International Journal of Engineering & Technology, 7 (1.1) (2018) 189-192

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Performance comparisons of the soft computing algorithms in lung segmentation and nodule identification

Afshan Khanum [1*], S. Purushothaman [2], P. Rajeswari [3]

[1] Research Scholar, Department of Computer Science and Engineering, VELS University, Chennai, India
[2] Chennai, India.
[3] Department of Computer Science, Girls Community College, King Khalid University, Abha, Saudi Arabia
**Corresponding author E-mail:afshankhanum2013@gmail.com*

Abstract

This paper presents the implementation back propagation algorithm (BPA) and fuzzy logic(FL) in lung image segmentation and nodule identification. Lung image database consortium (LIDC) database images has been used. Features are extracted using statistical methods. These features are used for training the BPA and FL algorithms. Weights are stored in a file that is used for segmentation of the lung image. Subsequently, texture properties are used for nodule identification.

Keywords: *Lung Image, Segmentation, Nodule Identification, Back Propagation Algorithm, Fuzzy Logic.*

1. Introduction

Computer aided diagnosis of lung nodule identification is a challenging task. As such human intervention is used for identifying nodules in the lungs. Proper segmentation of lung image can help in clearly identifying the presence of nodules and hence classification of nodules can be done without any ambiguity. Due to this reason, intelligent segmentation algorithms are used so that improvement in nodule identification and classification can be achieved.

This paper involves in segmentation of computed tomography (CT) lung images and identifying lung nodules in the images. Preprocessed images are obtained from the lung image database. Segmentation of the lung image is carried out by extracting features using statistical methods. The features are used to train back propagation algorithm/ Fuzzy logic algorithm for segmenting the lung image. Contrast should be low, energy should be maximum, entropy should be minimum and homogeneity should be maximum for identifying the presence of lung nodules. Graylevel co-occurrence matrix is used to find the texture properties of the nodules.

There are many works in the literature that discuss about lung image analysis. From these works, it is observed that first hand assessment of lung affected patients for radiologic diagnosis is by using cross-sectional and projectional imaging techniques.

According to Lee, et al., [5], Template-matching technique based on a genetic algorithm (GA) template matching (GATM) was used for detecting nodules existing within the lung area. In their work, GA was used to determine the target position in the observed image efficiently and to select an adequate template image from several reference patterns for quick template matching.

Wallace, et al., have discussed Ground-Glass Opacity (GGO). It is defined as increased attenuation of lung parenchyma without obscuration of the pulmonary vascular markings on CT images. It is considered one of the important features in lung cancer diagnosis using Computer Aided Diagnosis (CAD).

Anti-geometric diffusion has been used as preprocessing to remove image noise. Geometric shape features (such as shape index and dot enhancement), are calculated for each voxel within the lung area to extract potential nodule concentrations. Rule based filtering is then applied to remove False Positive regions.

A Neural network-based fuzzy model has been proposed by Lin, et al., [6] to detect lung nodules. The authors used a set of appropriate fuzzy inference rules, and the membership functions was refined through the steepest gradient descent-learning algorithm. The proposed method detects isolated, juxtavascular and peripheral nodules with sizes ranging from 2 to 20 mm diameter.

A multiscale joint segmentation and model fitting solution which extends the robust mean shift-based analysis to the linear scale-space theory was proposed by Okada, et al., [9]. In their work, an ellipsoidal (anisotropic) geometrical structure of pulmonary nodules in the multislice X-ray CT images have been used for target's center location, ellipsoidal boundary approximation, volume, maximum/average diameters.

Automated voxel-by-voxel classification of airways, fissures, nodules, and vessels from chest CT images using a single feature set and classification method have been done by Ochs, et al., [8].

ROI slices were combined to form 3D ROI image and a 3D template by Osman, et al., [10]. They used this template to find the structures with similar properties of nodules

Application of CAD in medical imaging has been presented by Giger, et al., [1].

