Published in final edited form as:

Phys Med Biol. 2017 January 21; 62(2): 612–632. doi:10.1088/1361-6560/62/2/612.

Learning with Distribution of Optimized Features for Recognizing Common CT Imaging Signs of Lung Diseases

Ling Ma[1,2], **Xiabi Liu**[2], and **Baowei Fei**[1,3]

[1]Department of Radiology and Imaging Sciences, Emory University, Atlanta, GA

[2]School of Computer Science, Beijing Institute of Technology, Beijing, China

[3]Department of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA

Abstract

Common CT Imaging Signs of Lung Diseases (CISLs) are defined as the imaging signs that frequently appear in lung CT images from patients. CISLs play important roles in the diagnosis of lung diseases. This paper proposes a novel learning method, namely learning with Distribution of Optimized Feature (DOF), to effectively recognize the characteristics of CISLs. We improve the classification performance by learning the optimized features under different distributions. Specifically, we adopt the minimum spanning tree algorithm to capture the relationship between features and discriminant ability of features for selecting the most important features. To overcome the problem of various distributions in one CISL, we propose a hierarchical learning method. First, we use an unsupervised learning method to cluster samples into groups based on their distribution. Second, in each group, we use a supervised learning method to train a model based on their categories of CISLs. Finally, we obtain multiple classification decision from multiple trained models and use majority voting to achieve the final decision. The proposed approach has been implemented on a set of 511 samples captured from human lung CT images and achieves a classification accuracy of 91.96%. The proposed DOF method is effective and can provide a useful tool for computer-aided diagnosis of lung diseases on CT images.

Keywords

Medical image classification; lung CT image; Common CT Imaging Signs of Lung Diseases (CISLs); features selection; hierarchical learning

1. Introduction

Lung cancer is the leading cause of cancer death among males and has surpassed breast cancer as the leading cause of cancer death among females in developed countries (Torre *et al.*, 2015). Cancers of the lung and bronchus account for more than one-quarter (27%) of all cancer deaths, and that makes lung cancer the leading cause of cancer deaths in the United States (Siegel *et al.*, 2015). One of the key issues in addressing this phenomenon is the fact that lung cancer is rarely diagnosed in the early stage. Hence, early detection and treatment of lung cancers are crucially important to improve survival. Computed tomography (CT) is a viable screening tool for lung cancer. It can provide valuable information to distinguish

between normal and diseased lungs. However, it is a time-consuming and error-prone task for radiologists to identify abnormal lesions from a large number of CT images. Therefore, the problem of automatically recognizing lesions in lung CT images for aiding radiologists in the diagnosis of diseases has received extensive attention in recent years.

There are three main purposes for developing lung lesion classification methods in previous works. The first purpose is to classify the tumors into malignant and benign.(Armato III *et al.*, 2003; Niki *et al.*, 2001; Shah *et al.*, 2005b; Way *et al.*, 2009; Lee *et al.*, 2010; Shah *et al.*, 2005a; Suzuki *et al.*, 2005; Iwano *et al.*, 2008) The second purpose is to classify lesion into lung diseases, such as the classification of obstructive lung diseases (Uppaluri *et al.*, 1999), chronic obstructive pulmonary disease (Chabat *et al.*, 2003), chronic lung disease (Uchiyama *et al.*, 2003), Emphysema(Fukushima *et al.*, 2004; Zheng *et al.*, 2004; Xu *et al.*, 2006; Prasad and Sowmya 2008), diffuse lung diseases (Thönnes *et al.*, 2010; Sorensen *et al.*, 2010; Gangeh *et al.*, 2010), interstitial lung diseases (Xu *et al.*, 2011), and pulmonary infectious disease (Yao *et al.*, 2011). The third purpose is for lung tissue classification. Some methods classified lung nodules from false positives (Farag *et al.*, 2004; Ge *et al.*, 2005; Kim *et al.*, 2005; Boroczky *et al.*, 2006; Murphy *et al.*, 2009). Armato *et al.* (Armato *et al.*, 2002) distinguished actual nodules from areas of normal anatomy. Sluimer *et al.* (Sluimer *et al.*, 2003) provided the classification for normal or abnormal lung tissue; and they (Sluimer *et al.*, 2006) classified tissue into six classes: normal, hyperlucency, fibrosis, ground glass, solid, and focal. Ochs *et al.* (Ochs *et al.*, 2007) developed a fully automated approach for the classification of multiple structures (airways, fissures, nodules, and vessels). Farag *et al.* (Farag *et al.*, 2010) classified the lung nodules into four categories: juxta, well-circumscribed, vascularized, and pleural-tail. Depeursinge *et al.* (Depeursinge *et al.*, 2011) proposed a novel texture classification method to categorize lung tissue patterns, including healthy and four pathological lung tissue types (ground glass, fibrosis, micronodules, and emphysema). Song *et al.* (Song *et al.*, 2013b) classified the lung tissue into five categories: normal, emphysema, ground glass, fibrosis, and micronodules. Liu *et al.* (Liu *et al.*, 2015) proposed a new feature selection method to tackle the common CT imaging signs of lung diseases (CISLs) recognition problem, and they developed a genetic optimization to find out candidate features and applied the Fisher criterion to evaluate these candidate features for selecting an optimal feature subset. Ma *et al.* (Ma *et al.*, 2015) proposed a new method of multiple classifier fusion to recognize the CISLs, which is based on the confusion matrices of the classifiers and the classification confidence values outputted by the classifiers.

Some studies improve the classification of lung CT images by selecting effective features or classifier. On the one hand, in order to select the discriminant features, they adopted the existing or proposed new feature selection methods. Shah *et al.* (Shah *et al.*, 2005a) used a feature selection method based on the stepwise model selection and Akaike Information Criterion. Way *et al.* (Way *et al.*, 2009) and Ge *et al.* (Ge *et al.*, 2005) used a stepwise feature selection method. Lee *et al.* (Lee *et al.*, 2010) used a two-step supervised learning system combining a genetic algorithm with the random subspace method. Fukushima *et al.* (Fukushima *et al.*, 2004) adopted a feature selection program based on the divergence measure. Farag *et al.* (Farag *et al.*, 2004) adopted the sequential forward floating selection (SFFS). Sluimer *et al.* (Sluimer *et al.*, 2006; Sluimer *et al.*, 2003) adopted the sequential forward search (SFS). Liu *et al.* (Liu *et al.*, 2015) proposed a new feature selection method

based on FIsher criterion and Genetic optimization (FIG). Wisniewski *et al.* (Wisniewski and Zielinski 2012) used the minimum Redundancy Maximum Relevance (mRMR) algorithm for feature selection. On the other hand, to achieve a good classification performance, various types of single classifier have been used in the past years for lung CT image recognition. For example, they used some single classifier: 1) Linear discriminant analysis (Armato III *et al.*, 2003; Niki *et al.*, 2001; Shah *et al.*, 2005b; Ge *et al.*, 2005; Sluimer *et al.*, 2003; Way *et al.*, 2009; Shah *et al.*, 2005a; Iwano *et al.*, 2008; Armato *et al.*, 2002; Farag *et al.*, 2010), 2) Artificial neural networks (Thönnes *et al.*, 2010; Kim *et al.*, 2005; Suzuki *et al.*, 2005; Gangeh *et al.*, 2010), 3) Quadratic discriminant (Shah *et al.*, 2005b; Sluimer *et al.*, 2003; Shah *et al.*, 2005a), 4) Logistic regression (Shah *et al.*, 2005b, a), 5) *K*-nearest neighbor (Sluimer *et al.*, 2006; Xu *et al.*, 2011; Sluimer *et al.*, 2003; Murphy *et al.*, 2009; Chabat *et al.*, 2003), 6) Support vector machine (SVM) (Way *et al.*, 2009; Boroczky *et al.*, 2006; Prasad and Sowmya 2008; Sluimer *et al.*, 2003; Sorensen *et al.*, 2010; Yao *et al.*, 2011), 7) Bayesian classifier (Farag *et al.*, 2004; Fukushima *et al.*, 2004; Uppaluri *et al.*, 1999; Zheng *et al.*, 2004), and 8) Delegated classifiers (Xu *et al.*, 2006). Some work used several ensemble classifiers, such as random subspace method (Lee *et al.*, 2010), AdaBoost (Ochs *et al.*, 2007), and a new fusion method in a weighted-sum form (Ma *et al.*, 2015).

Current methods for the classification of lung CT images suffer from two limitations. 1) The same category of CT imaging signs may be observed in the images corresponding to different diseases. Different categories of CT imaging signs may also appear in the images with the same disease. Hence, the correlation between CT imaging signs and diseases is complicated. It is difficult to achieve a high accuracy of diagnosis of lung diseases according to CT imaging sign in an automatic approach. To improve diagnosis accuracy, it is necessary to provide enough information of CT imaging signs for radiologists to diagnose diseases. The Common CT Imaging Signs of Lung Diseases (CISLs) are defined as the imaging signs that frequently appear in lung CT images from patients with lung diseases, which are really often encountered and widely used in the diagnosis of lung diseases (Han *et al.*, 2015). So it is useful to aid radiologists in the diagnosis of lung diseases by recognizing the categories of CT imaging signs in the regions of interests (ROIs). However, the classification of CISLs has not received much attention of researchers. 2) For the classification of CISLs, the existing feature selection methods do not consider the relationship between features. We call the relationship as context information of features. Existing classification methods do not consider the various distribution patterns in one category. Although they (Liu *et al.*, 2015; Ma *et al.*, 2015) pay attention to the classification of CISLs, neither the context knowledge nor the diversity of patterns was used for the classification.

