Development of Computer-aided Diagnosis System of Diffuse Lung Diseases on HRCT Images

(HRCT画像におけるびまん性肺疾患のコンピュータ診断支援システムの開発)

Wei Zhao

Applied Medical Engineering Science

Graduate School of Medicine

Yamaguchi University

2015.3

Contents

Chapter 1

Introduction

1.1 Background

The diffuse lung diseases (DLDs) refer to a series of disorders that affect many pulmonary tissues (parenchyma) in the lungs. Fig. 1.1 illustrates four example images of the DLDs. The expert meeting of American Thoracic Society and European Respiratory Society (ATS/ERS) produced a joint statement which divided the DLDs into two classes [1]: (1) disorders of known causes, such as collagen vascular diseases and environmental, occupational or drug-related diseases. (2) disorders of unknown causes which consist of three sub-classes: (2-a) idiopathic interstitial pneumonias (IIPs); (2-b) granulomatous diseases, such as sarcoidosis; (2-c) other forms including eosinophilic pneumonia. The common presence of the DLD patients can be divided into four types [2]: (1) exertional dyspnea or cough, (2) bilateral diffuse interstitial infiltrates on chest radiographs, (3) physiologic and gas exchange abnormalities and (4) histopathologic abnormalities of the pulmonary parenchyma. Fig. 1.2 illustrates several possible clinical effects of the DLDs on the body [3].

The pneumoconiosis is a kind of the DLDs, which is caused by prolonged inhalation and retention of industrial dust particles (20 years or more experience) in the lungs. So the workers in the industries like the mining and construction have a high risk of suffering from the pneumoconiosis. Fig. 1.3 gives the example image of the pneumoconiosis. It has been proven that the pneumoconiosis would lead to a severe lung function impairment, and there is a positive correlation

between the pneumoconiosis and lung cancer, especially for the smokers [4, 5]. At present, the pneumoconiosis is one of the most serious occupational diseases in some countries, especially in the developing countries. For example, in China, it was reported that there were total 527431 pneumoconiotic cases in the 2010 [6], and the number of new pneumoconiotic cases from 2010 to 2012 were 23812, 26401 and 24206 respectively [6, 7, 8].

For the diagnosis of the DLDs, it is recommended to start a consideration when the patient has a presence of the breathlessness and diffuse opacities on the chest radiographs [3, 9, 10, 11, 12]. Once the DLDs are suspected, the first step of the diagnosis should be a careful review of the clinical history, such as the medical and drug history, environmental and occupational exposures, history of the smoking and family, and any clues to the systemic disease that might involve the lungs. Then a comprehensive set of physical examinations should be operated, including a combination of chest radiographs (compared with previous films), pulmonary function test (such as spirometry), blood test, electrocardiogram and so on.

In order to provide a correct assessment and therapy program, the radiologists are required to make an accurate radiological interpretation of the chest radiographs. The correct identification of abnormalities and their distribution is useful to narrow the differential diagnosis of the DLDs. With the development of imaging technique, the high-resolution computed tomography (HRCT) is thought to be the best tool to assess the pulmonary patterns, because the nature of the three dimensional imaging can remove the superimposed of tissues on the images, and the high spatial resolution can provide sufficient anatomic details to detect the subtitle objects. Fig. 1.4 compares the images of the conventional CT (section thickness: 10mm) and HRCT (section thickness: 3mm). It can be found that the micro-structures can be clearly demonstrated on the HRCT image.

Furthermore, although the chest x-ray (CXR) is also an essential imaging modality in the clinic, it is thought to be non-specific to the DLDs due to the following two factors [13, 14, 15, 16]: 1. The CXR has a relative low resolution so that it is difficult to observe the subtle parenchymal abnormalities. 2. The two dimensional nature of the CXR imaging causes a superimposed of many parenchymal structures on the images. It was estimated that 40% of normal subjects on the CXR images would be obscured [14]. Fig. 1.5 compares the

example images of the CXR and HRCT. It is clearly that the HRCT image can provide a more accurate evaluation for the pulmonary tissues than the CXR image. The study [17] compared the diagnosis results based on the HRCT and CXR images. It is reported that 69% of subjects were correctly diagnosed by using CXR images, much lower than the HRCT images (80%). For the subjects recognized to be the normal case, 42% of the CXR and 18% of the HRCT were suffered from the DLDs. And the normal subjects which were correctly identified by using the CXR and HRCT were 82% and 96% respectively.

On the other hand, the main disadvantages of HRCT are the high expensive and high dose of radiation in the scanning. So, it is important to always consider the alternative techniques and adopt the HRCT examination only in the presence of the clear indications, especially for the children and pregnant women. In order to reduce the dose of the radiation, the HRCT scanning can be adapted to suit the examined body part and the body weight. Furthermore, researchers developed some elaborate techniques, for example, the low-dose HRCT by reducing the current during scanning. It was thought that although the low-dose HRCT cannot be employed for the initial evaluation, which may be failed to demonstrate the ground-glass opacities (2 of 10 cases) and emphysema (1 of 9 cases), it was valuable to screen the progression of the DLDs and evaluate the patients with some abnormalities [18, 19]

According to the appearances on the HRCT images, the abnormal opacities (pulmonary patterns) of the DLDs can be divided into four fundamental categories: reticular opacities (RET) shown as irregular lines, nodular opacities (NOR), increased lung opacities including honeycombing (HCM) and emphysema (EMP), and decreased lung opacities including consolidation (CON) and ground-glass opacity (GGO). Fig. 1.6 gives the example images of these six kinds of typical DLD opacities and normal tissues (NOR). Table 1.1 summarizes the typical abnormal pulmonary patterns and their distribution which may used to diagnose some DLDs .

