ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

Volume 3, Special Issue 3, March 2014

2014 International Conference on Innovations in Engineering and Technology (ICIET'14)

On 21st & 22nd March Organized by

K.L.N. College of Engineering, Madurai, Tamil Nadu, India

Cancer Detection by Cell Segmentation Using Clustering and Watershed Algorithms

C.Ramya[#1], V.Nirmala[*2]

[#1]Department of ECE, M.Kumarasamy college of Engineering, Karur, Tamilnadu, India

[*2]Department of ECE, M.Kumarasamy college of Engineering, Karur, Tamilnadu, India

ABSTRACT— Biopsy is one of the medical tests for skin cancer detection. A recent biopsy procedure requires invasive tissue removal from a living body. It is time consuming and complicated task. So non-invasive in-vivo virtual biopsy is preferable one, which is processed by automatic cell segmentation approach. The key component of the developed algorithms are Watershed transform that use the concept of morphological image processing and incorporate some principles of convergence index filter are used to segment cells in in-vivo virtual biopsy of human skin. This paper improves the success of automated cell segmentation for skin cancer diagnosis. This paper also presents different approaches involved in automated cell segmentation and identification of skin cancer at an earlier stage.

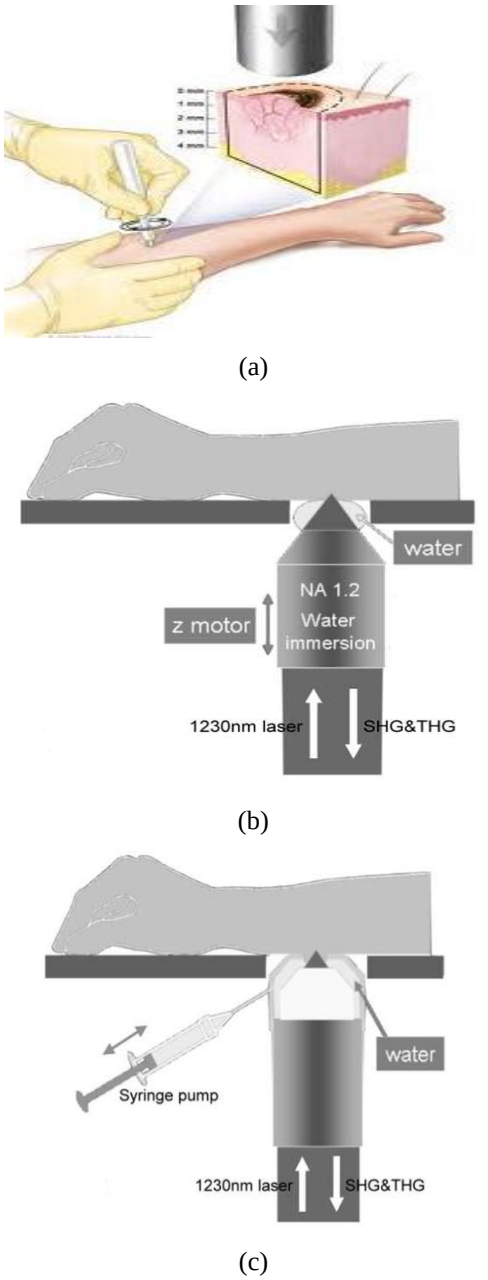
KEYWORDS— cell segmentation, watershed transform, skin cancer, convergence index filter.

I. INTRODUCTION

Today, cancer is a major health problem. In all over country, it is the second leading cause of death. Approximately one out of every two men and one out of every three women get cancer at some point during their lifetime. It is mainly occurs due to the change in the habits of People in our century such as the increase in usages of tobacco, under nutrition in food habits, and lacks of activities. However, the Chance of curing cancer primarily relies on its of the excised tissues is most commonly used today in Imaging modalities such as computed tomography (CT),magnetic resonance imaging (MRI), ultrasound imaging [1], are examples of well-developed technologies for visualizing internal biological structures in vivo. The costly instruments and slow acquisition speed of CT and MRI post limitation on in vivo developmental dynamic studies and none of them are capable of visualizing the cellular and subcelluar structures smaller than 10-nm.1230–1250 nm nonlinear excitation can benefit increased imaging depth along with reduced photo damages, while the generated THG signals are early diagnosis and the selection of its treatment depends on its malignancy level. Therefore, it is very difficult for us to find cancer structures from the benign and healthy ones and identify its malignant level. There are many biopsy techniques are available for diagnosing skin cancer. To accurately diagnose the skin conditions and diseases, the histopatholy microscopic analysis within the visible spectra for collection. The difference between manual and virtual biopsy is shown in fig (1)[7].

