Recognizing Common CT Imaging Signs of Lung Diseases through a New Feature Selection Method based on Fisher Criterion and Genetic Optimization

Xiabi Liu, Ling Ma, Li Song, Yanfeng Zhao, Xinming Zhao*, Chunwu Zhou

***Abstract*—Common CT Imaging Signs of Lung Diseases (CISLs) are defined as the imaging signs that frequently appear in lung CT images from patients and play important roles in the diagnosis of lung diseases. This paper proposes a new feature selection method based on FIsher criterion and Genetic optimization, called FIG for short, to tackle the CISL recognition problem. In our FIG feature selection method, the Fisher criterion is applied to evaluate feature subsets, based on which a genetic optimization algorithm is developed to find out an optimal feature subset from the candidate features. We use the FIG method to select the features for the CISL recognition from various types of features, including bag-of-visual-words based on the Histogram of Oriented Gradients, the wavelet transform based features, the Local Binary Pattern and the CT Value Histogram. Then the selected features cooperate with each of five commonly used classifiers including Support Vector Machine, Bagging, Naïve Bayes, k-Nearest Neighbor and AdaBoost to classify the Regions of Interests (ROIs) in lung CT images into the CISL categories. In order to evaluate the proposed feature selection method and CISL recognition approach, we conducted the 5-fold cross validation experiments on a set of 511 ROIs captured from real lung CT images. For all the considered classifiers, our FIG method brought the better recognition performance than not only the full set of original features but also any single type of features. We further compared our FIG method with the feature selection method based on classification Accuracy Rate and Genetic optimization (ARG). The advantages on computation effectiveness and efficiency of FIG over ARG are shown through experiments.**

***Index Terms*—Medical image classification, lung CT images, lung lesion classification, feature selection, Common CT Imaging Signs of Lung Diseases (CISL)**

Manuscript received October 20, 2013; revised April 16, 2014; accepted May 23, 2014. This work was partially supported by National Natural Science Foundation of China (Grant no. 60973059, 81171407) and Program for New Century Excellent Talents in University of China (Grant no. NCET-10-0044).

Xiabi Liu and Ling Ma are with Beijing Lab of Intelligent Information Technology, School of Computer Science, Beijing Institute of Technology, Beijing 100081, China (e-mail: liuxiabi@bit.edu.cn; maling0@bit.edu.cn).

Li Song was with Beijing Institute of Technology, Beijing 100081, China. She is now with Qihoo 360 Technology Co. Ltd., No. 6 Jiuxianqiao, Chaoyang District, Beijing 100015, China (e-mail: 522545061@qq.com).

Yanfeng Zhao and Chunwu Zhou are with the Department of Imaging Diagnosis, Cancer Institute & Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China (e-mail: zyf24@sina.com, cjr.zhouchunwu@vip.163.com).

Xinming Zhao* (corresponding author) is with the Department of Imaging Diagnosis, Cancer Institute & Hospital, Chinese Academy of Medical Sciences, Beijing 100021, China Beijing 100021, China (e-mail: xinmingzh@sina.com).

I. Introduction

COMPUTED tomography (CT) scan can provide valuable information in the diagnosis of lung diseases. We have been witnessing the enormous increase in CT images of the human lungs, which should be read in time. This challenge plus the difficulty of recognizing subtle lesions even for radiologists promote the research interests in the Computer-Aided Diagnosis (CAD) and the Content-Based Medical Image Retrieval (CBMIR) based on thoracic CT scans. To support CAD and CBMIR applications, the computer should have the abilities of detecting, classifying and quantifying CT findings of lung lesions. The CT findings denote what radiologists see in CT scans for diagnosing diseases, which are also often called "CT features" or "CT manifestation". This paper focuses on the problem of automatic classification of CT findings of lung lesions in CT scans.

There are two main purposes of developing lung lesion classification methods in previous works. The first one is to distinguish abnormal tissues from normal ones, usually for abnormality detection such as nodule detection. The second one is to identify visual patterns of a specific lung disease. In this paper, we try to achieve a slightly different purpose: classifying different types of CT findings of lung lesions under the ignorance of underlying diseases. To our knowledge, this problem has not received much attention of researchers. A radiologist relies on the analysis to CT findings of lesions for making decisions about the diagnosis. But the correlation between CT findings and diseases is complicated. On one hand, a same category of CT findings could be observed in the images corresponding to different diseases. On the other hand, different categories of CT findings could appear in the CT images from the patients with a same disease. Therefore, it is useful for CAD and CBMIR applications to recognize the categories of CT findings in the Regions of Interests (ROIs) in lung CT images under the ignorance of diseases. For example, we can apply this technique to retrieve historical CT scans containing the interested categories of CT findings from large

repositories and the retrieved results are valuable for not only diagnostics but also medical research and teaching.

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, which are illustrated in Fig. 1 and explained in the following. Notice that this taxonomy is neither complete nor widely accepted at present, but these CT signs are really often encountered and widely used in the diagnosis of lung diseases.

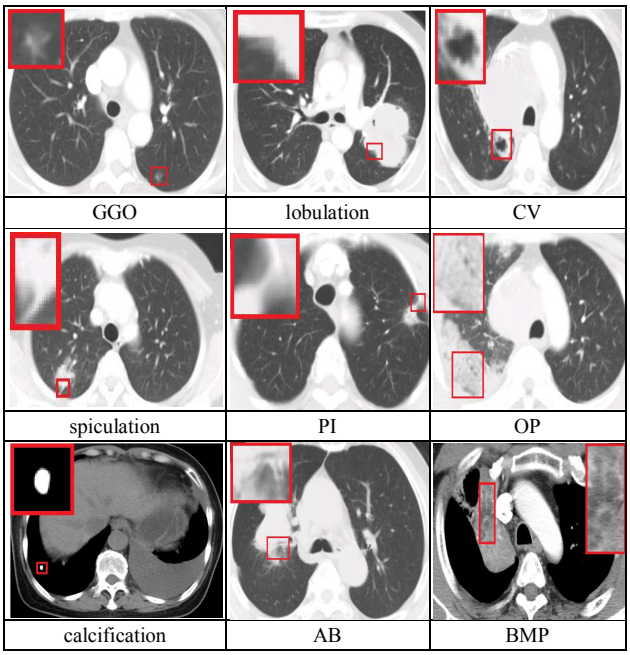

Fig. 1. The instances of nine categories of CISLs, which are indicated by the smaller rectangles in lung CT images and magnified to display clearer in the bigger rectangles overlapping on the images.

- Grand Grass Opacity (GGO). GGO can be characterized by areas of hazy increased attenuation of the lung with preservation of bronchial and vascular margins [1]. It is associated with the adenocacinoma of lung and bronchioloalveolar carcinoma [2], [3].
- Lobulation. Lobulation is dependent on the ingrowth of connective tissue septae containing fibroblasts derived from perithymic mesenchyme [4], which indicates a malignant lesion [5].
- Cavity & Vacuolous (CV). Both cavity and vacuolous are hollow spaces within the tissue. We can regard vacuolous as little cavity. Vacuolous is associated with the adenocarcinoma and bronchioloalveolar carcinoma, while cavity is associated with the tumors larger than 3cm [6], [7].
- Spiculation. Spiculation is a stellate distortion caused by the intrusion of cancer into surrounding tissue [8].
- Pleural Indentation (PI): PI is caused by the contraction of scar affected by the tumor, which is associated with most peripheral adenocarcinomas containing a central or subpleural anthracotic and fibrotic focus [9].
- Obstructive Pneumonia (OP): OP can be characterized by the following appearances: (1) alveolar septum has not been completely destroyed by tumor, (2) alveolar wall is thin, and (3) alveolus contains gas. This feature is associated with the alveolar carcinoma, lymphoma, pulmonary infarction and pulmonary edema [10].
- Calcification: Calcification is the deposition of insoluble salts of calcium and magnesium. Its morphology and distribution are important for discriminating between benign lung diseases and malignant ones. The coarse, dense, and popcorn-like calcification indicates benign lesions, while the calcification located in the center of lesions, spotted, and appearing irregularly suggests malign lesions [11].
- Air Bronchogram (AB): AB is an important radiologic sign of airspace consolidation, in which the normally invisible bronchial air column becomes visible. It usually accompanies with cavity. This feature is associated with the lung cancer, pulmonary pneumonia and lymphoma [12].
- Bronchial Mucus Plugs (BMP): BMP can be represented by focal opacities. Its density varies from liquefied density to higher than 100 Hounsfield Units (HU). It is associated with the allergic bronchopulmonary aspergillosis [13].

In a previous preliminary work [14] we began to investigate the problem of recognizing the CISLs contained in the ROIs in lung CT images, where four CISL categories including GGO, cavity, spiculation and calcification were considered. In this paper, we expand the number of CISL categories to nine and propose a new feature selection method based on FIsher criterion and Genetic optimization for tackling the problem. The proposed feature selection method is called FIG for short. It cooperates with each of five commonly used classifiers, including Support Vector Machine (SVM), Bagging (Bag), Naïve Bayes (NB), *k*-Nearest Neighbor (*k*-NN) and AdaBoost (Ada), to fulfill the CISL recognition task. We conducted the experiments to demonstrate the effectiveness of the proposed FIG feature selection method as well as CISL recognition approach.

