Lung Cancer Detection Using Image Processing Techniques

Mokhled S. AL-TARAWNEH

Computer Engineering Department, Faculty of Engineering, Mutah University, B.O.Box (7), Mutah 61710, Jordan
E-mail: mokhled@mutah.edu.jo

[*] Corresponding author: Phone: +962799717222 ; Fax: +962 3 2375540

Abstract

Recently, image processing techniques are widely used in several medical areas for image improvement in earlier detection and treatment stages, where the time factor is very important to discover the abnormality issues in target images, especially in various cancer tumours such as lung cancer, breast cancer, etc. 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 Gabor 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.

Keywords

Cancer Detection; Image processing; Feature extraction; Enhancement Watershed; Masking.

Introduction

Lung cancer is a disease of abnormal cells multiplying and growing into a tumour.

Cancer cells can be carried away from the lungs in blood, or lymph fluid that surrounds lung tissue. Lymph flows through lymphatic vessels, which drain into lymph nodes located in the lungs and in the centre of the chest. Lung cancer often spreads toward the centre of the chest because the natural flow of lymph out of the lungs is toward the centre of the chest. Metastasis occurs when a cancer cell leaves the site where it began and moves into a lymph node or to another part of the body through the blood stream [1]. Cancer that starts in the lung is called primary lung cancer. There are several different types of lung cancer, and these are divided into two main groups: Small cell lung cancer and non-small cell lung cancer which has three subtypes: Carcinoma, Adenocarcinoma and Squamous cell carcinomas.

The rank order of cancers for both males and females among Jordanians in 2008 indicated that there were 356 cases of lung cancer accounting for (7.7 %) of all newly diagnosed cancer cases in 2008. Lung cancer affected 297 (13.1 %) males and 59 (2.5%) females with a male to female ratio of 5:1 which Lung cancer ranked second among males and 10th among females [2]. Figure 1 shows a general description of lung cancer detection system that contains four basic stages. The first stage starts with taking a collection of CT images (normal and abnormal) from the available Database from IMBA Home (VIA-ELCAP Public Access) [3]. The second stage applies several techniques of image enhancement, to get best level of quality and clearness. The third stage applies image segmentation algorithms which play an effective rule in image processing stages, and the fourth stage obtains the general features from enhanced segmented image which gives indicators of normality or abnormality of images.
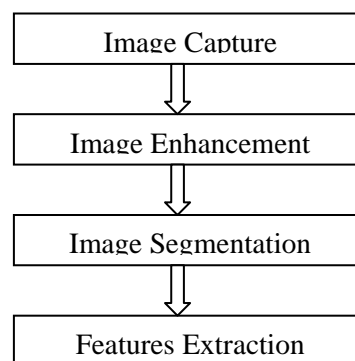
```
┌─────────────────────┐
│   Image Capture     │
└─────────────────────┘
          ⇓
┌─────────────────────┐
│  Image Enhancement  │
└─────────────────────┘
          ⇓
┌─────────────────────┐
│ Image Segmentation  │
└─────────────────────┘
          ⇓
┌─────────────────────┐
│ Features Extraction │
└─────────────────────┘
```

Figure 1. *Lung cancer image processing stages*

Lung cancer is the most dangerous and widespread cancer in the world according to stage of discovery of the cancer cells in the lungs, so the process early detection of the disease

plays a very important and essential role to avoid serious advanced stages to reduce its percentage of distribution.

The aim of this research was to detect features for accurate images comparison as pixels percentage and mask-labelling.

Material and Method

In this research, to obtain more accurate results we divided our work into the following three stages:

1. Image Enhancement stage: to make the image better and enhance it from noising, corruption or interference. The following three methods are used for this purpose: Gabor filter (has the best results), Auto enhancement algorithm, and FFT Fast Fourier Transform (shows the worst results for image segmentation).

2. Image Segmentation stage: to divide and segment the enhanced images, the used algorithms on the ROI of the image (just two lungs, the methods used are: Thresholding approach and Marker-Controlled Watershed Segmentation approach (this approach has better results than thresholding).