A. Manual biopsy & Virtual biopsy

Biopsy is a medical procedure involving the removal of tissues from a living subject for pathological examination to determine the presence or extent of a disease, especially life affecting diseases like cancer. The tissue is examined under a microscope by a pathologist, after extensive preparation procedure, including time consumption, embedding, sectioning and staining (Enhance contrast).The current methods used for cancer and precancer differential diagnosis require invasive tissue removal followed by pathological examination based on their staining pattern and morphological criteria of the cellular and nuclear features in the tissue sections [7]



(a)

(b)

(c)

Fig. 1: (a)manual biopsy (b) In vivo virtual biopsy of skin with an NA 1.2 water immersion objective. (c) In vivo virtual biopsy of skin with a syringe-pump objective

Biopsy is a medical procedure involving the removal of tissues from a living subject for pathological examination to determine the presence or extent of a disease, especially life threatening diseases like cancer. The tissue is generally examined under a microscope by a pathologist, after extensive preparation procedure, including time-consuming fixing, embedding, sectioning and staining (Enhance contrast).The current methods used for cancer and precancer differential diagnosis require invasive tissue removal followed by pathological examination based on their staining pattern and morphological criteria of the cellular and nuclear features in the tissue sections [7]. This biopsy procedure is not only invasive and painful, but also risks the sampling error as only few representative areas in a given lesion were taken for examination and only few sections were observed. Artifacts due to operation errors during biopsy and tissue processing may also lead to false diagnosis.

Besides, clinical difficulties and side effects also occur. These include pain, bleeding, scar formation, infection, and the risk of cancer cell spreading. Optical virtual biopsy techniques, for imaging cells and tissues at microscopic details capable of differentiating benign from malignant lesions noninvasively, are thus highly desirable. Without removing tissues, *in vivo* virtual biopsy avoids or minimizes the above mentioned disadvantages associated with the physically invasive biopsy procedure, reduces the cost and time for pathohistological processing. Optical virtual biopsy could also potentially provide a more comprehensive bedside noninvasive total lesion scanning for improved clinical disease classification and therapeutic guidelines, and a feasible way for continuous disease monitoring during and after treatment. Skin covers almost the whole body and provides many significant functions such as thermoregulation, protection, sensation, and social communication. To accurately diagnose the skin conditions and diseases, the histopathology microscopic analysis of the excised tissues is most commonly used today [7].

In dermatology, some skin conditions and diseases can be diagnosed by the naked eyes but frequently require immediate pathological confirmation. In some cases, especially in the early stage diagnosis of skin diseases, including early cancer diagnosis, painful physical biopsy should be avoided and a noninvasive virtual biopsy tool is more appropriate and is highly desired. For popular minor skin diseases like atopic dermatitis and vitiligo, where physical biopsy is not recommended; virtual biopsy could assist diagnosing and screening the skin conditions [1].Up to now, various noninvasive imaging techniques such as dermoscopy, ultrasound ,optical coherence tomography (OCT),Reflection confocal microscopy, two-photon fluorescence (2PF) microscope and second-harmonic generation (SHG) microscopy have been developed and applied for *in vivo* human skin diagnosis.

II. CELL SEGMENTATION ALGORITHMS

A. Watershed transform

Watersheds are one of the classics in the field of topography. A drop of water falling on one side of this line flows down until it reaches the Atlantic Ocean, whereas a drop falling on the other side flows down to the Pacific Ocean [2]-[3]. Now, in the field of image processing and more particularly in Mathematical Morphology (MM), grayscale pictures are often considered as topographic image. It means that topographic representation of a given image the numerical value (i.e., the gray tone) of each pixel stands for the elevation at that particular position.