This paper focuses on the classification of Common CT Imaging Signs of Lung Diseases (CISLs) based on the learning with Distribution of Optimized Features (DOF).

On the one hand, we select the discriminant features based on the minimum spanning tree (MST) algorithm. We construct a weighted graph where the node represents each dimension feature, the edge links each dimension feature, and the weight of edge represents the relationship between any two dimension features and their classification ability. Hence, we involve the context information of features through the graph structure. By computing the MST of the weighted graph, we can obtain the importance of each feature and select the

most important features to character the ROIs. Different from the other MST based feature selection methods (Liu *et al.*, 2014; Song *et al.*, 2013a), our method selects the optimized feature over the global scope, not at the local range. Our method selects features from the whole features, while Song *et al.* (Song *et al.*, 2013a) and Liu *et al.* (Liu *et al.*, 2014) cluster the whole features into several groups and then choose features from each group. That likely misses the best global features. Besides, we can minimize the feature-to-feature redundancy and maximize the feature-to-class relevance simultaneously, in order to find out an optimal feature subset for the original data. Not like the work (Song *et al.*, 2013a), they separated them and performed sequentially.

On the other hand, based on the selected features, we use a hierarchy method for the classification of CISLs. First, we cluster all the ROIs represented by optimized features into k groups based on the self-organizing map (SOM). Then, in each group, we train a classifier. The majority voting is adopted to fuse the labels predicted by k classifiers.

The main contributions of this work include several aspects. 1) We propose a new feature selection method using the MST. It involves not only the correlation among features but also the ability of selecting the discriminative feature. In addition, it takes the advantage of the graph structure to capture the context information for a good ranking of the features. 2) We propose a new hierarchy classification framework to learn the distribution of features. Since the same concept may have different distributions and same distribution may exist in the different concepts, it is desirable to decompose the features into classes with different distributions. The hierarchy classification method can divide the ROIs into k groups according to their distribution of features, and then train classifier in each cluster. By fusing the classification results from multiple classifiers with different distributions, it can improve the classification performance. 3) We demonstrate these techniques on lung imaging data from human clinical patients. We conducted the 5-fold cross validation experiments on a set of 511 ROIs from clinical lung CT images of 252 patients and achieved satisfactory classification performance.

2. Method

2.1 Learning with distribution of optimized feature (DOF)

We usually extract several different types of features to better character a ROI. Since these features may contain complementary or irrelevant information, we propose a new feature selection method based on graph theory for improving the efficiency and accuracy of classification. To learn the relationship between the distribution and class of samples for good classification, we propose a hierarchy learning method based on the unsupervised and supervised learning. We can predict the category of a new ROI by using the hierarchy models and majority voting. The flowchart of the proposed method is shown in figure 1, which includes two parts: 1) model training under different distributions of optimized features, and 2) recognition of a new ROI. The details of the two stages are described in the following.

2.1.1. Constructing a weighted graph—As shown in figure 2, we construct a connected, undirected graph (V, E), where V is the set of vertices and E is the set of edges,

to describe the original features and their relationships. The vertices represent the features and the weights of the edges represent their relationship and discrimination capability. Given an n-dimension feature vector, we can construct a graph with n vertices and $n(n-1)/2$ edges. For the weight of each edge, we simultaneously consider the discriminating ability of one feature and the correlation coefficients between two features. Let F_i and F_j be the i-th and j-th feature, respectively. Then the weight of edge linking F_i and F_j is computed by:

$$w(i,j)=\frac{\mathrm{coef}(F_i,F_j)}{\mathrm{ability}(F_i)+\mathrm{ability}(F_j)}, \quad (1)$$

where $coef(F_i, F_j)$ means the correlation coefficient between F_i and F_j, and $ability(F_i)$ is the classification accuracy of a classifier inputted with F_i feature. The coefficient can be computed according to the linear dependence of two features. If each feature contains L observations and the l-th observation of F_i is represented by F_{il}, then the $coef(F_i, F_j)$ can be defined as

$$\mathrm{coef}(F_i,F_j)=\frac{1}{L-1}\sum_{l=1}^{L}\left(\frac{\overline{F_{\mathrm{il}}-\mu_i}}{\sigma_i}\right)\left(\frac{F_{\mathrm{jl}}-\mu_j}{\sigma_j}\right), \quad (2)$$

where μ_i and μ_j are the means of F_i and F_j, and σ_i and σ_j are the standard deviation of F_i and F_j.

In (1), the numerator is the coefficient between two features. A high coefficient indicates more redundancy. While the denominator is represented by the classification accuracy based on the single feature F_i and F_j. If the accuracy is higher, it means that the two features have a higher ability to distinguish different concepts. If the two features have the less redundancy and higher classification ability, the weight value between them is smaller. If a graph has a weight which is the sum of the weights of the edges in that graph, then the optimized feature subset is corresponding to a subgraph with a smallest weight.

2.1.2. Selecting optimized features based on the MST—We can obtain the subgraph with the smallest weight by computing the MST of the weighted graph for selecting the optimized features.

A spanning tree of a graph is an acyclic subgraph which connects all the vertices together. It owns the same vertices with the graph, but the set of edges of the spanning tree is the subset of edges of the graph. A graph can have many different spanning trees. A MST (Graham and Hell 1985) is the spanning tree with a weight less than or equal to the weight of every other spanning tree. We adopt Prim's algorithm (Prim, 1957) to grow a MST. It starts with an empty graph and tries to add one edge owing minimum weight at a time. During this process, it guarantees that the tree remains acyclic. According to the MST, we can measure the discriminating ability of feature F_i, represented by $DA(F_i)$, as

$$\mathrm{DA}(F_i)=\sum_{j=1}^{R} w(E_{\mathrm{ij}}), \quad (3)$$

where E_{ij} is the edge between F_i and F_j in the MST, $w(E_{ij})$ is the weight of the edge, and R is the number of edges connecting to the vertex of F_i. We rank the features in an ascending order according to their $DA(\cdot)$.

We then adaptively choose the best size m of the feature subset based on feature coherency measured by the Fisher value. Let F_{ij}^m be the top m features of sample j in the class i, $\overline{F_i^m}$ be the mean of features in the class i, $\overline{F^m}$ be the mean of features in all the classes, D_{intra}^m be the intra-class scatter, and D_{inter}^m be the inter-class scatter. They are defined by

$$D_{\mathrm{intra}}^m=\sum_{i=1}^{C}\frac{1}{N_i}\sum_{j=1}^{N_i}\|F_{\mathrm{ij}}^m-\overline{F_i^m}\|, \quad (4)$$

$$D_{\mathrm{inter}}^m=\sum_{i=1}^{C}\|\overline{F_i^m}-\overline{F^m}\|^2,$$

where C is the number of classes and N_i is the number of samples in class i.

According to the intra-class scatter and the inter-class scatter, we can compute the Fisher value based on the subset of features with size m by

$$\mathrm{Fisher}(m)=\frac{D_{\mathrm{inter}}^m}{D_{\mathrm{intra}}^m}. \quad (5)$$

In (5), a high fisher value means a large inter-class scatter over the intra-class scatter. So a feature set owning a higher fisher value is more discriminate.

For obtaining a stable value, we calculate the mean and variance of the Fisher value of size m based on an interval of integer numbers $[m-2, m+2]$ as the value of the feature coherency:

$$f(m)=\frac{\mu(m)+\mathrm{var}(m)}{m}, \quad (6)$$

where $\mu(\cdot)$ and $var(\cdot)$ are the mean and variance function. They are defined by

$$\mu(m)=\frac{1}{(m+2)-(m-2)+1}\sum_{m_k=m-2}^{m+2}\mathrm{Fisher}(m_k)$$

$$\mathrm{var}(m)=\frac{1}{(m+2)-(m-2)+1}\sum_{m_k=m-2}^{m+2}(\mathrm{Fisher}(m_k)-\mu(m))^2. \quad (7)$$

Then the optimized m can be defined as follows:

$$m^*=\arg\ \max \nabla f(m). \quad (8)$$

2.1.3. Training with consideration of distributions—We use an unsupervised learning method to cluster the samples into several groups according to the distribution of features and under the ignorance of their true categories. Then we use a supervised learning method to train the model based on the samples in each cluster and their labels.

2.1.3.1. Unsupervised learning based on SOM: A Self-Organizing Map (SOM) (Kohonen and Somervuo 1998) is a type of artificial neural network, which is trained using unsupervised learning based on competitive learning. It implements the competition by the negative feedback paths between neurons.

Since SOM not only can divide the data into different groups but also can find out more particular details in the structure of the data, we adopt SOM as our unsupervised learning to separate the samples into k groups according to their distribution of features and ignoring their category. Please note that the k is not the number of classes of CISLs. It can be an integer between one to the number of training samples. To achieve a better cluster performance, we determine the best number of groups, k, by experiments.

2.1.3.2. Supervised learning based on AdaBoost: AdaBoost (Freund and Schapire 1996) is one of boosting methods. It is an ensemble classifier which combines many other types of learning algorithms to improve their performance. It gives the final output by fusing these outputs of the other learning algorithms ('weak learners') in a weighted sum. Since AdaBoost can combine a sequence of weak classifiers by adjusting the weights of each classifier dynamically according to the errors in the previous learning step, it can converge to a strong learner. Since a decision tree (Safavian and Landgrebe 1990) owns the fast speed and superior performance, we use the decision tree as the week classifier.