3. Features Extraction stage: to obtain the general features of the enhanced segmented image using Binarization and Masking Approach.

Results and Discussion

Image Enhancement

The image Pre-processing stage starts with image enhancement; the aim of image enhancement is to improve the interpretability or perception of information included in the image for human viewers, or to provide better input for other automated image processing techniques.

Image enhancement techniques can be divided into two broad categories: Spatial domain methods and frequency domain methods. Unfortunately, there is no general theory for determining what "good" image enhancement is when it comes to human perception. If it looks good, it is good. However, when image enhancement techniques are used as pre-

processing tools for other image processing techniques, the quantitative measures can determine which techniques are most appropriate [4]. In the image enhancement stage we used the following three techniques: Gabor filter, Auto-enhancement and Fast Fourier transform techniques.

Gabor Filter

Image presentation based on Gabor function constitutes an excellent local and multi-scale decomposition in terms of logons that are simultaneously (and optimally) localization in space and frequency domains [5].

A Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function. Because of the multiplication-convolution property (Convolution theorem), the Fourier transform of a Gabor filter's impulse response is the convolution of the Fourier transform of the harmonic function and the Fourier transform of the Gaussian function [6]. Figure 2 describes (a) the original image and (b) the enhanced image using Gabor Filter.
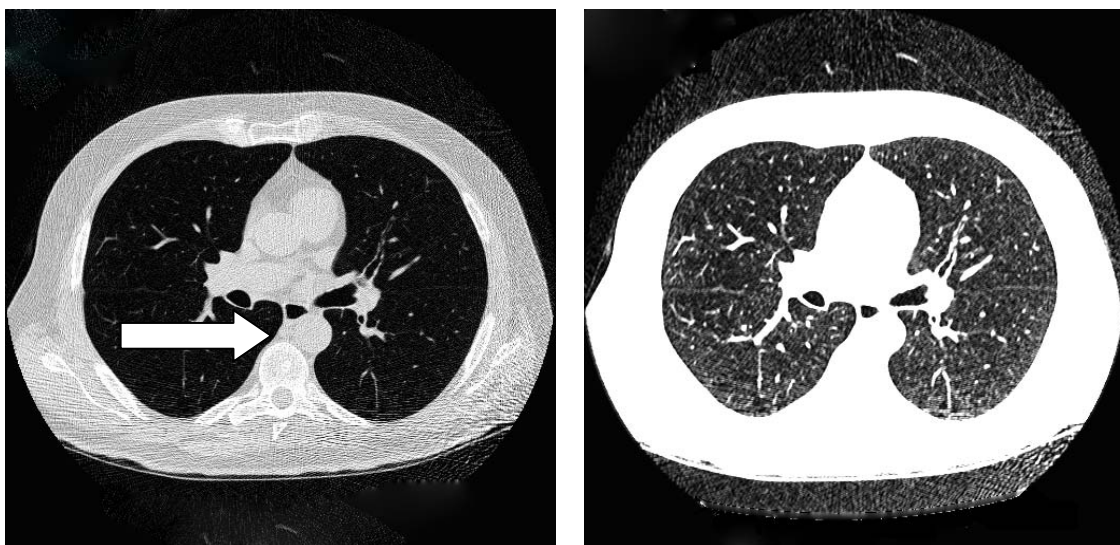
(a) Original Image   (b) Enhanced by Gabor

Figure 2. *The result of applying Gabor enhancement technique*

Auto enhancement method is strongly depends on subjective observation and statistical operations such as mean and variance calculation. The enhancement percentage in this research was equal to 38.025%.

