

Detecting Morphological Nature of Cancerous Cell Using Image Processing Algorithms

Jamali Firmat Banzi*, Xue Zhaojun**

* Masters Student, School of electronics engineering, Tianjin University of technology and education
** Associate Professor, School of electronics engineering, Tianjin University of technology and education

Abstract- The purpose of this research is to identify the morphological nature of cancerous cell by employing image processing algorithms so that they can be used to automate diagnostic techniques, reducing the cost and increasing the accuracy of epithelial biopsy procedures. The approach taken combines a mix of image processing techniques and methods such as (median filter, unsharp mask and threshold). Next an averaging was done to remove noise and to distinguish between the two cells so that it can be easily identified which one has some abnormality. Lastly an improved future work of line plot in which will clearly identify the abnormality of a cell and hence recognize a cancerous cell accurately was proposed in this research.

Image quality and accuracy is the core factors of this research, image quality assessment as well as improvement are depending on the enhancement stage where low pre-processing techniques is used based on Median filter within Gaussian rules. Following the segmentation principles, an enhanced region of the object of interest that is used as a basic foundation of feature extraction is obtained. Relying on general features, a normality comparison is made. In this research, the main detected features for accurate images comparison are pixels percentage and mask-labelling.

Index Terms- Image processing, Cancer detection, Image enhancement, unsharp mask.

I. INTRODUCTION

Cancer is a term used for diseases in which abnormal cells divide without control and are able to invade other tissues. Cancer cells can spread to other parts of the body through the blood and lymph systems. Cancer is not just one disease but many diseases. There are more than 100 different types of cancer.

Most cancers are named for the organ or type of cell in which they start for example, cancer that begins in the colon is called colon cancer; cancer that begins in melanocytes of the skin is called melanoma.

All cancers begin in cells, the body's basic unit of life. To understand cancer, it's helpful to know what happens when normal cells become cancer cells. The body is made up of many types of cells. These cells grow and divide in a controlled way to produce more cells as they are needed to keep the body healthy. When cells become old or damaged, they die and are replaced with new cells.

However, sometimes this orderly process goes wrong. The genetic material (DNA) of a cell can become damaged or changed, producing mutations that affect normal cell growth and division. When this happens, cells do not die when they should and new cells form when the body does not need them. The extra cells may form a mass of tissue called a tumor.

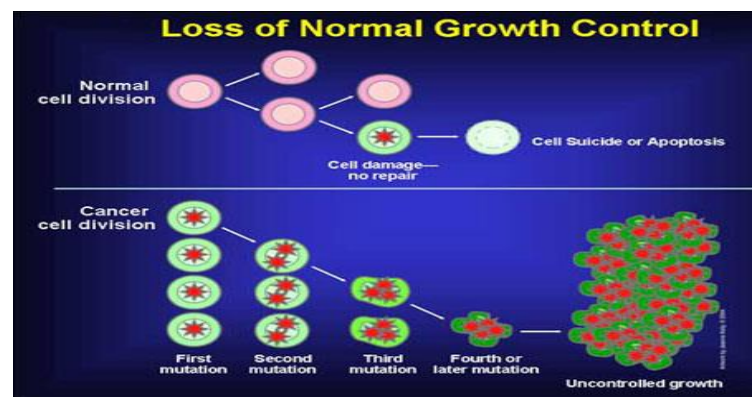


Figure 1: Cell division

Over half of all human cancers occur in stratified squamous epithelia. Approximately one million cases of non-melanoma cancers of the stratified squamous epithelia are identified each year [1]. Tissues with stratified squamous epithelia include the cervix, skin and oral cavity. Currently, the diagnosis of squamous epithelial cancers is carried out through visual inspection, followed by biopsy. In patients at high risk for malignancy, the entire epithelium may potentially be diseased. Therefore it is difficult to identify the best location to biopsy based on visual inspection alone. Techniques that can diagnose epithelial pre-cancers and cancers more accurately than visual inspection alone are needed to guide tissue biopsy. An instrument for this application would ideally use quantitative algorithms, thus decreasing the need for clinical expertise and waiting periods for diagnosis.

