Mass Detection in Lung CT Images using Region Growing Segmentation and Decision Making based on Fuzzy Systems

Hamid bagherieh\textsuperscript{1}

\textsuperscript{1} Department of computer, Islamic Azad University Hamedan branch, Iran
h.bagheri@iauh.ac.ir

Atiyeh Hashemi\textsuperscript{2}(Corresponding Author)

\textsuperscript{2} Department of computer, Islamic Azad University Hamedan branch, Iran
a.hashemi@iauh.ac.ir

Abdol Hamid Pilevar\textsuperscript{3}

\textsuperscript{3} Department of Computer Engineering, Bu-Ali Sina University Hamedan, Iran
pilevar@basu.ac.ir

Abstract — Lung cancer is distinguished by presenting one of the highest incidences and one of the highest rates of mortality among all other types of cancers. Detecting and curing the disease in the early stages provides the patients with a high chance of survival. In order to help specialists in the search and recognition of the lung nodules in tomography images, a good number of research centers have been developed in computer-aided detection (CAD) systems for automating the procedures. This work aims at detecting lung nodules automatically through computerized tomography images. Accordingly, this article aim at presenting a method to improve the efficiency of the lung cancer diagnosis system, through proposing a region growing segmentation method to segment CT scan lung images and, then, cancer recognition by FIS (Fuzzy Inference System).

The proposed method consists of three steps. The first step was pre-processing for enhancing contrast, removing noise, and pictures less corrupted by Linear-Filtering. In second step, the region growing segmentation method was used to segment the CT images. In third step, we have developed an expert system for decision making which differentiates between normal, benign, malignant or advanced abnormality findings. The FIS can be of great help in diagnosing any abnormality in the medical images. This step was done by extracting the features such as area and color (gray values) and given to the FIS as input. This system utilizes fuzzy membership functions which can be stated in the form of if-then rules for finding the type of the abnormality. Finally, the analysis step will be discussed and the accuracy of the method will be determined. Our experiments show that the average sensitivity of the proposed method is more than 95%.

Index Terms — Segmentation, Fuzzy Systems, lung cancer, tumor markers.

I. INTRODUCTION

Lung cancer is also called cancer of the bronchus which is produced as a result of uncontrolled growth of the lung tissues, especially the cells which line the air passages. The resulting cells will not develop into healthy ones; they divide to form tumors which are considered as the main cause of death from cancers. The symptoms begin to appear after almost a long period which explains why the cancer is difficult to detect [1].

Medical statistics shows that the number of the cases breast, prostate and colon cancer is totally less those of lung cancer [2]. More than one third of the British population will be diagnosed with the cancer in their lifetime and one fourth of the cases will die from the disease [3]. Approximately 60% of patients are men and 33% of all death-leading cancers in men are because of lung cancer. However, awareness of the common symptoms of this disease makes the treatment possible in the primary phases. A type of disease in which the malignant tissue grows in one or both lungs, lung cancer kills more than 80 % of patients within 5 years of the diagnosis. Many of them notice their disease when it’s too late and the surgery is not simply possible. In USA, lung cancer has been the main cause of death from cancers with 14% of cases in 2011, and the major cause of 27% of the death from cancers. In the recent three decades, we have failed to increase much to the live of the patients suffering from the cancer. While, in the late 1970s, around 37% of patients had the chance of survived for one year or longer after the diagnosis, today 42% of them have the chance [4-7].