Fast Fourier Transform

Fast Fourier Transform technique operates on Fourier transform of a given image. The frequency domain is a space in which each image value at image position F represents the amount that the intensity values in image "I" vary over a specific distance related to F. Fast Fourier Transform is used here in image filtering (enhancement). Figure 3 describes the effect of applying FFT on original images, where FFT method has an enhancement percentage of 27.51%.
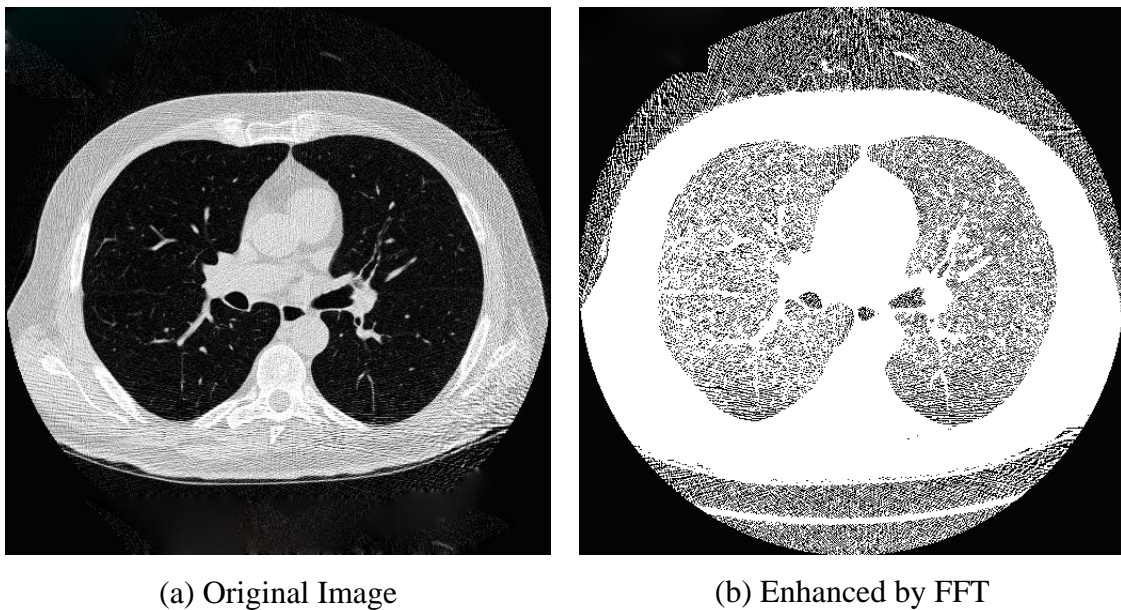

(a) Original Image   (b) Enhanced by FFT

Figure 3. Auto enhancement technique using FFT

Table 1 shows a comparison of the three mentioned techniques used for image enhancement. According to the values shown in the Table 1, we can conclude that the Gabor Enhancement is the most suitable technique for image enhancement. Observing the images enhanced by this method, we notice that new image details have appeared, in addition to good clearance and brightness shown by the enhanced images.

Table 1. Sub and final averages for three techniques used for image enhancement stage

| Subject	Auto Enhancement	Gabor Filter	FFT Filter
Sub1	37.95	80.975	27.075
Sub2	47.725	80	36.825
Sub3	36.825	79.5	25.625
Sub4	34.775	81.8	25.175
Sub5	32.85	81.4	22.85
Final Average	38.025	80.735	27.51

Image Segmentation

Image segmentation is an essential process for most image analysis subsequent tasks. In particular, many of the existing techniques for image description and recognition depend highly on the segmentation results [7]. Segmentation divides the image into its constituent regions or objects. Segmentation of medical images in 2D, slice by slice has many useful applications for the medical professional such as: visualization and volume estimation of objects of interest, detection of abnormalities (e.g. tumours, polyps, etc.), tissue quantification and classification, and more [8]. The goal of segmentation is to simplify and/or change the representation of the image into something that is more meaningful and easier to analyse. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics [9]. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image (edge detection). All pixels in a given region are similar with respect to some characteristic or computed property, such as colour, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

Segmentation algorithms are based on one of two basic properties of intensity values: discontinuity and similarity. The first category is to partition the image based on abrupt changes in intensity, such as edges in an image. The second category is based on partitioning the image into regions that are similar according to a predefined criterion. Histogram thresholding approach falls under this category.

