Automated Segmentation using Histopathology Images as a Diagnostic Confirmatory Tool in Detection of Bone Cancer

Vandana B S
Department of Computer Science& Engineering
KVG College of Engineering, Sullia, VTU, India

Antony P J
Department of Information Science & Engineering
St. Joseph Engineering College, Mangalore, VTU

ABSTRACT
This paper presents a method to extract cancer affected area from a histopatholical image of bone cancer. Existing approaches are manual, time-consuming and subjective. In the proposed approach, morphology technique is used to find the area affected in the bone cell and extract the same using adaptive threshold technique. To get more accurate segmentation, watershed algorithm is used which will separate the attached tissue cells. In this method we used nucleus size, area, orientation to define malignancy level. Experiment results show that, using the proposed method, the meaningful features in the background with heterogeneous intensities are appropriately segmented. Bone tissue samples contain several cell type and these cells including blood cells, normal cells, and cancerous cells. Nuclear size and shape are good visual descriptors which is used to differentiate normal and cancer cell. This method successfully demonstrated an automated image segmentation technique to overcome noise due to staining process from bone cancer microscopic images and provide accurate analysis of nuclear size and density with a comparable difference from normal bone histology. The automatic segmentation resulted in a sensitivity of 76.4%, defined as the percentage of hand segmented nuclei that were automatically segmented with good quality.

Keywords
Human pathology, Image Segmentation, Cancer Cell Images, nuclei, bone

1. INTRODUCTION
Analysis of microscopic images of cell and tissues has been a goal of human pathology and cytology which found potential use in the detection and analysis of cancerous cells. Previous work in this field consisted of manual measurements of cell and nuclear size, followed by calculation of cell and nuclear volume. Pathological examination of a biopsy is one of the reliable technique which is used to diagnose bone cancer. Historically Pathologists use histopathological images of biopsy samples removed from patients, examine them under a microscope. Based on their experience they will make judgments. However this visual study is not accurate often leading to considerable variability. In this field, accuracy is very essential for confirmatory diagnosis. Automation can improve the practice of Pathology by overcoming the limitations of manual microscopy. It has also been documented that the tumors of the bone are infrequent, when compared to all other tumors of the body. The wide spectrum of tumors of bone and their diverse origin from multiple cell type along with the tendency of these tumors to produce overlapping anatomical pattern make a complicated issue [1]. However it is a highly challenging field from the point of view of morphological diagnosis.

Automatically segment cell nuclei in histology images of bone tissue for cancer analysis are the main objective in this work. In the proposed method of automated cancer identification in the microscopic image of bone cancer includes the following steps. The first step is noise elimination to determine the focal area in the image. The noise arises from staining the biopsy samples. The second task is the nucleus/cell segmentation. Segmentation is one of the challenging task because some cells are attached to their neighboring cell (overlapping) and occurrence of noise. After segmentation process the next step is the feature selection.

2. RELATED WORK
Several automated cancer diagnosis tools have been reported in the literature. Depending on the feature set these tools use, they can be divided into five Categories; fractal[2,3,4], textural[5,6,7,8],topological[9],morphological[10], and intensity [11,12]. Different studies work on different types of cancer is given in table 1.They extract different types of features to represent the cells and the tissues of those cancer types. Since the cell and tissue structures may be different for different organs; a method that works on one cancer type may not work equally well on another. Moreover, some of these studies focus on distinguishing these different cancerous structures from their non-cancerous counterparts while some of them aim at classifying them into different grades. The experimental methods also may vary: (i) different lightening conditions and magnifications are used and (ii) different numbers of samples are collected from different numbers of patients.

3. PROPOSED METHODOLOGY
The images used in our experiments were tissue that is surgically extracted from bone which is stained with H&E. Hematoxylin stains the nuclei blue, while eosin stains cytoplasm and the extra cellular connective tissue matrix pink. There are hundreds of various other techniques which have been used to selectively stain cells. Biopsy or surgical specimens were examined by a Pathologist. After the specimen has been processed and histological sections were placed on glass slides. The tissue sections were observed under a microscope with a magnifying factor of 40x.The digital image was saved as a color 436×289 JPEG files for processing. The dataset consists of 96 samples.