Although the HRCT can provide an accurate assessment for the DLDs, there has not been an objective identification criterion to describe the complex DLD opacities in the current clinical protocol, and the interpretation of the abnormal

opacities mainly depends on the radiologists' individual expertise and clinical history. Therefore, the subjective differences between the radiologists' diagnosis are inevitable, which would lead to various treatments for the patients. In the study [20], the observer variations between a group of 11 radiologists' reports were estimated. Only 69% of the first choice diagnosis was made with no doubt or a high confidence (likelihood $> 70\%$), and the agreements on the first choice diagnosis were only moderate (the value of unadjusted kappa coefficients of agreement was 0.48). In order to improve the accuracy and reproducibility for the identification of the DLD patterns, an objective quantified is expected in the diagnosis of the DLDs. Additionally, the radiologists have to spend much time to review the large numbers of axial slices (about $300 \sim 500$ slices) in the HRCT scans. Due to the above reasons, a computer-aided diagnosis (CAD) system is required to facilitate the diagnosis of the DLDs by providing the radiologists with a "second opinion" and "objective criterion".

In the CAD methods, the calculation of image features is an important task which can represent the images or volumes of interest (VOIs) in a discriminative way. The conventional textural features, such as the statistical moments of the histograms and the moments based on the gray-level run-length matrices (GLRLM) [21] and gray-level co-occurrence matrix (GLCM) [22] have been successfully used in the classification of the DLD patterns. Wang et al. thought that the GLRLM could be partitioned into four areas with clear physical meaning [23]. For example, the top-left and bottom-right area of the GLRLM could be adopted to indicate the information of the GGO and NOR respectively. Park et al. used the GLRLM and the moments of the histograms to represent the texture information of the pulmonary patterns [24]. In the study [25], a total of 130 textural measures was calculated based on the 3D version of the GLCM. Besides, the adaptive multiple feature method (AMFM) [26] and its 3D version have been used in the CAD of the DLDs in the works [27, 28] respectively.

The features based on the textural information have an excellent performance on the classification of some DLD patterns. However, these features are difficult to distinguish the pulmonary patterns with inhomogeneous textures and special shapes. It can be found that the main finding of the NOD is the presence of miliary or centrilobular shapes, so the geometrical information based features such

as local binary pattern (LBP) [29, 30] and geometric fractal dimension (GFD) [31] would be more suitable than the textural information. The LBP can produce an same code to the structures with similar shapes by thresholding a local neighborhood with respect to the intensity of the center pixel. And the GFD refers to the index by measuring the complexity of structures as a ratio of the change in detail to the change in scale.

In order to improve the discriminative power, researches designed some sophisticated features by a combination of the textural and shape information based measures. Uppaluri et al. combined the textural features with the GFD, and the GFD was used to indicate the roughness of pulmonary textures [32]. In the study [33], the measures based on the histogram, gradient, GLCM and GLRLM were used for texture analysis, and the measures based on the top-hat transformation and clusters of low attenuation areas were used to analyze the shape information. In the work [34], the rotation-invariant Gabor-local binary patterns (RGLBP) and multi-coordinate histogram of oriented gradients (MCHOG) were used to describe the pulmonary texture and gradient information respectively. In additional, the LBP was employed to quantitatively measure the NOR and two subtypes of EMP [35], and the wavelet transform and Riesz transformed were also used for the classification in the works [36, 37, 38, 39].

1.2 Purpose of Our Research

In this thesis, by using a novel strategy named "sparse representation" to generate the features for the classification, we proposed two methods to classify the DLD patterns and pneumoconiosis respectively. Furthermore, considering that the operation of algorithms through the command-line interface (CLI) would be inconvenient for the radiologists, we implemented our methods as the plug-ins of a visualized CAD platform developed by our laboratory.

The goal of the sparse representation is to approximate the examples by a linear combination of few number of representative features (atoms) selected from an overcomplete dictionary, where the overcomplete dictionary means that the number of atoms in the dictionary is greater than the dimensionality of the examples. The application of the sparse representation approaches can improve

the performance of the image classification due to the following three reasons [40, 41, 42]. Firstly, the sparse representation of the image can naturally encode the semantic information of the image, because the image could be treated as a distribution of a set of representative examples. Secondly, the number of atoms is greater than the dimensionality of the input example. So the representation of each example is not unique, and it can search a relative better representation among the various combinations of atoms. Thirdly, the sparse representation with a high sparsity (few non-zero entries in the vector) was shown to be robust in the presence of noise [43]. Due to these advantages, the sparse representation approaches have been applied in the field of the CAD recently. Liu et al. developed a sparse representation based method to detect the colon polyp and lung nodule [44]. Kiet et al. used discriminative dictionaries to classify four kinds of pulmonary patterns [45]. In the work [46], the dictionary of the texton was trained and used to recognize the NOR and three subtypes of EMP.