After unsupervised learning based on SOM, the training ROIs can be divided into k groups. In each group, we train the AdaBoost with decision trees for distinguishing the ROIs which have the same distribution, but the different classes.

2.2. Classification of CISLs

2.2.1. CISLs—There are some well-known categories of CT findings of lung lesions that frequently appear in patients' lung CT images and play important roles in the diagnosis of lung diseases. We call this kind of CT findings as the common CT imaging signs of lung diseases (CISL). We summarized nine categories of CISLs, including Ground Glass Opacity (GGO), lobulation, cavity & vacuolous (CV), spiculation, pleural indentation (PI),

obstructive pneumonia (OP), calcification, air bronchogram (AB), and bronchial mucus plugs (BMP). These CISLs are really widely used in the diagnosis of lung diseases, and are illustrated in figure 3.

In general, GGO can be characterized by the areas of hazy increased attenuation of the lung with preservation of bronchial and vascular margins. It is associated with the adenocarcinoma of lung and bronchioloalveolar carcinoma. Lobulation is dependent on the ingrowth of connective tissue septa containing fibroblasts derived from peri-thymic mesenchyme. It indicates a malignant lesion. Cavity and Vacuolous (CV) is a gas-filled space, seen as a lucency or low-attenuation area. They are associated with the adenocarcinoma, bronchioloalveolar carcinoma and tumors larger than 3 cm. Speculation is a roughly set of lines radiating from a central point or region. It is caused by the intrusion of cancer into surrounding tissue. Pleural Indentation (PI) shows that the pleural is dragged toward the lung area by the spiculation. It is associated with most peripheral adenocarcinomas containing a central or subpleural anthracotic and fibrotic focus. Obstructive Pneumonia (OP) shows a flabellate or cuneate area with increased density. It is associated with adenocarcinoma and squamous cell carcinoma. Calcification is a punctuated or nodular area with high density. It suggests malign or benign lesion based on its appearance. Air Bronchogram (AB) is a tubular outline of an airway made visible by filling of the surrounding alveoli by fluid or inflammatory exudates. It is associated with the lung cancer, pulmonary pneumonia, and lymphoma. Bronchial Mucus Plugs (BMP) shows that the intrabronchial air is replaced by the mucus. It is associated with the allergic bronchopulmonary aspergillosis (Han *et al.*, 2015).

2.2.2. Feature extraction—To better characterize the ROI, we extract several different types of features, which are the local binary pattern (LBP), the bag-of-visual-words based on the HOG (B-HOG), the wavelet features, and the histogram of CT values (CVH).

LBP: For each pixel in a ROI, a binary code is produced by comparing a circularly symmetric neighborhood with the value of the center pixel and is transformed into an integer. The LBP feature vector can be obtained by figuring out the frequency of each integer. We can define the neighborhood in the LBP operator flexibly by (P, R), which means we evenly sample P neighbors evenly sampled on the circle of radius R around the center pixel. Then, the corresponding LBP features will be denoted as LBP (P, R) in the following descriptions. We consider multiple P (4 or 5) and R (1 or 2) to get multi-scale LBP features.

B-HOG: We partition a ROI into blocks of 8×8 pixels and divide each block into 4 cells of 4×4 pixels. Then, we compute the orientation histogram for each cell which contains 9 bins covering a gradient orientation range of 0°–180°. Finally, the HOG feature vector is extracted for each block by the linking of the orientation histograms of cells in it. However, this widely used strategy is not applicable in this work because the size of ROIs in lung CT images varies with different patients and different pathological lesions. Hence, we adopt the bag-of-visual-words on HOG features as the ROI representation. We employ the Expectation-Maximization (EM) and the Minimum Description Length (MDL) to learn a Gaussian Mixture Modeling (GMM) for generating more accurate visual words. The HOG feature vector of each block is mapped to the visual word corresponding to the highest

likelihood for it and the number of HOG feature vectors assigned to each visual word is accumulated. The final B-HOG feature vector is formed by normalizing the numbers.

Wavelet features: Wavelet feature is a common spectral texture feature, which is calculated from the image transformed into frequency domain. It can capture localized spatial and frequency information and multi-resolution characteristics effectively. In this paper, by using 2D symlets wavelet, the ROIs are decomposed to 4 levels. Then the horizontal, vertical and diagonal detailed coefficients are extracted from the wavelet decomposition structure. Finally, we get the wavelet features by calculating the mean and variance of these wavelet coefficients.

CVH features: CVH means the histogram of CT values. In lung CT images, the CT values of pixels are expressed in Hounsfield Units (HU). We compute the histogram of CT values over each ROI. The number of bins in the histogram is 40 because it can lead to the highest classification accuracy among the numbers from 20 to 60 at the step of 10 according to (Liu *et al.*, 2015).

2.2.3. Classification based on the DOF—Based on the four types of features, we select the optimized subset of features based on the MST. We cluster the optimized features into k groups and train an AdaBoost with 100 decision trees in each cluster. For a given ROI, we will use the k AdaBoost models to recognize its class. Let ROI_{test} be the testing ROI. Its final class, which is represented by $C_f(\mathrm{ROI}_{test})$ and recognized by our DOF, is defined by:

$$C_f(\mathrm{ROI}_{\mathrm{test}})=\mathrm{MV}(C_1(\mathrm{ROI}_{\mathrm{test}}),C_2(\mathrm{ROI}_{\mathrm{test}}),\ldots,C_k(\mathrm{ROI}_{\mathrm{test}})), \quad (9)$$

where MV(·) is the majority voting rule which fusing the classes from k trained models from k groups, $C_k(\mathrm{ROI}_{test})$, is the class identified by the k-th model for sample ROI_{test}.

2.2.4. Evaluation criterion—The performance of CISL recognition is evaluated by the sensitivity (SE), specificity (SP), and classification accuracy.

The sensitivity and specificity are widely used in the medical image classification community. They are essentially two measurements of performance of binary classifiers. We use them to reflect the ability of our method for discriminating one CISL category from other categories. If a positive example for a CISL category can be recognized correctly by the algorithm, we call it “true positive”; otherwise we call it “false negative”. The meaning of “true negative” and “false positive” is defined similarly, respectively. Let TP, TN, FP, FN be the number of true positives, true negatives, false positives and false negatives for one CISL category, respectively. Then the sensitivity and specificity of the classifier for one category are measured

$$\mathrm{Sensitivity}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \quad (10)$$

$$\text{Specificity}=\frac{\text{TN}}{\text{TN}+\text{FP}}$$

Our CISL recognition problem is also a multi-class classification problem. So we use the classification accuracy to give an overall measurement of classification performance. It is the ratio of the number of correctly classified examples to the number of all examples.

2.3. Databases

The instances of nine categories of CISLs from patients who were clinically selected from the Cancer Institute and Hospital at Chinese Academy of Medical Sciences were collected. The lung CT images were acquired by CT scanners of GE LightSpeed VCT 64 and Toshiba Aquilion 64, and saved slice by slice according to DICOM 3.0 standard. Their slice thickness is 5 mm, the image resolution is 512 × 512, and the in-plane pixel spacing ranges from 0.418mm to 1mm (mean: 0.664 mm).

The ground truth was provided by a qualified senior radiologist, a physician-in-charge with 11 years' experiences on radiography. The qualified radiologist manually labeled and annotated the rectangular ROIs wrapping CISLs in lung CT images. The resultant numbers of ROIs are 511. To conduct 5-fold cross validation experiments, we split the available instances into five disjoint subsets nearly evenly and guarantee that the instances in different subsets come from different patients for avoiding the bias in measuring classification performance. Table 1 lists the numbers of ROI instances in five data subsets, the numbers of patients for each CISL category, where D1-D5 denote the first to the fifth subsets, respectively, and NoP means the number of patients and the size range (mm × mm) of ROIs for each CISL. Actually, each of five data subsets is taken as the test set in turn. Then the four subsets in the remaining data are the training set.

3. Results

3.1. Parameter tuning

We performed ten-fold cross validation and a search on the training data to select the optimal value for the number of cluster, parameter k. We tested 9 numbers from 2 to 10. We cluster the ROIs into the tested number k and train k AdaBoost models. Based on the k models, we perform the CISL recognition in ten-fold cross validation on the five training dataset and record the average classification accuracy. The resultant results are listed in figure 4, from which we can see that the best number is 6, 6, 4, 5 and 7 for the five rounds.

3.2. Overall performance of the proposed DOF

We conducted the CISLs classification by using our DOF method in five-fold cross validation experiments. Table 2 shows that the cross validation results, where "SE" and "SP" mean "sensitivity" and "specificity", respectively. From table 2, we can see that our method achieves good classification performance. It achieves not only the high classification accuracy, but also the high sensitivity and specificity. We have the accuracy of 91.96%. Our method can give a high sensitivity of more than 90% for GGO, Calcification, CV, PI, BMP

and OP. Especially, for OP, we can get a sensitivity of 100%. The specificities for all CISLs are close to 100%, except the ones for CV and PI which are 94.23% and 97.48%, respectively.

