ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2019.0296

IMPROVED LUNG NODULE CLASSIFICATION USING MULTI-CLASS ARTIFICIAL NEURAL NETWORK WITH BACK PROPAGATION ALGORITHM

S. Vimalnath[1] and G. Ravi[2]

[1]*Department of Electronics and Communication Engineering, Paavai Engineering College, India*

[2]*Department of Electronics and Communication Engineering, Sona College of Technology, India*

Abstract

This paper designed a novel method to overcome the problem of the lung nodule overlapping adjacent structures. We developed a lobe segmentation algorithm for identifying lung lobes CT images. To find reliable method for nodule detection is an important problem in medicine. It requires efficient automatic method to perform segmentation and detection. The identification of tumor region involves extraction of lobar fissures from the input CT images which makes use of two phases. In the first phase the fracture region is identified. In the second phase the found fissure are extracted. There is some nodule-like object in testing data detected by algorithm and not included in ground truth information. These are probably nodules missed by human. We designed a novel method to overcome the problem of the lung nodule overlapping adjacent structures. The result Obtained show that the proposed work can help the surgeons to identify the lobar fissures correctly to locate the lung region before they plan for the surgery. It reduces the computation time and complexity. Our system was developed with Faculty Hospital, Motol and Prague and in future should be used there. In order to improve the performance of the proposed approaches some future enhancements could be necessary in the present research work.

Keywords:

CT image, Lung Nodule, Classification, ANN

1. INTRODUCTION

A digital image is fundamentally composed of a series of pixels, a word derived from combining image and element. By choosing and brightening these individual pixels, a digital image emerges. At face value, a digital image is nothing more than a slew of pixels set in some logical state. Three 8-bit numbers represent most color images with each octet corresponding to the amount of red, green, and blue a pixel embodies. A grayscale image typically contains a sole 8-bit number to signify the amount of gray in a pixel. In addition to the color depth an image contains, the number of pixels or resolution is an additional image attribute Common notation for an image's resolution is $M\times N$ where M represents the number of horizontal pixels and N represents the number of vertical pixels. Common examples include 800×600 or 2048×1536.

The total number of pixels in a particular digital image is calculated by multiplying both horizontal and vertical numbers. Accordingly, this would be the resolution of a digital image produced by a mega pixel digital camera. While the color depth and number of pixels represent a digital image, images are further classified by the particular image format chosen to store the image. Common image formats include BMP, TIFF, and JPEG. The BMP File format also known as Bit Map image file or device independent bit map file format are simply a bit map is a raster graphics image file format used to store bit digital images, independently of the display device (such as Graphic adapter) especially on Microsoft windows and OS/2 operating system. TIFF is a computer file format for storing raster graphic images. It is widely supported by scanning, faxing, word processing, optical character recognition, image manipulation and page layout applications. JPEG is a standard method of compressing photographic images. It is the most common all platforms. JPEG Compression artifacts blend well into photographs with detailed non-uniform textures allowing higher compression ratios. Each has its own pros and cons when choosing to represent a digital image. The selection of one format over another depends on the particular application of the digital image. One must consider file size, application on the web, and image quality. Image formats such as BMP and TIFF use a Lossless compression scheme. That is, they do not discard any information in the compression process, thus emphasizing quality over a smaller file size.

1.1 IMAGE PROCESSING

Image processing is any form of signal processing for which input is an image, such as a photography or video frame the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image treating the image as a two-dimensional signal and applying standard signal techniques to it. Image processing usually refers to digital image processing but optical and analog image processing also are possible. Image processing refers to processing of a 2D image by a computer. An image defined in the world is considered to be a function of two variables, for example, $a(x,y)$ with as the amplitude (e.g. brightness) of the image at the real coordinate position (x,y). Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers.

1.2 HISTOGRAM BASED METHODS

Histogram is an accurate representation of distribution of numerical data. It is an estimate of probability distribution of a continuous variable. The histogram based methods are very efficient when compared to other image segmentation methods because they typically require only one pass through the pixels. In this technique, a histogram is computed from all of the pixels in the image, colour or intensity can be used as the measures. A refinement of this technique is to recursively apply the histogram-seeking method to clusters in the image in order to divide them into smaller clusters. This is repeated with smaller and smaller clusters until no more clusters are formed.

