

LUNG NODULE SEGMENTATION IN COMPUTED TOMOGRAPHY IMAGES FOR PREOPERATIVE PLANNING AND DIAGNOSIS PURPOSES

M V SUDHAMANI¹, G T RAJU² & BHAVANISHANKAR K³

¹Professor, Department of ISE, RNS, Institute of Technology, Bangalore, Karnataka, India

²Professor, Department of CSE, RNS Institute of Technology, Bangalore, Karnataka, India

³Assistant Professor, Department of CSE RNS Institute of Technology, Bangalore, Karnataka, India

ABSTRACT

The precise segmentation of lung nodule is of great significance to identify tumors for the diagnosis of the lung cancer. Computed tomography (CT) is powerful evaluation technique for producing 2D cross sectional images of the organs under test by providing characteristics of the internal structure of an organ. Here, Segmentation concept is used to analyze the nodule in the CT lung image. In this proposed work, the lung CT image is segmented from which nodule is isolated and then classified as either benign or malignant. Segmentation is done using optimal threshold technique which is followed by Region of Interest (ROI) extraction. Then texture features of this ROI have been extracted using Gray Level Co-occurrence Matrix (GLCM) and classification is performed on these features. Experiments have been conducted to classify the nodules as benign or malignant. The proposed approach has fetched an accuracy of above 95%. CT images required for the work are collected from the diagnostic center.

KEYWORDS: CT Images, Segmentation, Lung Nodules

INTRODUCTION

In recent years lung cancer is proving to be a serious threat to mankind in this world. In reality the total number of deaths due to lung cancer is much higher than sum of deaths due to all other types of cancers. As per WHO's [1] fact sheet 1.6 million people are giving up their life to this killer disease. In its early stage lung cancer shows its presence only as small pulmonary nodules which are easily neglected by the radiologist and patients. But unfortunately most of the lung cancers are diagnosed in their later stages, by then the affected would have missed the best opportunity for treatment and hence the survival rate diminishes as the size/number of malignant tumor increases.

Anatomy of Lungs

The human respiratory system consists of the lungs and tubes associated with the lungs. It is located in the thorax or chest. The thorax is well protected by the ribs and the diaphragm forms the base of the thorax. Contractions of the diaphragm and the intercostal muscle change the size of the thorax and thus cause air to move in and out of the lungs. The main job of the respiratory system is to get oxygen into the

Body and get waste gases out of the body. It is the function of the respiratory system to transport gases to and from the circulatory system. The human respiratory system consists of the nose, nasal cavity, pharynx, larynx, trachea, smaller conducting passage way bronchi, bronchioles and air sacs alveoli.

CT Images

Computed tomography (CT) is among the mostly used imaging modalities in diagnosing and treatment of lung cancers. Recent day CT scanners have the capacity to produce a stack of isotropic images covering the lung cavity completely. In these CT images each pixel has its values expressed in Hounsfield Units (HU) which can be positive or negative. These scales are defined approximately as -1000HU for air, 0HU for water and 1000 HU for bone. With more advances in CT technology the slice thickness can be less than even 1mm, hence a CT data set which covers chest might contain more than 400 slices.

LITERATURE SURVEY

Cancer is an abnormal, continual multiplying of cells. The cells divide uncontrollably and may grow into adjacent tissue or spread to distant parts of the body. The mass of cancer cells will eventually become large enough to produce lumps, masses, or tumors that can be detected. These tumors can be further diagnosed as either benign or malignant:

Benign nodules: Benign nodules are not cancerous. They:

- Can usually be removed.
- Do not revisit in most cases.
- Do not spread to other parts of the body and the cells do not invade other tissues.

Malignant Nodules: Malignant nodules are cancerous. They

- Can invade and damage nearby tissues and organs.
- Metastasize (cancer cells break away from a malignant tumor and enter the bloodstream or lymphatic system to form secondary tumors in other parts of the body).