Table 1. Features used based on organ for cancer detection

Type of cancer	Feature extraction
Bladder	Morphological,Textural,Fractal-based,Topological,Choi(1997),Rajesh,Tasoulis (2003).
Brain	Textural, Topological,Demir (2004),Gunduz(2004).
Breast	Morphological,Textural,Fractal-based,Intensity-based,Anderson(1997),Dey and Mohanty(1993),Einstein(1998), Pornchai Phukpattaranont and Pleumjit Boonyaphiphat(2007).
Cervical	Textural,Topological,Fractal based,Sedivy(1999),Keenan(2000),Walker(1994).
Colorectal	Textural,Fractal-based,Intensity-based Esgiar(1998,2000).
Gastric	Morphological,Textural,intensity-based Blekas(1998).
Liver	Textural, Fractal-based Aibregtsen(1999,2000).
Lung	Morphological, Intensity-based Thiran and Macq (1996).
Skin	Textural,intensity-based Smolle(2000,2003),Vrushli Korde, M.S.,Hubert Bartels (2009)

3.1 Features of Cancer Cells

Figure 1 shows an example of microscopic image of tissue sample, malignant bone tumor. The main characteristic of malignant cell is irregular form and structure of nucleus. Nucleus of cancer cell has abnormal shape, larger in size compared to the normal cell. Scant cytoplasm is another feature of malignant cell. Cromatin is not uniformly distributed in the internal structure of the nuclei. Another important feature is reproduction and duplication of nuclear material. All these information is required for automation process.

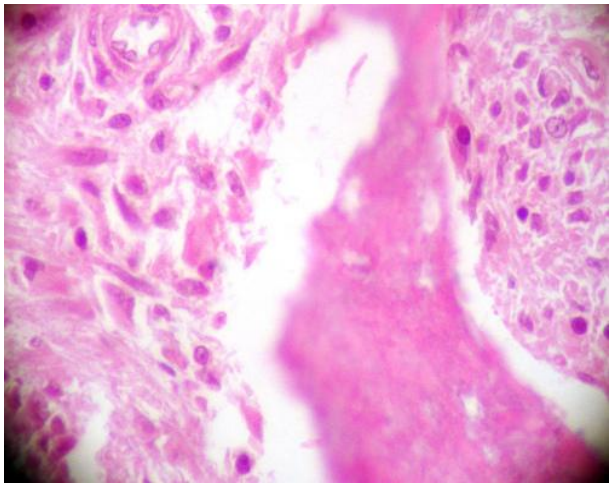


Fig 1: Microscopic image of tissue samples surgically removed from human bone tissues and stained with H&E

3.2 Segmentation Algorithm

In this study we have proposed a new approach for automatic cell nuclei segmentation of the image for cancerous analysis. The image was imported into MATLAB and converted to grayscale for faster processing using inbuilt functions. For each segmented regions, we extracted relevant features of size and shape that accurately describe the characteristics of the region which is needed for cancer analysis. The algorithm for cancer cell segmentation is as shown below.

An algorithm for cancer cell segmentation

Input : JPEG format of microscopic image of bone sample tissue.

Output : Segmentation of nucleus/cell from the image.

Method:

Step1: Read image from the file.

Step2: Convert the image into grayscale for faster processing.

Step3: // Noise elimination

a) Morphological opening by large square structuring element close to the size of the nucleus.

b) Morphological reconstruction for shape recovery was applied.

Step4: //Nucleus/cell Segmentation.

a) Thresholding of the reconstructed image to isolate light grains from the dark background.

b) The object (individual grains) with different colors were labeled.

c) Area from each detected object (size distribution) was extracted.

d) Non uniform background using morphology was characterized and background by image subtraction was removed.

e) The steps thresholding ,and label object was repeated for extract new area and updated size distribution.

f) Partial grains cut off by edges were removed which will remove the border objects.

g) Size distribution after border clearing was repeated.

h) First order statistics ,mean and standard deviation was computed.

e) Output the detected grains on the original image.

Step5: //Watershed segmentation to separate the attached cell

a) The size of individual object in the detected grains image were calculated.

b) IF size<small then // Small indicates noise,

Object ignored

IF size >big then

Objects were overlapped to each other.

Watershed algorithm was applied.

ENDIF

Step6: Output the result of watershed segmentation

3.3 Noise Removal

Mathematical morphology as a tool for noise filtering and shape simplification was adopted [10]. The operation is an opening of the image by a large flat square structuring element, close to the size of the nuclei. The result is the morphological filtering effect. Opening of an image A by a structuring element B, is, A oB denoted by

$$A \circ B = (A \ominus B) \oplus B \quad (1)$$

The entire element smaller than the structuring element disappears or roughly lighted. Thus its size has to be chosen according to the magnification factor used during the acquisition of the image. It is also noted that some cells do not have the complete shape due to an uneven distribution of intensity. To overcome this reconstruction of the image, morphological opening for spike noise removal. Subsequently, hole filling is performed using the algorithm based on morphological reconstruction.