In another work by Ozekes, et al., [11], the lung regions of the CTs were segmented using Genetic Cellular Neural Networks (GCNN). ROIs were specified with using the 8 directional search and +1 or -1 values were assigned to each voxel. Fuzzy rule based thresholding was applied and the ROI's were found.

Golosio, et al., [2] used MultiSlice Computed Tomography (MSCT) for lung cancer detection and to identify noncalcified

nodules of small size (from about 3 mm). The authors had used ROI based on a multithreshold surface-triangulation approach.
Diagnosis of pulmonary infections using texture analysis and Support Vector Machine (SVM) classification was carried out by Yao, et al., [15].
CAD can be used for identifying and diagnosing infectious pulmonary diseases; segmentation and registration of pulmonary anatomical structures, detection and classification using texture and shape analysis for respiratory tract infections Ulas, et al.,
The existing methods for Lung Nodule Detection (LND) have used different stages of image processing methods. Each intermediate stages of image processing have been discussed by Lee, et al., [4].
JinkeWang and HaoyanGuo, [3], implemented skin boundary detection, rough segmentation of lung contour, and pulmonary parenchyma refinement for lung segmentation. Wenjun et al., segmented lung image using hierarchical Dirichlet process.

2. Methodology: Implementation of the algorithms for lung image segmentation

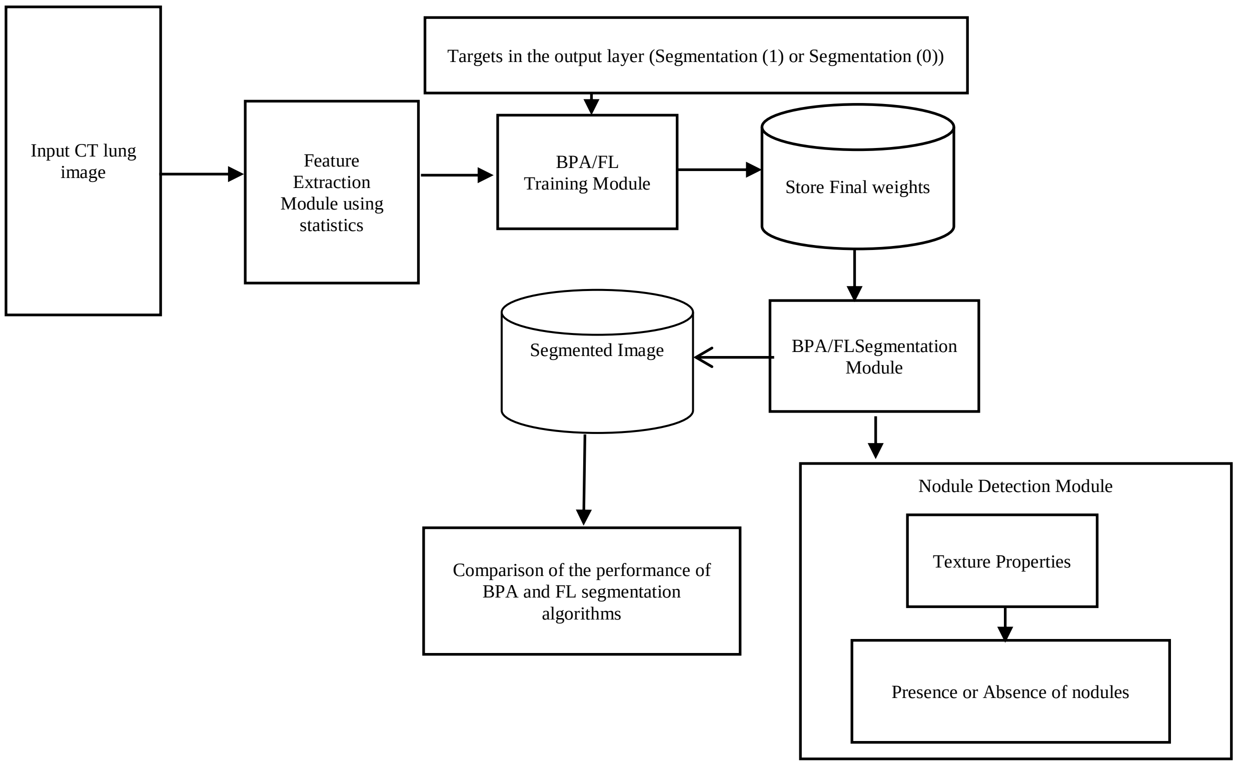


Fig. 1: Schematic flow for lung segmentation and nodule identifucation

Figure 1 presents the implementation of lung image segmentation and nodule identification. Computed Tomography (CT) is used as input for the algorithms for segmentation and nodule identification. Statistical method is applied for extracting features from the image. These features are used as inputs for the back propagation algorithm and fuzzy logic algorithm from training. At the end of training process, final weights are stored from BPA and fuzzy inference system (FIS) is stored from the Fuzzy logic. During the segmentation process, the final weights and the FIS are used for segmentation of the lung image.
Graylevel co-occurrence matrix (GLCM) process is applied on the segmented image using the ground truth image to identify the presence of nodules in the segmented image.

Data collection