1.3 IMAGE SEGMENTATION

Image segmentation is a process of partitioning a digital image into multiple segments. The goal of the segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. It will segment the image into edge-based, region-based. There are three basic types of gray level discontinuities in a digital image: points, lines, and edges. The most common way to look for discontinuities is to run a mask through the image. Region growing: Groups pixels or sub-region into larger regions.

1.4 LUNG NODULE

It is a small mass of tissue in the lung. The chest film throws-up all sorts of differential diagnoses, of which the foremost that comes to one’s mind is lung cancer, primarily because it is a difficult and challenging affliction to manage the solitary pulmonary nodule is the prototype of the mass lesion in the lung. In general, any patient whose chest X-ray shows a single, rounded, or ovoid lesion in the lung parenchyma which is not associated with any obvious adenopathy, atelectasis, or pneumonia is considered to have a PN. PNs are listed in the literature under various names as coin lesions, solitary intra-pulmonary tumors, isolated pulmonary nodules etc. The pulmonary nodule may be of any size. The shape may be round, oval, or slightly lobulated, but well-circumscribed. The lesion should be more-or-less homogeneous in density, but may contain calcium. The shadow of the PN could be in contact with the diaphragm or chest wall, but not to involve or appear to fuse with these structures.

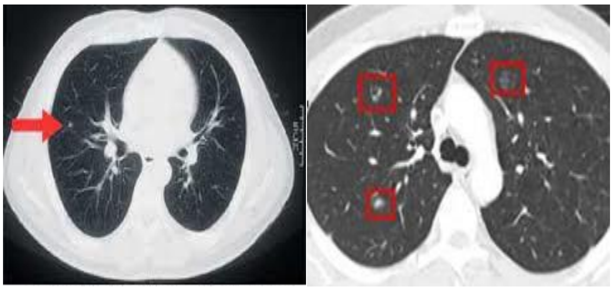


Fig.1. Images for Lung Nodule

Any nodule whose shadow shall appear to be continuous with that of the hilum is excluded. The lesion should have a smooth contour or only moderately irregular or fuzzy borders. There should be no adjacent inflammatory or atelectasis component and no region in anal lymphadenopathy.

Classification in a pulmonary nodule typically indicates benign disease. Benign nodules often contain dense, central, laminated, popcorn, and punctuate patterns of calcification. Malignant pulmonary nodules can also have a classified appearance. However, these are typically small and contain central nodules with speckled patterns, or eccentric punctuate foci. Since the patterns of classification typically do not overlap, a benign pattern is one of the few radiographic features that justify a decision that no further evaluation of the lung nodule is warranted. The calculated LR for malignancy in a pulmonary nodule with a benign pattern of classification approaches. While benign patterns of classification can be highly specific for the absence of malignancy. The nodule most commonly represents a benign tumor such as a granuloma or hamartoma, but in around 20% of cases it represents a malignant cancer, especially in older adults and smokers. Conversely, 10 to 20% of patients with lung cancer are diagnosed in this way. Thus, the possibility of cancer needs to be excluded through further radiological studies and interventions, possibly including surgical resection.

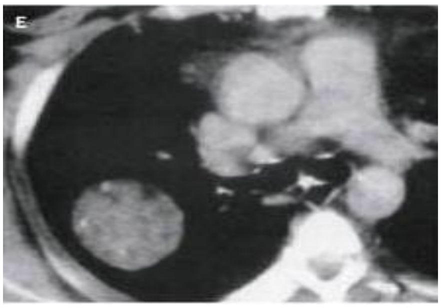


Fig.2. Image for Hamartoma Nodule

1.5 LUNG NODULES WITH VARIATIONAL METHODS

These classes of methods are primarily effective for pulmonary nodules, however, fail in separating nodules from juxtaposed surrounding structures, such as the pleural wall (i.e. Juxta-Pleural and Pleural-Tail nodules) and vessels (Vascular), due to their similar intensities. More sophisticated approaches have been proposed to incorporate nodule-specific geometrical and morphological constraints to address this issue. However, juxta-pleural, or wall-attached, nodules still remain a challenge because they can violate geometric an assumptions and appear frequently. Robust segmentation of the juxta-pleural cases can be addressed in two approaches: global lung or rib segmentation local non-target removal or avoidance. The first can be effective but also computationally complex and dependent on the accuracy of the whole-lung segmentation.