We presented the main content of this thesis in the following five chapters. In the Chapter 2, we briefly reviewed the sparse representation approaches. In the Chapter 3, we proposed and optimized a sparse representation based method to classify the DLD patterns. In the Chapter 4, we designed a bag-of-features based method to classify the pneumoconiosis. The bag-of-features could be treated as a special version of the sparse representation. In the Chapter 5, we introduced a visualized CAD platform developed by our laboratory, and the proposed methods were implemented as the plug-ins of this platform. Finally, we concluded the thesis in the Chapter 6.

1.3 Computer-aided diagnosis (CAD) of Diffuse lung Disease (DLD)

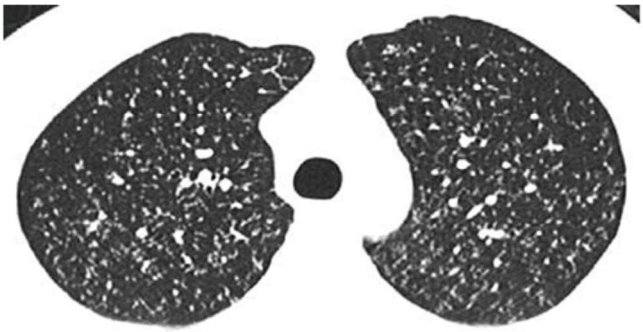
The goal of the CAD system is "to aid the radiologists for the medical image interpretation process by using computer algorithms" [47]. The output of CAD methods could be used a "second opinion" and "objective criterion" to help the radiologists to make the final diagnosis. There are four requirements for the application of the CAD methods in the clinic: (1) improve diagnosis accuracy

and reproducibility, (2) save review time, (3) seamless integrated into the workflow, (4) regulatory approval and cost efficiency. Now the major bottleneck of most CAD systems is the performance of the methods [48]. So far, some CAD algorithms have received approval or clearance from the U.S. Food and Drug Administration (FDA) in the United States.

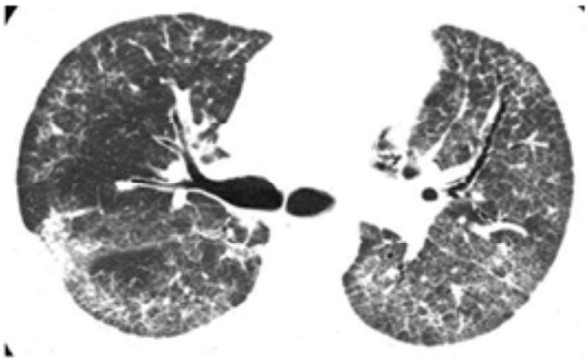
Over the last decade, several methods have been proposed for the CAD of the DLDs and achieved good results. For example, Uchiyama et al. designed 6 kinds of the physical measures to classify the NOR and six kinds of DLD patterns [49]. The features consist of three measures based on the CT values (mean and standard deviation of CT values, air density components) and other three measures based on the geometrical information (nodular components, line components and multilocular components). The three-layered artificial neural network (ANN) with the back propagation algorithm was adopted as the classifier. The accuracy of this method for the classification of pulmonary patterns was 100.0% (CON), 99.2% (GGO), 100% (HCM), 100% (RET), 95.8% (EMP), 88.0% (NOD) and 88.1% (NOR) respectively. In the work [50], a set of 12 kinds of measures based on the visible structures was calculated and combined with the measures based on the GLCM and GLRLM. The support vector machine (SVM) with a radial basis function (RBF) kernel was used as the classifier. The classification accuracy of the CON, GGO, HCM, RET, EMP, NOD and NOR was 95.2%, 90.2%, 97.9%, 65.4%, 92.9%, 86.7% and 96.3% respectively.

For the CAD of the pneumoconiosis, because the current diagnosis criterion made by International Labour Organization (ILO) is based on the CXR, most methods are developed to analyze the CXR images, such as the studies [51, 52, 53, 54, 55]. A recent work by Zhu et al. [56] designed a method to classify the first two stages of the pneumoconiosis on the digital chest radiographs (DRs). A set of 28 wavelet-based texture features was calculated, and the support vector machine and decision trees (DTs) were adopted as the classifiers. For the SVM with the RBF kernel and DTs wit algorithm C5.0, the areas under the receiver operating characteristic (AUCs) were 0.94 ± 0.02 and 0.86 ± 0.02 respectively. The work [57] designed three enhancement filters based on the window function, top-hat transformation and GLCM to remove the false positive of nodules. The tool combined the rule-based and ANN was adopted as the classifier. For the

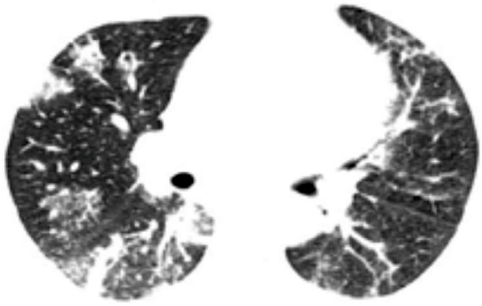
classification of the severe and early pneumoconiosis, the AUCs were 0.93 ± 0.02 and 0.72 ± 0.03 respectively. However, according to our knowledge there is no work on the CAD of the pneumoconiosis utilizing the HRCT images.