In [2], a knowledge-based system for segmenting and labeling lung nodule on CT images was proposed. In [3], a two stage CAD system was developed where a pre-processing technique was applied to obtain a better quality image to enable higher success rate on detection following which the cancerous nodule region is segmented. In [4], a Medical Digital Image Analysis System (MDIAS) such as Computer Aided Detection (CAD) technology was built which deals with another source of uncertainty that is inherent in an image-based practice of medicine. In [5] the author describes the efficient implementation of the Hessian based filters. These filters are commonly used in medical image analysis and are employed in the Voxel Based Neural Approach (VBNA), lung CAD system for lung nodule detection. In [6], the system exploits a thorax anatomical model in order to distinguish the 3D anatomical structures corresponding to chest wall, the trachea and two lung lobes in order. In [7], a computerized system for lung nodule detection in CT scan images is proposed. The automated system consists of two stages i.e. lung segmentation and enhancement, feature extraction and classification. In [8], the research work aims to develop a CAD system to detect pulmonary lung nodules from Low Dose CT (LDCT) scan images using template matching algorithm integrated with multi-resolution feature analysis technique in order to enhance the false positive detection rate.

134 out of 165 nodules were correctly detected by the proposed scheme with a detection rate of 81.212%.

PROPOSED SYSTEM

The system takes the lung CT image as input and is converted to grayscale image. A median filter is applied on this to get smooth surfaces. Then the filtered image is cropped to extract the lung area in the image and threshold segmentation separate foreground from the background of the image. In the ROI extract phase the initial morphological opening is used to eliminate the small object inside the lung. ROI is extracted from the image then nodule is identified which is overlapped on the original image. The nodule feature is extracted to differentiate them as either benign or malignant. The nodule feature is stored in the database and retrieved whenever required for diagnosis proposes by the experts. Advantages of proposed system are, level of over segmentation is reduced. Different types of noise present in the CT image can be investigated and restored in the preprocessing of images. These methods are not time consuming and expensive.

System Architecture

Architecture of the system provides an overview of an entire system, identifying all its elements at some level of abstraction. Systems are typically designed as object, involving heterogeneous components (e.g. hardware, software, humans) working together to perform a mission. The complexity of a system will mainly come from two aspects:

- **Integration of Components:** There are many interrelations between a possibly huge number of components, and there are recursive levels of integration
- **Heterogeneity of Components:** Several specialized fields are involved in the design of a complex system, making it difficult to keep a unified vision of this system and to manage its design.

The proposed system takes lung CT image as an input image and applies segmentation on this image to remove back ground and extracts the nodules from image. The ROI extraction phase applies post-processing to remove as many false regions as possible. The remaining candidate nodules are classified based on their properties which are extracted in feature extraction phase.

Threshold Segmentation

Segmentation is the division of an image into meaningful structures, image segmentation, is often an essential step in image analysis, object representation, visualization, and many other image processing tasks. A great variety of segmentation methods has been proposed. A disjunct categorization does not seem to be possible though, because even two very different segmentation approaches may share properties that defy singular categorization. Thresholding and slicing techniques are used to segment the image.

They may be applied directly to an image, but can also be combined with pre- and post-processing techniques. Thresholding is the simplest and most commonly used method of segmentation. An adequate threshold value T is selected,

$$G(x,y) = \begin{cases} 1 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) \leq T \end{cases}$$

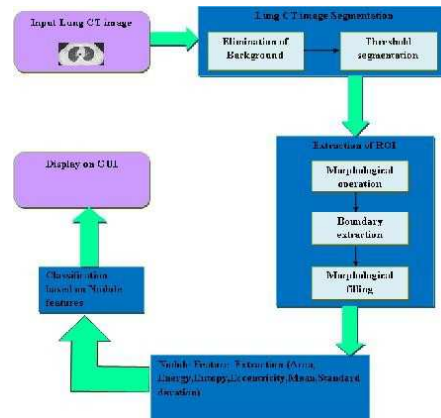


Figure 1: Architecture of the System

- Extraction of ROI
- Erosion

The **erosion** of a binary image f by a structuring element s (denoted $f \ominus s$) produces a new binary image $g = f \ominus s$ with ones in all locations (x, y) of a structuring element's origin at which that structuring element s fits the input image f , i.e. $g(x, y) = 1$ if s fits f and 0 otherwise, repeating for all pixel coordinates (x, y) . Erosion removes small-scale details from a binary image but simultaneously reduces the size of regions of interest.

Dilation

The **dilation** of an image f by a structuring element s (denoted $f \oplus s$) produces a new binary image represented by $g = f \oplus s$ with ones in all locations (x, y) of a structuring element's origin at which that structuring element s hits the input image f , i.e. $g(x, y) = 1$ if s hits f and 0 otherwise, repeating for all pixel coordinates (x, y) . The holes enclosed

$$Energy(j) = \sum_K Intensity(K)^2$$

- Eccentricity ($F3$) is a measure of circularness inside a candidate object.
- Entropy ($F4$) is a statistical measure of randomness that can be used to characterize the texture of the candidate region.

$$Entropy = -\sum_i p_i * \log_2(p_i)$$

-where p contains the histogram count.