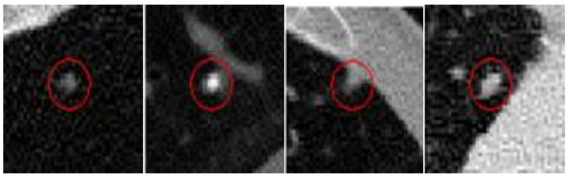


Fig.3. Variational Types of Nodule

1.6 LUNG CANCER OR CARCINOMA

Lung cancer is also known as carcinoma of the lung or pulmonary carcinoma, is a malignant lung tumor characterized by uncontrolled cell growth in tissues of the lung. If left untreated, this growth can spread beyond the lung by process of metastasis into nearby tissue or other parts of the body. Most cancers that start in the lung, known as primary lung cancers, are carcinomas that derive from epithelial cells. The main primary types are small-cell lung Carcinoma (SCLC) and Non-small-cell lung carcinoma (NSCLC). The most common symptoms are coughing (including coughing up blood), weight loss, shortness of breath, and chest pains.

The vast majority (80-90%) of cases of lung cancer are due to long-term exposure to tobacco smoke. About 10-15% of cases occur in people, who have never smoked. These cases are often caused by a combination of genetic factors and exposure to radon gas, asbestos, or other forms of air pollution, including second-hand smoke. Lung cancer may be seen on chest radiographs and computed tomography (CT) scans. The diagnosis is confirmed by biopsy which is usually performed by bronchoscopy or CT-guidance.

The proposed method is to locate the accurate nodule detection by using neural network classifier with MR-16 filters. To support multi classification application a neural network classifier is used instead of SVM. The need to go for neural network can classify the nodule into an accurate detection image and also to eliminate intersection i.e., multiclass application. The proposed technique uses as neural network with MR-16. The MR-16 (Maximum response filter) represents a bit map encoding levels. It is a powerful tool to analyze and modulate make sense of complex clinical data across a broad range of application. The block diagram of proposed system is given in Fig.4.

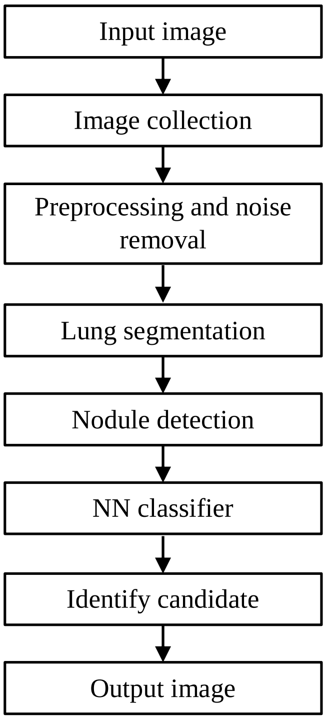


Fig.4. Block Diagram of Proposed System

2. THEORETICAL BACKGROUND

A lung nodule is a spot on the lung that is less than 3cm (or 1½ inches) in diameter. If a spot is larger than 3cm, it is considered a lung mass, rather than a lung nodule. The overall chance that a lung nodule is cancer is 40%, but that risk varies a lot depending on factors. As nodules are the most common sign of lung cancer, nodule detection in chest images is a main diagnostic problem. Conventional projection radiography is a simple, cheap, and widely used clinical test. Unfortunately, its capability to detect lung cancer in its early stages is limited by several factors, both technical and observer-dependent. Lesions are relatively small and usually contrast poorly with respect to anatomical structures. This partially explains why radiologists are commonly credited with low sensitivity in nodule detection, ranging from 60 to 70%.