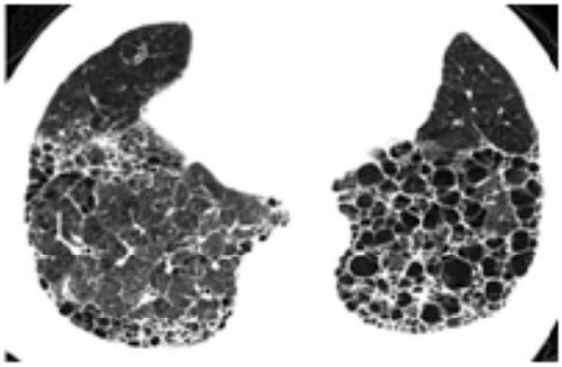
(a)



(b)



(c)



(d)

Figure 1.1: Example images of DLDs

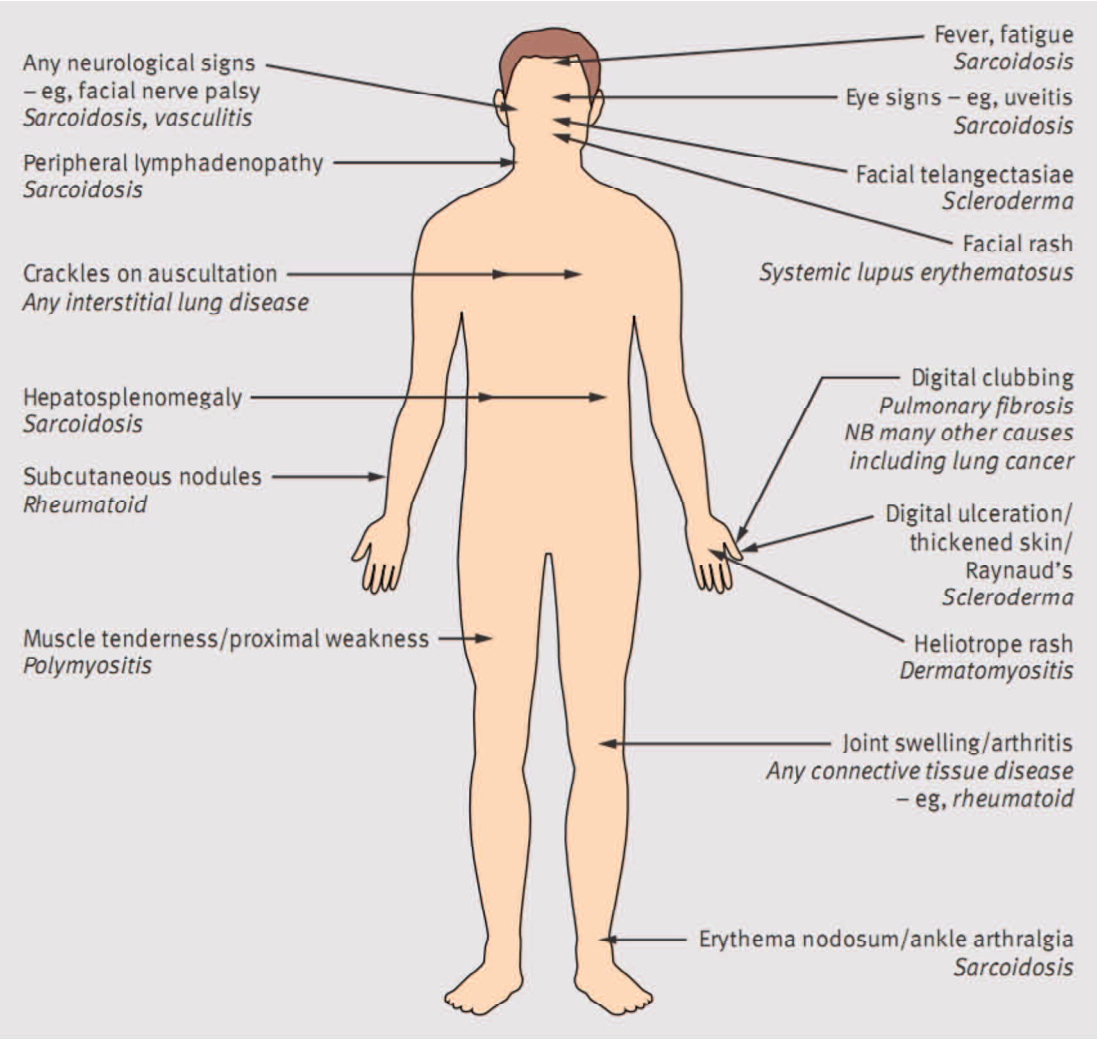


Figure 1.2: Possible clinical effects of DLDs [3]