This type of representation is highly useful, since it first allows one to better appreciate the effect of a given transformation on the image under study [5]. It is also more accurate and behaves well in all the particular pixel configurations where many algorithms produce incorrect results [16]. Furthermore, the present algorithm is very general: its adaptation to any kind of underlying grid is straightforward, and it can be fairly easily extended to n-dimensional images and even to graphs [6]. Mainly watershed transform is used in nuclei segmentation. It may be combined with marker controlled strategy which identifies regional minima. By using watershed transform we can avoid photo damage and produce over segmentation.

B. Marker controlled strategy

By using marker controlled watershed approach we can identify or mark foreground object and back ground locations in a better way. The following steps are involved.

- Compute segmentation. Where image dark regions are objects which one is need to segment.
- Compute foreground markers. It represents the connected blobs of pixels within each of the objects.
- Compute background markers. These are pixels that are not part of any object.
- Modify the segmentation so that it has minima at the foreground and background marker locations.
- Compute the watershed transform for the modified segmentation function.

C. Convergence index filter

Detection of edges has been one of the basic problems in the area of image processing as edges give important clues for image recognition and image understanding. There have been many proposals of edge detection methods, most of which are based on the magnitude of spatial differences since considerable intensity differences can be expected at the boundary between an object and its background [7]. However, there are objects that have very weak contrast to their backgrounds, such as abnormal opacities on X-ray images wherein boundaries of cancerous tumors are vague. Characteristics of tumor boundaries supply important clues in discriminating between malignant tumors and benign ones, hence an effective method to enhance and detect vague boundaries is required.

Edge enhancement can be performed by several methods including discrete convolution by a high-pass mask and statistical differencing [3]. However, these methods are based on the magnitude of spatial differences where edge christening effects are not sufficient if contrasts are weak. The proposed filter is applied not to intensity images but to gradient vector fields. It evaluates the degree of convergence of the gradient vectors in the neighborhood of each pixel of interest, i.e., the distribution of the orientations of the gradient vectors with respect to the pixel of interest is evaluated. And additionally it changes its size and shape adaptively to the distribution of gradient vectors in the neighborhood of the pixel of interest. This adaptively makes it possible to enhance indistinct boundaries. It is also effective in detecting rounded convex regions in an image regardless of their contrast to their background [4].

D. Image thresholding

Image thresholding is the simplest method. Though it provides high speed for cell segmentation it produces good results only for image with high contrast between object and background. From a grayscale image, thresholding can be used to create binary images. Thresholding methods [14]

E. Histogram shape based

Here the peaks, valleys and curvatures of the smoothed histogram were analyzed. The multispectral remote-sensing images are applied into KLT transform to produce first principal component image. From this output co-occurrence matrix is computed. This output used to construct a histogram [15].By using this threshold values are automatically determined through Otsu algorithm.

F. Entropy based

It uses the entropy of the foreground and background regions, and also the cross-entropy between the original and binary image [14].

G. Spatial methods

It uses higher-order probability distribution and correlation between pixels.

H. Clustering methods

Here the grey levels are clustered into foreground and background [13]-[14].

I. Spectral clustering

It maps the segmentation problem into group. By using this can find meaningful objects. Algorithms that cluster points using eigenvectors of matrices derived from the data. To obtain data representation in the low-dimensional space that can be easily clustered [14].

J. K means clustering

These algorithms are unsupervised learning algorithm. It also generates a no of disjoints and non-hierarchical clusters. It generates globular clusters. Clusters are not overlapped. Its computational speed is high than hierarchical clusters [13].

K. Sliding band filter

Much of cell biology experimental research is based on microscopy image analysis of cell culture. In many cases the use of different fluorescence [1] dies or proteins is used to enable the collection of multivariate images which contain information on different aspects of each cell. Although analysis of such images can be performed manually, it is time consuming, exhausting and prone to human error, requiring frequent repetitions to validate results. These factors motivate the development of automatic cell analysis tools, to identify each

individual cell and extract relevant cell characteristics. The SBF filter[8] is based on gradient convergence and not intensity and as such can detect low contrast cell information which otherwise would be lost in the background noise. Additionally, the convergence evaluation in a regional band allows the reduction of uncertainty caused by noise. As an additional advantage, the SBF filter's parameters are directly related to simple observable image characteristics like cell size and shape, leading to an easy setup of parameters even by someone with no knowledge of the underlying image processing details. To apply the SBF filter to multivariate images we assume that the gradient convergence center of the cell's nuclei and cytoplasm images is the same.