3. METHODOLOGY

An input CT image is taken from the form of JPEG image. It is a loss compression. The CT images normally contain artefacts, noise which will be not be suitable for processing and hence it has to be pre-processed to reduce the noise using Wiener filter. The segmentation of lung regions plays an important role to speed up the process of detection and analysis of lung nodule. Input image is the segmented lung region. Apply intensity threshold, shape and area features to get nodule mask in binary image format. Superimpose the mask with the input image to extract lung nodules with original intensities. The lung is segmented from the CT images using morphological operations. The gray scale image is first converted to binary image. All the pixels in the input image with the intensity greater than a threshold level is replaced with value ‘1’ and all pixel values with an intensity less than threshold level is replaced with value ‘0’. To identify the actual fissure location and curvatures, the lobe segmentation algorithm uses a fissure search technique to look for the longest continuous lines crossing the fissure regions, which signifies the actual fissures. In order to detect the nodules, we extract feature vectors from the detected nodule candidates. Next, the feature vectors become the input for classifier and classified into nodules and non-nodules by using canny edge. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data. The MR-16 (Maximum response filter) represents a bit map encoding levels. The nodule candidates are considered as nodules or non-nodules using annotation provided by chest radiologist. The results from the experiments on the ELCAP dataset showed promising performance of our method.

3.1 ARTIFICIAL NEURAL NETWORK

An Artificial Neural Network is based on the collection of connected units or nodules called artificial neurons, which loosely model the neurons in the biological area. Each connection like the synapses in the biological part can transmit a signal one artificial neuron to another. An Artificial neuron that receives a signal can process it and then a signal additional artificial neuron connected to it. In common ANN Implementations, the signal at a connection between artificial neuron is a real number and the output of each artificial neuron is computed by some non-linear function of the some of its input. The connection between artificial neurons is called edges. Artificial neurons and edges typically have a weight that adjust as learning proceeds the weight increases or decreases the strength of the signal at the connection. Artificial neurons may have a threshold such that the signal is only send if the aggregate signal crosses the threshold.

Typically, artificial neurons are aggregated into layers. Different layers may perform different kinds of transformation on their inputs. Signal travel from first layer to last layer possibly after traversing the layers multiple times. The original goal of ANN approach was to solve problems in the same way that is human brain would. However, over time attention has moved on performing a specific task, which leads to deviation from biology. This is often just called a neural network, is a mathematical model inspired by biological neural networks.

A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases a neural network is an adaptive system that changes its structure during a learning phase. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data. A neuron is an information-processing unit that is fundamental to the operation of a neural network.

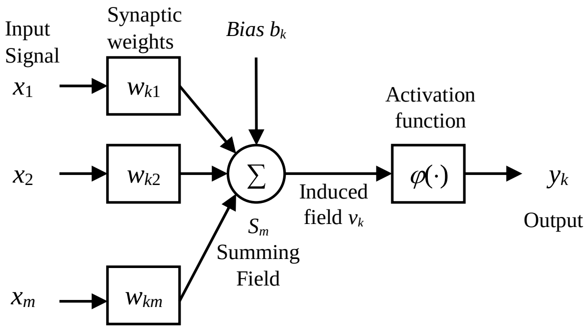


Fig.5. Non-linear model of a neuron

A set of synapses, each of which is characterized by a weight or strength of its own. Specifically, a signal x_j at the input of synapse j connected to neuron k is multiplied by the synaptic weight w^k_j. It is important to make a note of the manner in which the subscripts of the synaptic weight w^k_j are written. The first subscript refers to the neuron in question and the second subscript refers to the input end of the synapse to which the weight refers. The weight w^k_j is positive if the associated synapse is excitatory; it is negative if the synapse is inhibitory. An adder for summing the input signals, weighted by the respective synapses of the Neuron. An activation function for limiting the amplitude of the output of a neuron.

3.1.1 Modes of Learning and Algorithm:

There are two modes of learning to choose from: One is on-line (incremental) learning and the other is batch learning. In on-line (incremental) learning, each propagation is followed immediately by a weight update. In batch learning, many propagation occur before weight updating occurs. Batch learning requires more memory capacity, but online learning requires more update.