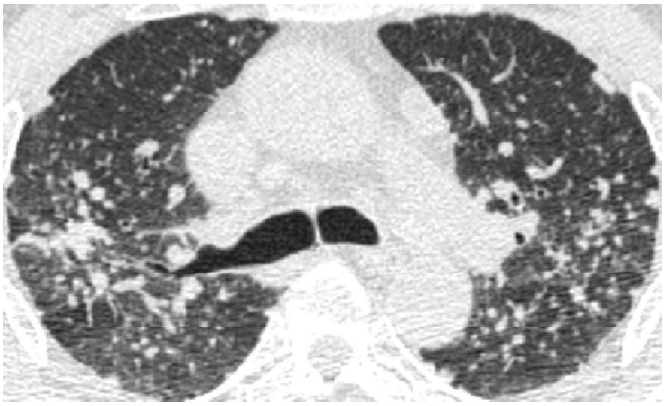
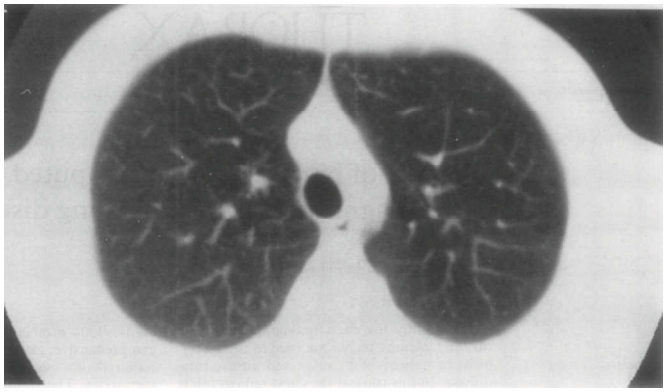
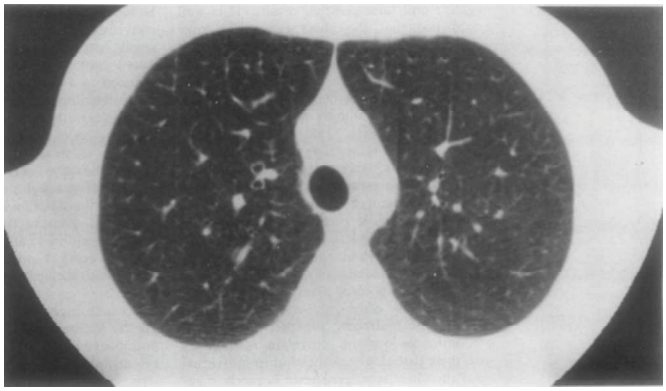


Figure 1.3: Example image of pneumoconiosis

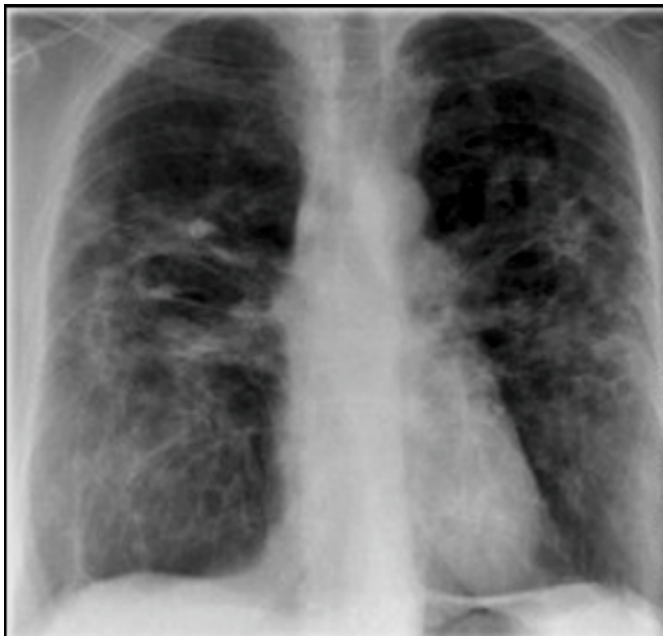


(a) Conventional CT image with a section thickness of 10mm



(b) HRCT image with a section thickness of 3mm

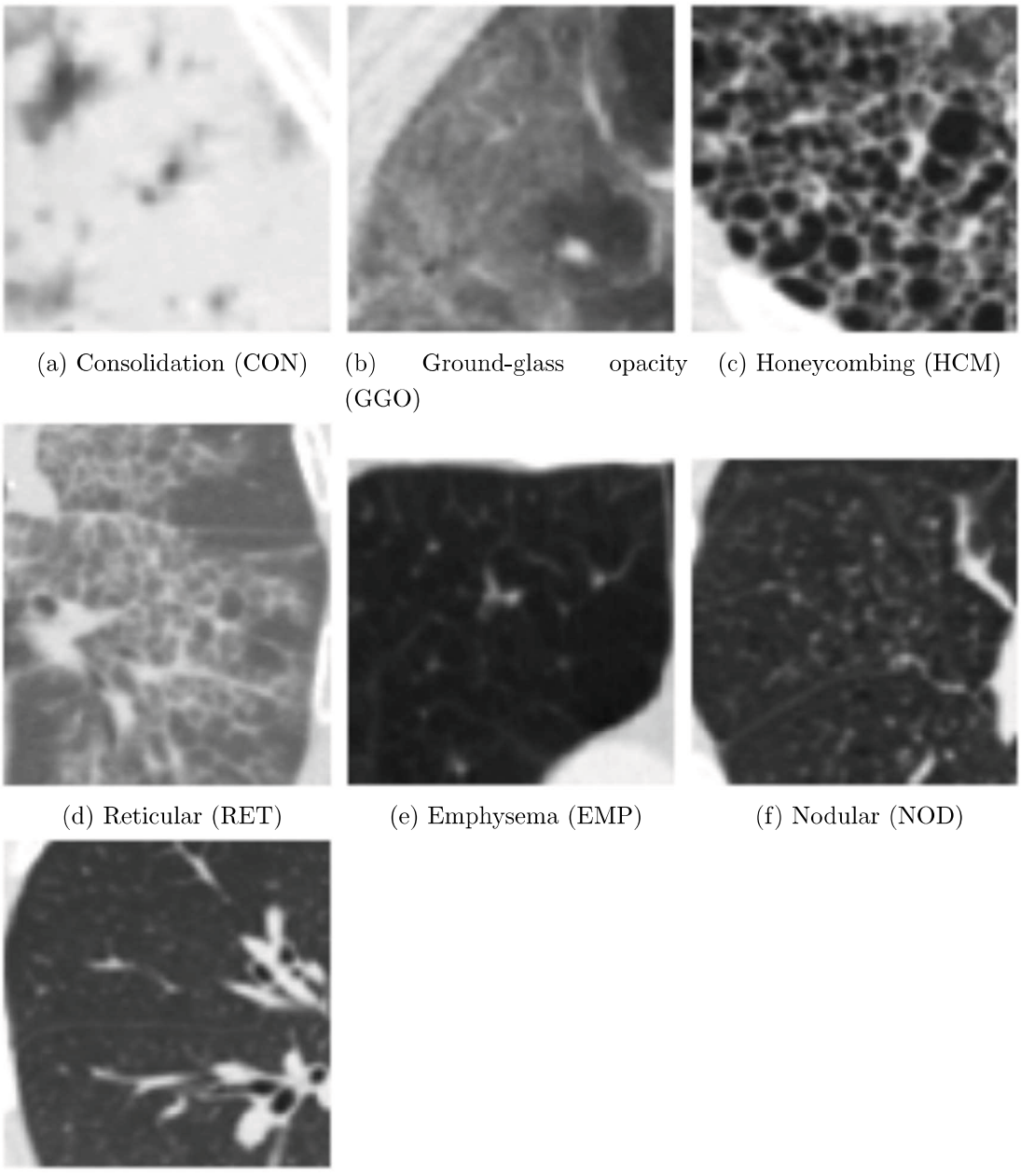
Figure 1.4: Comparison of conventional CT and HRCT images [14];