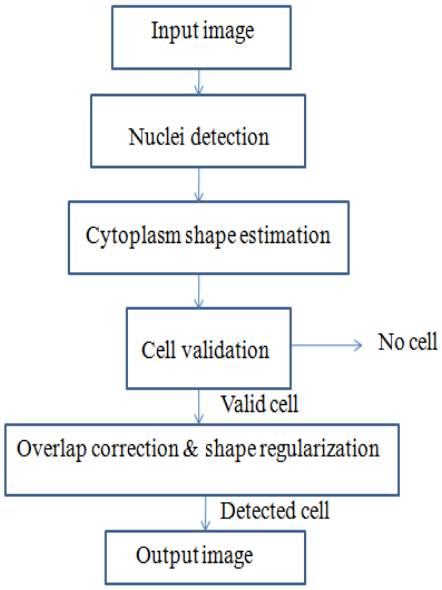


Fig. 3 Cell detection method

This enables the use of nuclei detection information to guide the cytoplasm shape segmentation, and by doing so we are able to avoid errors caused by noise in the images. Contrary to state of the art methods like active contours, this approach constrains the final shape, limiting unlikely shape estimation deviations. Using the same assumption, we propose a detection criterion that jointly evaluates the nuclei and cytoplasm convergence for the estimated cell center and removes cell detections which do not meet a minimum convergence level.

In this way errors in nuclei detection do not imply erroneous final cell detection. To limit the allowed cell shape and cell overlap in the final cell detection results, we introduce two correction steps. First, overlap is corrected by highly irregular cytoplasm shapes that greatly overlap other cells were incorrectly estimated. Second, cell shapes are regularized to eliminate strong discontinuities.

III. PROPOSED METHOD

The proposed method consists of two parts. These are nuclei segmentation and cytoplasm segmentation. By using watershed transform, we can compute nuclei segmentation. The cytoplasm segmentation followed by nuclei segmentation. A local filter is used to obtain segmented cytoplasm. Finally the evaluated nc ratio is obtained. The detailed explanation of above methods was given below. The steps involved are

1. Nuclei segmentation
2. Cytoplasm segmentation
3. NC ratio evaluation

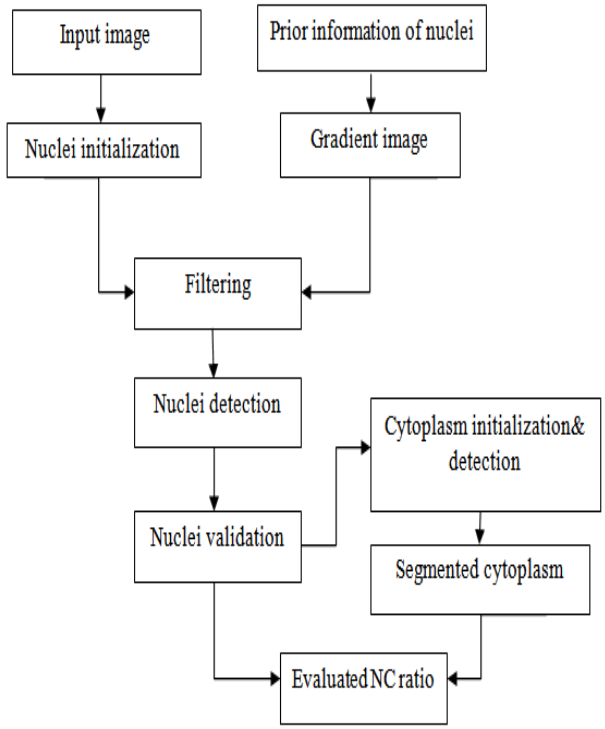


Fig. 4 Flow chart of proposed method

A. Nuclei segmentation

By using gradient watershed transform with marker controlled strategy nuclei segmentation was performed. To get accurate segmented nuclei blob detection and shape descriptors were used. Watershed transforms identifies regional minima, which was used for nuclei segmentation. Before performing nuclei segmentation the input image is converted into gradient map [6].We can apply direct watershed transform on gradient map. But it will give over segmentation. To avoid this nuclei initialization is performed with marker controlled strategy.