The rest of this paper is organized as follows. Section II reviews related works on lung CT image classification and feature selection in medical imaging. Section III presents our FIG method for feature selection. Section IV describes our CISL recognition approach. The experiments are discussed in Section V. We conclude in Section VI.

II. Related Works

We review the previous works on the image classification and the feature selection in the medical image community. For the former problem, we restrict our discussions on lung CT images. For the latter problem, since there is not much related work specific to lung CT images, we expand our view to include other types of medical images.

A. Lung CT image classification

As described in Section Ⅰ, the works on lung CT image classification can be divided into three categories according to their purposes: (1) the discrimination between normal and abnormal lung tissues, (2) the identification among visual

patterns of specific lung diseases, and (3) the classification of different types of lung lesions. In the first category of works, many methods are presented for nodule detection and GGO detection. They are usually adopted in the final stage of detection systems to decide whether a candidate is true or false. In the second category of works, the explored lung diseases include Diffuse Parenchyma Lung Disease (DPLD), Chronic Obstructive Pulmonary Disease (COPD) and Interstitial Lung Disease (ILD). Although the purposes of three categories of works are different, the frameworks of classification systems are similar in principle, which are usually composed of two components: feature extractor and classifier.

For the classifier, the researchers have tried two strategies: single classifier and classifier fusion. The main single classifiers have been explored, such as rule-based [15], Linear Discriminant Analysis (LDA) [16], Artificial Neural Networks (ANN) [17]-[20], Bayesian classifier [21]-[24], *k*-NN [25], [26], and SVM [16]. Sluimer et al. [27] evaluated linear discriminant classifier, quadratic discriminant classifier, SVM, and *k*-NN. The *k*-NN classifier performed best according to their experimental results. Nuzhnaya et al. [28] compared *k*-NN, SVM and ANN. They showed that the *k*-NN achieved the best average performance, and the SVM performed fairly well on some of individual datasets. Depeursinge et al. [29] compared five common classifiers, including NB, *k*-NN, decision tree, ANN and SVM, in their abilities to categorize six lung tissue patterns in high-resolution computed tomography images of patients affected with ILD. The results revealed that the SVM constitutes the best tradeoff between the error rate on the training set and the generalization. In the classifier fusion strategy, we have witnessed the applications of various combinations, such as the rule-based classifier and LDA [30], the rule-based and ANN [31], *k*-NN and ANN [32], and multiple SVMs [33].

For the feature extractor, there are three main types of features for lung CT image classification. The first one is the geometric features, such as geometric shape features [15], radius features and profile features [16], the boundary and circularity information [22], major and minor axes and their ratio [27], the eccentricity of a fitted ellipse [27]. The second type of features are textural features, such as run-length features [20, 23], Local Binary Patterns (LBP) [21], co-occurrence features [23]-[24], [27], [32], multiple texton-based features [33], vector quantization generating texture descriptor [28], Histogram of Oriented Gradients (HOG) features [21], and wavelets [29]. The third type of features is intensity based ones. We have gradient magnitude features [16], edge-gradient features [31], CT value histogram (CVH) [21] and intensity distributions [27]. Among the three types of features, the geometric features are mainly used on the lesions having the fixed geometrical properties. The other two types of features, especially textural features, are used more often.

B. Feature Selection for Medical Image Classification

In order to achieve good classification results, we usually use several types of features at the same time. Since the different types of features may contain complementary information, it could brings better classification performance through selecting discriminative features from various feature spaces. This idea has attracted a lot of attention in the related fields, including the medical image classification [34]. According to Guyon and Elisseff [35], the feature selection techniques can be organized into mainly three categories: filter, wrapper and embedded methods. We follow their taxonomy to review the feature selection methods for medical image classification.

Filter techniques rank the features by the intrinsic properties of the data, independent of the choice of the classifier. The features are selected based on their ranking. Zuluaga et al. [36] used three different strategies, including F-score, Random Forest (RF) and SVM-Recursive Feature Elimination (SVM-RFE), to rank the features and take the top 10 features for vascular anomaly detection. The classification results based on the features selected by using F-score and RF are pretty close, while the ones from SVM-RFE present higher sensitivity and specificity. Nithya and Santhi [37] proposed a feature filter that is called maximum difference feature selection. They used the dissimilarity between the features in normal and abnormal patterns as the criterion function and then selected the top five features as the most discriminative ones. Silva et al. [38] proposed two filter-based feature selection algorithms for medical image classification: the silhouette-based greedy search and the silhouette-based Genetic Algorithm (GA) search, in which the simplified silhouette statistic is calculated and used to evaluate the features. Huang et al. [39] employed the information entropy and the sequential backward selection algorithm to determine the importance degrees of features for breast cancer diagnosis.

Wrapper techniques take the optimal subset of features as the one that lead to the best performance of the classifier, but the learning of the classifier is invisible to the feature selection. The crucial factors in wrapper techniques are the search algorithm and the criterion for evaluating feature subsets. Firpi and Vogelstein [40] used the misclassification error and the particle swarm optimization search algorithm to select features for cognitive state detection in a brain-computer interface system. Dy et al. [41] used the trace ratio of the Expectation-Maximization (EM) clustering result in the feature space and the sequential forward selection search algorithm to select feature subsets. The resultant algorithm was applied to the CBMIR of lung CT images. Park et al. [42] applied the *k*-NN classifier to detect the pulmonary embolisms depicted on CT images. They preselected an optimal feature set by using the GA and the evaluation criterion of the normalized area under a Free-response Receiver Operating Characteristics (FROC) of the classifier. Zheng et al. [43] used the sensitivity and specificity of the classifier as the evaluation criterion and the GA as the search algorithm to select the features for colonic polyp detection. Hupse and Karssemeijer [44] used the mean sensitivity of the classification system in a predefined range of FROC and the sequential floating forward selection search algorithm to select features for detecting malignant masses in mammograms. Wu et al. [45] used the GA search algorithm and

the criterion involving the classification rate and the number of selected features to select feature subsets. The method is firstly performed to select features from each feature space, respectively, and then performed again on the resultant features from all the considered feature spaces to select the final feature subset for ultrasonic liver tissue characterization. Zhu et al. [46] employed the GA search algorithm and the misclassification rate of the classifier to select multiple groups of feature subsets with different numbers of features and tested them for discriminating benign solitary pulmonary nodules from malignant ones.

Embedded methods integrate the feature selection into the process of classifier training. Ozcift [47] firstly made use of a linear SVM to rank the features. Then the feature vector was determined for each of other classifiers by adding the features one by one and in order until the accuracy of the classifier discontinues increasing. Finally, the rotation forest ensemble classifier was established for improving the diagnosis of Parkinson disease. Maggio et al. [48] evaluated the performance of the feature subset at increasing set size. For different cardinalities, the subset which maximizes the min-redundant max-relevance measure was selected and the performances of the Fisher linear discriminant classifier trained on this subset were computed. The best cardinality and consequently the best subset were chosen according to the minimum misclassification error.

III. Feature Selection Method based on Fisher Criterion and Genetic Optimization

In essence, the feature selection problem is to find out the best feature subset in the power set of features. Therefore, it involves two sub-problems: (1) how to evaluate feature subset and (2) how to implement search. For the search algorithm, the GA is a popular and good choice. But most of GA based feature selection algorithms measure the quality of feature subset by its Classification Accuracy Rate (CAR). In the following descriptions, a feature subset is called an individual, and the quality of it is called its fitness, according to GA's terminology. Using the CAR as the individual fitness has two disadvantages. First, it makes the feature selection method depend on the underlying classifier. The optimal feature subset generated for one classifier may not be necessarily appropriate to another one. Second, for getting the individual fitness, the classifier must be re-trained with the corresponding feature subset and then used to perform classification on the data set to obtain the CAR. This procedure of fitness evaluation is obviously time-consuming and leads to the unsatisfactory efficiency of GA search. In order to solve the two shortcomings above, our FIG method introduces the Fisher discriminative criterion [49] to measure the individual fitness in the GA based optimum search. Although both the Fisher criterion and the GA algorithm have been explored in previous works on feature selection, respectively, this strategy of ours for combining them is the first one to our knowledge.

Furthermore, in most of GA based feature selection methods, the feature selection result is represented by a binary string. Each bit in the string corresponds to a feature, where the value 1 indicates that the feature is selected and 0 indicates that the feature is discarded. Different from these methods, we assign a weight in [0, 1] to each feature and evolve the weights. It is more reasonable and more accurate for measuring the importance degree of a feature than the hard value of 0 or 1. After the weight evolution is completed, the feature whose weight exceeds a threshold is chosen as a member of the optimal feature subset. The threshold is determined adaptively according to training data, as explained in the last paragraph of the first sub-section below.

