new candidates for joining introduce unacceptable errors. The algorithm is stopped when all pixels get labels, i.e., the object in the image is separated from the background. The images we deal with may contain more than a single object per image. Additionally, the assumption of the project is that the segmentation process has to be fully automatic (there is no human operator which manually initializes the method). These two factors force us to modify the algorithm to meet the stated requirements. Thus, the algorithm, which in our case is based on the fast marching method (FMM) [25] must have a multilabel extension (Steć and Domański, 2005) and the seeding process has to be done without the end-user's interaction.[3]

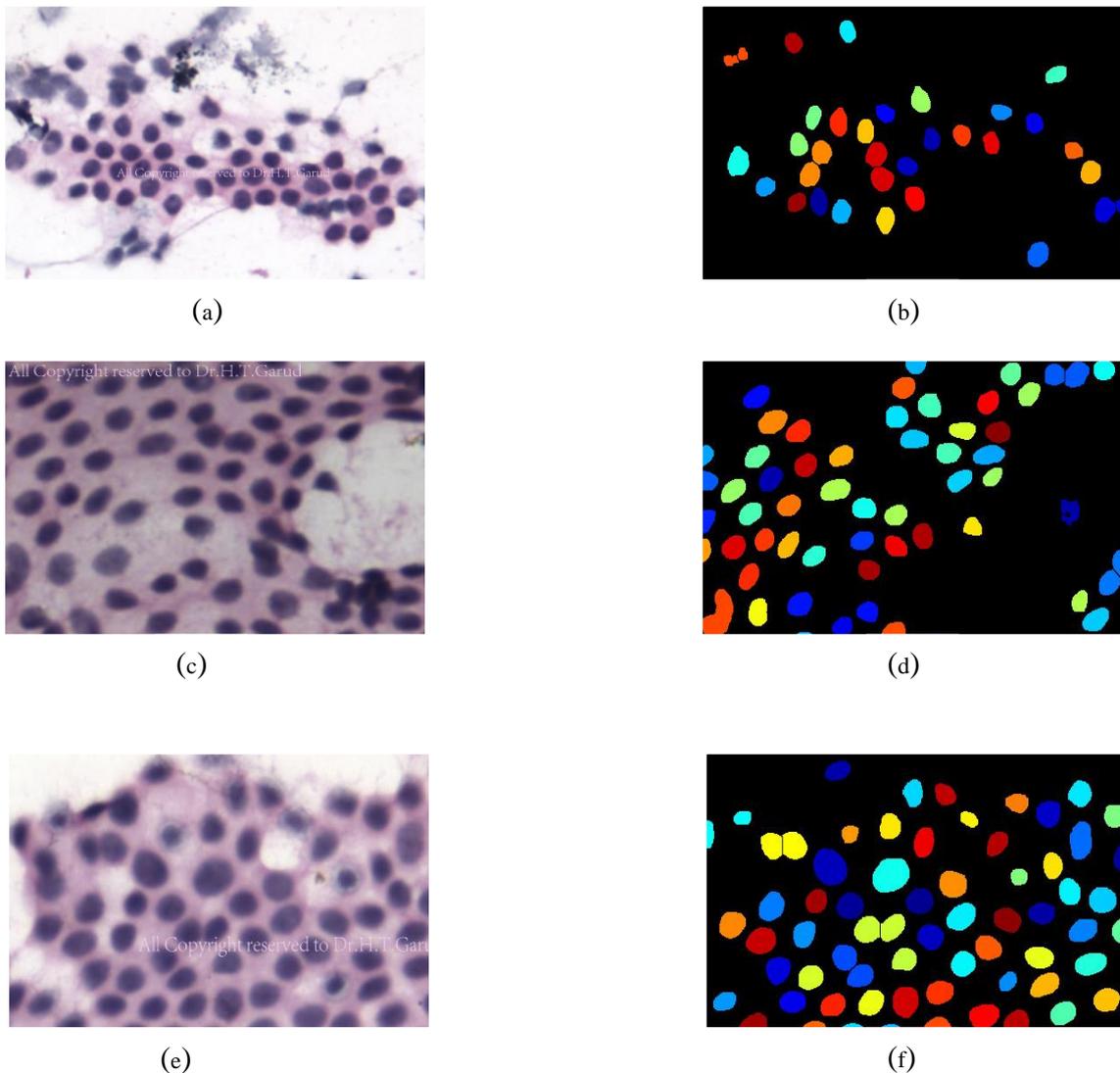


Fig8. Exemplary results of the watershed segmentation. [Images from Database of Dr. H. T. Garud]



Fig.9 Exemplary results of the Active contour segmentation.