Today several techniques are used to diagnosis of the disease, such as Computerized Tomography (CT), Chest Radiograph (x-ray), Magnetic Resonance Imaging (MRI scan). Nevertheless, these techniques can detect the disease only in its advanced stages, which leads to the death of a number of the patients. Hence, we need the
Mass Detection in Lung CT Images using Region Growing Segmentation and Decision Making based on Fuzzy Systems

In recent study, three tracking algorithms are given to find the best path between the nodes of the graph and a comparison is made with others. The study concludes that the method based on the Minimum Cost Path (MCP) method is the best and creates the most optimal route [4]. Despite the great need for such a technique to evaluate large 3D datasets, unfortunately little study has been done in this area. Another study has been done to segment the wire frame of the stent in 2D slices [11]. In another study the stent’s frame is found by means of a statistical model. The researcher report that the threshold at 2000 Hounsfield units is appropriate to obtain a graph consisting of points related to the stent. Unfortunately, due to difference in the applied doses or the stent frame materials, our data had a lot of noise. In this regard, further studies have been carried out to segment blood vessels in 3D, which have similarities in common with the wire of the frame of the stent. These investigations may be importance to our study (see Lesage [6-8]). The two-step approach is one such method [12-15], segment the vessel by a vessel measure [16] and then centerline extraction. Methods which fit a number of ellipsoids or spheres to the vessel, and methods which segment the contour in slices perpendicular to the vessel centerline suppose a solid vessel with a diameter of several vertices. Due to the small diameter of the stent’s wire (1–3 vertices) and its sharp corners, these methods are also inappropriate. Region growing methods have problems with leaks and gaps and need to be followed by a second stage to find the geometry from the segmented vertices. Another study focuses on two types of stent grafts commonly used in our work to treat patients with AAA made of nitinol and stainless steel, respectively. This article uses endovascular aortic replacement (EVAR) and stent grafts [8]. In another study a complete system for computer-aided detection of GGNs is presented consisting of initial segmentation steps, candidate detection, feature extraction and a two-stage classification process. A rich set of intensity, shape and context features is constructed to describe the appearance of GGN candidates and apply a two-stage classification approach using a linear discriminative classifier and a gentle boost classifier to classify candidate regions efficiently [17]. Another article a computer-aided detection (CAD) system has been developed for identification of small pulmonary nodules in low-dose and thin-slice CT scans. And the FROC curve obtained on this data set exhibits high sensitivity to lung nodules (80-85% range) at an acceptable level of false positive findings as for each patient [18]. This study analyzes four geo-statistical functions- semivariogram, semimadogram, covariogram, and correlogram. The aim was to characterize lung nodules as malignant or benign in computerized tomography images. Stepwise discriminant analysis was applied to see which measures could distinguish the benign and malignant nodules. ROC curve was presented for evaluating the results with an accuracy between good and excellent [19]. In this study a computer assisted automatic diagnostic system was used for lung cancer. The system consists of two steps: analytic and diagnostic. In the analysis step, it extracts the lung and the blood vessel regions by the fuzzy clustering algorithms, analyzes the features of these. In the diagnosis step, the system defines diagnostic rules using the extracted features which support the determination of the nodule locations [20]. Another study tries to enhance the efficiency of the lung cancer diagnosis system by proposing the segmentation of the abnormal lung nodules, and this is carried out by region based segmentation and cancer identification using FIS. The accuracy of this proposed system is estimated 92.3% [9]. Another study applied the same method and arrived...
at an accuracy of 90% [21]. In the first step, the input image was preprocessed and the cancerous nodule areas were segmented and in the second step, involved the diagnosis of the nodal points based on the fuzzy system. Another researcher [1] presented two segmentation methods, FCM (Fuzzy C-Mean) and HNN (Hopfield Neural Network clustering algorithm. These two methods are designed to classify the image of X pixels among Y classes. The results showed that the HNN segmentation results are more reliable than FCM clustering. Another researcher [3] is formed FCM as one of the most applicable fuzzy clustering algorithms. Nevertheless, this paper mentions two limitations with the methods: a predefined number of clusters which have to be given in advance and the FCM property which can be get stuck in sub-optimal solutions. Another researcher [22] attempt to evaluate the diagnostic power of a fuzzy classifier and a marker panel for the detection of lung cancers in comparison to asbestosis patients at high-risk of developing lung cancer. Finally the positive predictive value was 77.7%, and the negative predictive value amounted to 94.8%. Conclusion: With the fuzzy classifier and a marker panel, a reliable diagnostic tool for the detection of lung cancers in a high risk population is available. Another article [23] presents a genetics-based machine learning approach that performs feature extraction on data from a lab to help increase the classification performance of an existing classifier. The classifier was built using the data from a different laboratory which uses the same protocols, while learning about the shape of the fractures between data that motivate the bad behaviors.