A. Fitness Function based on Fisher Criterion

A reasonable objective of feature selection for pattern classification is to maximize classification accuracy. The Fisher criterion measures the distance among all the classes and the divergence within the members of each class. Thus it reflects the classification accuracy under the absence of classifiers.

Let d be the number of considered feature elements, $\boldsymbol{w} = (w_1, w_2, \cdots, w_d)$ be the feature-weight vector, where $w_i\big|_{i=1}^{d}$ reflects the importance of the i-th feature. According to GA's terminology, a $\boldsymbol{w}$ is an individual required to be evaluated in this paper. We complete the evaluation task based on Fisher criterion. Accordingly, the fitness of individuals are computed as follows.

Let $\boldsymbol{X}^{i,j} = \left(x_1^{i,j}, x_2^{i,j}, \cdots, x_d^{i,j}\right)$ be the full feature vector of the j-th example of the i-th class, n_i be the number of examples of the i-th class, C be the number of classes. Firstly, we calculate the mean of feature vectors belonging to the i-th class as

$$\boldsymbol{m}^i = \frac{1}{n_i}\sum_{j=1}^{n_i}\boldsymbol{X}^{i,j}, \tag{1}$$

and the mean of feature vectors of all the training examples as

$$\boldsymbol{m} = \frac{\sum_{i=1}^{C}\sum_{j=1}^{n_i}\boldsymbol{X}^{i,j}}{\sum_{i=1}^{C} n_i}. \tag{2}$$

Suppose the resultant $\boldsymbol{m}^i = \left\{m_1^i, m_2^i, \cdots, m_d^i\right\}$, and the resultant $\boldsymbol{m} = \{m_1, m_2, \cdots, m_d\}$. Secondly, we get the average weighted distance between all the training examples and the corresponding class mean as

$$S_W = \sum_{i=1}^{C}\frac{1}{n_i}\sum_{j=1}^{n_i}\sum_{k=1}^{d} w_k\left(x_k^{i,j} - m_k^i\right)^2, \tag{3}$$

and the weighted distance between classes as

$$S_B = \sum_{i=1}^{C}\sum_{k=1}^{d} w_k\left(m_k^i - m_k\right)^2. \tag{4}$$

Finally, the Fisher criterion can be formulated as maximizing S_B and minimizing S_W simultaneously. Thus the fitness function for evaluating $\boldsymbol{w}$ is designed as

$$f(\boldsymbol{w}) = \frac{S_W}{S_B}. \tag{5}$$

The optimal $\boldsymbol{w}$ is taken as the one that minimizes (5). Then

we select the features whose weights in the optimal $\boldsymbol{w}$ are larger than a threshold. Here we use k-NN classifier examination to obtain a data-driven threshold. Actually, the nine thresholds from 0.1 to 0.9 are used to select the features, respectively. Each resultant subset of features is employed in a k-NN classifier to perform the classification. The feature subset leading to the best CAR is taken as the final selection result and the corresponding threshold as the optimal one. This final feature selection result is unchanged in the subsequent classification stage, no matter what classifier is used.

B. Genetic Optimization for Feature Selection

Under GA optimization framework, the main components of our FIG algorithm include population initialization, fitness evaluation, selection, crossover, mutation, and termination judgment. The corresponding flowchart of the algorithm is illustrated in Fig. 2, where "Fisher Fitness Evaluation" means "the fitness evaluation based on Fisher criterion". The fitness evaluation method has been presented in the last sub-section. The details of other components are given as follows.

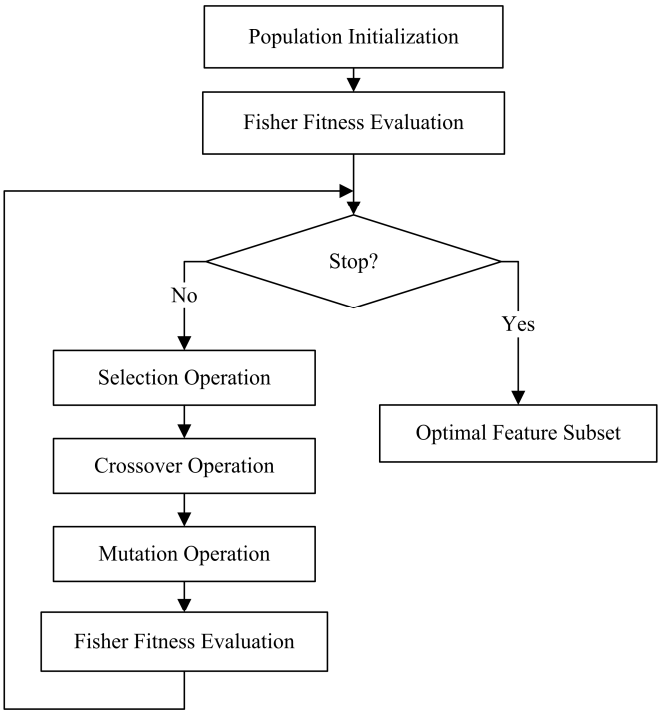

Fig. 2. The flowchart of the proposed FIG algorithm

1) Population initialization

In the GA algorithms, all the individuals in each generation construct the population. Each individual is encoded as a binary string, which is thought to be the individual's chromosome. As described above, an individual in the FIG algorithm is a feature-weight vector. Suppose the weights are required to be accurate to p decimal places, then the closed interval [0, 1] needs to be divided into 10^p equal parts. If

$$2^{q-1} < 10^p < 2^q , \tag{6}$$

then the length of the binary string for each weight should be q-bit and the chromosome will be encoded as $\boldsymbol{C} = \{\underbrace{c_1 \cdots c_q}_{w_1}, \underbrace{c_{q+1} \cdots c_{2q}}_{w_2}, \cdots, \underbrace{c_{(d-1)q+1}, \cdots, c_{dq}}_{w_d}\}$.

2) Selection operator

The selection operator is used to select the parent individuals which will participate in producing offsprings for the next generation. Here the commonly used roulette wheel selection technique [50] is used. Actually, the probability of selecting an individual is calculated as $p(\boldsymbol{C}_i) = f(\boldsymbol{C}_i) \Big/ \sum_{k=1}^{M} f(\boldsymbol{C}_k)$, where $\boldsymbol{C}_i$ is the chromosome of the i-th individual in the population, $f(\boldsymbol{C}_i)$ is the fitness value corresponding to $\boldsymbol{C}_i$, and M is the number of individuals in the population.

3) Crossover operator

The crossover operator is used to create new individuals by recombining the genes of the chromosomes of the selected two parents. Considering that there are different types of features for selection and at least one feature in each type should be selected, the multi-point crossover is performed. Actually, we divide the chromosome of an individual into several parts, each of which is corresponding with a type of features. Then we perform the single-point crossover in each part of the chromosome, respectively.

The probability of crossover affects the search ability and the convergence speed of GA. In this work, we follow Yang et al. [51] to adopt the adaptive probability of crossover. Let it is denoted as P_c. Initially, a large P_c is used to strengthen the search ability. As the evolution goes on, P_c is decreased to improve the convergence speed gradually. Formally, let P_{c_0} be the initial crossover probability, g be the number of generation; $\boldsymbol{C}_i$ and $\boldsymbol{C}_j$ be the chromosomes of parent individuals, then P_c is adjusted by

$$p_c = \begin{cases} \dfrac{p_{c_0}}{\log_{10}(g+1)}, & f_{\max} \geq \overline{f} \\ p_{c_0}, & f_{\max} < \overline{f} \end{cases}, \tag{7}$$

where

$$f_{\max} = \max(f(\boldsymbol{C}_i), f(\boldsymbol{C}_j)) \tag{8}$$

and

$$\overline{f} = \frac{1}{M} \sum_{i=1}^{M} f(\boldsymbol{C}_i) . \tag{9}$$

Notice that in (7), we use $\log_{10}(g+1)$ instead of $\log_2(g+1)$ which is used in [51]. The reason is that this change makes the crossover probability drop more slowly and thus leads to better results in our experiments.

4) Mutation operator

The mutation occurs right after the crossover is completed. It

is performed by inversing one bit in each part of an individual's chromosome to create a child. Similar to the processing in the crossover, each part of the chromosome is corresponding with a type of features, and the mutation probability is also adjusted adaptively. The adaptive equation is

$$p_m = \begin{cases} \dfrac{p_{m_0}}{\log_{10}(g+1)}, & f \geq \overline{f} \\ p_{m_0}, & f < \overline{f} \end{cases} \quad (10)$$

where P_{m_0} is the initial mutation probability, f is the fitness of the individual mutated; g and $\overline{f}$ have the same meaning as those in (7).