3.4 Find the Affected Area in the Image

The After noise removal we looked at the intensity distribution in the smoothed image. Applied threshold to segment light grains from dark background and individual grains with different colors were labeled. This area values from each detected object was extracted. Non uniform background using morphology was characterized and background was removed by image subtraction. After background removal, thresholding to segment cells was applied and new areas were extracted. Partial grains removed by cut off edges. Redo size distribution again after border clearing shown in fig 2(a). After detecting grain ,first order statistics, mean and standard deviation was computed.

If y is a matrix, matlab function mean(y) treats the columns of A as vectors, returning a row vector of mean values. There are two common definitions for the standard deviation s of a data vector y.

$$S = \left[\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \overline{y})^2 \right]^{\frac{1}{2}} \quad (2)$$

$$S = \left[\frac{1}{n} \sum_{i=1}^{n} (y_i - \overline{y})^2 \right]^{1/2} \quad (3)$$

Where $\overline{y} = \sum_{i=1}^{n} y_i$ and n is the number of elements in the sample. The two forms of the equation differ only in n versus n-1 in the divisor. If y is a matrix, the matlab function std(y) returns a row vector containing the standard deviation of the elements of each column of y. If y is a multidimensional array, std(y) is the standard deviation of the elements along the first nonsingleton dimension of y.

Finally, detected grains on the original image were indicated. These segmented cells are shown in fig 2(b). Thus by using this information total area of nucleus was calculated.

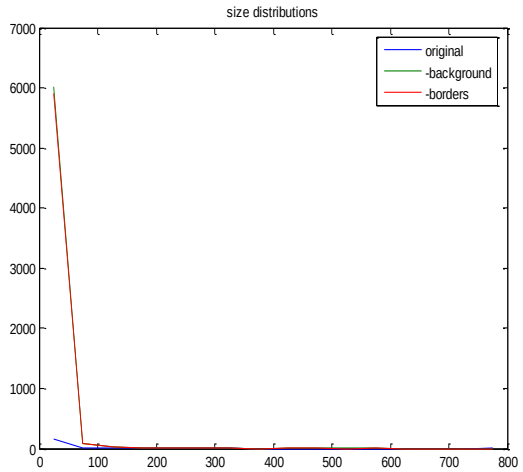


Fig 2: (a) Size distribution after background removal and border clearing

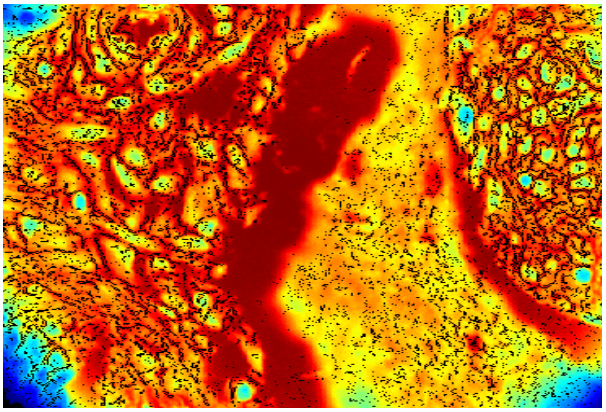


Fig 2: (b) Final detected grains on the original image

3.5 Watershed Segmentation

The In order to obtain more accurate cell segmentation, to separate the attached cancer cell watershed method was used [18].The result of watershed segmentation is shown in 3(a).The watershed algorithm performs very accurate segmentation, which is beneficial in case when objects overlap and their borders are hardly detectable. It splits the image area into disjoint regions and the watershed lines are closed curves. Disk-shaped structure element was created bottom hat filter and top hat filter were performed. Images were subtracted to enhance contrast (lower right) and intensity valleys detected for watershed segmentation. Separate regions were labeled shown in fig 3(b) and shapes of detected features were assessed by exploring aspect ratio.

Watershed algorithm tends to achieve better results, but these algorithms tend to over-segmentation especially when the images are noisy. The excellent segmentation results from the proposed algorithm are demonstrated with microscopic images under histological noise conditions.