B. Cytoplasm segmentation

For cytoplasm segmentation, a convergence index filter is used based on valid nuclei. Because it considers gradient vectors instead of intensity of images, a convergence index filter is suitable for low-contrast and noisy microscopy images. Such a filter makes unnecessary the pre-processing to enhance contrast and remove irregular noise in biomedical images and also preserves the information needed for clinical diagnosis.

C. Nuclei detection and validation

The watershed transform is used again to calculate the watersheds of the filtered gradient map and obtain the segmented nuclei from original skin cell images. To help ensure that the nuclei have been accurately segmented [7], we also consider their shape. Shape descriptors are the crucial measures in the applications of computer vision and pattern recognition, and especially in microscopy imaging analysis, where they help exclude undesired objects. Compactness, which indicates irregularity associated with cancer cells, is utilized in this stage of nuclei validation and is defined as

$$\text{Compactness} = \frac{A}{P^2}$$

Here A is the area of the object and P is the perimeter of the object. The proposed algorithm is applied for skin cell image shown in fig4.

This image consists of no of cells. It may be valid cells and invalid cells. By using shape descriptors invalid cells are removed. Normally cell has circular or elliptical structure [4].

D. NC ratio evaluation

Cell size and nucleus size are indicators of some diseases. Normally the value of NC ratio is larger in cancer affected cells compared with normal cells. Hence this ratio provides developing stages of some skin diseases. So we can detect any diseases at an earlier stage.

IV. SIMULATION

A. Nuclei and cytoplasm segmentation

We have taken normal skin cell image as input image. The image is first converted into gradient map. It is used to find out boundary of the nucleous.It will be compared with the original image. Through this we can obtain valid nuclei regions. The result of nuclei segmentation is shown in fig.7.Following nuclei segmentation cytoplasm segmentation takes place using convergence index filter.

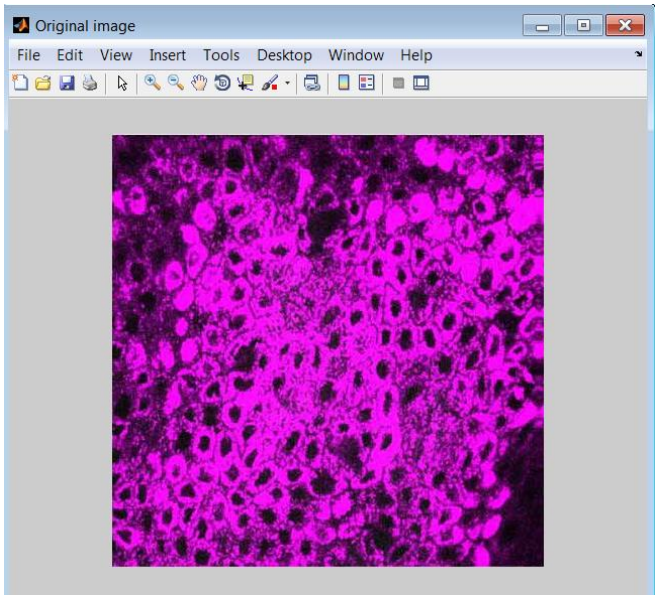


Fig.5 Skin cell image

The above figure shows normal skin cell image. It will be applied as the input image.The fig.7 shows valid cells after conversion of gradient map .The blue mark indicates about valid nuclei.The remaining shows unwanted nuclei