5) Termination judgment

The algorithm will be terminated when it converges or the predefined maximum number of generations is reached. The condition that we use to judge whether the algorithm converge is: the difference between the maximum fitness values of adjacent two generations does not exceed an infinitesimal (denoted as ε) after m generations.

IV. CISL Recognizer

Our approach of recognizing CISLs in ROIs in lung CT images consists of two components: feature extraction and ROI classification. Firstly, the features are extracted from each ROI and some of them are selected by using the proposed FIG method to form a feature vector for representing the ROI. Then the ROI is classified into the corresponding CISL category by using some classifiers.

A. Feature Extraction

We consider four types of ROI features, including the Bag-of-visual-words based on the HOG (B-HOG), the wavelet features, the LBP and the CVH. We have 18-D B-HOG features, 26-D wavelet features, 96-D LBP features and 40-D CVH features. Total 180 features are extracted. The details of each type of features are given as follows.

1) B-HOG

The HOG feature is a texture descriptor describing the distribution of image gradients in different orientations. Following the HOG feature extraction scheme of Dalal et al. [52], we divide a ROI into smaller rectangular blocks of 8×8 pixels and further divide each block into 4 cells of 4×4 pixels. An orientation histogram which contains 9 bins covering a gradient orientation range of $0^\circ - 180^\circ$ is computed for each cell. Then a block is represented by the linking of the orientation histograms of cells in it. This means a 36-D HOG feature vector is extracted for each block.

The commonly used image representation based on HOG features is to join the feature vectors of all the blocks in the image in sequence. This kind of HOG based image representation strategy requires that all the images have the same size, or else the dimensions of resultant feature vectors will be diverse for different images. But the size of ROIs in lung CT images varies with different patients and different pathological lesions. So this widely used strategy is not applicable in this work. To solve this problem, we adopt the bag-of-visual-words [53] on HOG features as the ROI representation. However, different from the original bag-of-visual-words method, we use a clustering algorithm based on Gaussian Mixture Modeling (GMM) [54], instead of the *k*-means algorithm, to generate more accurate visual words. In this paper, total 18 visual words are obtained.

The 36-D HOG feature vector of each block is mapped to the visual word corresponding to the highest likelihood for it. Then the number of HOG feature vectors assigned to each visual word is accumulated and normalized by the number of all the HOG feature vectors to form a 18-D histogram representation of the ROI.

2) Wavelet features

Wavelets are important and commonly used feature descriptors for texture analysis, due to their effectiveness in capturing localized spatial and frequency information and multi-resolution characteristics [55]. In this paper, the ROIs are decomposed to 4 levels by using 2D symlets wavelet because the symlets wavelet has better symmetry than Daubechies wavelet and more suitable for image processing [56]. Then the horizontal, vertical and diagonal detail coefficients are extracted from the wavelet decomposition structure. Finally, we get the wavelet features by calculating the mean and variance of these wavelet coefficients.

3) LBP

The LBP feature is a compact texture descriptor in which each comparison result between a center pixel and one of its surrounding neighbors is encoded as a bit [57]. In this way we can get an integer for each pixel. Then the frequency of each integer is figured out on the ROI level to obtain the corresponding feature vector.

The neighborhood in the LBP operator can be defined very flexibly by using circular neighborhoods and the bilateral interpolation of pixel values. These kinds of neighborhoods can be denoted by (P, R), which means we evenly sample P neighbors on the circle of radius R around the center pixel. The corresponding LBP features will be denoted as LBP(P, R) in the following descriptions. We consider multiple P and R to get multi-scale LBP features.

4) CVH features

CVH means the histogram of CT values. In lung CT images, the CT values of pixels are expressed in HU. We compute the histogram of CT values over each ROI. The number of bins in the histogram is determined by experiments. In fact, we obtain various CVHs with different numbers of bins. Each CVH is tested for classification under *k*-NN classifier and the corresponding CAR is calculated. Then the number of bins, which brings the highest CAR, is adopted. This choice will keep unchanged for all the experiments.

B. ROI Classification

Five classifiers, including SVM, Bag, NB, *k*-NN, and Ada, are respectively tested for cooperating with the selected features to classify ROIs into CISL categories. These classifiers are implemented by using the corresponding functions in WEKA [58], a machine learning library in java. The name of these functions are: 1) "SMO" (SVM), 2) "Bagging" (Bag), 3) "NaïveByes" (NB), 4) "IBk" (*k*-NN, $k=1$ and Euclidean distance are adopted), 5) "AdaBoostM1" (Ada, using REPTree as weak learner).

Each function provides two execution modes: training and testing. We call the function with the training mode on the training data to obtain the corresponding classifier. Then it is evaluated on the test data by calling the function with the testing mode.

V. Experiments

A. Experimental Setup

1) Dataset

The instances of nine categories of CISLs were collected from the Cancer Institute and Hospital at Chinese Academy of Medical Sciences. The lung CT images were acquired by CT scanners of GE LightSpeed VCT 64 and Toshiba Aquilion 64 and saved in DICOM 3.0 format. The 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 rectangular ROIs wrapping CISLs in these lung CT images are manually labeled and annotated by qualified radiologists to produce a gold standard. The resultant numbers of ROIs are 511. The set of all these available instances are split into 5 disjoint subsets nearly evenly, in order that 5-fold cross validation experiments can be conducted. Furthermore, the data in different subsets are guaranteed to come from different patients, so that the bias in measuring classification performance is avoided. Table Ⅰ lists the numbers of ROI examples in 5 data subsets and the numbers of patients for each CISL category, where S1-S5 denote the first to the fifth subsets, respectively, and NoP means "the number of patients".

TABLE Ⅰ
THE DISTRIBUTION OF ROIS USED IN 5-FOLD CROSS-VALIDATION EXPERIMENTS

CISL	S1	S2	S3	S4	S5	Total	NoP
GGO	9	9	9	9	9	45	25
lobulation	9	8	8	8	8	41	21
calcification	10	10	9	9	9	47	20
CV	30	30	29	29	29	147	75
spiculation	6	6	6	6	5	29	18
PI	16	16	16	16	16	80	26
AB	5	5	5	4	4	23	22
BMP	17	16	16	16	16	81	29
OP	4	4	4	3	3	18	16
Total	106	104	102	100	99	511	252

2) Evaluation criterion

The performance of CISL recognition is evaluated by the sensitivity (SE) and specificity (SP), CAR, and confusion matrix (CM).

- The SE and SP are widely used in the medical image classification community. They are essentially two measurements of performance of binary classifiers. In this paper we use them to reflect the ability of our CISL recognizer for discriminating one CISL category from any 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 meanings of 'true negative' and 'false positive' are defined similarly. Let TP, TN, FP, FN be the number of true positives, true negatives, false positives and false negatives for a CISL category, respectively. Then the SE and SP of the classifier for this category are measured as *TP/(TP+FN)* and *TN /(TN+FP)* , respectively.
- Our CISL recognition problem is actually a multi-class classification problem. So we use the CAR to give an overall measurement of performance of our CISL recognizer. It is the ratio of the number of correctly classified examples to the number of all examples.
- The CM is used to summarize the tendency for our CISL recognizer to classify a pattern into a correct class or any of other wrong classes.

3) Parameter Setting

Two groups of parameters of our approach were set up through experiments. The first group of parameters are those in the proposed FIG feature selection method. Table Ⅱ lists the values of this group of parameters, which correspond to the experimental results reported in the following. The reasons behind these values are explained as follows. (1) The population size should be designed on the basis of the dimension of original feature vector. The small population size will weaken the search ability of our FIG; while the large population size will slow down the speed of the algorithm. Since the dimension of original feature vector in this research is 180, we assign a moderate value, i.e. 60, to population size. (2) The initial probabilities of crossover and mutation, i.e., P_{c_0} and P_{m_0}, are set by following Yang et al. [51]. (3) ε and m for terminating the algorithm are set by observing the change of maximum fitness values of adjacent generations in the experiments. We found that the maximum fitness values will keep stable after the converge condition configured with these two values is satisfied.

TABLE Ⅱ
THE PARAMETERS OF OUR FIG METHOD USED IN THE EXPERIMENTS

Parameters	Population size	P_{c_0}	P_{m_0}	ε	m
Values	60	0.8	0.8	0.001	50

The second group of parameters are those in LBP and CVH feature extraction. The ranges of *P* and *R* for calculating LBP feature are set to be $\{4,5\}$ and $\{1,2\}$, respectively. They are enough to adapt to the sizes of ROIs encountered in the experiments. As for the number of bins in CVH computation,

we tested 5 numbers from 20 to 60. For each tested number, the corresponding CVH features were extracted and cooperated with *k*-NN to perform the CISL recognition. The resultant CARs are listed in Table Ⅲ, from which we can see that the best number is 40 and it was used in all the following experiments.