VIII. EXPERIMENTAL RESULTS

VIII.1 WATERSHEDS

The conducted experiments show that, on average, the watershed algorithm gives a 80.74% agreement with the hand-prepared templates using a simple XOR metric. Most errors are located at boundaries (see, e.g., Fig. 9(a)) of nuclei where the average distance between the edges of segmented and reference objects is about 3 pixels. The XOR metric is underestimated as a consequence of a moderate level of water flooding the modelled terrain, but the shape of the nuclei seems to be preserved, which is important in our future work—the estimation of morph metric parameters of segmented nuclei.



Fig.10 Exemplary XOR results with a fragment of the Hand-prepared segmentation mask for: (a) watershed algorithm, (b) active contouring technique.

VIII.2 ACTIVE CONTOURS

The conducted experiments show that the modified multilabel FMM algorithm is very stable and robust to initialization errors. Visually, segmentation quality is promising and yields a good detection of even small objects. Unfortunately, the algorithm has problems with connected nuclei and detects them as a single object, which is erroneous. Despite the above-mentioned problems, the shape of segmented nuclei seems to be represented accurately, and most errors are located at the boundaries of the segmented objects (see, e.g., Fig. 10(b)). This illustrates that a proper selection of the merging threshold and the detection of overlapping nuclei is still a challenge and has to be improved in future works and it gives 85% accuracy.

IX. CONCLUSION

The problem concern to fake circles created by spots of fat and unnecessary effects it gives as the final output should be considered and eliminated in future work. Images with a mixed nucleus type still constitute a challenge because it seems to be impossible to detect only one type without the end-user's interaction and when there should not be any dependencies and assumptions concerning the color of contrasting pigments used to prepare cytological material. The presented solutions should also be extended to perform better on poor quality images, or a fast classifier should be constructed to reject too poor.

All of the above-mentioned drawbacks of the described algorithms should also be considered and faced in the future. The described segmentation methods needs about 2-3 minutes on today's machines for an image to give a final response. All the simulations were performed in the Matlab environment so the authors believe that the time consumption can be significantly reduced. Computationally expensive stages and the nature of the algorithms used force and allow us at the same time to implement them using a high level of parallelism to perform more effectively on today's core 2 duo machines and thread oriented operating systems. One can also consider a dedicated hardware which can perform even more effectively then the appropriate software implementation. Give a good base for our further research in the area of cytological image segmentation. Additionally, all preparation steps including pre-segmentation and the automatic nuclei localization stage can be reused with other segmentation algorithms which need such information.

Summarizing the presented solutions, promise and give a good base for further research in the area of cytological image segmentation. Additionally, all preparation steps including pre-segmentation and the automatic nuclei localization. Stage can be reused with other segmentation algorithms which need such information .In this approach each and every process could be done automatically hence our ultimate aim can be achieved.

ACKNOWLEDGMENT

We would like to thank the following people for their help Dr. Hrushikesh T. Garud.