The existing Lung image database from early lung cancer action program (ELCAP-http://www.ielcap.org/) and lung image database consortium (LIDC-http://imaging.cancer.gov/programs and resources/ information systems/ lidc) public image database provides a set of CT images for comparing different computer-aided diagnosis (CAD) systems. The database consists of an image set of 50 low-dose documented whole-lung CT scans for detection. The whole-lung dataset consists of 50 CT scans obtained in a single breath hold with a 1.25 mm slice thickness. The locations of nodules detected by the radiologist are also provided. Case id is W0001-1 till W0001-50. The size of the image is 512 X 512 in the x-y direction. The number of slices (276 or 280 or 249 or 243 or 273) varies for each case id. The resolution in the slice varies from 0.59 x 0.59, or, 0.62 x 0.62, or, 0.67 x 0.67, or, 0.7 x 0.7, and 0.76 x 0.76. Similarly, the numbers of slices that contain nodules also vary in each case id.

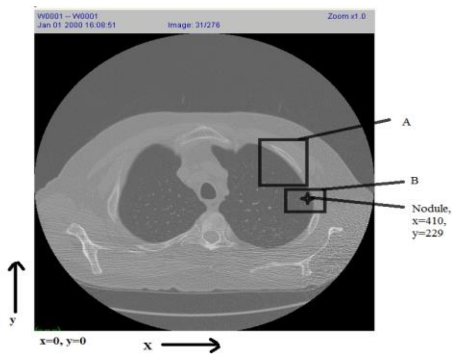


Fig. 2: Sample lung LIDC CT image slice number 31

We have used case id W0001-1.2.826.0.1.3680043.2.656.1.138.1 to 276. The slice numbers that have nodules are 18, 31, 42, 52, 54, 84, 130, 169, 170, 229, 239. In this paper, the 31st slice is presented exclusively. In this slice, the nodule is located at x=410 and y=229. The x increases from left to right and y increases from bottom to top. Figure 2 shows the 31st slice with rectangular box 'B' showing the nodule location at x=410 and y=229. Block 'A' is also shown for considering the texture information.

Nodule template

Nodule template is a collection of information about different types of nodules. These nodules have a different texture appearance. The texture appearances can be represented by GLCM features. The file contains a set of numerical values which represents the GLCM features. Totally 4 columns are present the template file in which the Column 1 represents energy feature, Column 2 represents Homogeneity feature, column 3 represents contrast feature and column 4 represent correlation feature respectively. These features can be used to identify the presence of nodule in an image. The number of rows present in the file corresponds to nodules of 192 slices which belongs to group 1.

Sample nodules

The template data creation is performed to identify nodules from the CT lung image. Figure 3 provides nodules of different CT lung slices. Each nodule is cropped with 40 X 40 around the location of the nodule. The location of nodule in each slice is given in the LIDC database. In most of the images presented here, the size of the nodule is small. Give a suitable title on each column head.