TABLE Ⅲ
THE NUMBER OF BINS IN CVH FEATURE EXTRACTION AND THE RESULTANT CAR

Number of Bins	CAR（%）
20	44.1
30	47.1
40	49.1
50	45.1
60	45.1

B. *Experimental Results*

1) Results of feature selection and CISL recognition

We conducted feature selection and ROI classification experiments. Table Ⅳ shows the numbers of features selected from original 180 features and the determined weight threshold for selecting features in each round of 5-fold cross-validation experiments.

TABLE Ⅳ
THE NUMBERS OF SELECTED FEATURES AND WEIGHT THRESHOLD FOR SELECTING FEATURES IN EACH ROUND OF TESTS

Test Round	Num of Selected Features	Weight Threshold
1	92	0.5
2	132	0.3
3	145	0.2
4	146	0.2
5	141	0.2

Table Ⅴ lists the average CISL recognition performance over nine categories of CISLs by using selected features and each of five classifiers. In each round of tests, each classifier with selected feature vector is trained and tested by using the corresponding routines in WEK library. According to the CAR, the best CISL recognition performance came from the SVM classifier. Thus we further show the SE and SP from the combination of selected features and SVM for each category of CISLs in Table Ⅵ. Notice that in this situation, the recognition performance is evaluated for each category of CISLs, separately and respectively. For an input pattern, we need to determine whether this pattern belongs to the specific category of CISLs or not. Thus this is a binary classification problem and the values of resultant SE and CAR are equal with each other.

We carefully analyzed the reasons behind wrong classification results from the recognizer established by combining selected features and the SVM. The reasons are illustrated in Fig. 3 and explained as follows, where the lesions are indicated by the smaller rectangles in lung CT images and magnified to display clearer in the bigger rectangles overlapping on the images. (1) Some CISLs are noised by blood vessels surrounding them, as shown in Fig. 3a. (2) Some CISLs are so small and hazy that it is difficult to recognize them even by radiologists, as shown in Fig. 3b. (3) The visual appearance of some CISLs are very similar with each other, which can be seen by comparing Fig. 3c and d as well as Fig. 3e and f. Especially for the CISL "AB", it is very visually similar with the CISL "CV". Furthermore, the training examples of "AB' are far less than those of 'CV'. Consequently, most of 'AB' instances were classified into 'CV', which leads to unsatisfied SE of 0 for 'AB' category.

TABLE Ⅴ
THE AVERAGE CISL RECOGNITION PERFORMANCE OF OUR FIG FEATURE SELECTION METHOD

Classifiers	Classification Results (%)		
	SE	SP	CAR
SVM	70.2	97.2	**80.26**
Bag	71.8	96.9	77.88
NB	79.4	97.1	77.84
k-NN	68.4	96.4	73.58
Ada	68.1	96.7	75.70

TABLE Ⅵ
THE SE AND SP FROM SVM AND SELECTED FEATURES FOR EACH CATEGORY OF CISLS

CISLs	Classification Results of SVM (%)	
	SE	SP
GGO	100	99.4
lobulation	80	99.6
calcification	89.3	99.6
CV	89.3	86.5
spiculation	18.2	99.8
PI	79.8	91.7
AB	0	100
BMP	95.0	98.9
OP	80.0	99.4
Average	70.2	97.2

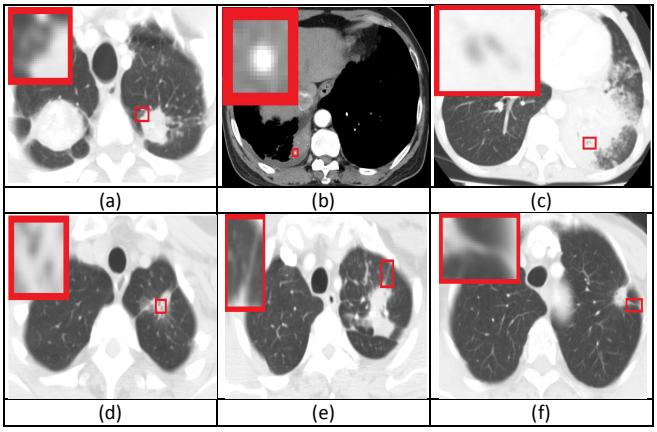

Fig. 3. The examples of wrong classified CISLs: (a) lobulation noised by blood vessel; (b) calcification identified difficultly; (c-d) easy confused AB and CV; (e-f) easy confused spiculation and PI.

2) Comparisons with independent feature space and original full set of original features

In order to prove the necessity of feature selection, we further conducted the CISL recognition by using each type of original features and the full set of original features, respectively. The corresponding average results on all the categories of CISLs are

shown in Table Ⅶ, where LBP(P, R) means the LBP feature vector configured with P neighbors and radius R, as described in Section Ⅳ.

TABLE Ⅶ
THE CLASSIFICATION PERFORMANCE FROM EACH OF SINGLE FEATURE SPACES AND THE FULL SET OF FEATURES

Feature Spaces	SVM (%)			Bag (%)			NB (%)			*k*-NN (%)			Ada (%)		
	SE	SP	CAR	SE	SP	CAR	SE	SP	CAR	SE	SP	CAR	SE	SP	CAR
B-HOG	47.9	95.3	**67**	48.4	94.9	**65**	62.4	95.2	**63**	54.5	94.5	**59.5**	54.4	95.1	**63.4**
LBP(5,1)	32.7	93.0	52.3	34.7	93.2	51.3	45.5	93.6	50.5	36.8	93.1	48.9	33.1	92.7	45.6
LBP(5,2)	33.1	92.7	51.1	37.4	93.2	52.4	49.4	93.6	50.5	44.4	93.7	53.2	33.2	92.5	46.8
LBP(4,1)	27.7	91.8	47.4	38.9	93.6	54.4	43.0	93.3	49.1	38.8	92.9	46.8	33.6	92.4	45.2
LBP(4,2)	30.4	92.4	49.3	39.5	93.2	51.7	48.4	93.4	49.9	40.1	93.2	48.5	33.5	92.3	44.2
CVH	39.1	93.4	54.2	38.3	93.4	51.7	43.7	92.2	36.8	40.0	93.1	49.1	34.5	92.9	47.2
Wavelet	34.1	93.2	52.6	40.2	93.7	55.0	40.6	93.8	44.2	49.2	94.2	56.6	36.4	92.8	48.3
Full	65.3	96.6	76.32	66.1	96.5	74.74	78.4	97.0	**77.26**	66.9	96.0	69.66	67.4	96.4	73.82

According to Table Ⅶ, (1) the best classifier is NB for the full set of features; (2) the best single type of features is B-HOG; and (3) the combination of different types of features can really improve the classification performance, since the SE, SP and CAR from the full set of features are all better than those from each single type of features for all the classifiers. But through comparing the data in Table Ⅴ and Ⅶ, we can see that all the measurements of recognition performance from the full set of features are behind those from the selected features by using our FIG method. In Table Ⅷ, We list the increase rates of SE, SP, and CAR brought by our selected features for each of classifiers, compared with the full set of features and B-HOG, respectively. These data confirms the effectiveness of our FIG method. It leads to better classification results and is independent of used classifiers. However, the increase rates for NB classifier and the full set of features are not very impressive. A possible reason is that the NB classifier is established based on the assumption that the features are statistically independent with each other, thus the influence of negative features may be weakened greatly after the training of NB, similar as the effect of feature selection.

To demonstrate the advantage of our selected features over the full set of features more clearly, we further calculate the difference between CMs for the combination of each classifier and our selected features and those for each classifier and the full set of features. The results are shown in Fig. 4. The fact that most of diagonal elements in differential CMs are positive and most of the others are negative proves that the use of selected features can increase the possibility of classifying the patterns into its true class and lower the possibility of confusing between different classes.

TABLE Ⅷ
THE INCREASE RATES OF SE, SP AND CAR BROUGHT BY OUR SELECTED FEATURES, COMPARED WITH THE FULL SET OF FEATURES AND THE BEST SINGLE TYPE OF FEATURES, RESPECTIVELY

Classifier	Increase Rates (%)					
	Full Set of Features			B-HOG		
	SE	SP	CAR	SE	SP	CAR
SVM	7.50	0.62	5.16	46.56	1.99	19.79
Bag	8.62	0.41	4.20	48.35	2.11	19.82
NB	1.28	0.10	0.75	27.24	2.00	23.56
k-NN	2.24	0.42	5.63	25.50	2.01	23.66
Ada	1.04	0.31	2.55	25.18	1.68	19.40

3) Comparisons with ARG feature selection method

To further prove the performance of our FIG feature selection method, we compare it with the commonly used GA feature selection method based on CAR. We call it ARG for short. The ARG method is similar to our FIG method in the framework, the main difference between them is the design of fitness function. The fitness in the FIG is computed based on the Fish criterion, while it is computed based on CAR in the ARG. We recorded the CISL recognition performance and computation time of FIG and ARG algorithm, respectively. All the experiments were performed on a computer with 2.33GHz CPU and 4GB Memory.

The comparisons of CAR for each considered classifier between FIG and ARG methods are shown in Fig. 5, where we can see that the classification accuracy brought by the ARG is slightly behind that by the FIG for all the consider classifiers.