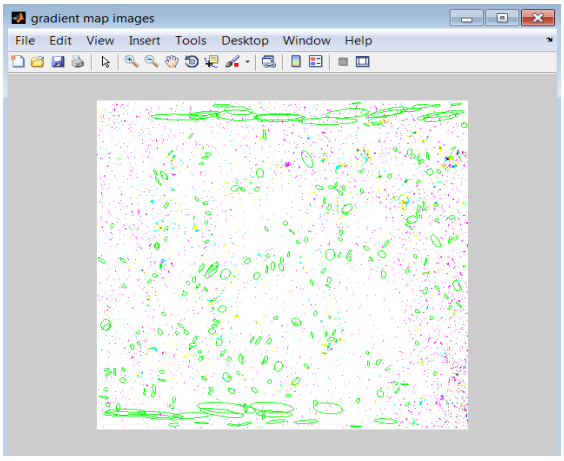


Fig.6 Gradient map image

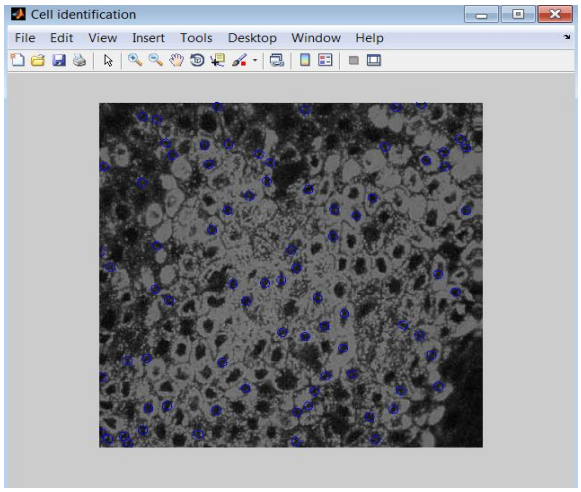


Fig.7 Cell identification

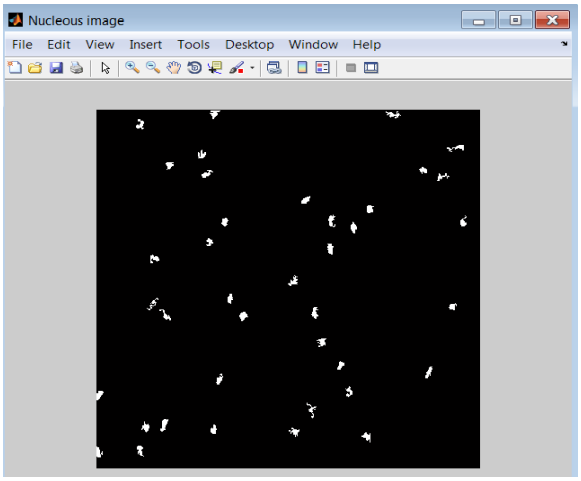


Fig.8 Segmented nuclei

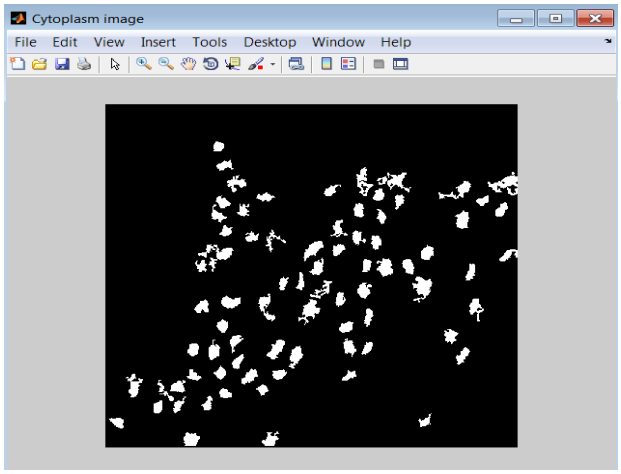


Fig.9 Segmented cytoplasm

B. Evaluated nc ratio

The table shown in below result of the segmented cells. This result is for normal cells. The nc ratio for the normal cells is smaller than the cancer affected cells. By using this nc values the medical pathologist can easily diagnosed the intensity of the disease. It also helpful to identify the disease at an early stage.

Cell index	Cell area	Nuclear area	Cytoplasmic area	NC ratio	Position row	Position column
1	193	54	139	0.3855	13	400
2	294	60	234	0.2564	50	354
3	183	56	127	0.4409	55	381
4	173	60	112	0.5446	62	361
5	220	61	165	0.3333	76	355
6	166	55	120	0.3833	76	378
7	283	46	231	0.2251	89	421
8	172	52	120	0.4333	92	341
9	212	58	144	0.4722	92	308
10	216	49	167	0.2934	101	255
11	159	51	108	0.4722	100	203
12	431	61	370	0.1649	117	189
13	163	46	117	0.3932	107	377
14	251	66	185	0.3568	113	310
15	200	53	147	0.3505	117	140
16	168	50	118	0.4237	118	53
17	184	51	133	0.3835	120	338
18	189	49	140	0.3500	122	123
19	207	62	145	0.4276	124	77
20	212	66	146	0.4521	126	279
:	:	:	:	:	:	:
:	:	:	:	:	:	:
Total	8354	2214	6140	15.2794	-	-
Average	214.2051	56.7692	157.4359	0.3918	-	-