REFERENCES

- [1] Marek Brejl and Milan Sonka, "Object Localization and Border Detection Criteria Design in Edge-Based Image Segmentation: Automated Learning from Examples" *Ieee Transactions On Medical Imaging*, Vol. 19, No. 10, October 2000.
- [2] C.Ortiz De Solo', Rzano,E., Garcí'A Rodriguez,A. Jones,D. Pinkel, J. W. Gray, D. Sudar & S. J. Lockett, "Segmentation of confocal microscope images of cell nuclei in thick tissue sections" *Journal of Microscopy*, Vol. 193, Pt 3, March 1999, pp. 212–226.
- [3] Maciej Hrebien', Piotr Stec', Tomasz Nieczkowski, Andrzej Obuchowicz, "Segmentation Of Breast Cancer Fine Needle Biopsy Cytological Images" *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 2, 159–170.
- [4] Theodoros Mouroutis, Stephen J.Roberts, Anil A.Bharath, "Robust Cell Nuclei Segmentation Using Statistical Modelling" *Bioimaging*,1998,citeseer.Pg.No 79—91,vol. 6.
- [5] Łukasz Jelen', Thomas Fevens, Adam Krzyż' Ak, "Classification Of Breast Cancer Malignancy Using Cytological Images Of Fine Needle Aspiration Biopsies" *Int. J. Appl. Math. Comput. Sci*, 2008, ol.18,No.1,75–83.
- [6] Marcin Smereka, Ignacy Dule, Ba, "Circular Object Detection Using A Modified Hough Transform" *Int. J. Appl. Math. Comput. Sci.*, 2008, Vol. 18, No. 1, 85–91.
- [7] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D. Fox, "Level Set Evolution Without Re-initialization: A New Variational Formulation" *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* 1063-6919/05 \$20.00 © 2005 IEEE.
- [8] F. Yan, X. Shao, G. Li, Z. Sun, and Z. Yang, "Edge detection of tank level IR imaging based on the auto-adaptive doublethreshold canny operator," *Intell. Inform. Technology Applicat. Research Assoc.*, vol. 3, pp. 366 - 370, Dec. 2008.
- [9] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. the ACM*, vol. 15, pp. 11-15, Jan. 1972.
- [10] L. Xua, M. Jackowskia, A. Goshtasbya,*, D. Rosemanb, S. Binesb, C. Yuc, A. Dhawand, A. Huntleye, "Segmentation of skin cancer images" *Image and Vision Computing* 17 (1999) 65–74,ELSEVIER.

- [11] W. N. Street, “Xcyt: A System for Remote Cytological Diagnosis and Prognosis of Breast Cancer”.
- [12] Borovec, Jiří. (2013). Fully automatic segmentation of stained histological cuts. In L. Husník (Ed.), 17th International Student Conference on Electrical Engineering (pp. 1–7). Prague: CTU in Prague.
- [13] Arabas J.. “*Lectures on Evolutionary Algorithms*”, WNT, Warsaw (in Polish) (2004).
- [14] Ballard D. “*Generalizing the Hough transform to detect arbitrary shapes*”, *Pattern Recognition* **13**(2): 111-122 (1981).
- [15] Duda R. and Hart P. “*Use of the Hough transformation to detect lines and curves in picture*”, *Communications of Association for Computing Machinery* **15**(1): 11-15 (1972).
- [16] Marciniak A., Obuchowicz A., Monczak R. and Kołodziński M. “*Cytomorphometry of fine needle biopsy material from the breast cancer*”, *Proceedings of the 4th International Conference on Computer Recognition Systems CORES'05*, Springer, Rydzyna, Poland, pp. 603-609. (2005).
- [17] Otsu N. “*A threshold selection method from grey-level histograms*”, *IEEE Transactions on Systems, Man and Cybernetics* **9**(1): 62-66 (1979).
- [18] Pratt W. *Digital Image Processing*, Wiley, New York. (2001).
- [19] Russ J. *The Image Processing Handbook*, CRC Press, Boca Raton, FL (1999).
- [20] Gonzalez R. and Woods R. (2002). *Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ.
- [21] en.wikipedia.org/wiki/YCbCr.
- [23] HGalera_Davidson,R.Gonzalez-Campora,A.Mora-Marin “*Cryptographic DNA measurement in medullary thyroid carcinoma*”. *Cancer*,65:2255-2260,1990.
- [24] R.Bowman and G.Angilini, “*Grading of astrocytomas using neural network and linear discriminant functions based on morphometric measurement collected using static image cytometry*”. *American Journal of clinical pathology*,102:502,1995.
- [25] Setiono R.“*Extracting rules from pruned neural networks for breast cancer diagnosis*”, *Artificial Intelligence in Medicine* **8** (1): 37-51. (1996).

AUTHORS

1. Mr.Aditya P. Pise, Mtech Student, Department of Computer Science and Engineering, G. H. Rasoni Academy of Engineering and Technology, Nagpur, India.
2. Mr.Rushi Longadge, Assistant Professor, Department of Computer Science and Engineering, G. H. Rasoni Academy of Engineering and Technology, Nagpur, India.