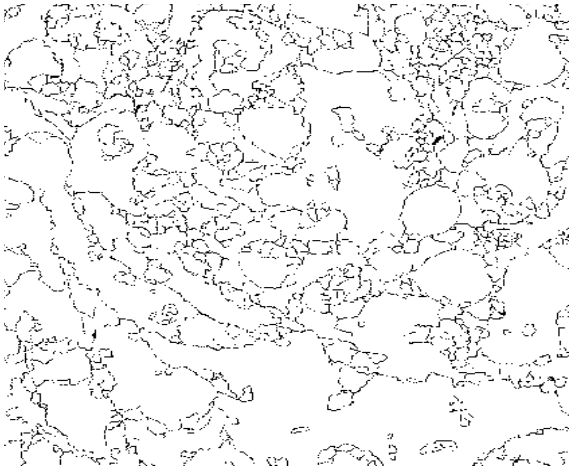


Fig.3: (a) Result of watershed algorithm

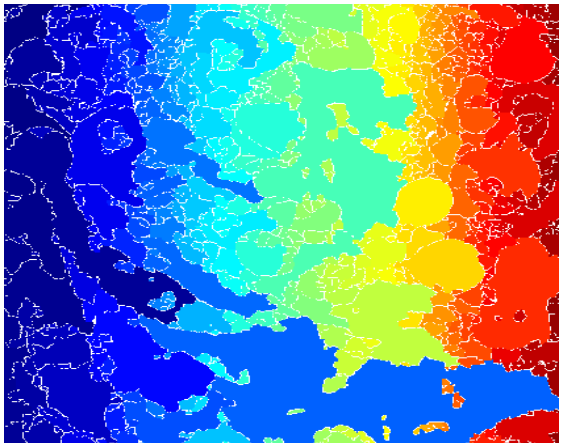


Fig.3: (b) Labeled regions

3.6 Smoothing the Image to Avoid Over-segmentation

The morphological operation opening and reconstruction is used for noise removal, the method is quite robust. But the watershed produces an over segmentation of the image. Watershed algorithm is very sensitive to noisy pictures and hence difficult to trace out relevant information. Each local minimum becomes a seed of a new region, but this minimum is not necessarily a point inside a candidate region. The advantage of the median filter is that it preserves the borders of the segments, which is beneficial to proper shape extraction. The median filter removes the impulsive noise from the image which prevents over segmentation. We considered low-pass filter and Gaussian filter applied as convolution masks. Result of smoothing is shown in figure 4.

4. RESULT ANALYSIS AND PERFORMANCE COMPARISON

In this method we used nucleus size, area, orientation to define malignancy level. Fig 5 illustrates the relationship between particles orientation and area. Fig 6 gives the nuclear size which is expressed in terms of major axis and minor axis length. The method is evaluated on a database of 96 tissue-section histological images captured under same environment such as subjective magnification, exposure and data formatted.

The results obtained for healthy and malignant of the database are summarized in table 2. In each group 6 images were taken (n=6).A1 is the result of background subtraction from the original image.A2 is the result of border clearing of image A1.So we can compare mean(A1),standard deviation(A1) with mean (A2) ,standard deviation(A2). The statistical significance i.e. Mean nuclear area and standard deviation shows small variability in values.

The diagnostic feature is nuclear area expressed in pixels. The figure 7 below shows significant difference between these values for healthy and malignant bone image. The feature of nucleus is modeled such that higher values associated with malignancy.

5. DISCUSSION

Segmentation process is the critical domain in image analysis. It is typically used to locate objects and boundaries in image. It was studied that segmentation result based on mathematical morphology cannot remove stains clearly. In the proposed method preprocessing steps includes noise removal followed by segmentation by using adaptive threshold and watershed. It gives good result but the watershed produces over segmentation of the image. Gaussian filter and median filter is used to smooth the image, which avoids over segmentation, leads to better result.

Cell nuclear morphology is a preferred and confirmatory method of research in biology. Several biological functions seem to be related to major changes in the geometry of the nucleus. Visual glance of nuclear size and shape depicts nucleus abnormality leads to quick diagnosis.

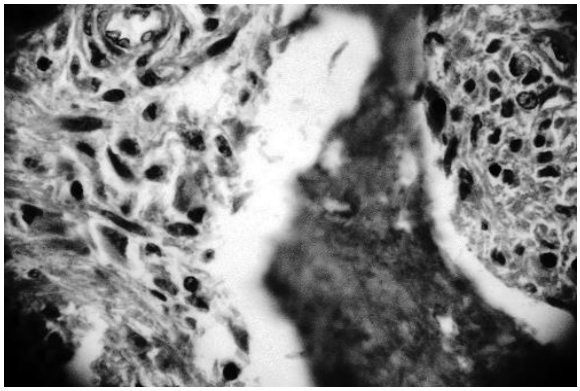
Mathematical morphology technique is applied on set of image samples. The value of nucleocytoplasmic ratio is used to detect malignancy. But the segmentation algorithm based on mathematical morphology not suitable for the image which doesn't have uniform background leads to inaccurate result. The result is shown in figure 8 (a). In figure 8 (b), noise is also treated as object of interest and it is segmented.