3.3. Effectiveness of the proposed DOF

Our DOF considers both the distribution and optimized features. To illustrate the effectiveness of our learning framework, we test the performance of all individual parts of our method. We call them method NDNOF, NDOF, and DNOF for short. The method NDNOF (no distribution and no optimized features) is a classification framework without considering the optimized features and their distribution, the method NDOF which doesn't consider the distribution but involves optimized features, and the method DNOF which considers the distribution and uses the original features. All of these methods classify the CISLs based on the same classifier AdaBoost with 100 decision trees. Their differences lie at the features selected and learning strategies, which are shown in table 3.

We perform the four methods on the classification of CISLs and show the classification accuracy in figure 5. From figure 5, we can see that our DOF can achieve the highest accuracy. NDOF and DNOF can improve the accuracy compared with NDNOF. That proves the significance of feature selection based on the MST and the hierarchy learning based on SOM and AdaBoost. It also illustrates that our DOF can achieve a better performance by considering the two parts together.

3.4. Advantage of our feature selection method

In our feature selection method, we extract multiple types of features, combine them into an integrated feature vector, and select the discriminant features from the integrated feature vector. We list the dimension of original features and selected features for each type of feature in our 5-fold cross validation experiments in table 4. In table 4, we can see that our feature selection method reduces the dimension of features, and some features in each type of feature are selected in each fold. That indicates not only our method is effective, but also, each type of feature contains discriminant information and there exists complementary information among them.

Since method FIG (Liu *et al.*, 2015) performed the feature selection method on the CISLs data, and mRMR (Peng *et al.*, 2005) is well known feature selection framework focusing on the redundancy and relevance, we compare our feature selection method with them. The two compared methods are conducted on the same dataset and in five rounds. FIG and our method can adaptively choose the number of optimized features while mRMR needs user to provide a size of selected features.

Firstly, we show the number of optimized features selected by FIG and our method in table 5. From table 5, we can see our method can get more compact features than FIG method in most situations, which means our selected features have less redundancy and can be implemented efficiently.

Secondly, we use AdaBoost as classifier to classify the CISLs based on the features selected by mRMR, FIG, and our method, where we test five different numbers (30, 50, 80, 120 and

150) of features selected for mRMR. The compared results are shown in figure 6. From figure 6, we can see our method can obtain the best performance.

3.5. Advantage of our classification method

For a good classification of CISLs, Ma et al (Ma *et al.*, 2015) used a parallel strategy to combine various types of single classifier in a weighted-sum form. Compared with it, our DOF combine the unsupervised and supervised learning in a serial structure to give full consideration to the distributions of optimized features. We conducted the compared method and our method on the same dataset and in five rounds. We record the average classification accuracy and show them in figure 7. From figure 7, we can say our DOF can improve the classification performance greatly. Our method brought increase rates of 20.7% in classification accuracy compared with Ma et al (Ma *et al.*, 2015). That proves the distributions of optimized features is important to the classification of CISL and our method is effective and promising.

3.6. Robustness of our method

To test the robustness of our method, we ran our method on the noisy data, on the combined data from lesion and normal tissues, and on the data from different patient's age groups.

Firstly, we add the salt & pepper noise with the noise density of 0.02 on each ROI. Then, we extract and select the features on the noisy ROIs. We conduct 5-fold cross validation experiments on the noisy ROIs and show the average results in Table 6. We give the example of original and noisy image for each CISL in the second and third columns, respectively, where the CISLs are indicated in a red-bordered box around for a clear view. We can achieve a high sensitivity of more than 90% for four CISLs, specificity with nearly 100%, and accuracy with 90.41%. Hence, the good classification performance proves that our method is robust to the noise.

Secondly, we add some normal tissue into our database for changing the classification of CISLs into the classification of abnormal and normal tissue, in order to test the performance of our method. The radiologist who produced the gold standard selects 85 rectangular ROIs wrapping normal tissue with a size from 12 mm×18 mm to 70 mm×53 mm from 85 different patients. We split these normal ROIs into five disjoint parts evenly and insert them into the five disjoint subsets described in section 2.3 for the 5-fold cross validation experiments. We show the average results in Table 7. We obtain the classification accuracy of 91.44%, a sensitivity of 97.65%, and a specificity of 99.03% for the normal tissue. These good classification results prove that our method can be used to distinguish between the CISLs and the normal tissue.

Finally, since the lung structures are dependent on the patient age and the age of patients in our database is from 30 to 89, we divide the instances in the database into two parts, Part A and Part B, according to their ages. Part A includes the patients from 30 to 54 years old while Part B includes the patients from 55 to 89 years old. We use the instances in Part A as the training data and the instances in Part B as the testing data to conduct 5-fold cross validation experiments for evaluating the robustness of our algorithms to the patient's age. We give the details (the patient age and number of ROIs in each CISL in training data and

testing data) and average results in Table 8. Our method can achieve good classification performance under the different patient ages. Hence, our method is robust to the patient age.

4. Discussion

A novel learning method with Distribution of Optimized Feature (DOF) is put forward to improve the recognition of Common CT Imaging Signs of Lung Diseases (CISLs). It consists of two parts: (i) feature selection based on the MST; (ii) hierarchy learning based on the unsupervised learning (SOM) and supervised learning (AdaBoost). The proposed approach has been tested on a set of 511 samples captured from human lung CT images and our proposed method can obtain a classification accuracy of 91.96%.

4.1. Common CT Imaging Signs of Lung Diseases (CISLs)

In clinical practice, generally, radiologists relay on the analysis of CT findings for the diagnosis of diseases. However, the same CT finding may be within the different types of disease and the different categories of CT findings may within the same type of disease. Hence, it is significant not only for diagnostics but also for medical research and teaching to provide a classification method for the CT findings.

This paper focuses on the classification of Common CT Imaging Signs of Lung Diseases (CISLs). Nine types of CT findings are selected as the CISLs, by the radiologists in the Cancer Institute and Hospital at Chinese Academy of Medical Sciences, according to the diagnosis of lung diseases clinically. They are Ground Glass Opacity (GGO), lobulation, cavity & vacuolous (CV), spiculation, pleural indentation (PI), obstructive pneumonia (OP), calcification, air bronchogram (AB), and bronchial mucus plugs (BMP). Although they are not the complete set of CT findings, they frequently appear in the lung CT images from patients and play important roles in the diagnosis of lung diseases.

4.2. Feature selection

In this study, we select the optimized features based on the MST. We use the graph structure to involve the ability, correlation, and the context information of features to obtain a good ranking of features. And select the top m features as our optimized features where m is computed adaptively based on the Fisher value.

Different from other MST based feature selection methods (Liu *et al.*, 2014; Song *et al.*, 2013a), we just use the degree of each node in the MST to obtain the ranking order of each feature, and have not involved any clustering procedure for the feature selection. Compared with the well-known mRMR feature selection method (Peng *et al.*, 2005), our method not only adaptively determines the number of selected features, but also can select the more effective features in discriminating among ROIs from different classes of CISLs. We use the AdaBoost as the classifier to classify the features selected by mRMR and our method on the same database and experimental results show that our method can give higher classification accuracy. Compared with the FIG method (Liu *et al.*, 2015) developed for the CISLs, our method not only can give a better performance, but also can select the less features to improve the efficiency.

Hence, our MST based feature selection method can choose the discriminative features for the improvement of the classification of CISLs.

4.3 Hierarchy learning

Our CISLs have unique character that the intra-class variation is large on account of the severity of disease and the location located in the lung, and the inter-class variation varies drastically across different pairs of classes. For example, GGO is similar to calcification but is different from air bronchogram. To solve these problems, we propose a hierarchy learning method. For the first level, we use SOM to separate ROIs into groups under the ignorance of their categories of CISLs. For the second level, we train the AdaBoost model in each cluster to distinguish similar ROIs with different categories of CISLs.

We compare our hierarchy learning method with the combined classification method (Ma *et al.*, 2015) which fuse multiple various types of classifier in a weighted-sum form for the classification of CISLs. Our advantages on efficiency over theirs are shown through experiments.

Hence, our hierarchy learning can capture the character of CISLs, and achieve good classification performance.

4.4 Robustness

Since the features of lung extracted from CT images may vary with different patient ages, the quality of the imaging and so on, we perform our method on the different data for a proof of robustness. In order to test the influences of the patient age, we use the instances with different patient ages as the training and testing data. Comparing the results in table 2 and table 8, we can see the performance of our method is not affected by patient age. Although we add the noise and normal tissue into our classification framework, we can obtain relatively stable results. These results prove our method is not only effective, but also robust.

5. Conclusions

This paper proposed a new learning framework based on the distribution of optimized features for the classification of Common CT Imaging Signs of Lung Diseases (CISLs). The framework represents a new hierarchy classification method with a set of efficient features. It captures the context information of features to select the discriminate features based on a graph structure (minimum spanning tree). Before training classifier, it clusters the samples according to their distribution under the ignorance of the class of CISLs using the neural network architecture (self-organizing map). When compared with those algorithms without the considerations of context information or hierarchy learning, the proposed method improves the classification performance. In addition, our classification method can perform better for the CISLs classification than some other methods. Our optimized features can be consistent with the target concept by exploring the context information of features in a graph structure. Our models can distinguish the samples with same concepts in the different distribution by training the classifiers in a hierarchy learning framework. Hence, our method which considers the distribution of optimized features can offer significantly enhanced

abilities to characterize lung lesions for an improved classification of CISLs. As CISL is closely related to lung diseases, our proposed method has the potential to aid radiologists in decision making during the clinical practice.