V. CONCLUSION AND FUTURE WORK

Cell segmentation is valuable in the field of computer integrated surgery. In this paper, we presented a method for finding nc ratio for normal skin cell image. But the value of nc ratio in cancer affected cells is larger when compared with normal one. Mainly it is used to identify severity and stage of disease. It is also used to find skin aging, other skin diseases and affected skin due to using of cosmetic products. This technique provides less time consumption and high speed. In future the cell segmentation will improve accuracy, precision, computational speed of segmentation methods as well as reduce the amount of over segmentation.

REFERENCES

[1] Chi-Kuang Sun,a,*Shi-Wei Chu,a Szu-Yu Chen,a Tsung-Han Tsai,a Tzu-Ming Liu,aChung-Yung Lin,b and Huai-Jen Tsaib 'Higher harmonic generation microscopy for developmental biology',2005.

[2] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," IEEE Trans. Pattern Anal.Machine Intell., vol. 13, pp. 583–598, 1991

[3] Bleau and L. J. Leon, "Watershed-based segmentation and region merging," Computer Vis.Image Understand., vol. 77, no. 3, pp.317–370, 2000.

[4] H. Kobatake and S. Hashimoto,'Convergence index filter for vector fields', *IEEE* Trans. Image Process., vol. 8, no. 8, pp. 1029–1038, 1999.

[5] J. M. Gauch,'Image segmentation and analysis via multiscale gradient watershed hierarchies', *IEEE* Trans. Image Process., vol. 8, pp. 69–79,1999.

[6] J. Park and J. M. Keller, "Snakes on the watershed," *IEEE* Trans. Pattern Anal. Mach. Intell., vol. 23, no. 10, pp. 1201–1205, 2001

[7] S.-Y. Chen, S.-U. Chen, H.-Y.Wu, W.-J. Lee, Y.-H. Liao, and C.-K.Sun, "In vivo virtual biopsy of human skin by using noninvasive higher harmonic generation microscopy," *IEEE* J. Sel. Topics Quantum Electron., vol. 16, no. 3, pp. 478–492, 2010.

[8] P. Quelhas, M. Marcuzzo, A. M. Mendonca, and A. Campilho, "Cell nuclei and cytoplasm joint segmentation using the sliding band filter,"*IEEE* Trans. Med. Imag., vol. 29, no. 8, pp. 1463–1473, 2010.

[9] H.S.Wu, J. Barba, and J. Gil, "A parametric fitting algorithmfor segmentation of cell images*," IEEE* Trans Biomed Eng, vol. 45, pp. 400–407, 1998.

[10] M. Silveira andA.Monteiro, "Automatic recognition andmeasurement of butterfly eyespot patterns," *Biosyst.*, vol. 95, no. 2, pp. 130–136,2009.

[11] P.Soille,Morphological Image Analysis: Principles and Applications.New York, NY, USA: Springer-Verlag, 2003.

[12] C.-K. Sun, "Higher harmonic generation microscopy," *Adv. Biochem. Eng./Biotechnol.*, vol. 95, pp. 17–56, 2005.

[13] N. Bolshakova and F. Azuaje, "Clustering validation techniques for genome expression data," Genomic Signal Process., vol. 83, no. 4,pp. 825–833, 2003.

[14] Mehmet Sezgin and Bu lent Sankur 'Survey over image thresholdingtechniques and quantitative performance evaluation'**,** 13(1), 146–165 (January 2004).

[15] Pornphan Dulyakarn, Yuttapong Rangsanseri, and Punya Thitimajshima'Histogram transformation based threshold selection for Image Segmentation'2000.

[16] R.C.Gonzalez and R.E.Woods, Digital Image Processing.Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.

[17] Sameer Ruparelia,' Implementation of watershed based algorithm inFPGA'mar-2012.