Application of watershed algorithm produces over segmentation, but this drawback can be overcome by post processing technique. From this method the cancerous cells are extracted from the background are used as a preliminary step before extracting cell features. The excellent segmentation results from the proposed algorithm are demonstrated with microscopic images of bone tissue under histological noise conditions. A wide variety of histological tumor type was included in this study. Selection of nuclei for analysis as performed in our study has been shown to be robust. Nuclear area across sample (table 1) indicates that statistical assessment for tumor identification was included in the study. The variation in accuracy is largely caused by the subjectivity that is inherent in visual interpretation.

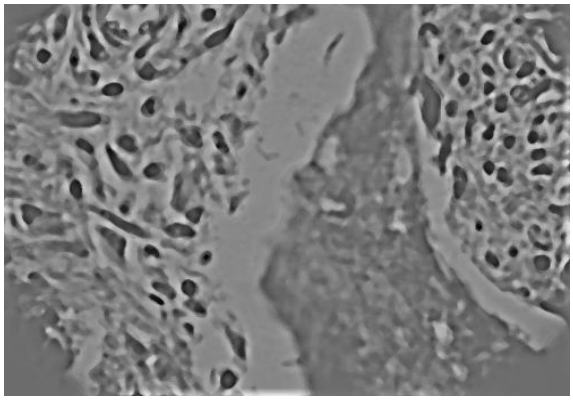
The process of manual calculation of size shape and texture features of nucleus to detect malignancy is now automated. This method provides good accurate automated diagnostic systems for bone cancer identification.

6. CONCLUSIONS

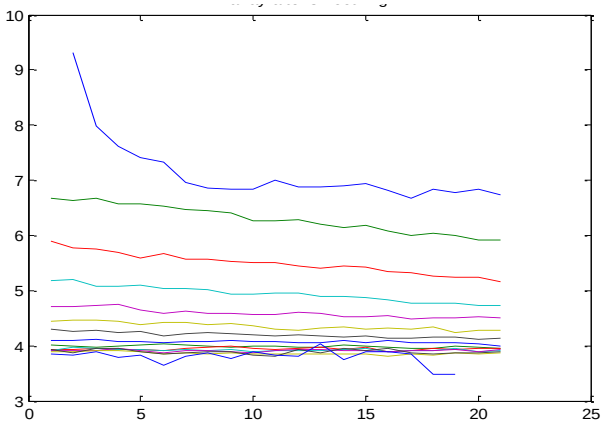
In Digital pathology can be used for accurate diagnose by overcoming the limitations of manual microscopy. Manual microscopic observation may be tedious got chances of bias. The role of the pathologist is augmented by a tool that allows



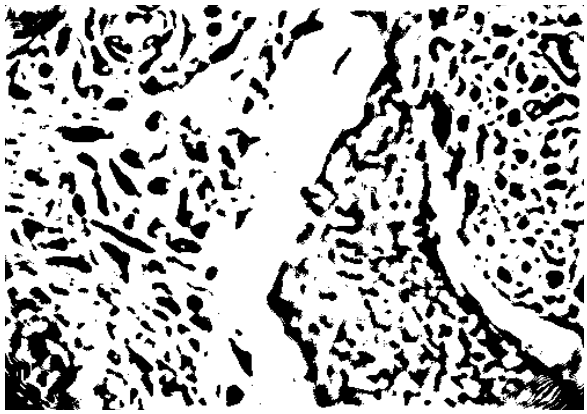
(a)



(b)

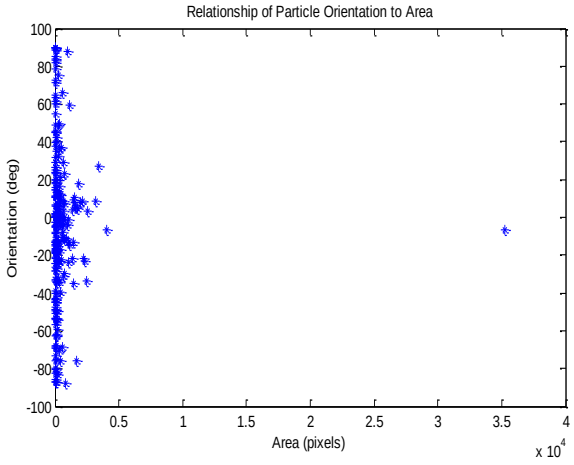


(c)