Multiphoton laser scanning microscopy (MPLSM) is a potentially attractive technique for the diagnosis of epithelial pre-cancers and cancers. This technique enables the visualization of cellular and sub-cellular structures with exceptional resolution. Visualization of these structures is important because it is well-known that the development of pre-cancers and cancers is accompanied by changes in cellular and subcellular morphology [2]. For example, as tissue progresses toward cancer, cells become increasingly disorganized.

An automated method of classifying normal from abnormal tissue is required so the images do not have to immediately

interpreted by a trained pathologist. The goal of this research is to exploit differences in cell organization between normal and cancerous tissues using image processing techniques, thus allowing for automated diagnosis.

Previous studies have used Fourier analysis of *in vivo* human corneal endothelial cells to correlate cell structure with patient age [3]. It was found that the Fourier transforms provided quantitative descriptions of population cell size and organization. Fourier transform analysis will be applied to MPLSM images of normal and cancerous tissues to determine whether automated diagnosis based on tissue morphology is feasible.

II. RELATED WORK

Different researches proposed the Segmentation method of nucleus and cytoplasm. However segmenting nucleus and cytoplasm of leukocytes from bone marrow images is a very difficult task, as the images show heterogeneous staining and high-cell population. [7]

Our approach involves employing different image processing techniques and methods to recognize a cancerous tissue through its cell organization. A normal tissue usually has a defined cell organization but a cancerous tissue cells are disorganized

III. METHODOLOGY

The first steps of our algorithm preprocess the images to remove graininess (median filter), enhance contrast between the cytoplasm, and nucleus and extracellular components (unsharp mask, threshold). After the pre-processing steps, the relative disorganization of the images was determined with Fourier transform analysis [3, 4].

Averaging was then used to reduce the noise of the Fourier domain image and a 1-D line plot was advised. From this line plot, normal and cancerous tissue could be differentiated by a machine

For each step of the algorithm, the output of the previous step is the input of the subsequent step. The parameters of the algorithm were optimized for maximum contrast between normal and cancerous images at the last step of the algorithm. Each step of the algorithm, along with the output from each step is described in detail below.

Flowchart

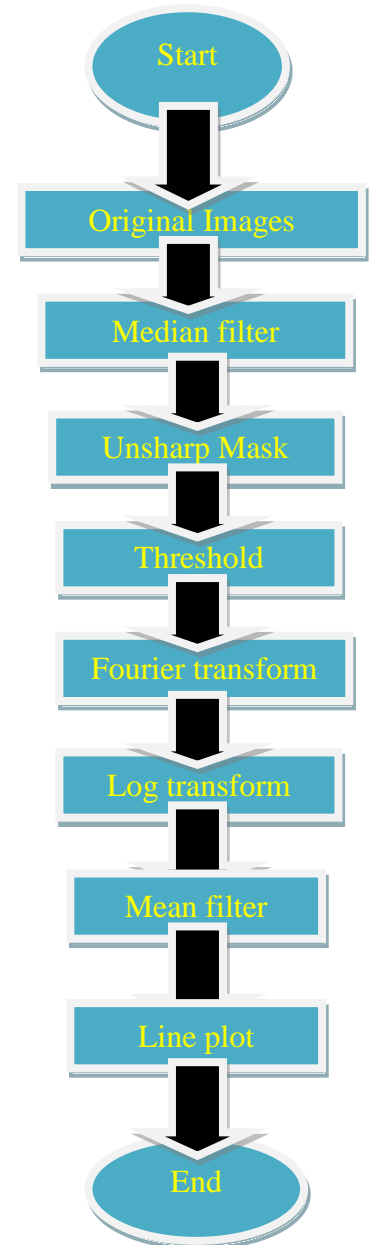


Figure 2: shows a flowchart of the algorithm.

IV. EXPERIMENT AND RESULTS

The image processing algorithm was applied to a total of two images from normal tissues and two images from cancerous tissues for a total of four images using MATLAB. The results from one normal and one cancerous image are shown in this report. The results for the remaining one normal and one cancerous image are similar.