Thresholding approach

Thresholding is one of the most powerful tools for image segmentation. The segmented image obtained from thresholding has the advantages of smaller storage space, fast processing speed and ease in manipulation, compared with gray level image which usually contains 256 levels. Therefore, thresholding techniques have drawn a lot of attention during the past 20 years [10]. Thresholding is a non-linear operation that converts a gray-scale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold value. In this research, Otsu's method that uses (gray thresh) function to compute global image threshold is used. Otsu's method is based on threshold selection by statistical criteria. Otsu suggested minimizing the weighted sum of within-class variances of

the object and background pixels to establish an optimum threshold. Recalling that minimization of within-class variances is equivalent to maximization of between-class variance. This method gives satisfactory results for bimodal histogram images. Threshold values based on this method will be between 0 and 1, after achieving the threshold value; image will be segmented based on it. Figure 4 shows the result of applying thresholding technique.
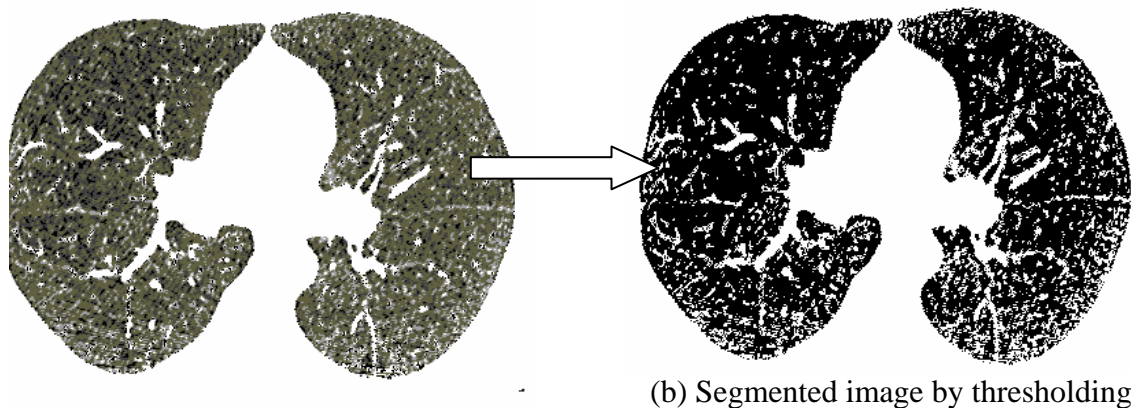
(a) Enhanced image by Gabor (b) Segmented image by thresholding

Figure 4. *Normal enhanced image by Gabor filter and its segmentation using thresholding approach*

Marker-Controlled Watershed Segmentation Approach

Marker-driven watershed segmentation technique extracts seeds that indicate the presence of objects or background at specific image locations. Marker locations are then set to be regional minima within the topological surface (typically, the gradient of the original input image), and the watershed algorithm is applied [11]. Separating touching objects in an image is one of the most difficult image processing operations, where the watershed transform is often applied to such problem. Marker-controlled watershed approach has two types: External associated with the background and Internal associated with the objects of interest.