3.1.2 Actual Algorithm for a 3-layered Network (One Hidden Layer):

Step 1: Initialize the weights in the network (often randomly): Do

Step 2: For each example *e* in the training set, *O* is the neural-net-output (network, *e*); forward pass and *T* is the teacher output for *e*

Step 3: Calculate error (*T*-*O*) at the output units

Step 4: Compute *delta*_w_h for all weights from hidden layer to output layer; backward pass

Step 5: Compute delta_w_i for all weights from input layer to hidden layer; backward pass continued

Step 6: Update the weights in the network

Step 7: Until all examples classified correctly or stopping criterion satisfied

Step 8: Return the network

As the algorithm name implies, the errors propagate backwards from the output nodes to the inner nodes. Technically speaking, back propagation calculates the gradient of the error of the network regarding the network modifiable weights. Backpropagation usually allows quick convergence on satisfactory local minima for error in the kind of networks to which it is suited.

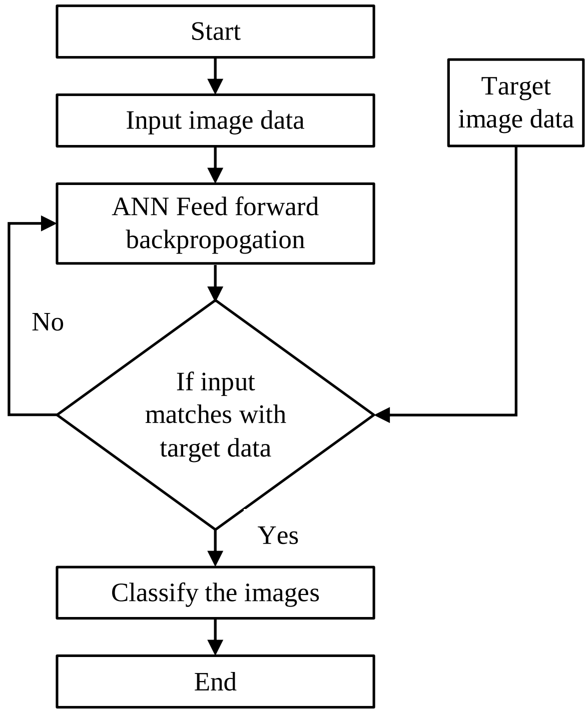


Fig.6. ANN-BP

3.1.3 Back Propagation Algorithm:

Back propagation is the method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be in the networks. Back propagation is shorthand for the backward propagation of errors, since an error is computed at the output and distributed backwards throughout the networks layer. It is commonly used to train deep neural networks. Backpropagation is a special case of a more general technique called Automatic differentiation. In the context of learning, backpropagation is commonly used by the gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function. The motivation of the backpropagation is to train a multi-layered neural network such that it can learn the appropriate internal representations to allow it to learn any arbitrary mapping of input to output. In BPN, weights are initialized randomly at the beginning of training. There will be a desired output, for which the training is done. Supervisory learning is used here. During forward pass of the signal, according to the initial weights and activation function used, the network gives an output. That output is compared with desired output. If both are not same, an error occurs. During reverse pass, the error is back-propagated and weights of hidden and output layer are adjusted. The whole process then continues until error is zero. The network is trained with known values. After training, network can perform decision making

3.1.4 Flowchart for ANN-BP:

The desired image is taken to compare with the target image so that the nodule can be detected whether It is cancerous cell or not. To start the process, the image has been taken as the input data and this is further processed with the ANN feed forward back propagation method. After the process is completed, it is then compared with the target data. If it matches with the target image it is then moved the next process known as Image classification in which the interpreted results is shown and the process is end and the nodule is detected its cancerous or not. If it not matches with the target it is detected that it is non-cancerous cell.