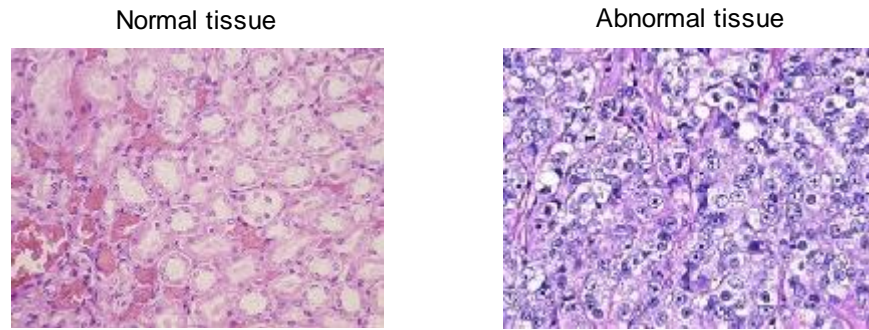


Figure 3: shows the original images for normal and abnormal (cancerous cell) resulted after implementing MATLAB code for reading image

4.1 Median Filter

The first step of the algorithm aims to attenuate noise without blurring the images. A 2dimensional median filter was applied using the 'medfilt2' function in Matlab. Each output pixel contains the median value in the 5-by-5 neighborhood

around the corresponding pixel in the input image. 'Medfilt2' pads the image with zeros on the edges, so the median values for the points within 3 pixels of the edges may appear distorted.

The result is shown in the figure below

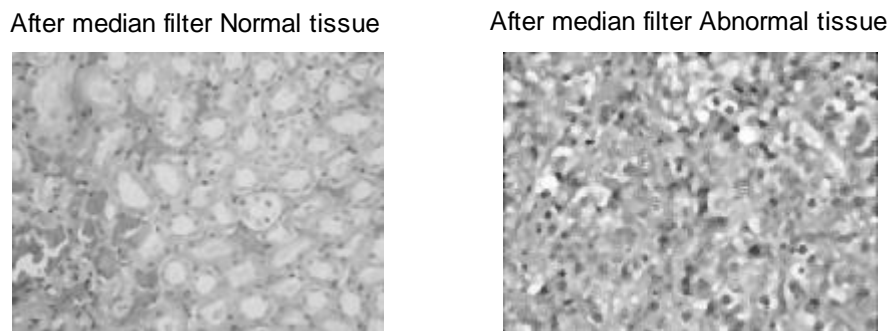


Figure 4: Normal and abnormal tissue after median filter

4.2 Unsharp Mask

Next, contrast between the cytoplasm, and nuclei and extracellular components were enhanced using an unsharp filter. The filter was applied to the image by subtracting the gaussian filtered input image, multiplied by a scaling factor, from the

input image. The gaussian filter was created using the built-in Matlab functions 'fspecial' and 'gaussian'.

A rotationally symmetric Gaussian lowpass filter with a standard deviation of 10 pixels was used, with a total filter size of 15-by-15 pixels. The scaling factor was 0.9.

The result of this step is shown in Figure 5:

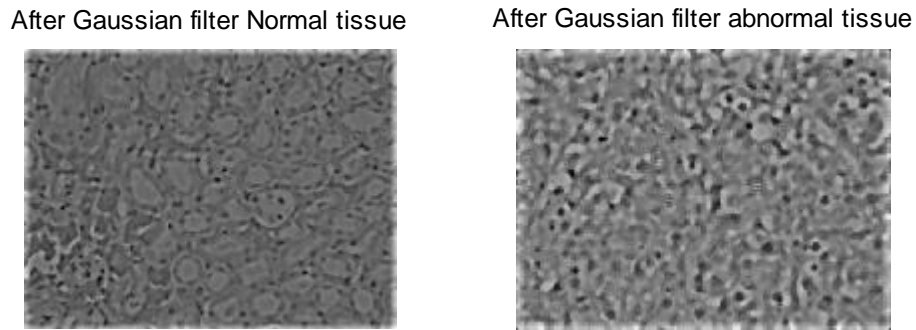
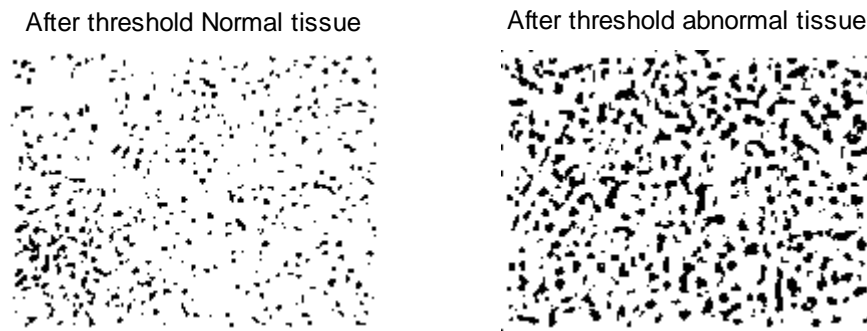


Figure 5: Normal and abnormal cell after Gaussian filter

4.3 Threshold

The built-in Matlab function 'graythresh' was used to threshold all images so that the cytoplasm was white (or 1) and

the nucleus and extracellular components were black (or 0). The Matlab function 'graythresh' computes the global image threshold using Otsu's method.