Image Segmentation using the watershed transforms works well if we can identify or "mark" foreground objects and background locations, to find "catchment basins" and "watershed ridge lines" in an image by treating it as a surface where light pixels are high and dark pixels are low. Figure 5 shows a segmented image by watershed.
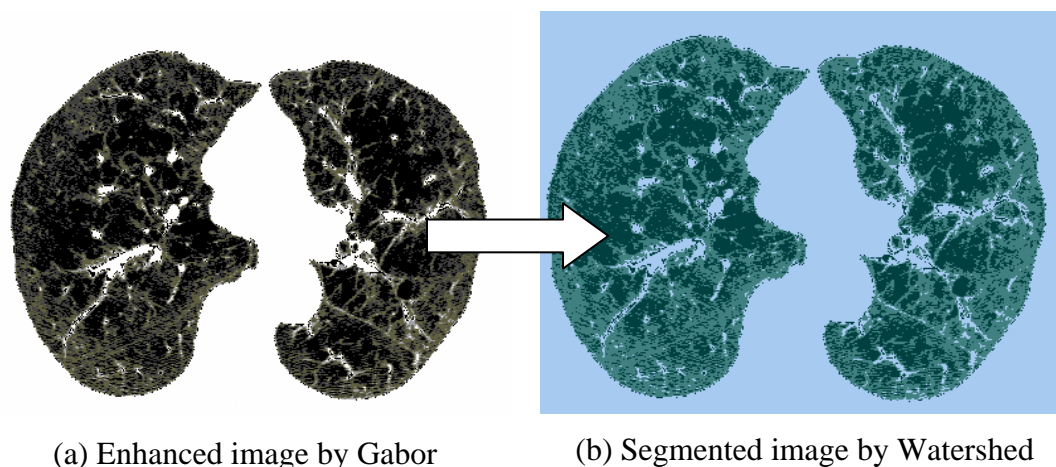
(a) Enhanced image by Gabor   (b) Segmented image by Watershed

Figure 5. *Normal Enhanced image by Gabor filter and its Segmentation using Marker-Controlled Watershed approach*

According to the experimental subjective assessment during the segmentation stage as shown in table 2, Marker-Controlled Watershed Segmentation approach has more accuracy (85.165%) and quality than Thresholding approach (81.835%).

Table 2. Image Segmentation experimental result

Subject	Thresholding	Watershed Filter
Sub1	81.625	85.375
Sub2	82.2	85.25
Sub3	82.125	85.55
Sub4	81.725	84.75
Sub5	81.5	84.9
Final Average	81.835	85.165

Features Extraction

Image features Extraction stage is an important stage that uses algorithms and techniques to detect and isolate various desired portions or shapes (features) of a given image. To predict the probability of lung cancer presence, the following two methods are used: binarization and masking, both methods are based on facts that strongly related to lung anatomy and information of lung CT imaging.

Binarization Approach

Binarization approach depends on the fact that the number of black pixels is much greater than white pixels in normal lung images, so we started to count the black pixels for normal and abnormal images to get an average that can be used later as a threshold, if the

number of the black pixels of a new image is greater that the threshold, then it indicates that the image is normal, otherwise, if the number of the black pixels is less than the threshold, it indicates that the image in abnormal. The threshold value that is used in this research is 17178.48 and the True acceptance rate (TAR) is (92.86%) and False acceptance rate (FAR) is (7.14%). Figure 6 shows the binarization method procedure and figure 7 shows binarization check method flowchart.
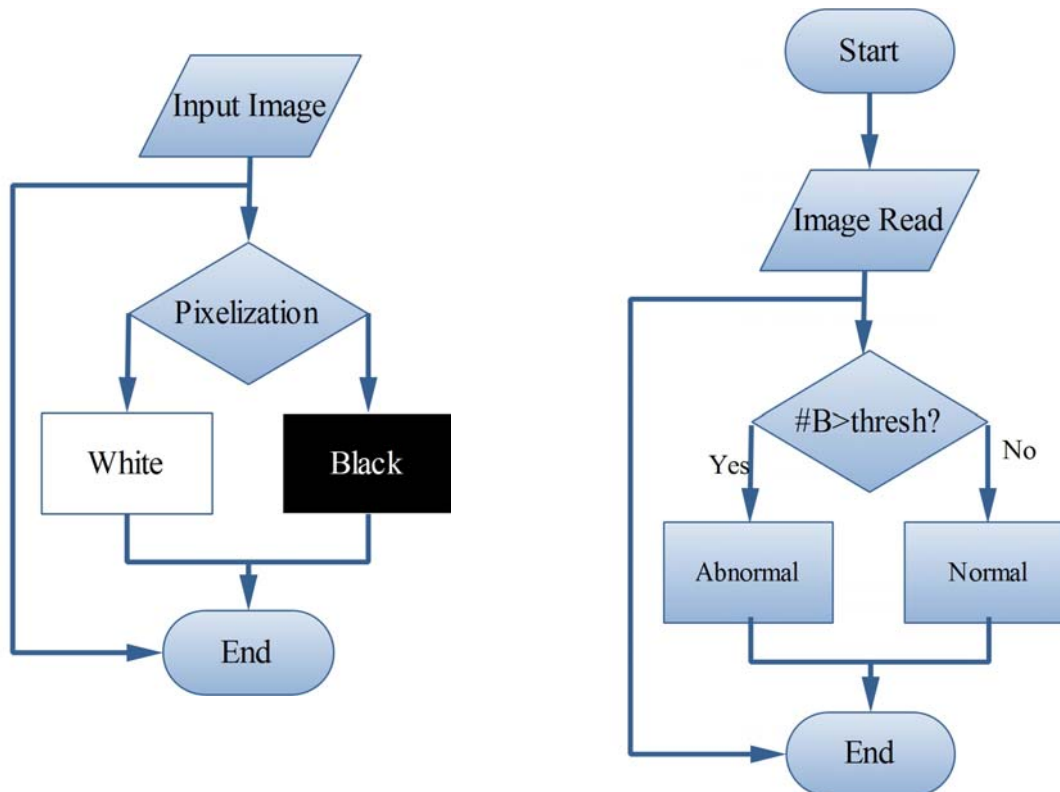