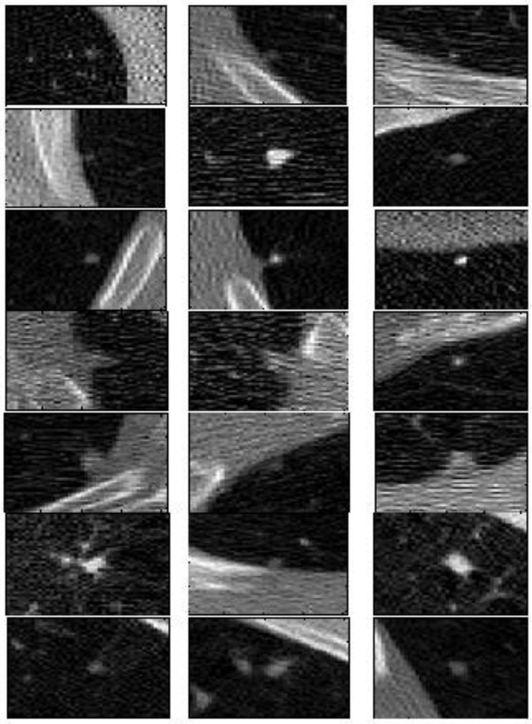


Fig. 3: Nodule cropped around its location (40×40) window

Back propagation algorithm for lung image segmentation

The training of Backpropagation algorithm was developed by Werbos 1974. It is a supervised algorithm that uses steppest descent procedure for minimizing the error between network output and the target output. The input-output data is presented to the algorithm during the training process. The training is stopped when a minimum squared error (MSE) is reached. The features obtained from the statistical method is presented to the BPA. The training of the BPA is stopped once the mean squared error (MSE) is reached. The final weights are stored in a file. In the lung segmentation process, features are extracted from the lung image and resented to the BPA. The final weights are used to compute outputs in the hidden layer and in the output layer. The output is thresholded and a values of '0' or '1' is allotted to the image that is called the segmented lung image.

Fuzzy logic for lung image segmentation

Zadeh 1965 proposed fuzzy sets explaining how a given linguistic variable can be clearly quantified using membership functions. The concept developed by him is called fuzzy logic. In fuzzy logic approach proper decisions can be obtained. In this work, the following steps are followed to implement lung segmentaiton.

1) Training fuzzy logic using the features obtained from the statistical method to obtain fuzzy inference system (FIS).
2) Segmenting the lung image using the features obtained from the statistical method and the FIS.

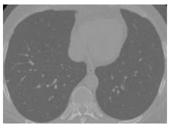
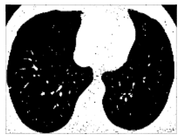
Fuzzy logic maps input data with target output by forming a transfer function called fuzzy inference system (FIS). Different models for FIS were developed by Mamdani 1977, Takagi &Sugeno. Exact mapping of input data with output data is not possible, and it is an approximate mapping by FIS. The accuracy of mapping is increased by modifyingthe parameters of FIS membership function, rules, defuzzification to obtain the best mapping of input-output data by FIS.

Texture related feature extraction using GLCM for lung nodule identification

Texture plays an important role in image processing. Various texture properties can be extracted from the binarized image to know the area of the object, its centroid, its orientation, its perimeter and many more. These information can be used in combination with the Graylevel cooccurence matrix (GLCM) properties. In general GLCM creates graycomatrix by calculating how often a pixel with gray-level (grayscale intensity) value '*I'* occurs horizontally adjacent to a pixel with the value '*j'* Each element (*i,j*) in glcm specifies the number of times that the pixel with value '*I'* occurred horizontally adjacent to a pixel with value '*j'*. Subsequently, correlation, energy, contrast, homogeneity can be obtained. This helps especially in identifying the lung nodules.

3. Results of the segmentation algorithms

Table 1: Presents Lung Segmentation by the Back Propagation Algorithm and Fuzzy Logic

Algorithm	Original image	Segmented result
Back propagation algorithm segmentation		
Fuzzy logic segmentation		

4. Conclusions

This work has considered CT Lung image segmentation and lung nodule identification. The conclusions of this work are as follows:

1. The size of the window for feature extraction in statistical method is preferred to be 3X3. The higher the size of the window, the segmentation appears to artifacts.
2. The training of back propagation algorithm involves more iterations for convergence. The accuracy of segmentation is less when compared to that fuzzy logic algorithm.
3. Fuzzy logic provides better segmentation when compared to that of BPA.
4. GLCM properties are used for confirming the presence of nodules.