III. THE PROPOSED METHOD

The main purpose of the suggested model is to diagnose the lung cancer diseases trough image segmentation and decision making by using Fuzzy Inference System (FIS) or fuzzy rules. The methodology proposed for lung nodule detection consists of the acquisition of computerized tomography images of the lung.

The method starts with pre-processing the image to enhance contrast and remove noise in order to make the image suitable for further processing. Linear filtering was used to remove the noise, in this study.

After pre-processing the process of image segmentation was accomplished. The process of segmentation of the lung images is also a result of the region growing algorithm.

Next, decision making was done by using FIS to classify tumors (abnormal areas) as normal, benign, malignant or advanced. Figure 1 shows the steps involved in the diagnostic process.

A. Pre-Processing

The images were pre-processed to adjust the colors for better segmentation. Pre-processing deals with techniques for increasing contrast and removing undesirable noises [21]. Filtering is a neighborhood operation, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel. In this article Linear-Filtering technique is used for enhancement of the input CT images.

Linear filtering is a type of filtering in which the value of the output pixel is a linear combination of the values of the pixels in the input pixel's neighborhood.
So that, \( B = \text{Linear-Filtering} (A, H) \) filters the multidimensional array \( A \) with the multidimensional filter \( H \). The array \( A \) can be a member of any class or dimension and can also be logical or an array of any non sparse numeric type. The result \( B \) is the same as \( A \) in terms of size and class. After preprocessing, the image segmentation is done by using region growing segmentation.

**B. Image segmentation**

Segmentation refers to the process of partitioning a digital image into different classes, based on the application, different methods have been used for image segmentation, such as edge detection, histogram threshold, ANN, stochastic models, region growing and clustering techniques [10]. In medical imaging, segmentation has a major role, since it helps in the delineation of anatomical structures. A region growing based approach is utilized in segmentation process in the existing works. Region Growing is an approach in which pixels neighbors are examined and added to a region class as long as no edges are found between them [21]. We start with segmentation by finding boundaries through finding coherent regions.

This approach has certain merits over boundary based methods [24]:

- It is guaranteed to produce coherent regions.
  - Linking edges and gaps produced by missing edge pixels, and the like, are not an important issue.
- It works from the inside to the outside. Notice that, selecting of the initial point is important.
  - However, it also has disadvantages.
- Choosing region membership is more difficult than applying edge detectors.
- It can’t search objects that span different disconnected regions.

We start with a pixel \( p \) and want to expand from that pixel to fill a coherent region. Let’s define a certain measure \( D(i, j) \) which it produces a high result, if pixels \( x \) and \( y \) are similar and a low one otherwise. First, consider a pixel \( m \) adjacent to pixel \( n \). We can add pixel \( m \) to the region of pixel \( n \), if \( D(n, m) < T \) for some threshold \( T \). We can also proceed to the other neighbors of \( n \). Suppose that \( D(n, m) < T \), and we can added pixel \( m \) to the region of pixel \( n \). We can now also consider the neighbors of \( m \) and add them again if they are similar.

1. The proposed algorithm
   The learning algorithm of the proposed algorithm consists of 7 phases:
   - Phases 1: In the region growing segmentation, the first aim is to determine the initial seed points and threshold value.
   - Phases 2: The size of the image is calculated for checking the condition of the loop.
   - Phases 3: Compare the first initial seed point with neighboring pixels according to threshold value.
   - Phases 4: We can add neighboring pixels to the region, if these are similar to seed point. For example, we can add pixel \( q \) to the region of pixel \( p \), if \( (p-q) < T \) for some threshold \( T \).
   - Phases 5: Region growing should stop when no more pixels are found to meet the criteria for inclusion in that region. Then, another seed point is selected that does not belong to any other region.
   - Phases 6: This process is continued until all pixels (equal to the size of the image) belong to one region.
   - Phases 7: Finally, the area of the regions is calculated. Indeed, segmentation is a process that partitions an image like \( R \) into ‘\( n \)’ sub regions, like \( R_1, R_2, R_3, \ldots, R_n \), such that:

\[
\bigcup_{i=1}^{n} R_i = R
\]

\[
R_i \cap R_j = \emptyset \text{ for all } i \text{ and } j \text{, } i \neq j
\]

The first condition indicates that the segmentation is completed, that occurs when each pixel belongs to a region. The second condition indicates that the region must be disjoint. The presented region growing segmentation algorithm is applied for four different lung images and the outcome results are displayed in Figure 3.

![Figure 3](image-url)

**Figure 3. The results of the region growing segmentation algorithm**

**C. Decision making based on fuzzy systems**

Clustering is the process of dividing the data into sub classes. Different kinds of clustering algorithms have been developed in the literature. Fuzzy Clustering has been utilized in various fields such as pattern recognition and Fuzzy identification.
Diagnosis plays a major role in medical care. The aim of this step is to determine the diagnostic power of a fuzzy classifier and a marker panel for detecting lung cancers in comparison to asbestosis patients at high-risk of developing lung cancer [22].

A FIS is regarded as a method of mapping an input space to an output one with the aid of fuzzy logic (by IF-THEN rules). FIS can be designed by using different architectures to enhance the performance of the system. Three points must be considered: finding the optimal number of the rules, selecting the appropriate membership functions, and tuning of both [25].

Figure 4 illustrates the architecture of our designed FIS. This Fuzzy system has two inputs of ‘color’ and ‘area’, and one output of ‘diagnoses’.

As is shown in Figure 5, the range of the ‘color’ is taken from 1 to 255 which are separated into two sections, ‘low’ and ‘high’ that are in the form of triangle and trapeziums. So, the range of the ‘area’ is taken from 1 to 5000 that are also separated into two sections, ‘low’ and ‘high’, in the form of trapeziums. the range of the output ‘diagnosis’ is taken from 1 to 2 that are again separated into four sections ‘None’, ‘Benign’, ‘Malignant’ and ‘Advanced’, in the form of triangle, as shown in Figure 6.

One of the most important problems in fuzzy clustering algorithms is how to design membership functions. According to research done, one comes across with some the membership functions (IF-THEN rules) that are used to detect the lung nodules are shown in Figure 7 and the rule viewer in Figure 8.

The rule viewer (Figure 8) shows that the given CT scan image is of the malignant type, because the output is 1.75, which is in the range of 1.5 to 2. The area input of fuzzy inference system is 2800 (high) and the color (grey value) is 127.5 (low), hence, the output is of malignant type.

In this article, an expert system has been developed which works with fuzzy rules. An expert system is a system which simulates the knowledge of a human expert. An expert system:

- Works with a large database of knowledge.
- Allows the database to be accessed.
- Has a set of rules and, with its inference engine, makes deductions [26].

For example, in this research, expert systems are being used for lung cancer diagnosis. The details of the image such as color and area are input, and the output of the system is the diagnostic type of the tumor, like Normal, Benign, Malignant or Advanced.

The system cannot to get facts wrong or forget. The computer can have access to the expertise that a doctor may not possess or may be wrong.