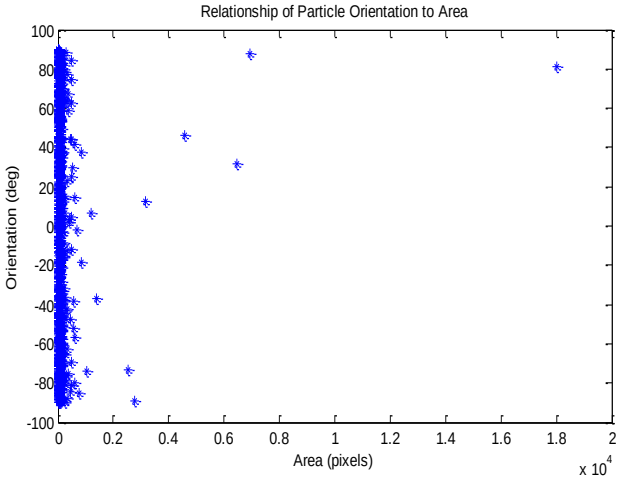


(d)

Fig.4: (a) Enhance the contrast of image. (b) Result of median followed by Gaussian filter (c) Smoothing the array. (d) Result of Segmentation.

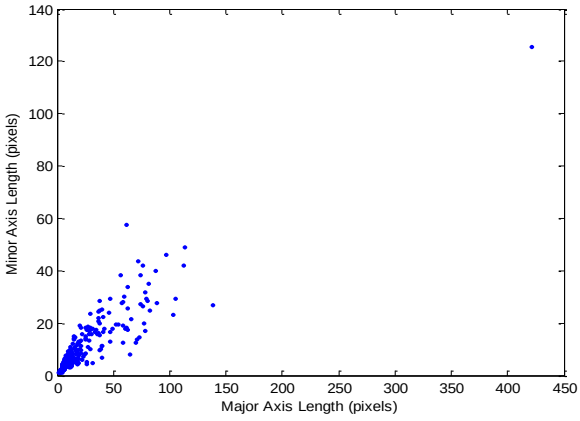


Healthy bone image

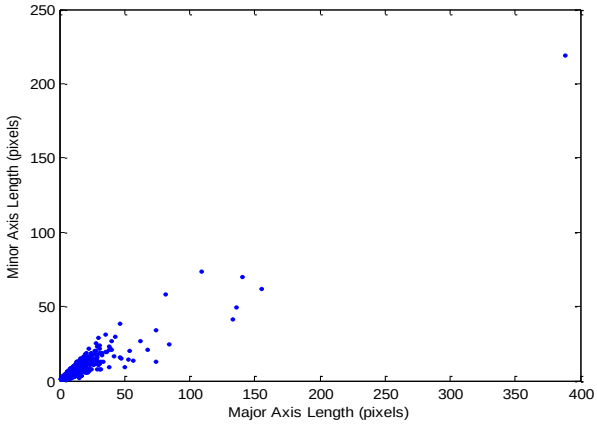


Cancerous bone image

Fig 5: Figure shows relationship of particles orientation to area



Healthy bone image



Cancerous image

Fig 6: Nucleus size is expressed in terms of major axis and minor axis length.

Table 2. Shows nucleus area for healthy and cancerous segmented image

Tumor type	Compare first order statistics; A1=original-background ; A2=original-background-border				Total area of an original image (pixels)	Total nucleus area in the segmented image(pixels)
	Mean(A1)	Mean(A2)	Standard deviation (A1)	Standard deviation (A2)		
Healthy bone tissue	0.0022	0.0020	0.0082	0.0081	389784	8731
	0.0023	0.0022	0.0077	0.0077	398672	7689
	0.0021	0.0021	0.0081	0.0081	388630	8001
	0.0020	0.0020	0.0078	0.0078	377534	7980
Osteosarcoma	0.0044	0.0044	0.0108	0.0108	384552	21016
	0.0060	0.0060	0.0154	0.0155	384975	22472
	0.0055	0.0055	0.0114	0.0115	392931	36954
Ewing's sarcoma	0.0076	0.0075	0.0258	0.0258	376704	22509
	0.0103	0.0101	0.0288	0.0280	381906	64631
	0.0144	0.0143	0.0299	0.0298	399618	30711
Chondrosarcoma	0.0011	0.0010	0.0019	0.0017	328308	11952
	0.0013	0.0012	0.0025	0.0025	419520	19342
	0.0150	0.0145	0.0423	0.0396	375396	36875
Malignant fibrous Histocytoma	0.0006	0.0006	0.0014	0.0014	379320	35758
	0.0055	0.0055	0.0114	0.0115	392931	36954
	0.0059	0.0059	0.0104	0.0104	337110	29647