(a) CXR image

(b) HRCT image

Figure 1.5: Comparison of CXR and HRCT images [58];



(a) Consolidation (CON) (b) Ground-glass opacity (GGO) (c) Honeycombing (HCM)

(d) Reticular (RET) (e) Emphysema (EMP) (f) Nodular (NOD)

(g) Normal tissues (NOR)

Figure 1.6: The images of seven typical kinds of pulmonary patterns

1.3 Computer-aided diagnosis (CAD) of Diffuse lung Disease (DLD)

Table 1.1: Several findings of DLDs on the HRCT images[12]

Disease	Pulmonary patterns	Distribution
Idiopathic pulmonary fibrosis, usual interstitial pneumonia	reticular,honeycombing, traction bronchiectasis	Peripheral, subpleural, lower lung zones
Nonspecific interstitial pneumonia	Ground-glass opacity, reticular lines	Peripheral
Cryptogenic organizing pneumonia	Consolidation	Peripheral, peribronchial
Acute interstitial pneumonia	Ground-glass opacity, consolidation	Diffuse
Desquamative interstitial pneumonia	Ground-glass opacity, reticular lines	Lower lung zones
Hypersensitivity pneumonitis	Ground-glass opacity, nodules, air trapping	Upper and mid lung zones

Chapter 2

Sparse Representation Approaches

2.1 Introduction

Let the example and overcomplete dictionary be $\boldsymbol{y} \in \mathbb{R}^n$ and $\boldsymbol{D} \in \mathbb{R}^{n \times k}, n \ll k$ respectively, the sparse representation of m examples $\boldsymbol{a}_i \in \mathbb{R}^k, i = 1, 2, ..., m$ can be formulated as the solution of

$$\min_{\boldsymbol{D},\boldsymbol{a}} \sum_{i=1}^{m} ||\boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{a}_i||_2^2 \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq T \tag{2.1}$$

where the $||\cdot||_0$ means the ℓ^0 norm indicating the sparsity of the vector (number of non-zero entries in the vector), and T is the threshold of the sparsity. Define the reconstruction error vector as $\boldsymbol{e} \in \mathbb{R}^n$, the overall reconstruction of m examples can be calculated by Eq.2.2

$$\|\boldsymbol{E}\|_F^2 = \|[\boldsymbol{e}_1, \boldsymbol{e}_1, ..., \boldsymbol{e}_m]\|_F^2 = \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{A}\|_F^2, \boldsymbol{e}_i = \boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{a}_i \tag{2.2}$$

where the $\|\boldsymbol{.}\|_F$ means the Frobenius Norm.

The calculation of the sparse representation involves two essential components: (1) calculating the sparse approximation of the examples (coefficients of the atoms) $\boldsymbol{a}$ according to a pre-defined dictionary (it is also named sparse coding); (2) training an overcomplete dictionary $\boldsymbol{D}$. For the task of the image classification, the sparse representation approaches can be used to calculate the

sparse representation of image features, and these representation vectors can be adopted as the input of the classifier. If the input of the sparse representation is a set of local features (calculated at local regions such as the interest points or nodes in a dense grid instead of the global regions), such as the famous scale invariant feature transform (SIFT) [59] or histogram of orient gradients (HOG) [60], the process named "spatial pooling" should be operated to summarize the approximations of local features over local regions into a global descriptor before the classification.