Acknowledgments

This research was supported in part by U.S. National Institutes of Health (NIH) grants (CA176684 and CA156775). LM was partially supported by International Graduate Exchange Program of Beijing Institute of Technology. XL was partially supported by National Natural Science Foundation of China (Grant no. 60973059, 81171407) and the Program for New Century Excellent Talents in Universities of China (Grant No. NCET-10-0044). The work was also supported in part by the Georgia Research Alliance (GRA) Distinguished Cancer Scientist Award to BF.

References

Armato SG III, Altman MB, Wilkie J, Sone S, Li F, Doi K, Roy AS. Automated lung nodule classification following automated nodule detection on CT: A serial approach. Medical Physics. 2003; 30:1188–97. [PubMed: 12852543]

Armato SG, Li F, Giger ML, MacMahon H, Sone S, Doi K. Lung cancer: Performance of automated lung nodule detection applied to cancers missed in a ct screening program 1. Radiology. 2002; 225:685–92. [PubMed: 12461246]

Boroczky L, Zhao L, Lee KP. Feature subset selection for improving the performance of false positive reduction in lung nodule CAD. IEEE Transactions on Information Technology in Biomedicine. 2006; 10:504–11. [PubMed: 16871718]

Chabat F, Yang G-Z, Hansell DM. Obstructive lung diseases: texture classification for differentiation at CT 1. Radiology. 2003; 228:871–7. [PubMed: 12869685]

Depeursinge, A., Foncubierta-Rodriguez, A., Van de Ville, D., Müller, H. Lung texture classification using locally–oriented Riesz components; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2011. p. 231-8.

Farag, A., Elhabian, S., Graham, J., Farag, A., Falk, R. Toward precise pulmonary nodule descriptors for nodule type classification; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2010. p. 626-33.

Farag, AA., El-Baz, A., Gimel'farb, GG., Falk, R., Hushek, SG. Automatic detection and recognition of lung abnormalities in helical CT images using deformable templates; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2004. p. 856-64.

Freund Y, Schapire RE. Experiments with a new boosting algorithm. Icml, 1996. 1996; 96:148–56.

Fukushima A, Ashizawa K, Yamaguchi T, Matsuyama N, Hayashi H, Kida I, Imafuku Y, Egawa A, Kimura S, Nagaoki K. Application of an artificial neural network to high-resolution CT: usefulness in differential diagnosis of diffuse lung disease. American Journal of Roentgenology. 2004; 183:297–305. [PubMed: 15269016]

Gangeh, MJ., Sørensen, L., Shaker, SB., Kamel, MS., De Bruijne, M., Loog, M. A texton-based approach for the classification of lung parenchyma in CT images; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2010. p. 595-602.

Ge Z, Sahiner B, Chan H-P, Hadjiiski LM, Cascade PN, Bogot N, Kazerooni EA, Wei J, Zhou C. Computer-aided detection of lung nodules: false positive reduction using a 3D gradient field method and 3D ellipsoid fitting. Medical physics. 2005; 32:2443–54. [PubMed: 16193773]

Graham RL, Hell P. On the history of the minimum spanning tree problem. Annals of the History of Computing. 1985; 7:43–57.

Han G, Liu X, Han F, Santika INT, Zhao Y, Zhao X, Zhou C. The LISS—A Public Database of Common Imaging Signs of Lung Diseases for Computer-Aided Detection and Diagnosis Research and Medical Education. IEEE Transactions on Biomedical Engineering. 2015; 62:648–56. [PubMed: 25330480]

Iwano S, Nakamura T, Kamioka Y, Ikeda M, Ishigaki T. Computer-aided differentiation of malignant from benign solitary pulmonary nodules imaged by high-resolution CT. Computerized Medical Imaging and Graphics. 2008; 32:416–22. [PubMed: 18501556]

Kim KG, Goo JM, Kim JH, Lee HJ, Min BG, Bae KT, Im J-G. Computer-aided diagnosis of localized ground-glass opacity in the lung at CT: initial experience 1. Radiology. 2005; 237:657–61. [PubMed: 16192320]

Kohonen T, Somervuo P. Self-organizing maps of symbol strings. Neurocomputing. 1998; 21:19–30.

Lee MC, Boroczky L, Sungur-Stasik K, Cann AD, Borczuk AC, Kawut SM, Powell CA. Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction. Artificial intelligence in medicine. 2010; 50:43–53. [PubMed: 20570118]

Liu, Q., Zhang, J., Xiao, J., Zhu, H., Zhao, Q. A Supervised Feature Selection Algorithm through Minimum Spanning Tree Clustering; IEEE 26th International Conference on Tools with Artificial Intelligence; 2014. p. 264-71.

Liu X, Ma L, Song L, Zhao Y, Zhao X, Zhou C. Recognizing common CT imaging signs of lung diseases through a new feature selection method based on Fisher criterion and genetic optimization. IEEE journal of biomedical and health informatics. 2015; 19:635–47. [PubMed: 25486652]

Ma L, Liu X, Song L, Zhou C, Zhao X, Zhao Y. A new classifier fusion method based on historical and on-line classification reliability for recognizing common CT imaging signs of lung diseases. Computerized Medical Imaging and Graphics. 2015; 40:39–48. [PubMed: 25453465]

Murphy K, van Ginneken B, Schilham AM, De Hoop B, Gietema H, Prokop M. A large-scale evaluation of automatic pulmonary nodule detection in chest CT using local image features and k-nearest-neighbour classification. Medical Image Analysis. 2009; 13:757–70. [PubMed: 19646913]

Niki N, Kawata Y, Kubo M. A CAD system for lung cancer based on CT image. International Congress Series. 2001; 1230:631–8.

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. Medical Image Analysis. 2007; 11:315–24. [PubMed: 17482500]

Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence. 2005; 27:1226–38. [PubMed: 16119262]

Prasad, M., Sowmya, A. Multi-level classification of emphysema in HRCT lung images using delegated classifiers; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2008. p. 59-66.

Prim RC. Shortest connection networks and some generalizations. Bell system technical journal. 1957; 36:1389–401.

Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. 1990

Shah SK, McNitt-Gray MF, Rogers SR, Goldin JG, Suh RD, Sayre JW, Petkovska I, Kim HJ, Aberle DR. Computer-aided Diagnosis of the Solitary Pulmonary Nodule 1. Academic radiology. 2005a; 12:570–5. [PubMed: 15866129]

Shah SK, McNitt-Gray MF, Rogers SR, Goldin JG, Suh RD, Sayre JW, Petkovska I, Kim HJ, Aberle DR. Computer Aided Characterization of the Solitary Pulmonary Nodule Using Volumetric and Contrast Enhancement Features 1. Academic Radiology. 2005b; 12:1310–9. [PubMed: 16179208]

Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA: a cancer journal for clinicians. 2015; 65:5–29. [PubMed: 25559415]

Sluimer IC, Prokop M, Hartmann I, van Ginneken B. Automated classification of hyperlucency, fibrosis, ground glass, solid, and focal lesions in high-resolution CT of the lung. Medical Physics. 2006; 33:2610–20. [PubMed: 16898465]

Sluimer IC, van Waes PF, Viergever MA, van Ginneken B. Computer-aided diagnosis in high resolution CT of the lungs. Medical physics. 2003; 30:3081–90. [PubMed: 14713074]

Song Q, Ni J, Wang G. A fast clustering-based feature subset selection algorithm for high-dimensional data. IEEE transactions on knowledge and data engineering. 2013a; 25:1–14.

Song Y, Cai W, Zhou Y, Feng DD. Feature-based image patch approximation for lung tissue classification. IEEE transactions on medical imaging. 2013b; 32:797–808. [PubMed: 23340591]

Sorensen L, Shaker SB, De Bruijne M. Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE transactions on medical imaging. 2010; 29:559–69. [PubMed: 20129855]

Suzuki K, Li F, Sone S, Doi K. Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network. IEEE Transactions on Medical Imaging. 2005; 24:1138–50. [PubMed: 16156352]

Thönnes E, Bhalerao A, Parr DG. Classification of lung disease in HRCT scans using integral geometry measures and functional data analysis. Proceedings of Medical Image Understanding and Analysis. 2010:25–30.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians. 2015; 65:87–108. [PubMed: 25651787]

Uchiyama Y, Katsuragawa S, Abe H, Shiraishi J, Li F, Li Q, Zhang C-T, Suzuki K, Doi K. Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. Medical Physics. 2003; 30:2440–54. [PubMed: 14528966]

Uppaluri R, Hoffman EA, Sonka M, Hunninghake GW, McLennan G. Interstitial lung disease: a quantitative study using the adaptive multiple feature method. American Journal of Respiratory and Critical Care Medicine. 1999; 159:519–25. [PubMed: 9927367]

Way TW, Sahiner B, Chan H-P, Hadjiiski L, Cascade PN, Chughtai A, Bogot N, Kazerooni E. Computer-aided diagnosis of pulmonary nodules on CT scans: improvement of classification performance with nodule surface features. Medical physics. 2009; 36:3086–98. [PubMed: 19673208]

Wisniewski, M., Zielinski, TP. MRMR-based feature selection for automatic asthma wheezes recognition; Signals and Electronic Systems (ICSES), 2012 International Conference on; 2012. p. 1-5.

Xu, R., Hirano, Y., Tachibana, R., Kido, S. Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach; International Conference on Medical Image Computing and Computer-Assisted Intervention; 2011. p. 183-90.