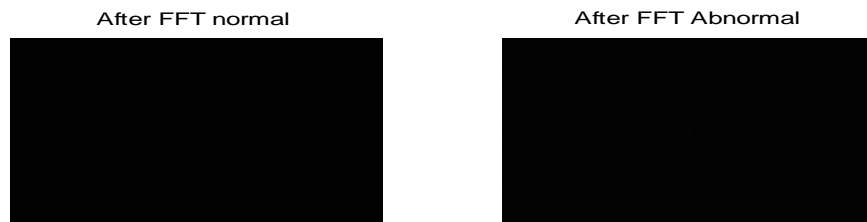


The result is shown in Figure 6:

4.4 First Fourier Transform

The built-in Matlab function 'fft2' was used to convert the binary images into the spatial frequency domain using the two-

dimensional discrete Fourier transform. The image was shifted before the Fourier transform so that the zero frequency component was at the center of the frequency space.



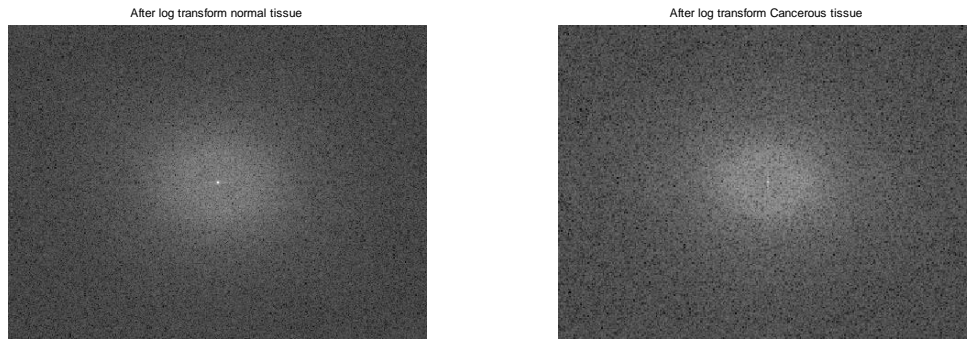
The result is shown in Figure 7:

4.5 Log Transform

Next, the log transform of the image in Fourier space was performed using the equation

$s = \log(r + 1)$. The log transform compressed the values of the light pixels of the image and expanded the values of the dark pixels of the image. This reduced the DC values relative to the

rest of the pixel values, allowing the details of the transform to become visible (Fig. 8). At this point a feature starts to become apparent which might be used to automatically separate the normal from the cancer. In the cancer samples, the low frequency bright spot is fairly uniform. Looking closely at the normal samples, a dark ring is visible.



(a) Normal tissue after log transform

(b) A cancerous tissue after log transform

4.6 Mean Filter

These images are fairly noisy which may make automatic detection schemes challenging.

To reduce the noise a 5 by 5 pixel mean filter was implemented. This filtered averaged 25 points thus reducing the noise. The results can be seen in Figure 9:

noise by 5. Because a single pass of this filter did not seem to provide sufficient noise reduction, the image was passed through the filter a second time.