Figure 6. *Binarization method procedure*   **Figure 7.** *Binarization check method flowchart*
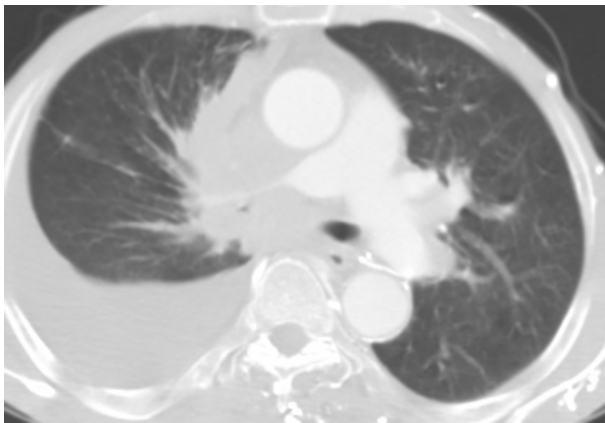
Masking Approach

Masking approach depends on the fact that the masses are appeared as white connected areas inside ROI (lungs), as they increase the percent of cancer presence increase. The appearance of solid blue colour indicates normal case while appearance of RGB masses indicates the presence of cancer, the TAR of this method is (85.7%) and FAR has (14.3%). Figure 8 shows normal and abnormal images resulted by implementing Masking approach using MATLAB.
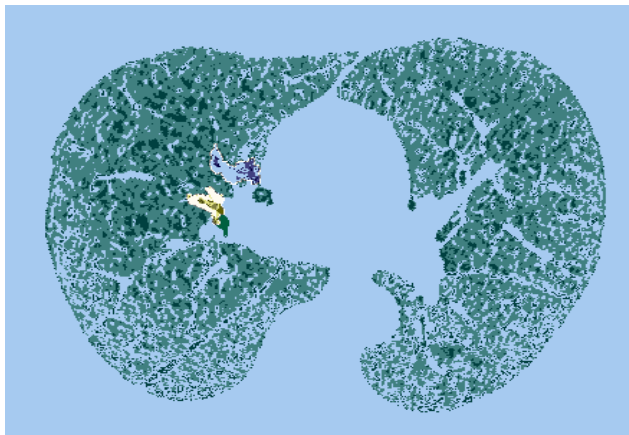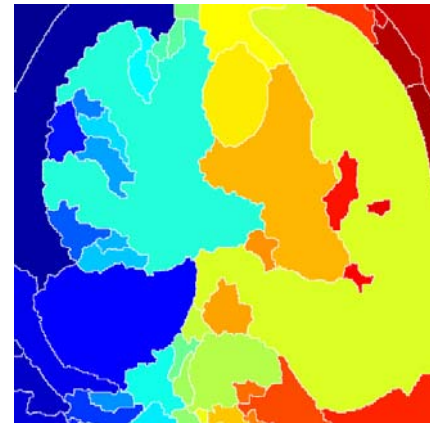
(a) Normal Image enhanced by Gabor, segmented by watershed
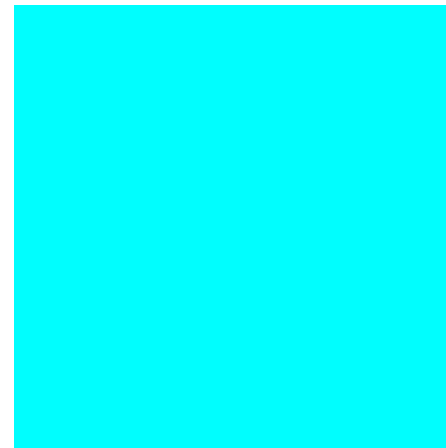
(b) The resulted image indicates normality