Xu Y, Sonka M, McLennan G, Guo J, Hoffman EA. MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. IEEE transactions on medical imaging. 2006; 25:464–75. [PubMed: 16608061]

Yao J, Dwyer A, Summers RM, Mollura DJ. Computer-aided diagnosis of pulmonary infections using texture analysis and support vector machine classification. Academic radiology. 2011; 18:306–14. [PubMed: 21295734]

Zheng B, Leader JK, Fuhrman CR, Sciurba FC, Gur D. Automated detection and classification of interstitial lung diseases from low-dose CT images. Medical Imaging 2004, International Society for Optics and Photonics. 2004:849–56.

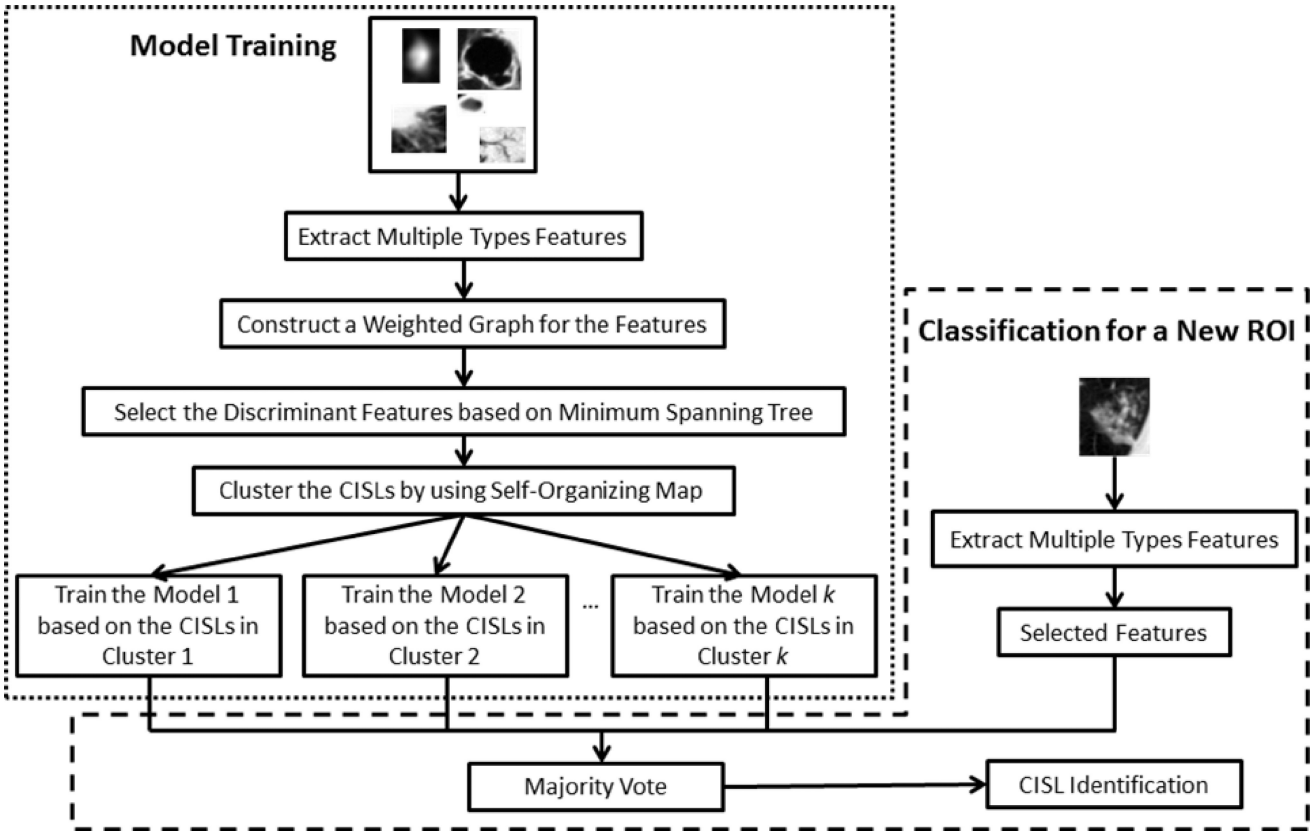


Figure 1.
The flowchart of our method for the classification of CISLs.

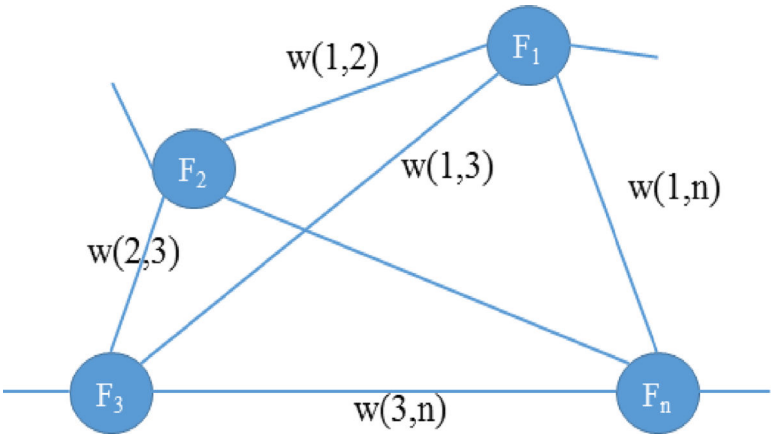


Figure 2.
The illustration of the weighted graph.

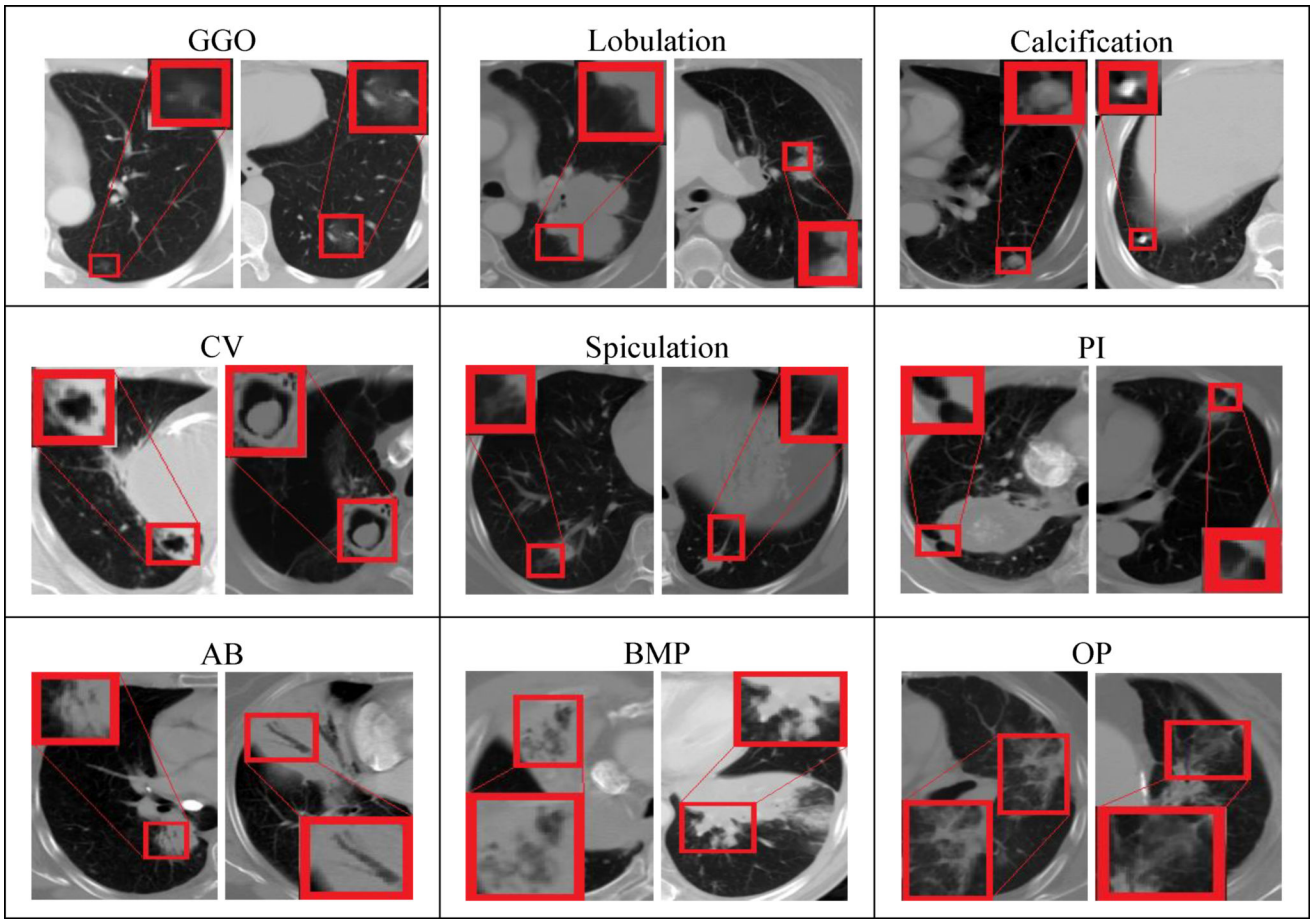


Figure 3.
The instances of nine CISL categories. The smaller rectangular boxes in lung CT images are magnified to show the details of the images