2.2 Sparse coding

The sparse coding can be formulated as an optimization problem P_0 with a sparsity-inducing regularization

$$P_0 : \quad \min_{\boldsymbol{a}} \|\boldsymbol{a}\|_0 \quad s.t. \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \leq \varepsilon \tag{2.3}$$

The L_0-norm minimization has been proven to be a NP-hard (non-deterministic polynomial-time hard) problem [61], which means that the problem is hard to be solved precisely. Therefore, it is recommend to use the greedy heuristic strategy to get an approximate solution.

In order to make the sparse coding problem more tractable, a variation of the sparsity optimization is to replace the constraint of L_0-norm with the L_1-norm (P_1)

$$P_1 : \quad \min_{\boldsymbol{a}} \|\boldsymbol{a}\|_1 \quad s.t. \quad \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \leq \varepsilon \tag{2.4}$$

By using an appropriate Lagrange multiplier λ, the L_1-norm minimization can be tuned into an unconstraint convex function given by

$$F(\boldsymbol{a}) = \lambda\|\boldsymbol{a}\|_1 + \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \tag{2.5}$$

It has been proven that the L_1-norm minimization is convex and equivalent to the L_0-norm minimization if the solution is sufficient sparse [62]. So conventional optimization tools such as the steepest descent could be adopted to calculate the sparse approximation of input examples. However, these approaches would

be inefficient for the high dimensional image processing tasks, which requires too many iteration and calculation. Researchers have proposed several kinds of approaches to address the problem efficiently, for example, the classes of FOcal Underdetermined System Solver (FOCUSS) [63, 64, 65] and iterative-shrinkage algorithms. Besides, the greedy strategy and the principle of basis pursuit [66, 67] were widely used to solve the sparse coding problem.

It should be noticed that although the L2-norm minimization (P_2) is strictly convex ($\nabla^2\|\boldsymbol{a}\|_2^2 = \boldsymbol{I} \geq 0$) and have an unique optimal solution given by Eq.2.7, it cannot produce a sparse solution.

$$P_2: \quad \min_{\boldsymbol{a}} \|\boldsymbol{a}\|_2 \quad s.t. \quad \tfrac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \leq \varepsilon \tag{2.6}$$

$$\Rightarrow \quad \boldsymbol{a} = \boldsymbol{D}^{+}\boldsymbol{y} = (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}\boldsymbol{y} \tag{2.7}$$

2.2.1 Greedy algorithms

The calculation of the greedy algorithms can be divided into two steps: (1) searching the suitable atoms as the support of the solution and (2) calculating the coefficients of the selected atoms. The matching pursuit (MP)[68] and orthogonal matching pursuit (OMP) [69] are two of the most popular choices to solve the sparsity optimization problem (P_0). In the beginning, the support of the solution was empty ($\|\boldsymbol{a}\|_0 = 0$). Then the atom was identified which can maximally reduce the residual error, and the approximation of input example was updated at each iteration, until the reconstruction error was small enough or the number of used atoms was greater than the desired sparsity. There were two main differences between the MP and OMP. Firstly, in each iteration, the MP identified the atom from the set of all atoms, but the OMP searched the atom that had not been used in the previous iteration(s). Secondly, the MP updated the coefficient of the atom selected in the current iteration (the coefficients of other atoms were kept), but the OMP calculated the coefficients of all selected atoms in each iteration. Therefore, the OMP had a fast convergence speed and lower reconstruction error than the MP.

The least angle regression stagewise (LARS) [70] was a greedy method for the sparse coding with the L_1-norm minimization. In the calculation, the LARS

applied the derivative of unconstraint L_1-norm function with respect to the coefficient vector $\boldsymbol{a}$ (Eq.2.8).

$$\frac{\partial F}{\partial \boldsymbol{a}}(\boldsymbol{a}) = \boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}) + \lambda \boldsymbol{z}, \quad \boldsymbol{z}[i] = \begin{cases} 1 & \boldsymbol{a}[i] > 0 \\ [-1,\ 1] & \boldsymbol{a}[i] = 0 \\ -1 & \boldsymbol{a}[i] < 0 \end{cases} \tag{2.8}$$

Let the $\boldsymbol{a}_s$ be the non-zero portion of the $\boldsymbol{a}$, and the $\boldsymbol{D}_s$ be the sub-matrix of $\boldsymbol{D}$ constructed by the atoms used to represent the $\boldsymbol{a}_s$, the non-zero part of the coefficient was obtained by solving the restricted zero-derivative function

$$\frac{\partial F}{\partial \boldsymbol{a}_s}(\boldsymbol{a}_s) = \boldsymbol{D}_s^T(\boldsymbol{y} - \boldsymbol{D}_s\boldsymbol{a}_s) + \lambda \boldsymbol{z}_s = 0 \tag{2.9}$$

$$\Rightarrow \boldsymbol{a}_s = (\boldsymbol{D}_s^T\boldsymbol{D}_s)^{-1}(\boldsymbol{D}_s^T\boldsymbol{y}_s - \lambda \boldsymbol{z}_s) \tag{2.10}$$