3.1.5 Applications and Advantages of Neural Network:

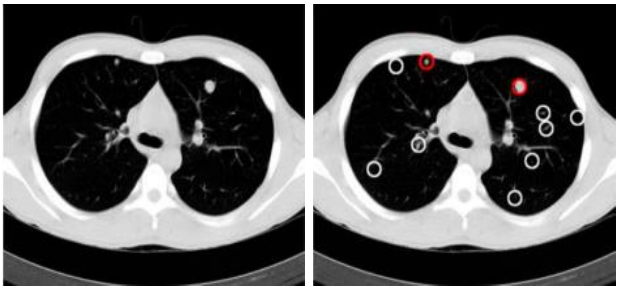
- Neural networks have emerged as an important tool for classification. The recent vast research activities in neural classification have established that neural networks are a promising alternative to various conventional classification methods.
- The effectiveness of neural network classification has been tested empirically. Neural networks have been successfully applied to a variety of real world classification tasks in industry, business and science.

4. RESULTS AND DISCUSSIONS

The proposed method is simulated in matlab tool and the results obtained through the proposed method is given below:

4.1 INPUT CT IMAGE OF LUNG NODULE

The lung nodule in CT images appears to be blurred boundaries with low contrast. The automatic and semi-automatic fissure extraction methods were proposed by the different groups. To detect the lobar fissures on a cost image, which was computed from a combination of the original data and the distance map performed on a previously generated vessel mask. Train the neural network using the input data and then target area should be identified whether the nodule is FPs and Red IPs.



(a) (b)

Fig.7(a). Input slice and (b) non-pleural nodule candidate detected (white - FP, Red - TP)

4.2 GRADIENT CT IMAGE FOR NODULE

In this, nodules to be detected near the bronchus and find false nodules.

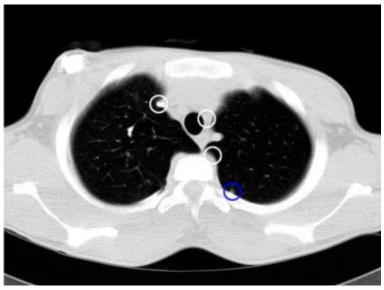


Fig.8. NN classifier used for classification. Colors and FN added manually: TP (red), FPs (white), FN (blue)

4.3 LEVEL SET SEGMENTATION

The Fig.9 denotes that the CT images normally contain artefacts, noise which will be not be suitable for processing and hence it has to be pre-processed to reduce the noise using Wiener filter.

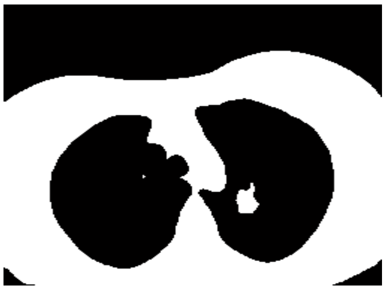


Fig.9. Segmented Image

Fig.10. HOG feature extraction image using SIFT

4.4 NODULE DETECTION

Neural Network classifier used for classification. Object marked by green color is probably Juxta-Pleural nodule missed by human and detected by algorithm

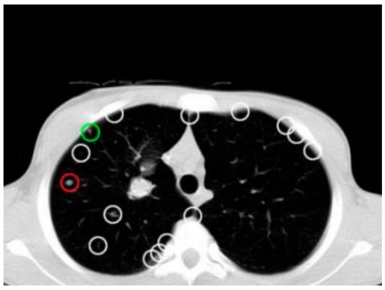


Fig.11. Nodule Detection

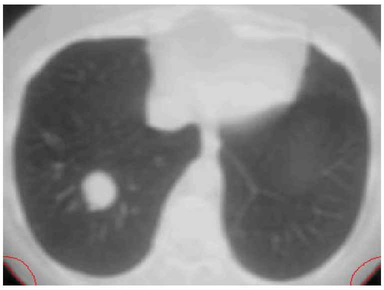


Fig.12. Iteration image for well-circumscribed nodule

5. CONCLUSIONS

A supervised classification method for lung nodule LDCT images in this paper. The four main categories of lung nodules well-circumscribed, vascularized, Juxta-Pleural, and Pleural-Tail were the objects to be differentiated. We designed a novel method to overcome the problem of the lung nodule overlapping adjacent structures. In conclusion, we developed a lobe segmentation algorithm for identifying lung lobes CT images. To find reliable method for nodule detection is an important problem in medicine. It requires efficient automatic method to perform segmentation and detection. The identification of tumor region involves extraction of lobar fissures from the input CT images which makes use of two phases. In the first phase the fracture region is identified. In the second phase the found fissure are extracted. There is some nodule-like object in testing data detected by algorithm and not included in ground truth information. These are probably nodules missed by human. We designed a novel method to overcome the problem of the lung nodule overlapping adjacent structures. The result Obtained show that the proposed work can help the surgeons to identify the lobar fissures correctly to locate the lung region before they plan for the surgery. It reduces the computation time and complexity.