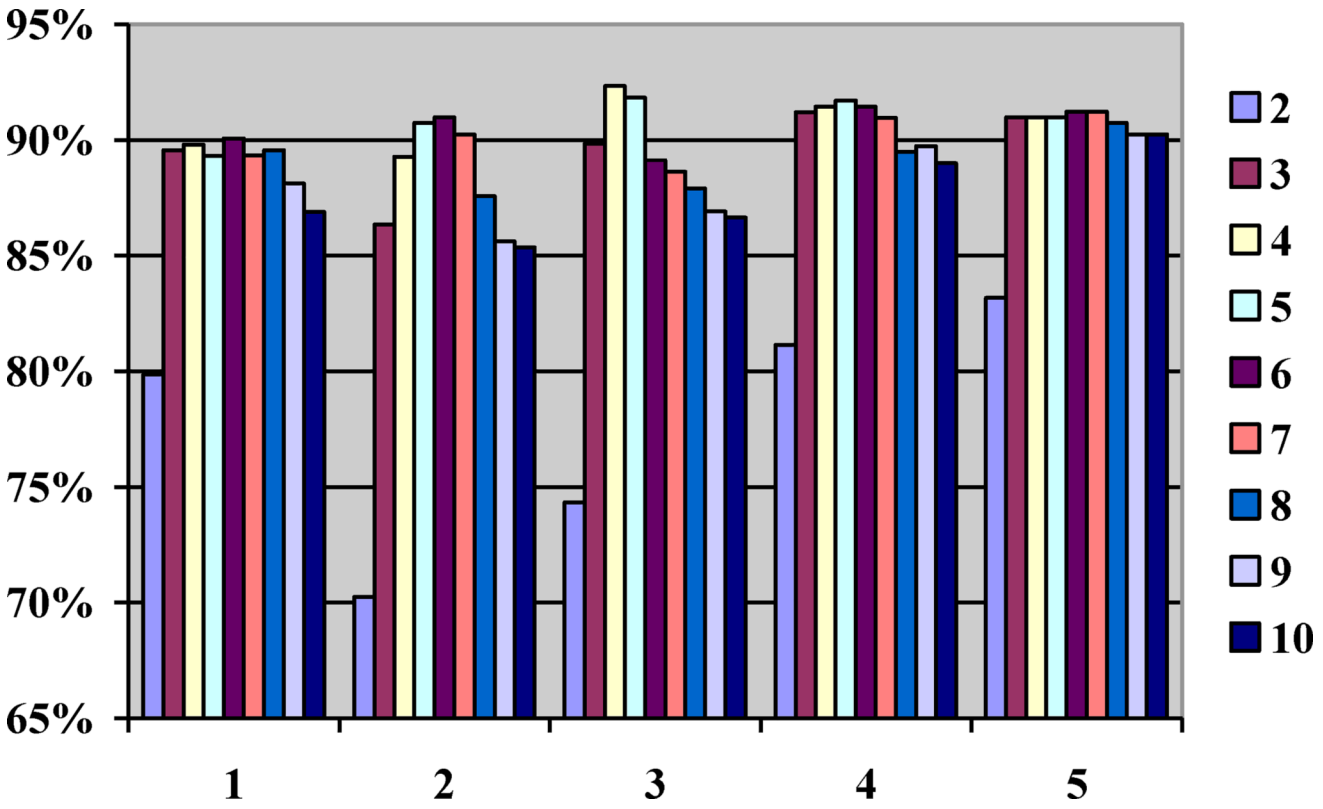


Figure 4.
The average classification accuracy in five rounds for different numbers of cluster.

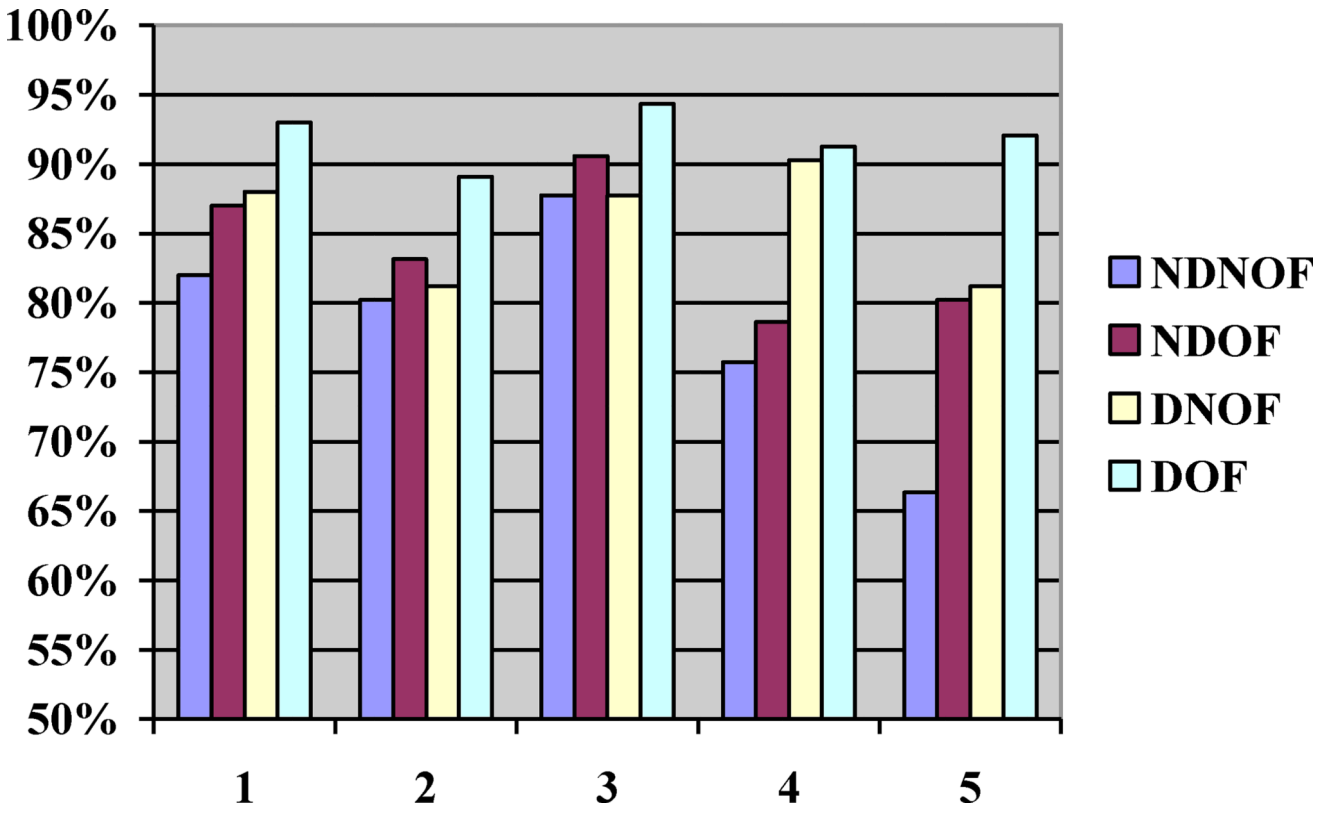


Figure 5.
The classification accuracy obtained by our DOF, and three compared methods which are NDNOF, NDOF and DNOF.

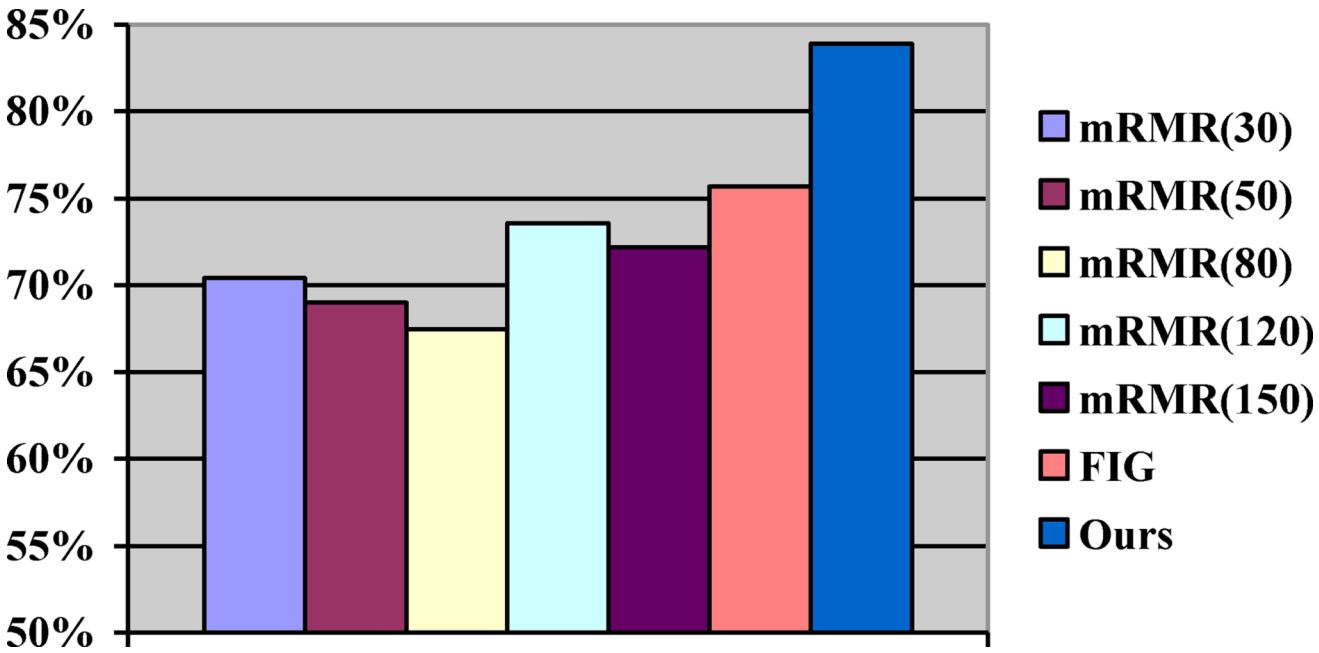


Figure 6.
The classification accuracy obtained by mRMR with different number of selected features, FIG (Liu *et al.*, 2015), and our feature selection method.