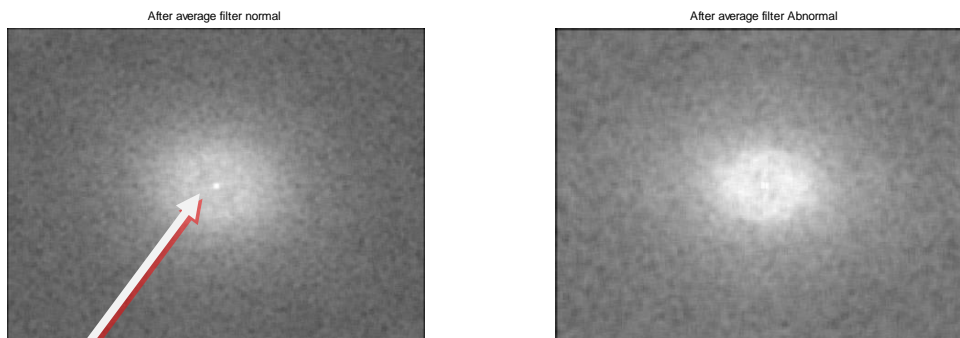


Figure 9 : (a) Normal

(b) Cancer cell

pixels was plotted against their positions to quantitatively analyze the normal tissue from the cancerous one.

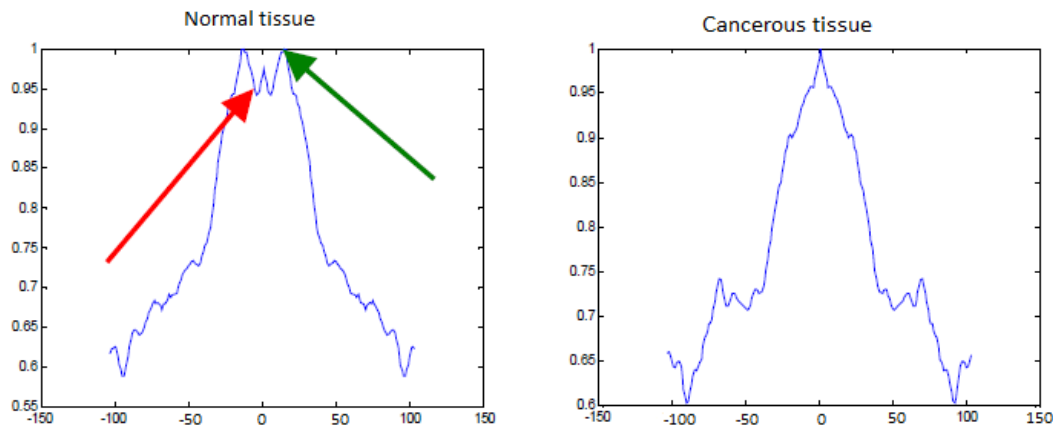
Here the dark ring in the low frequency area of the normal tissue is still visible but the noise is reduced.

From this point is where we can start to distinguish and recognize cancerous cell distinctively. However the methodology is not ample to accurately identify a cancerous cell from a normal one until with the help of machine. In addition to that we employed the use of Line plot in which the center row of the

4.7 Line Plot

To get these two dimensional images such that they could easily and quantitatively analyzed by 1 dimensional signal processing techniques, the center row of pixels was extracted and their values plotted against their positions.

The results can be seen in Figure 10:



From the plots in Fig. 10 it can be seen that there is a local minimum in the normal images at approximately the 7th pixel from the center, or at a frequency of approximately 55 mm-1 (red arrow). Possibly more telling is that there is a local maximum at approximately the 16th pixel from the center or, 127 mm-1 (green arrow). This would indicate that normal cells contain regular features which repeat at 7.8 μm , whereas the cancerous cells do not contain this repeating nature.

An alternative approach was also implemented starting with the log transformed Fourier space image. As already described,

there were two things that we wanted to do to this picture. The first was to reduce the noise. The second was to reduce the image to a plot that could be quantitatively analyzed. To accomplish both goals simultaneously the radial symmetry of the image was exploited, and pixels were averaged according to their radius. For example, the value of all pixels, 5 pixels from the center of the image were averaged. Then the values of all 6 pixels from the center were averaged.

This was done along the entire radius of the image. This function was then mirrored around DC to make the result more intuitive (Fig. 11).