- The mean intensity value indicates the average intensity value of all the pixels that belong to the same region. by a single region and gaps between different regions become 1 N smaller, and small intrusions into boundaries of a region are filled in.
- **Opening Operation**

The **opening** of an image f by a structuring element s

(denoted by $f \oslash s$) is an erosion followed by a dilation:

$$f \oslash s = (f \ominus s) \oplus s$$

Opening is so called because it can open up a gap between objects connected by a thin bridge of pixels.

- **Closing Operation**

The **closing** of an image f by a structuring element s

(denoted by $f \bullet s$) is a dilation followed by an erosion:

$$f \bullet s = (f \oplus s) \ominus s$$

Closing is so called because it can fill holes in the regions while keeping the initial region sizes.

- **Sobel Edge Detection**

In Sobel operators, for a 3×3 neighbourhood, each simple central gradient estimate is vector sum of a pair of orthogonal vectors. Each orthogonal vector is a directional derivative estimate multiplied by a unit vector specifying the derivative's direction. The vector sum of these simple gradient estimates amounts to a vector sum of the 8 directional derivative vectors. Thus for a point on Cartesian grid and its eight neighbours the density values are shown.

- **Nodule Feature Extraction**

The most difficult step in computerized lung cancer detection system is the lung nodule detection. In CT images of lungs, (i) the lung nodules are mostly attached to blood vessels or to the pleura; (ii) the grey level of blood vessels is almost similar to nodules. Hence varying features for the lung nodule detection are considered and the feature vector is formulated as $FV = \{F1, F2, F3, F4, F5, F6\}$. These features are:

- Area ($F1$) is a total number of pixels within a nodule. This feature helps to isolate small objects.
- Energy ($F2$) is used to describe measure of information in an image and is represented in equation.

$$Mean(g) = \sum Intensity(K)$$

- Standard deviation is a measure of how much that grey levels differ from mean.

$$Std(g) = \sqrt{\frac{1}{N} \sum_{K=1}^N (Mean(K) - Intensity(K))^2}$$

EXPERIMENTAL RESULTS AND DISCUSSIONS

The segmentation and classification have been carried out by using the options mentioned in the menu shown in the figure 1. The combined features of the region and its texture features are considered to classify the segmented nodule as benign or malignant using GLCM method. Features like Area, Energy, Eccentricity, Entropy, Mean and Standard deviation are extracted from segmented nodules. The threshold values for area, energy, entropy, eccentricity, mean and standard deviation are set as 0.0020, 0.6172, 0.00501, 0.0234, 0.050 and 0.0145. Depending on these values the nodules are classified accordingly.

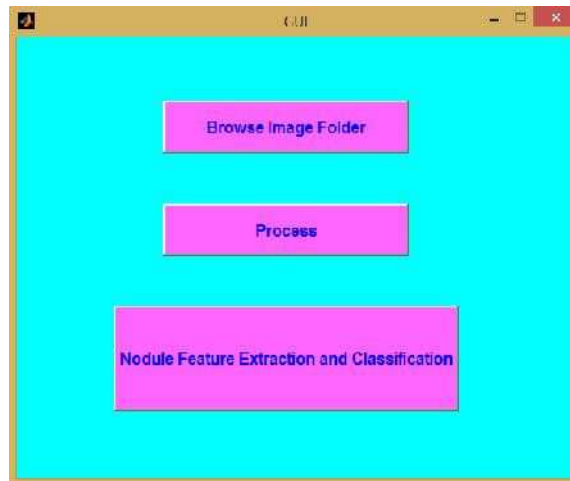


Figure 1: Graphical User Interface

Table 1: GLCM Features Extracted From Segmented Module

Image Name	Area	Energy	Eccentricity	Entropy	Mean	Standard Deviation	Category
Imgb_59	0.0015	0.9990	0.0035	0.0062	0.0281	0.013859	Benign
Imgb_101	0.017	0.9989	0.0028	0.0065	0.0251	0.012515	Benign
Imgm_32	0.017	0.9993	0.0026	0.0044	0.0222	0.012694	Malignant
Imgb_126	0.027	0.9983	0.0032	0.0090	0.0494	0.018776	Benign
Imgm_12	0.07	0.9995	0.0012	0.0025	0.0092	0.07520	Malignant
Imgb_140	0.081	0.9948	0.0040	0.0298	0.1268	0.027165	Benign
Imgm_96	0.032	0.9910	0.0056	0.0371	0.0460	0.03091	Malignant
Imgm_379	0.019	0.9335	0.00223	0.0064	0.0428	0.03271	Malignant
Imgb_142	0.0026	0.9508	0.00360	0.0049	0.0421	0.06291	Benign
Imgb_25	0.0071	0.9103	0.00381	0.0071	0.0294	0.0395	Benign

TABLE 1 summarizes the experimental results based on the feature values obtained by different feature vectors. The last column lists the result of classification of the extracted nodules as either benign or malignant.