Our system was developed with Faculty Hospital, Motol and Prague and in future should be used there. In order to improve the performance of the proposed approaches some future enhancements could be necessary in the present research work. The main aim of the future enhancement would be to increase the sensitivity and specificity of the system. The other future enhancement would be to incorporate nanotechnology into the system for the better performance.

REFERENCES

[1] S.G. Armato, M.L. Giger and H. MacMaho, "Automated Detection of Lung Nodules in CT Scans: Preliminary Results", *Medical Physics*, Vol. 28, No. 8, pp. 1552-1561, 2001.

[2] C. Bhuvaneswari, P. Aruna and D. Loganathan, "A Novel Shape Based Feature Extraction Technique For Diagnosis of Lung Diseases using Evolutionary Approach", *ICTACT Journal of Soft Computing*, Vol. 4, No. 4, pp. 804-810, 2014.

[3] F. Zhang, Y. Song, W. Cai, M.Z. Lee, Y. Zhou, H. Huang, S. Shan, M.J. Fulham and D.D. Feng, "Lung Nodule Classification with Multilevel Patch-Based Context Analysis", *IEEE Transactions on Biomedical Engineering*, Vol. 61, No. 4, pp. 1155-1166, 2014.

[4] H. Takizawa, Y. Shinji and S. Tsuyoshi, "Recognition of Pulmonary Nodules Thoracic CT Scans using 3D Deformable Object Models of Different Classes", *Algorithms*, Vol. 3, No. 2, pp. 125-144, 2010.

[5] A. Farag, A. Ali, J. Graham, S. Elshazly and R. Falk, "Evaluation of Geometric Feature Descriptors for Detection and Classification of Lung Nodules in Low Dose CT Scans of the Chest", *Proceedings of International Symposium on Biomedical Imaging from Nano to Macro*, pp. 169-172, 2011.

[6] Amal A. Farag, Shireen Y. Elhabian, Salwa A. Elshazly and Aly A. Farag, "Quantification of Nodule Detection in Chest CT: A Clinical Investigation Based on the ELCAP Study", *Proceedings of International Conference on Image Processing*, pp. 149-160, 2010.

[7] Y. Lee, T. Hara, H. Fujita, S. Itoh and T. Ishigaki, "Automated Detection of Pulmonary Nodules in Helical CT Images based on an Improved Template Matching Technique", *IEEE Transactions on Medical Imaging*, Vol. 20, No. 7, pp. 595-604, 2001.

[8] S.L.A. Lee, A.Z. Kouzani and E.J. Hu, "Automated Detection of Lung Nodules in Computed Tomography Images: A Review", *Machine Vision and Applications*, Vol. 23, No. 1, pp. 151-163, 2012.

[9] Y. Song, W. Cai, Y. Zhou and D. Feng, "Feature based Image Patch Approximation for Lung Tissue Classification", *IEEE Transactions on Medical Imaging*, Vol. 32, No. 4, pp. 797-808, 2013.

[10] L. Srensen, S.B. Shaker and M. DeBurjin, "Quantitative Analysis of Pulmonary Emphysema using Local Binary Pattern", *IEEE Transactions on Medical Imaging*, Vol. 29, No. 2, pp. 559-569, 2010.

[11] X. Ujiong, Y. Xinjulin, J. Dehmesk, G. Slabaugh and G. Beddoe, "Shape based Computer-Aided Detection of Lung Nodules in Thoracic CT Images", *IEEE Transactions on Biomedical Engineering*, Vol. 56, No. 7, pp. 1810-1820, 2009.