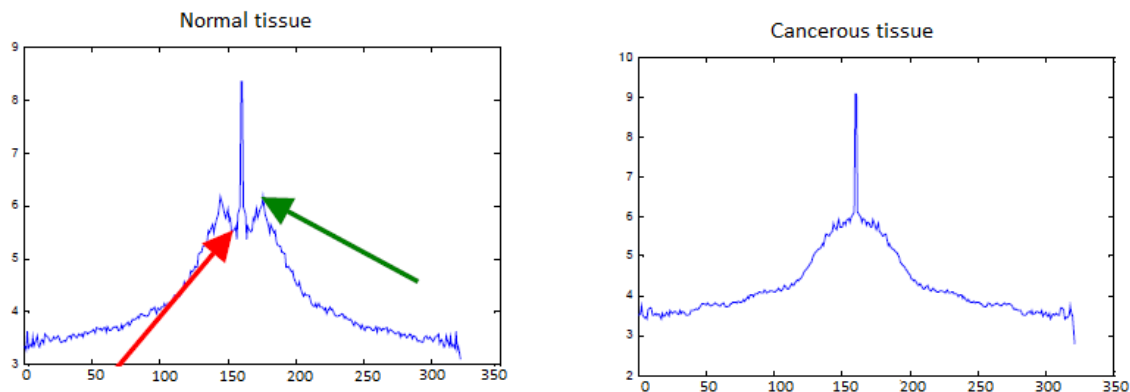


Figure 11: a line plot

Noise reduction by averaging is the square root of the number of pixels averaged, thus the noise reduction changes as a function of radius. However, it may be argued that this retains the radial features of the image better than applying a uniform averaging filter.

Again in these plots, it can clearly be seen there is a local minimum in the normal images at approximately the 7th pixel from the center or at a frequency of approximately 55 mm-1 (red arrow) and a local maximum at approximately the 16th pixel from the center or, 127 mm-1 (green arrow).

V. CONCLUSION

An image improvement technique is developing for earlier disease detection and treatment stages; the time factor was taken

in account to discover the abnormality issues in target images. Image quality and accuracy is the core factors of this research, image quality assessment as well as enhancement stage where were adopted on low pre-processing techniques based on filtering within Gaussian rules.

However multiple image enhancement steps were needed to exaggerate the differences between the frequency-domain images of normal and cancerous tissues (median filter, unsharp mask, and threshold).

Additional enhancements were needed to improve contrast between the power spectrum of normal and cancerous tissues (averaging). After these enhancements, clear differences could be seen between the normal and cancerous power spectrum. For example, in Figs. 10 and 11 there are frequency peaks as indicated by the green arrows in the normal tissue spectrum,

which are not present in the cancerous tissue spectrum. With more extensive testing, we believe we may be able to use this local maximum to quantify the organization of the tissue structure. This could be investigated in the future and potentially lead to automated diagnostic techniques, reducing the cost and increasing the accuracy of epithelial biopsy procedures.

REFERENCES

- [1] Cancer Facts and Figures. 2003, American Cancer Society.
- [2] White, F.H., K. Gohari, and C.J. Smith, Histological and ultra structural morphology of 7, 12 dimethylbenz (alpha)-anthracene carcinogenesis in hamster cheek pouch epithelium. *Diagn Histopathol*, 1981. 4(4): p. 307-33.
- [3] Fitzke, F.W., Fourier Transform Analysis of Human Corneal Endothelial Specular Photomicrographs. *Exp Eye Res*, 1997. 65.
- [4] Tarawneh M., Nimri O., Arqoub K., Zaghal M., Cancer Incidence in Jordan 2008,
- [5] National Cancer institute US <http://www.cancer.gov/> accessed on 23rd October 2013
- [6] P. Soile, "Morphological image analysis, principles and applications", Springer, Berlin, 2003
- [7] S.Jagadeesh, Dr.E.Nagabhooshanam, Dr.S.Venkatachalam," image processing based approach to cancer cell prediction in blood samples", *International Journal of Technology and Engineering Sciences* Vol.1 (1), 2013, pp2320-8007.

AUTHORS

First Author – Jamali Firmat Banzi. He received Bachelor's degree of science in information systems from University of Dodoma, Tanzania. He then he worked as a Tutorial assistant for some months before he joined Tianjin University of Technology and Education, Tianjin, China for his further education.



His research areas of interest include signal and information processing; Image processing, Pattern recognition, neural networks, communication and Information security.

Second Author – Prof. Zhaojun Xue received the bachelor's degree in mechanical design and automation at Shandong University, Jinan, China, received his master's degree in biomedical engineering from Tongji University, Shanghai, China, and completed his doctor's degree in biomedical engineering from Tianjin University, Tianjin, China. He is presently a postdoctoral fellow in Tianjin University. His research interests include bioelectronics, biomedicine information detection, image processing, and pattern recognition.