Testing for Benign Nodule

An image is selected from the data set for processing and to classify as either benign or malignant.

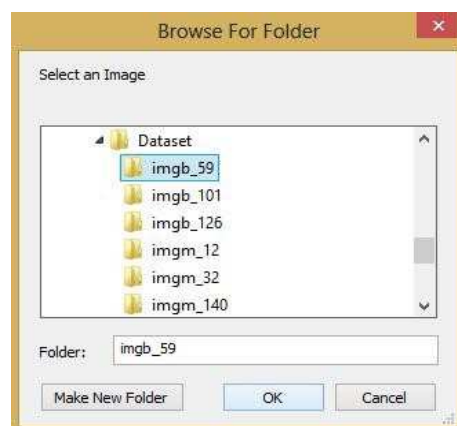


Figure 2: Selection of an Image from a Dataset

The figure 3 is the original CT scan image, which is viewed with the help of DICOM image viewer software.

A single orientation is taken and exported to jpeg image format on which further processing is carried out. The median filter is used to remove noise present if any in the image. The threshold for different intensity of the image is identified by optimal thresholding, which helps for segmentation of the filtered image. The segmented image is converted to binary image for further processing i.e. ROI extraction and nodule feature extraction. The morphological operations open and close are applied on thresholded image for eliminating the small objects inside and outside the lungs, enhance borders and fill the gaps in the border repeatedly. The discontinuous region in the binary image is identified and original grayscale image is overlapped on it, so that only lung tissue is extracted. The features such as area, centroid etc. is extracted and nodule is isolated from lung tissue. The nodule is isolated and overlapped on the original image so that nodule is marked prominently. The nodule identified with red mark is isolated on the RGB image and the features of the nodule are extracted based on which nodule it is classified as benign. The results of these operations are depicted in series from figure 4 through figure 10.

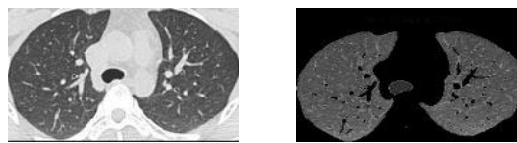


Figure 3: Original Image



Figure 4: Segmented Image

Eliminating the small objects inside and outside the lungs and enhances borders and fills the gaps in the border repeatedly. The discontinuous region in the binary image is identified and original grayscale image is overlapped on it, so that only lung tissue is extracted. The features such as area, centroid etc. is considered and nodule is isolated from lung tissue. The nodule is isolated and overlapped on the original image so that nodule is marked prominently. The nodule identified with red mark is isolated from the RGB image and features of the nodule are extracted based on which nodule is classified as malignant. The results of these operations are depicted in series from figure 13 through figure 19.

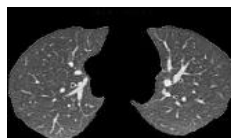


Figure 5: Binary Image



Figure 6: After Morphological Operations

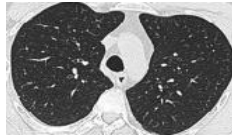


Figure 7: Extracted Lung Tissue

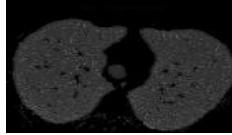


Figure 8: Isolated Lung Nodule

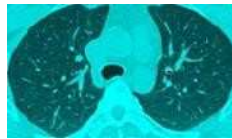


Figure 9: Identified Nodule



Figure 10: Classification Result

Testing for Malignant Nodule

An input image is selected from the repository to be classified as benign or malignant as shown in the figure 11.