(c) Abnormal image

(d) The resulted indicates abnormality

Figure 8. *Normal and abnormal images using Masking approach*

Combining Binarization and Masking approaches together will lead us to take a decision whether the case is normal or abnormal according to the mentioned assumptions in the previous two approaches, we can conclude that image that has number of black pixels greater than white ones, indicates normality, and otherwise it indicates abnormality.

Conclusions

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 Gabor filter within Gaussian rules. The proposed technique is efficient for segmentation principles to be a region of interest foundation for feature extraction obtaining. The proposed technique gives very promising results comparing with other used techniques. Relying on general features, a normality comparison is made. The main detected features for accurate images comparison are pixels percentage and mask-labelling with high accuracy and robust operation.

References

1. *Non-Small Cell Lung Cancer*, Available at: http://www.katemacintyrefoundation.org/ pdf/non-small-cell.pdf, Adapted from National Cancer Institute (NCI) and Patients Living with Cancer (PLWC), 2007, (accessed July 2011).

2. Tarawneh M., Nimri O., Arqoub K., Zaghal M., *Cancer Incidence in Jordan 2008*, Available at: http://www.moh.gov.jo/MOH/Files/Publication/Jordan%20Cancer%2 0Registry_2008%20Report_1.pdf, 2008, (accessed July 2011).

3. *Lung Cancer Database,* Available at: https://eddie.via.cornell.edu/cgi-bin/datac/signon.cgi, (accessed July 2011).

4. Gonzalez R.C., Woods R.E., *Digital Image Processing*, Upper Saddle River, NJ Prentice Hall, 2008.

5. Cristobal G., Navarro. R., *Space and frequency varient image enhancment based in Gabor representation*, Pattern Recognition Letters, Elsevier, 1994, 15, p. 273-277.

6. Krishan A., *Evaluation of Gabor filter parameters for image enhancement and segmentation,* in Electronic Instrumentation and Control Engineering, Master. Punjab: Thapar University, 2009, p. 126.

7. Nunes É.D.O., Pérez M.G., *Medical Image Segmentation by Multilevel Thresholding Based on Histogram Difference*, presented at 17th International Conference on Systems, Signals and Image Processing, 2010.

8. Venkateshwarlu K., *Image Enhancement using Fuzzy Inference System*, in Computer Science & Engineering, Master thesis, 2010.

9. Shapiro L.G., Stockman G.C., *Computer Vision: Theory and Applications,* Prentice Hall, 2001.

10. Huang Q., Gao W., Cai W., *Thresholding technique with adaptive window selection for uneven lighting image*, Pattern Recognition Letters, Elsevier, 2004, 26, p. 801-808.

11. Levner I., Zhang H., *Classification-Driven Watershed Segmentation*, IEEE Transactions on Image Processing, 2007, 16(5), 1437-45.