We further conducted the paired t-test analysis [59] to determine whether there is a significant difference in effectiveness between FIG and ARG. The resultant two-tailed p values for SVM, Bag, NB, *k*-NN and Ada are 0.823, 0.334, 0.319, 0.957 and 0.858, respectively. Usually $p < 0.05$ is accepted as significant. So we conclude that although the FIG behaved a little better than the ARG on the average, the difference in the effectiveness between them is not significant.

However, the FIG is much better than the ARG on the computation efficiency. This can be demonstrated by comparing the average running time of one generation in the FIG and that in the ARG, as shown in Fig. 6. Since our FIG method is independent of the classifiers and is performed only once for all the classifiers, the computation time of it does not vary with the classifiers. Only 0.16s is needed for the FIG to complete a computation of one generation. In contrast, the computation time of ARG varies from 1.86s to 684.40s according to the classifier complexity. The big difference of the efficiency between the ARG and the FIG exists in that the ARG must re-train the classifier with the feature subset and perform the data classification in each iteration of fitness evaluation.

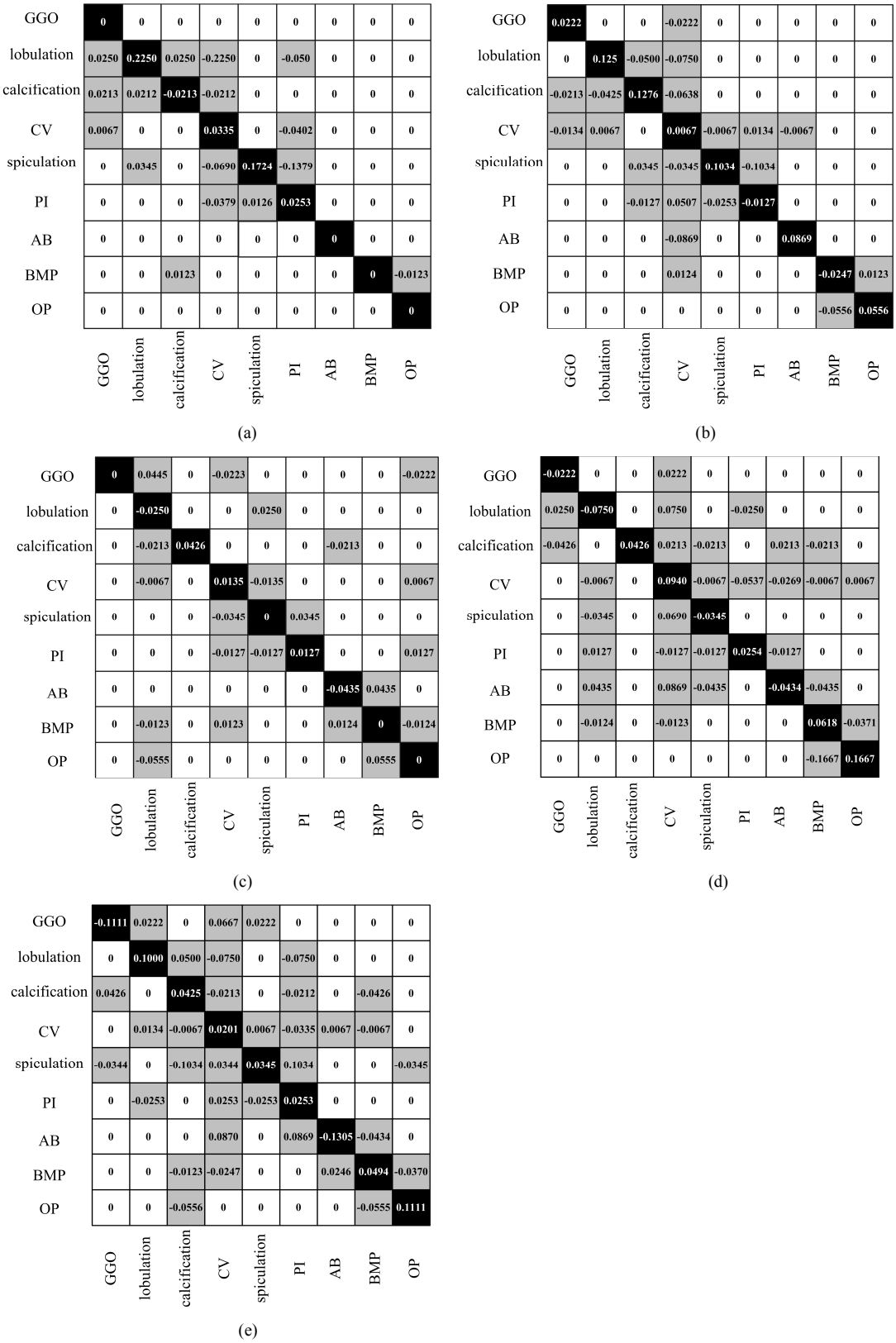

Fig. 4. The difference between the CMs for each classifier and our selected features and those for each classifier and the full set of original features: (a) for SVM; (b) for Bag; (c) for NB; (d) for *k*-NN; (e) for Ada.

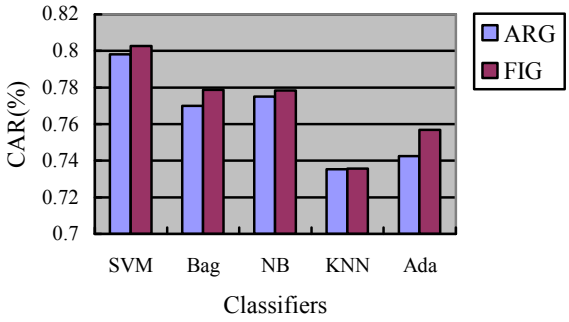

Fig. 5. The comparisons of CARs for each considered classifier between ARG and FIG feature selection methods.

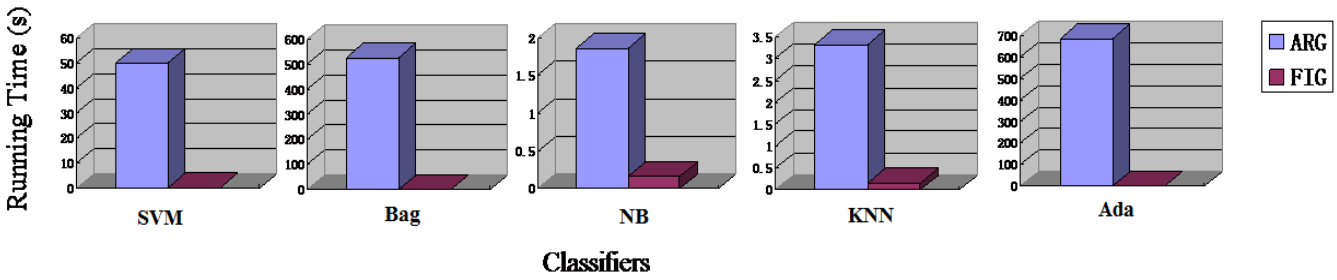

Fig. 6. The comparison of running time between ARG and FIG feature selection methods.

VI. Conclusions

This paper has proposed a new feature selection method based on Fisher criterion and genetic optimization for recognizing Common CT Imagining signs of Lung diseases (CISLs). The main contributions of this paper are summarized as follows.

1) The problem of recognizing nine categories of CISLs in lung CT images is put forward, which is important for the Computer-Aided Diagnosis (CAD) and the Content-Based Medical Image Retrieval (CBMIR) based on thoracic CT scans. To our knowledge, this problem has not received much attention of researchers. The previous works on lung tissue classification mainly concern about how to distinguish abnormal tissues from normal ones or identify among different visual patterns of a specific lung disease.

2) A feature selection method is presented based on FIsher criterion and Genetic optimization, which is called FIG for short. The Fisher criterion is applied to evaluate feature selection results, based on which a genetic optimization algorithm is developed to find out the optimal feature subset from candidate features. As demonstrated by the experimental results, our FIG method can bring more effective recognition results at the satisfactory computation costs, compared with single type of features and the full set of original features. Furthermore, it brought slightly better recognition performance and much better computation efficiency than the commonly used genetic feature selection method based on classification accuracy rate. Another advantage of the FIG is that it is independent of the classifiers; it is required to be performed only once to select the features suitable for all the considered classifiers.

3) The FIG method and each of five commonly used classifiers are combined to establish CISL recognizers, respectively, among which the SVM classifier behaved best. In 5-fold cross validation experiments on 511 ROIs which are manually extracted from real lung CT images, the cooperation of FIG and SVM achieved the average sensitivity of 70.2%, the average specificity of 97.2%, and the classification accuracy rate of 80.26%.

In the future, we want to add some image preprocessing steps to further improve the performance of our CISL recognizer. We can filter the blood vessels to get rid of the confusion between vessels and CISLs. We can also enhance the regions wrapping CISLs to make the visual appearance of CISLs be clearer and thus increase the possibility of correct classification.