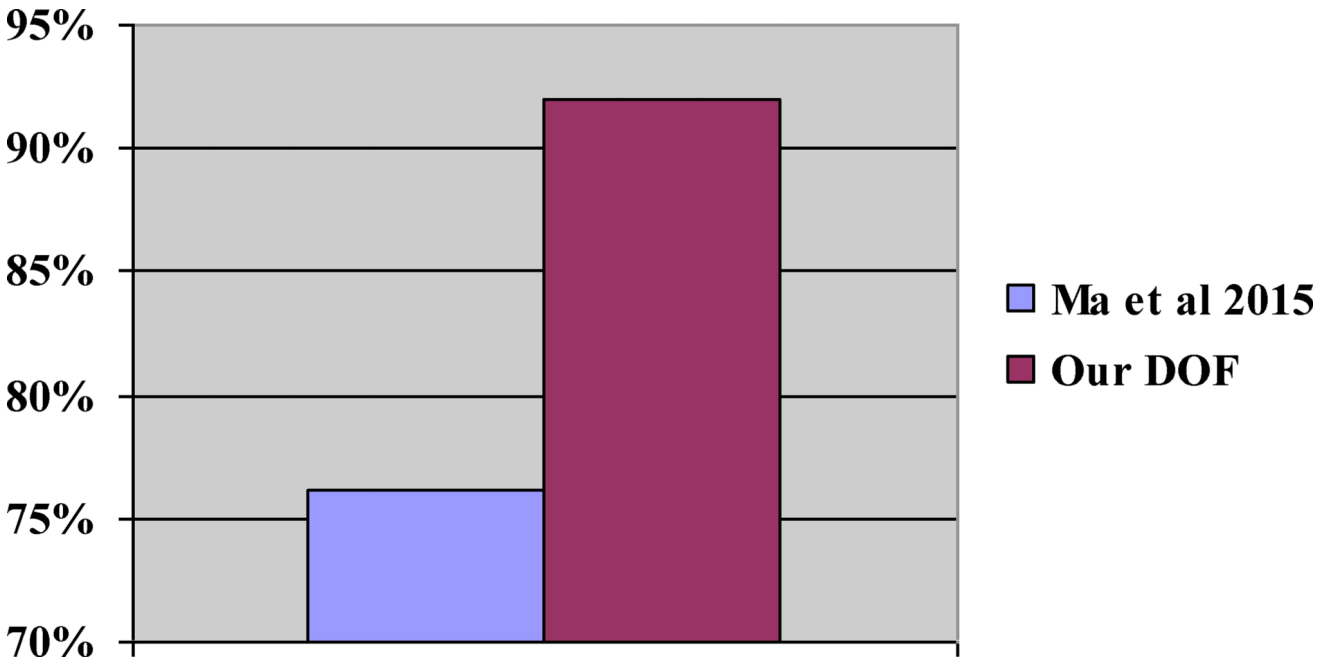


Figure 7.
The classification accuracy obtained by Ma et al (Ma *et al.*, 2015) and our DOF.

Table 1

The distribution and size range of CISLs (GGO, lobulation, CV, spiculation, PI, OP, calcification, AB, and BMP). D1–D5 are the different subsets for 5-fold cross validation experiments. NOP is the number of patients.

CISLs	D1	D2	D3	D4	D5	Size Range (mm × mm)	Total	NoP
GGO	9	9	9	9	9	6×6 ~ 26×20	45	25
Lobulation	9	8	8	8	8	8×7 ~ 23×19	41	21
Calcification	10	10	9	9	9	7×3 ~ 49×49	47	20
CV	30	30	29	29	29	16×15 ~ 37×40	147	75
Spiculation	6	6	6	6	5	4×5 ~ 53×56	29	18
PI	16	16	16	16	16	12×8 ~ 26×27	80	26
AB	5	5	5	4	4	9×17 ~ 32×40	23	22
BMP	17	16	16	16	16	10×10 ~ 30×35	81	29
OP	4	4	4	3	3	8×8 ~ 78×42	18	16
Total	106	104	102	100	99	4×5 ~ 78×42	511	252

Table 2

Cross validation results of five partitions of the training and test subjects for the CISLs classification. (SE and SP mean the sensitivity and specificity. GGO, lobulation, Calcification, CV, spiculation, PI, AB, BMP, and OP are the nine categories of CISLs)

	1		2		3		4		5		Average		Accuracy
	SE	SP	SE	SP	SE	SP	SE	SP	SE	SP	SE	SP	
GGO	100.0	100.0	100.0	100.0	100.0	98.97	100.0	97.87	88.89	100.0	97.78	99.37	
Lobulation	75.0	100.0	75.0	100.0	87.5	100.0	100.0	100.0	87.5	100.0	85.0	100.0	
Calcification	100.0	100.0	90.91	100.0	100.0	100.0	88.89	100.0	77.78	100.0	91.52	100.0	
CV	93.55	97.18	100	84.72	100.0	97.40	89.66	96.10	100.0	95.71	96.64	94.23	
Spiculation	80.0	100.0	40.0	100.0	80.0	99.02	88.89	100.0	60.0	100.0	69.78	99.80	91.96
PI	93.33	96.51	93.75	100.0	93.75	97.80	81.25	97.78	100.0	95.29	92.42	97.48	
AB	100.0	100.0	50.0	100.0	57.14	100.0	100.0	97.98	50.0	100.0	71.43	99.6	
BMP	93.75	98.82	87.5	100.0	100.0	100.0	93.75	100.0	100.0	98.82	95.0	99.53	
OP	100.0	98.97	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.79	

Table 3

Differences between our DOF and three other classification methods, including NDNOF, NDOF, and DNOF.

	NDNOF	NDOF	DNOF	DOF
Feature selection based on MST		√		√
Hierarchy learning based on SOM and AdaBoost			√	√

Table 4

The dimension of original feature and selected features for each type of feature and full features in 5-fold cross validation

Feature Types		Dimension of Original Features	Dimension of Selected Features				
			1	2	3	4	5
LBP(P,R)	LBP(5,1)	32	21	28	17	10	15
	LBP(5,2)	32	11	29	12	11	14
	LBP(4,1)	16	6	5	3	3	5
	LBP(4,2)	16	8	9	5	8	4
CVH		40	13	31	12	9	13
wavelet		26	17	22	15	15	17
B-HOG		18	16	15	14	13	15
FULL		180	92	139	78	69	83

Phys Med Biol. Author manuscript; available in PMC 2018 January 21.

Table 5

The dimension of original features and selected features by FIG (Liu *et al.*, 2015) and our method in each round.

Fold	Dimension of original features	Dimension of FIG's selected feature (Liu *et al.*, 2015)	Dimension of our selected features
1	180	92	92
2	180	132	139
3	180	145	78
4	180	146	69
5	180	141	83

Table 6

Average classification results in 5-fold cross validation for the CISLs classification on the noisy data.

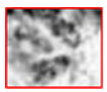
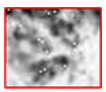
CISLs	Original ROI	Noisy ROI	Sensitivity (%)	Specificity (%)	Accuracy (%)
GGO			91.11	99.78	
Lobulation			88.06	99.37	
Calcification			93.56	99.78	
CV			98.64	92.60	
Spiculation			76.67	99.80	90.41
PI			93.75	97.47	
AB			65.00	99.80	
BMP			88.75	99.77	
OP			66.67	99.39	

Table 7

Average results in 5-fold cross validation for the classification of CISLs and normal tissue.

CISLs	Sensitivity (%)	Specificity (%)	Accuracy (%)
GGO	88.89	98.20	
Lobulation	80.49	99.64	
Calcification	95.74	99.46	
CV	99.32	96.44	
Spiculation	68.97	100.00	91.44
PI	96.25	98.65	
AB	73.91	99.83	
BMP	87.65	99.24	
OP	72.22	99.49	
Normal	97.65	99.03	

Table 8

Average results in 5-fold cross validation for the classification of CISLs under different patient age groups.

CISLs	Training (Age Group: 55 – 89)		Testing (Age Group: 30 – 54)		Sensitivity (%)	Specificity (%)	Accuracy (%)
	Age	Number	Age	Number			
GGO	55 – 75	26	39 – 54	19	89.47	100.00	
Lobulation	55 – 73	32	46 – 54	9	88.89	99.09	
Calcification	57 – 83	42	44 – 48	5	100.00	99.12	
CV	55 – 90	130	30 – 54	17	94.12	96.08	
Spiculation	56 – 74	18	33 – 53	11	90.91	98.15	90.68
PI	55 – 75	59	43 – 53	21	90.48	98.99	
AB	56 – 74	14	41 – 54	9	77.78	99.10	
BMP	55 – 89	56	30 – 53	25	92.00	100.00	
OP	55 – 74	16	48 – 54	2	100.00	99.14	