perfect analysis and generates greater confidence in results. This project successfully demonstrated an automated image segmentation technique to overcome noise due to staining process from bone cancer microscopic images and provide accurate analysis of nuclear size, volume. All images were segmented successfully. Visual examination by the pathologist is still required to identify abnormal cell. However, the accuracy can be improved further by increasing our limited data size and a more accurate segmentation algorithm is required since bone tissue samples includes several cell types, and these cells are functionally different from each other.

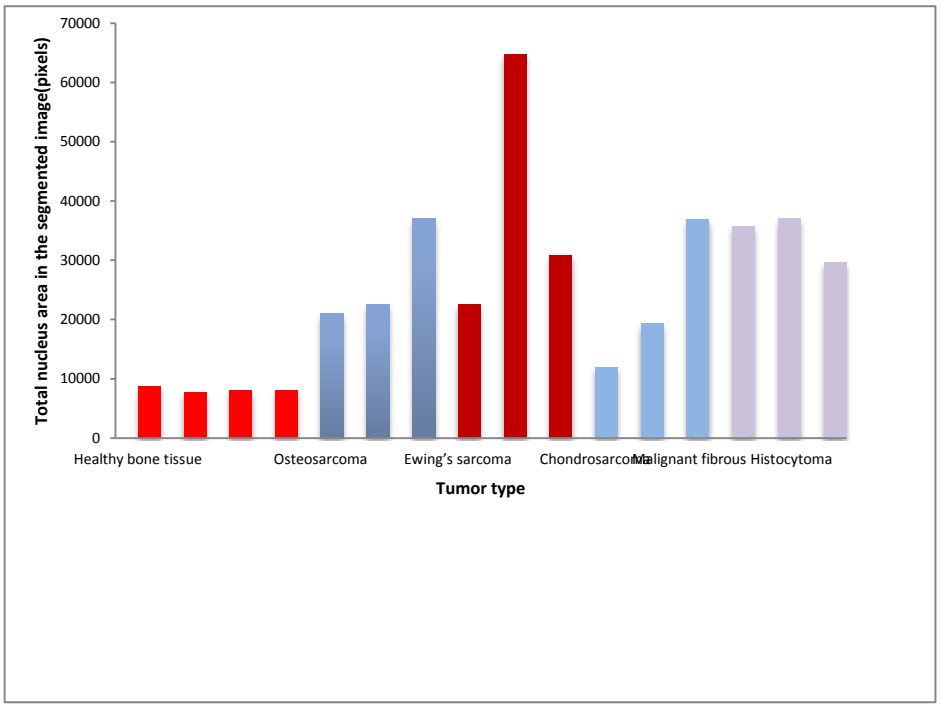
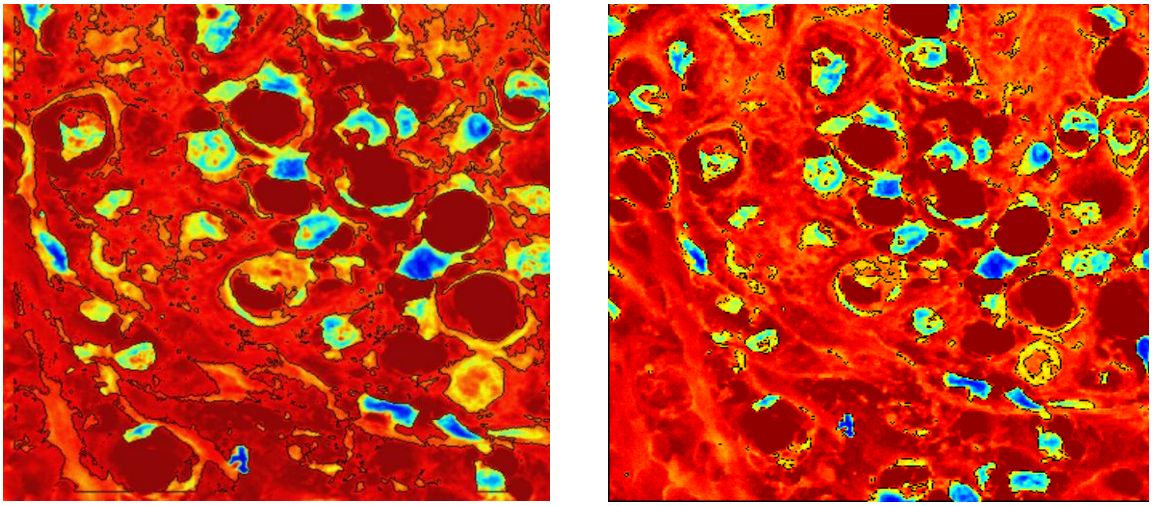