References

[1] G. Battista, C. Sassi, M. Zompatori, D. Palmarini, and R. Canini, "Ground-glass opacity: interpretation of high resolution CT findings," *La Radiologia Medica*, vol. 106, pp. 425, 2003.

[2] Z.G. Yang, S. Song, and S. Talcashima, "high-resolution CT analysis of small lung adenocarcinoma revealed on screening helical CT," *Am. J. Roentgenol*, vol. 176, no. 6, pp.1399-1407, 2001.

[3] T. Aoki, Y. Tomada, H. Watanabe, et al, "Peripheral lung adenocarcinoma: correlation of thin-section findings with histologic factors and survival," *Radiology*, vol. 220, pp. 803-809, 2001.

[4] J. J. T. Owen, D. E. McLoughlin, R. K. Suniara, and E. J. Jenkinson, "The role of mesenchyme in thymus development," *Current Topics in Microbiology and Immunology*, vol. 251, pp. 133-137, 2000.

[5] M. R. Melamed, B. J. Flehinger, M. B. Zaman, et al, "Screening for lung cancer: Results of the Memorial Sloan-kttering study in New York", *Chest*, vol. 86, no. 1, pp. 44-53, 1984.

[6] C. V. Zwirewich, S. Vedal, R. R. Miller, et al, "Solitary pulmonary nodule: high-resolution CT and radiologic-pathologic correlation," *Radiology*, vol. 179, no. 2, pp, 469-476, 1991.

[7] S. F. Huang, R. F. Chang, D. R. Chen, and W. K. Moon, "Characterization of spiculation on ultrasound lesions,", *IEEE Trans. Medical Imaging*, vol. 23, pp. 111-121, 2004.

[8] M. Noguchi and Y. Shimosato, "The development and progression of adenocarcinoma of the lung," *Cancer Treatment and Research*, vol. 72, pp. 131-142, 1995.

[9] T. V. Colby and C. Lombard. "Histiocytosis X in the lung," *Human Pathology*, vol. 14, no. 10, pp. 847-856, 1983.

[10] V. J. Lowe, J. W. Fletcher, L. Gobar, et al, "Prospective investigation of positron emission tomography in lung nodules," *Journal of Clinical Oncology*, vol. 16, no. 3, pp. 1075-1084, 1998.

[11] K. S. Lee, Y. Kim, and S. L. Primack, "Imaging of pulmonary lymphomas," *American journal of roentgenology*, vol. 168, no.2, pp.339-345, 1997.

[12] J. W. Gurney, "Determining the likelihood of malignancy in solitary pulmonary nodules with Bayesian analysis: Part 1. Theory," *Radiology*, vol. 186, no. 2, pp. 405-413, 1993.

[13] J. J. Erasmus, H. I. McAdama, and J. H. Connolly, "Solitary pulmonary nodules: Part II. Evaluation of the indeterminate nodule," *Radiographics*, vol. 20, no. 1, pp. 59-66, 2000.

[14] L. Song, X. Liu, L. Ma, et al. "Using HOG-LBP features and MMP learning to recognize imaging signs of lesions," in *Proc. Computer-Based Medical System (CBMS)*, 2012, pp: 1-4.

[15] X. Ye, X. Lin, G. Beddoe, and J. Dehmeshki. "Efficient Computer-Aided Detection of Ground-Glass Opacity Nodules in Thoracic CT Images," in *Proc. 29th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBS)*, 2007. pp: 4449-4452.

[16] T. W. Way, B. Sahiner, H. P. Chan, et al, "Computer-aided diagnosis of pulmonary nodules on CT scans: Improvement of classification performance with nodule surface features," *Med. Phys.*, vol. 36, no. 7, pp. 3086-3098, 2009.

[17] H. Chen, Y. Xu, Y. Ma, and B. Ma. "Neural network ensemble-based computer-aided diagnosis for differentiation of lung nodules on CT images clinical evaluation," *Acad Radiol*, vol. 17, no. 5, pp.595-602, 2010.

[18] H. U. Kauczor, K. Heitmann, C. P. Heussel, et al, "Automatic detection and quantification of ground-glass opacities on high-resolution CT using multiple neural networks: Comparison with a density mask," *Am. J. Roentgenol*, vol. 175, no.5, pp.1329-1334, Nov. 2000.

[19] K. G. Kim, J. M. Goo, J. H. Kim, et al. "Computer-aided Diagnosis of Localized Ground-Glass Opacity in the Lung at CT: Initial Experience," *Radiology*, vol. 237, no. 2, pp. 657-661, 2005.

[20] S. C. Park, J. Tan, X. Wang, et al, "Computer-aided detection of early interstitial lung diseases using low-dose CT images," *Phys Med Biol*, vol. 56, no. 4, pp: 1139-1153, 2011.

[21] L. Song, X. Liu, A. Yang, et al. "A novel approach of computer-aided detection of focal ground-glass opacity in 2D lung CT images". in *Proc. SPIE 8670, Medical Imaging 2013: Computer-Aided Diagnosis*.

[22] R. Shen, I. Cheng, and A. Basu. "A hybrid knowledge-guided detection technique for screening of infectious pulmonary tuberculosis from chest radiographs," *IEEE Trans. Biomedical Engineering*, vol. 57, no. 11, pp. 2646–56, 2010.

[23] R. Uppaluri, T. Mitsa, M. Sonka, E.A. Hoffman, and G. McLennan, "Quantification of Pulmonary Emphysema from Lung Computed Tomography Images," *Amer. J. Respir. Crit. Care Med*, vol. 156, pp. 248-254, Jul. 1997.

[24] F. Chabat, G.Z.Yang, and D.M. Hansell, "Obstructive Lung Diseases: Texture Classification for Differentiation at CT," *Radiology*, vol. 228, pp. 871-877, Jul. 2003.

[25] P. Korfiatis, A. Karahaliou, C. Kalogeropoulou, A. Lazamtzo, and L. Costaridou, "Texture Based Identification and Characterization of Interstitial Pneumonia Patterns in Lung Multidetector CT," *IEEE Trans. Inf. Technol. Biomed*, vol. 14, pp: 675-680, May 2010.

[26] L. Sørensen, P. Lo, H. Ashraf, J. Sporring, M. Nielsen, and M. de Bruijne, "Learning COPD Sensitive Filters in Pulmonary CT," in *Proc. MICCAI*, vol. 5762, pp. 699-706, 2009.

[27] I. Sluimer, P. F. van Waes, M. A. Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Med. Phys.*, vol. 30, no. 12, pp.3081-3090, Dec. 2003.

[28] T. Nuzhnaya, V. Megalooikonomou, H. Ling, et al, "Classification of Texture Patterns in CT Lung Imaging," in *Proc. SPIE 7963, Medical Imaging 2011: Computer-Aided Diagnosis,* 2011.

[29] A. Depeursinge, J. Iavindrasana, and A. Hidki, "A classification framework for lung tissue categorization," in *Proc. SPIE 6919, Medical Imaging 2008: PACS and Imaging Informatics*.

[30] M. N. Gurcan, B. Sahiner, N. Petrick, et al. "Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system," *Med. Phys.*, vol. 29, no. 11, pp. 2552-8, 2002.

[31] J. Shiraishi, F. Li, and K. Doi. "Computer-aided diagnosis for improved detection of lung nodules by use of PA and lateral chest radiographs," *Radiology*, vol. 237, no. 2, pp. 657-661, 2005.

[32] M. Huber, M. Nagarajan, G. Leinsinger, et al, "Classification of Interstitial Lung Disease Patterns with Topological Texture Feature," in *Proc. SPIE 7624, Medical Imaging 2010: Computer-Aided Diagnosis*.

[33] J. G. Mehrdad, S. Lauge, B.S. Saher, et al, "Multiple Class System in Texton-Based Approach for the Classification of CT Images of Lung," *Medical Computer Vision*, vol. 6533, pp.153-163, Feb. 2011.

[34] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507-2517, 2007.

[35] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," *Journal of Machine Learning Research*, vol. 3, pp:1157-1182, 2003.

[36] M. A. Zuluaga, E. J. F. D. Leyton, M. H. Hoyos, M. Orkisz, "Feature Selection for SVM-Based Vascular Anomaly Detection," in *Proc. MICCAIW-MCV*, 2010, pp. 141-152.

[37] R. Nithya and B. Santhi, "Mammogram classification using maximum difference feature selection method," *Journal of Theoretical and Applied Information Technology*, vol. 33, pp. 197-204, 2011.

[38] S.F da Silva, B. Brandoli, D.M. Eler, et al, "Silhouette-based feature selection for classification of medical images," in *Proc. Computer-Based Medical System (CBMS)*, 2010, pp. 315-320.

[39] M. L. Huang, Y. H. Hung, and W. Y. Chen, "Neural Network Classifier with Entropy Based Feature Selection on Breast Cancer Diagnosis," *Journal of Medical Systems*, vol. 34, no. 5, pp 865-873, October 2010.