As is shown in Figure 4, the fuzzy system has three layers. The first layer displays the input variables. The second displays fuzzy rules and the third displays the output variables. Hence, Fuzzy inference system is used to recognize the nodule based on fuzzy membership functions, the input and output. This Fuzzy system (in our article) has two inputs, ‘color’, and ‘area’ with one output, ‘diagnosis’, and the rule base containing four fuzzy if-then rules.
IV. CONCLUSIONS AND PROSPECTS

In this section, we present the results obtained from our proposed method. To prepare the image for segmentation, pre-processing was carried out by contrast enhancement and linear filtering. Linear-filtering was used for noise removal. After pre-processing of the image, image segmentation was done by using region growing segmentation for detection of lung cancer (abnormal area in images). The method was employed to segment an image into different regions, using a set of seeds. The results of segmentation are shown in Figure 5. As we know, Diagnosis plays a major role in medical care. Also differentiating between normal, benign, malignant and advanced abnormality is difficult. The aim of this work is to minimize the human errors in catching and interpreting points that may be neglected by the radiologist. The FIS technology was utilized for this purpose. This step was done by extracting the features such as area and color (gray values) and was given as the input to the FIS. The Fuzzy system employs fuzzy membership functions that may be expressed in the form if-then rules to find the type of the abnormality. This method is applied to 21-tumor contained 250 CT slices from 21 lung tumor patients.

The experimental results for 4-tumor images are shown in Table 1. The accuracy of the proposed system in calculation of the masse area was more than 95%, which means that the proposed method is efficient and can help the radiologists to increase their diagnostic confidence.

In the future, we are trying to offer the more accurate method for image segmentation. Also we can focus on the appropriate filtering method and better membership functions in FIS.

<table>
<thead>
<tr>
<th>Image</th>
<th>Segmented image</th>
<th>area</th>
<th>gray level</th>
<th>Decision</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>4600</td>
<td>135</td>
<td>1.75 (Advanced)</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4200</td>
<td>85</td>
<td>1.47 (Malignant)</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>986</td>
<td>70</td>
<td>0.25 (None)</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2031</td>
<td>120</td>
<td>0.75 (Benign)</td>
<td>7</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENT

This article is the result of a research project entitled "Mass detection in lung CT images using region growing segmentation and decision making based on Fuzzy Inference System and artificial neural network" which was supported by the Islamic Azad University of Hamedan branch (Iran). We thank from responsible of related university that have the necessary cooperation.

REFERENCES

Mass Detection in Lung CT Images using Region Growing Segmentation and Decision Making based on Fuzzy Systems


Hamid bagherieh was born in Hamedan, Iran on Mar 17, 1978. He received the B.Sc. degrees from Islamic Azad University of Lahijan (Iran), in 2003. In 2011, he joined the Islamic Azad University, Department of Computer Engineering, in Malayer, Iran as a student. His interesting areas are image processing, Fuzzy Inference System (FIS), artificial neural network and Genetic Algorithms.

Atiyeh Hashemi was born in Hamedan, Iran on June 18, 1988. She received the B.Sc. degrees from PNU University of Hamedan (Iran), in 2010 and received M.Sc. degrees from Islamic Azad University, Department of Computer Engineering, in Malayer, Iran, in 2013. Her interesting areas are image processing, Fuzzy Inference System (FIS), Neuro Fuzzy Inference Systems (ANFIS) and artificial neural network.

Abdol Hamid Pilevar, Assistant professor in Computers Engineering Department, Bu Ali Sina University, Hamedan, Iran. He received his B.Sc. and MSc. degrees in Computer Systems from Florida Atlantic University, Florida, U.S.A. Dr. Pilevar received the PhD degree in Computer Science from the University of Mysore, Mysore, India in 2005. He was a Research Associate and Post-Doctoral Fellow at the Indian Institute of Science (Bangalore, India), and Mediscan Prenatal Diagnosis and Fetal Therapy Center (Chennai, India) in 2005-2006. His fields of interest include Medical Intelligence and Image Processing, 3D Modeling, and Speech and Natural Language Processing. He has published more than 60 articles in international and national journals, and conferences.

How to cite this paper: Hamid bagherieh, Atiyeh Hashemi, Abdol Hamid Pilevar, "Mass Detection in Lung CT Images using Region Growing Segmentation and Decision Making based on Fuzzy Systems", IJIGSP, vol.6, no.1, pp.1-8, 2014.DOI: 10.5815/ijigsp.2014.01.01