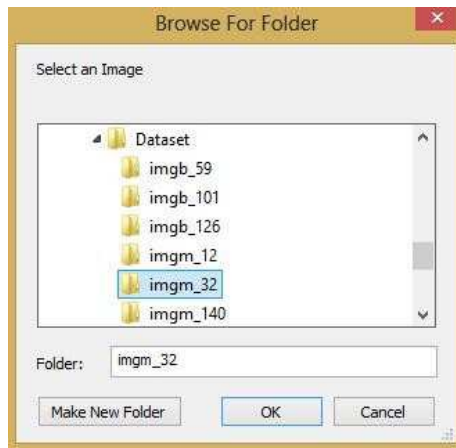


Figure 11: Selection of AN Image from the Dataset

The figure 12 is the original CT scan image, which is viewed with the help of DICOM image viewer software. A single orientation is taken and exported to jpeg image format on which further processing is carried out. The median filter is used to remove noise present if any in the image. The threshold for different intensity of the image is identified by optimal thresholding, which helps for segmentation of the filtered image. The threshold segmented image is converted to binary image for further processing i.e. ROI extraction and nodule feature extraction. The morphological operation open and close is applied to the thresholded image for

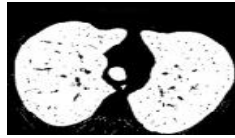


Figure 12: Original Image



Figure 13: Segmented Image

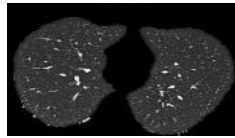


Figure 14: Binary Image

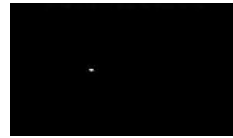


Figure 15: After Morphological Operations

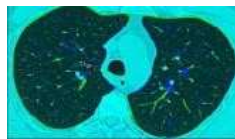


Figure 16: Extracted Lung Tissue

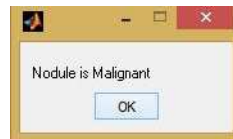


Figure 17: Isolated Lung Nodule

Figure 18: Identified Nodule

Figure 19: Classification Result

CONCLUSIONS AND FUTURE SCOPE

In this paper, Segmentation based approach has been used to separate the nodules in the lung image and the corresponding ROI is extracted for better results. The proposed system differentiates background and foreground regions by applying post-processing techniques. After this, selective features are extracted from ROI using GLCM and nodules are then classified as *benign* or *malignant*.

The proposed methodology is tested on a dataset of 50 lung CT images. Experimental results show that better categorization of lung nodules as benign and malignant when compared with the domain expert's knowledge. These results may help in preoperational planning and early diagnosis of lung tumors.

The results of threshold segmentation created a space for discussion on some of the limitations. These could be over segmentation and poor ability to detect areas with low contrast boundaries. To overcome these, Fuzzy Logic and Statistical Shape Model based segmentation can be used.

REFERENCES

1. <http://www.who.int/mediacentre/factsheets/fs297/en/>; accessed March 2015
2. HongshunSu, Wei Qian, Ravi Sankar and Xuejun Sun A, "New Knowledge -based Lung Nodule Detection System", IEEE ICASSP, pp.465-468, 2004.
3. S. Ashwin, J. Ramesh, S. Aravind Kumar, K. Gunavathi, "Efficient and Reliable Lung Nodule Detection using a Neural Network Based Computer Aided Diagnosis System", IEEE, pp.135-142, 2012.
4. Rahil Hosseini, Jamshid Dehmeshki, Sarah Barman, Mahdi Mazinani, Anne-Marie Jouannic, Salah Qanadli, "A Fuzzy Logic System for Classification of the Lung Nodule in Digital Images in Computer Aided Detection", IEEE Fourth International Conference on Digital Society, pp. 255-259, 2010.

5. Niccolo Camarlinghi, Francesco Bagagli, Piergiorgio Cerello T, Alessandra Retico, Maria Evelina Fantacci, "Chest CT automatic analysis for lung nodules detection implemented on a GPU computing system", IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), pp. 2008-2011, 2012.
6. M. Antonelli, G. Frosini, B. Lazzerini and F. Marcelloni A, "CAD System for Lung Nodule Detection based on an Anatomical Model and a Fuzzy Neural Network", IEEE, pp. 469-472, 2006.
7. Anam Tariq, M. Usman Akram and M. Younus Javed, "Lung Nodule Detection in CT Images using Neuro Fuzzy Classifier", IEEE, pp. 49-55, 2013
8. Mickias Assefa, Ibrahima Faye, Aamir Saeed Malik, "Lung Nodule Detection Using Multi-Resolution Analysis", IEEE International Conference on Complex Medical Engineering May 25 - 28, Beijing, China, pp.457-462, 2013.