[40] H. A. Firpi and R. J. Vogelstein, "Particle Swarm Optimization-based Feature Selection for Cognitive State Detection," in *Proc. Engineering in Medicine and Biology Society (EMBC)*, 2011, pp. 6556-6559.

[41] J. G. Dy, C. E. Brodley, A. Kak, L. S. Broderick, and A. M. Aisen, "Unsupervised Feature Selection Applied to Content-Based Retrieval of Lung Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25 , no. 3, pp. 373 - 378, Mar. 2003.

[42] S. C. Park, B. E. Chapman, and B. Zheng, "A multistage approach to improve performance of computer-aided detection of pulmonary

embolism depicted on CT images: Preliminary Investigation," *IEEE Trans. Biomedical Engineering*, vol. 58, no. 6, pp. 1519-1527, 2011.

[43] Y. Zheng, X. Yang, M. Siddique, and G. Beddoe, "Simultaneous feature selection and classification based on genetic algorithms: an application to colonic polyp detection," in *Proc. SPIE 6915, Medical Imaging 2008: Computer-Aided Diagnosis*.

[44] R. Hupse and N. Karssemeijer, "The effect of feature selection methods on computer-aided detection of masses in mammograms," *Physics in Medicine and Biology*, vol.55, no.10, pp. 2893-2893, 2010.

[45] C. C. Wu, W. L. Lee, Yung-Chang Chen, et al, "Ultrasonic liver tissue characterization by feature fusion," *Expert Systems with Applications*, vol. 39, pp. 9389-9397, Aug. 2012.

[46] Y. Zhu, Y. Tan, Y. Hua, et al, "Feature selection and performance evaluation of support vector machine (SVM)-based classifier for differentiating benign and malignant pulmonary nodules by computed tomography," *Journal of Digital Imaging*, vol. 23, no. 1, pp.51-65, 2010.

[47] A. Ozcift, "SVM Feature Selection Based Rotation Forest Ensemble Classifiers to Improve Computer-Aided Diagnosis of Parkinson Disease," *Journal of Medical Systems*, vol. 36, no. 4, pp. 2141-2147, August 2012.

[48] S. Maggio, A. Palladini, L. De Marchi, et al., "Predictive deconvolution and hybrid feature selection for computer-aided detection of prostate cancer," *IEEE Trans. Medical Imaging*, vol. 29, no.2, pp.455-464, 2010.

[49] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley, New York, 2007.

[50] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," *Urbana*, vol. 51, pp. 61801-2996, 1991.

[51] W. Yang, D. Li, and L. Zhu, "An improved genetic algorithm for optimal feature subset selection from multi-character feature set," *Expert Systems with Application*, vol. 38, pp. 2733-2740, May 2011.

[52] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *Proc. CVPR*, 2005, pp.886-893.

[53] J. Wang, Y. G. Jiang, A. G. hauptmann, and C. W. Ngo, "Evaluating bag-of-visual-words representations in scene classification," in *Proc. Intel. Workshop Multimedia Information Retrieval*, 2007, pp.197-206.

[54] Y. Wan, X. Liu, K. Tong, et al. "GMM-ClusterForest：A Novel Indexing Approach for Multi-Features Based Similarity Search in High-Dimensional Spaces," in *Proc. Intel. Conf. Neural Information Processing (ICONIP2012)*, 2012, pp. 210-217.

[55] Y. Tao, L. Lu, M. Dewan, et al. "Multi-level ground glass nodule detection and segmentation in CT lung images," in *Proc. MICCAI*, 2009, pp: 715-723.

[56] A. Wang, H. J. Sun, and Y. Y. Guan. "The application of wavelet transform to multi-modality medical image fusion," in Proc. IEEE Intel. Conf. Networking, Sensing and Control (ICNSC), 2006, pp. 270-274.

[57] T. Ojala, M. Pietikainen, and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," *Pattern Recognition*, vol.29, no.1, pp. 51-59, 1996.

[58] H. Geoffrey, D. Andrew, and H.W. Ian, "WEKA: a machine learning workbench," in *Proc. the Second Australia and New Zealand Conf. Intelligent Information Systems*, 1994, pp.357-361.

[59] D. B. Rubin. "Matching to remove bias in observational studies," *Biometrics*, vol. 29, no. 1, pp. 159-183, 1973.

Xiabi Liu was born in Jingzhou city, Hubei province, China, in 1972. He received the B.S. degree in computer application from Huazhong University of Science and Technology, Hubei, China, in 1998 and the Ph.D. degree in computer science from Beijing Institute of Technology, Beijing, China, in 2005.

From 2005-2008, he was a lecturer in School of Computer Science at Beijing Institute of Technology. Since 2008, he has been an associate professor in School of Computer Science at Beijing Institute of Technology. He found and leads the Machine Learning and Multimedia Retrieval (MLMR) lab at Beijing Institute of Technology. He was the principal investigator of several research grants from National Natural Science Foundation of China and Ministry of Education in China. He is the author or coauthor of one book and more than 30 papers. He holds 4 patents. His current research interests include multimedia retrieval, machine learning, pattern recognition, and computer vision.

Ling Ma received the B.S. and M.S. degrees in computer science and technology from Jilin University, Changchun, China, in 2008 and 2011.

She is currently working toward the Ph.D. degree in the school of Computer Science at Beijing Institute of Technology, Beijing, China. Her research interests include computer-aided diagnosis and content-based image retrieval.

Li Song was born in Datong, Shanxi province, China, in 1987. She received the B.S. degree in Computer Science and Technology from Qufu Normal University, Shandong, China, in 2010 and the M.S. degree in Computer Science and Technology from Beijing Institute of Technology, Beijing, China, in 2013. She is currently working for Qihoo 360 Technology Co. Ltd as a Safety Engineer.

From 2010 to 2013, she pursued her M.S. degree in the MLMR Lab at Beijing Institute of Technology. Her research interest included pattern recognition, machine Learning, computer vision, et al. During the study, she had published several papers related to her research fields.

Yanfeng Zhao was born in Beijing, China, in 1978. He received the B.S. degree in medicine from Capital Medical University, in Beijing, China, in 2001 and the M.S. degree in Imaging and Nuclear Medicine from Chinese academy of medical sciences (CAMS) & Peking Union Medical College (PUMC) in Beijing, China, in 2012.

From 2001 to 2014, he worked as a resident physician and attending doctor in the department of diagnostic radiology in CAMS. His research interest is mainly on medical imaging diagnosis of oncology. He is the author of more than 10 articles.

Xinming Zhao (corresponding author) was born in Xiangyang city, Hubei province, China, in 1964. He received the B.S. degree in medicine from Tongji Medical University, in Wuhan, China, in 1988.

Since 2005, he has been a professor in the department of diagnostic radiology at the cancer hospital of Chinese Academy of Medical Sciences (CAMS). He has published 6 books and more than 70 papers.

Prof. Zhao is a member of abdominal group in Chinese Society of Radiology and pancreatic oncology group in China Society of Oncology. He is on two journals' editorial boards including Chinese Journal of Cancer and Chinese Journal of Oncology Prevention and Treatment. He was a recipient of the second class prize of Science and Technology Progress Award in Higher School, nominated by Ministry of Education in China, in 2012, and the third class prize of Beijing Municipal Science and Technology Award, in 2012.

Chunwu Zhou received the M.D. from the school of clinical medicine, China Medical University, Shenyang, China, in 1978.

From 1995 to 2006, he served as the clinical vice director of the Cancer Hospital, Chinese Academy of Medical Sciences. Currently he is a professor, chief physician, doctoral tutor, and the director of the Department of Diagnostic Imaging, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. Thirty-year's work on cancer imaging diagnosis, research and teaching accumulate extensive clinical experience for Dr. Zhou, and he is a renowned expert in the diagnostic imaging and differential diagnostic imaging of tumors and other diseases. Dr. Zhou was also involved in many important national and international research projects. He was the principal investigator of the National "9th Five-year" and "11th Five-year" Science & Technology Construction Program, and National "863" High Tech research project. He also worked as a co-investigator of two national "9th Five-year" projects and two national "10th Five-year" Projects. Throughout his career, Dr. Zhou has published more than 80 papers in the international and Chinese journals. He was alo the chief editor of five books and the deputy editor of one book.

Dr. Zhou is a member of the Standing Committee of Chinese society of Radiology, Chinese Medical Association, the leader of the breast research group of the 12th committee of the Chinese Society of Radiology, Chinese Medical Association; the vice chairman of Radiologist Association, Chinese Medical Association, the Chairman of Beijing branch of Chinese Association of Radiologists. He is the deputy editor of Chinese Journal of medical imaging technology and Chinese Journal of Oncoradiology. He is also on several other journals' editorial boards including Chinese Journal of Radiology and Journal of Clinical Radiology. He accepted the special government allowances of the state council in China, in 1997.