References

[1] Giger ML, Chan HP & Boone J, “Anniversary paper: history and status of CAD and quantitative image analysis: the role of medical physics and AAPM”, *Medical Physics*, Vol.35, No.12, (2008), pp.5799–5820.

[2] Golosio B, Masala GL, Piccioli A, Oliva P, Carpinelli M & Cataldo R, “A novel multi-threshold method for nodule detection in lung CT”, *Medical Physics*, Vol.36, No.8, (2009), pp.3607–3618.

[3] Wang J & Guo H, “Automatic Approach for Lung Segmentation with Juxta-Pleural Nodules from Thoracic CT Based on Contour Tracing and Correction”, *Computational and mathematical methods in medicine*, (2016).

[4] Lee SLA, Kouzani AZ & Hu EJ, “Automated detection of lung nodules in computed tomography images: a review”, *Machine Vision and Applications*, Vol.23, No.1, (2012), pp.151-163.

[5] Lee Y, Hara T, Fujita H, Itoh S & Ishigaki T, “Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique”, *IEEE Trans. Med. Imaging*, Vol.20, (2001), pp.595–604.

[6] Lin DT, Yan CR & Chen WT, “Autonomous detection of pulmonary nodules on CT images with a neural network-based fuzzy system”, *Comput. Med. Imaging Graph*, Vol.29, (2005), pp.447–458.

[7] Mamdani EH, “Application of fuzzy logic to approximate reasoning using linguistic synthesis”, *IEEE Trans. on Computers*, Vol.26, No.12, (1977), pp.1182–1191.

[8] Ochs RA, Goldin JG, Abtin F, Kim HJ, Brown K, Batra P, Roback D, McNitt-Gray MF & Brown MS, “Automated classification of lung bronchovascular anatomy in CT using Adaboost”, *Med. Image Anal.*, Vol.11, (2007), pp.315–324.

[9] Okada K, Comaniciu D & Krishnan A, “Robust Anisotropic Gaussian Fitting for Volumetric Characterization of Pulmonary Nodules in Multislice CT”, *IEEE Trans. Med. Imaging*, Vol.24, No.3, (2005), pp.409–423.

[10] Osman O, Ozekes S & Ucan ON, “Lung nodule diagnosis using 3D template matching”, *Comput. Biol. Med*, Vol.37, No.8, (2007), pp.1167–1172.

[11] Ozekes S, Osman O & Ucan ON, “Nodule detection in a lung region that’s segmented with using genetic cellular neural networks and 3D template matching with fuzzy rule based thresholding”, *Korean J. Radiol.*, Vol.9, (2008), pp.1–9.

[12] Ba˘gcı U, Bray M, Caban J, Yao J & Mollura DJ, “Computer-assisted detection of infectious lung diseases: A review”, *Computerized Medical Imaging and Graphics*, Vol.36, (2012), pp.72–84.

[13] Millar Jr WT & Shah RM, “Isolated diffuse Ground–Glass opacity in Thoracic CT, causes and clinical presentations”, *American journal of Roentgenology*, Vol.184, (2005), pp.2613-2622.

[14] Cheng W, Ma L, Yang T, Liang J & Zhang Y, “Joint Lung CT Image Segmentation: A Hierarchical Bayesian Approach”, *PLOS ONE*, Vol.11, No.9, (2016), pp.1-13.

[15] Yao J, Dwyer A, Summers RM & Mollura DJ, “Computer-aided diagnosis of pulmonary infections using texture analysis and support vector machine classification”, *Academic Radiology*, Vol.18, No.3, (2011), pp.306–314.

[16] Zadeh LA, “Fuzzy sets”, *Information and Control*, Vol.8, (1965), pp.338-353.