Fig 7: Bar-chart 1



(a) (b)

Fig 8: Comparative segmentation results employing a) Mathematical morphology b) Adaptive threshold and watershed

7. REFERENCES

[1] GiorgioV.Scagliotti, Management of bone metastases in Cancer, Reviews in Oncology/Hematology 56(2005):365-378.

[2] L. Rajesh, P. Dey, Fractal dimensions in urine smears: a comparison between benign and malignant cells, Anal. Quant. Cytol. Histol.25(2003) 181-182.

[3] A. J. Einstein, H.S. Wu, M. Sanchez, J. Gil, Fractal characterization of chromatin appearance for diagnosis in breast cytology, J. Pathol.185(1998) 366-381.

[4] R. Sedivy, C. Windischberger, K. Svozil, E. Moser, G. Breitenecker, Fractal analysis: an objective method for identifying atypical nucleiin dysplastic lesions of the cervix uteri, Gynecol. Oncol. 75(1999) 78-83.

[5] H.-K. Choi, T. Jarkrans, E. Bengtsson, J. Vasko, K. Wester, P.-U. Malmstrom, C. Busch, Image analysis based grading of bladder carcinoma.Comparison of object, texture and graph based methods and their reproducibility, Anal. Cell. Pathol. 15(1997) 1-18.

[6] R.F. Walker, P.T. Jackway, B. Lovell, Classification of cervical cell nuclei using morphological segmentation and textural feature extraction,Proc of the 2nd Australian

and New Zealand Conference on Intelligent Information Systems (1994) 297-301.

[7] S. Baheerathan, F. Albregtsen, H.E. Danielsen, New texture features based on complexity curve, Pattern Recogn. 32(1999) 605-618.

[8] P.W. Hamilton, P.H. Bartels, D. Thompson, N.H. Anderson, R. Montironi, Automated location of dysplastic fields in colorectal histology using image texture analysis, J. Pathol. 182(1997) 68-75.

[9] C. Demir, S.H. Gultekin, B.Yener, Learning the topological properties of brain tumors, Rensselaer Polytechnic Institute Technical Report TR-04-14, Troy, (2004).

[10] J.-P. Thiran, B. Macq, Morphological feature extraction for the classification of digital images of cancerous tissues, IEEE T. Bio-Med.Eng. 43(1996) 1011-1020.

[11] J. Smolle, Computer recognition of skin structures using discriminant and cluster analysis, Skin Res. Technol. 6(2000) 58-63.

[12] Vrushli R. Korde, M.S.,Hubert Bartels, Automatic Segmentation of Cell Nuclei in Bladder and Skin Tissue for Karyometric Analysis, Anal Quant Cytol Histol. 31(2009) 83–89.

[13] C. Gunduz, B. Yener, S.H. Gultekin, The cell graphs of cancer, Bioinformatics 20(2004) i145-i151.

[14] P.Phukpattaranont ,P. Boonyaphiphat,Automatic classification of cancer cells in Microscopic images: vol.5, no.2, (2007).

[15] S. J. Keenan, J. Diamond, W.G. McCluggage, H. Bharucha, D. Thompson, B.H. Bartels, P.W. Hamilton, An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN), J. Pathol. 192(2000) 351-362.

[16] A.N. Esgiar, R.N.G. Naguib, B.S. Sharif, M.K. Bennett, A. Murray, Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa, IEEE T. Inf. Technol. B. 2(1998) 197-203.

[17] K. Blekas, A. Stafylopatis, D. Kontoravdis, A. Likas, P. Karakitsos, Cytological diagnosis based on fuzzy neural networks, J. Intelligent Systems. 8(1998) 55-79.

[18] N. Malpica, C. Ortiz de Solorzano ,Applying watershed algorithms to the segmentation of clustered nuclei, Cytometry,28(1997) 289-297.

[19] D.K. Tasoulis, P. Spyridonos, N.G. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, M.N. Vrahatis, Urinary bladder tumor grade diagnosis using on-line trained neural networks, Proc. Knowl. Based Intell. Inform. Eng. Syst. (2003) 1999-206.