The work [71] improved the LARS algorithm by adding a step of searching the "feature-sign" of each entry in the approximation vector. And the solution was updated by using an discrete line search between the current approximation and $\boldsymbol{a}_s$. In the study [72], the locality-constraint of the atoms was introduced in the calculation of the approximation, because it was thought that the locality of the atom can lead to the sparsity of the approximation. The object function can be re-written as

$$\begin{aligned} \min_{\boldsymbol{a}} \|\boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{a}_i\|_2^2 + \lambda \|\boldsymbol{b}_i \odot \boldsymbol{a}_i\|_2^2 \\ \boldsymbol{b}_i = exp(\frac{dist(\boldsymbol{y}_i, \boldsymbol{D})}{\sigma}), \\ dist(\boldsymbol{y}_i, \boldsymbol{D}) = [dist(\boldsymbol{y}_i, \boldsymbol{d}_1), ..., dist(\boldsymbol{y}_i, \boldsymbol{d}_k)]^T \end{aligned} \tag{2.11}$$

where $\odot$ stands for the element-wise multiplication, and the $dist(\boldsymbol{y}_i, \boldsymbol{d}_j), j = 1, 2, ..., k$ is the Euclidean distance between the vector $\boldsymbol{y}_i$ and $\boldsymbol{d}_j$. The LLC can be solved by

$$\widetilde{\boldsymbol{a}}_i = (\boldsymbol{C}_i + \lambda diag(\boldsymbol{d})) \setminus \boldsymbol{1} \tag{2.12}$$

$$\boldsymbol{a}_i = \widetilde{\boldsymbol{a}}_i / \boldsymbol{1}^T \widetilde{\boldsymbol{a}}_i \tag{2.13}$$

2.2.2 FOcal Underdetermined System Solver (FOCUSS)

The core of the FOCUSS based approaches is to apply an iteratively reweighed least squares (IRLS) strategy to update the approximation [73]. It was thought

that the solution of the relaxed optimization (L_1-norm) problem had the tendency to spread the energy among a large number of entries in the $\boldsymbol{a}$ instead of choosing just few non-zero entries. So in order to avoid this bias, a diagonal weighted matrix $\boldsymbol{W}, \boldsymbol{W} = diag(|\boldsymbol{a}|)$ was introduced to scale the entries in the minimization [74]. Because $\|\boldsymbol{a}\|_1 = \boldsymbol{a}\boldsymbol{W}^{-1}\boldsymbol{a}$, the unconstraint function of L_1-norm with weighted matrix can be expressed by

$$F_{\boldsymbol{W}}(\boldsymbol{a}) = \lambda \boldsymbol{a}\boldsymbol{W}^{-1}\boldsymbol{a} + \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \tag{2.14}$$

The derivative of Eq.2.14 with respect to $\boldsymbol{a}$ is given by

$$\frac{\partial F_{\boldsymbol{W}}}{\partial \boldsymbol{a}}(\boldsymbol{a}) = 2\lambda \boldsymbol{W}^{-1}\boldsymbol{a} + \boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}) \tag{2.15}$$

At k-th iteration, the approximation was calculated by

$$\boldsymbol{a}_k = (2\lambda \boldsymbol{W}_{k-1}^{-1} + \boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{y} \tag{2.16}$$

If $\|\boldsymbol{a}_k - \boldsymbol{a}_{k-1}\|_2$ was smaller than a pre-defined threshold, the $\boldsymbol{a}_k$ was output as the resulting approximation. In the study [75], it was thought that the non-negative coefficients would be advantageous in some cases in image processing. So a non-negative FOCUSS algorithm was proposed, which forced all all negative entries to 0.

2.2.3 Iterative-Shrinkage Algorithms

The main idea of the iterative-shrinkage based algorithms is to operate a shrinkage (soft-threshold) step at each iteration. The shrinkage operator was original designed in the work [76] expressed by

$$Shrinkage\ operator: \qquad S_v(\boldsymbol{a})_i = sign(\boldsymbol{a}_i) \cdot (|\boldsymbol{a}|_i - v) \tag{2.17}$$

where v is a constant. For the sparse coding problem, even if the original function was not convex, the obtained solution can be the globally optimal if the operator was properly designed [77].

The work [78] proposed an algorithm named "iterative shrinkage thresholding" (IST) by introducing an additional term $\varphi(\boldsymbol{a}, \boldsymbol{a}_0)$.

$$\varphi(\boldsymbol{a}, \boldsymbol{a}_0) \quad = \frac{c}{2}\|\boldsymbol{a} - \boldsymbol{a}_0\|_2^2 - \frac{1}{2}\|\boldsymbol{D}\boldsymbol{a} - \boldsymbol{D}\boldsymbol{a}_0\|_2^2 \tag{2.18}$$

Adding this term into the Eq.2.5, the surrogate object function of the sparse coding problem with L_1-norm is given by

$$\widetilde{\boldsymbol{F}}(\boldsymbol{a}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 + \lambda \mathbf{1}^T\rho(\boldsymbol{a}) + \frac{c}{2}\|\boldsymbol{a} - \boldsymbol{a}_0\|_2^2 - \frac{1}{2}\|\boldsymbol{D}\boldsymbol{a} - \boldsymbol{D}\boldsymbol{a}_0\|_2^2 \tag{2.19}$$

where the function $\rho(\boldsymbol{a}) = \|\boldsymbol{a}\|_1$ performs entry-wise on the vector $\boldsymbol{a}$. The choice of the parameter c should satisfy the condition of $c\boldsymbol{I} - \boldsymbol{D}^T\boldsymbol{D} > 0$, which means that the function $\varphi(\cdot)$ is strictly convex. By opening and re-organizing the terms, Eq.2.19 can be expressed as

$$\widetilde{\boldsymbol{F}}(\boldsymbol{a}) = constant + \frac{1}{2}\|\boldsymbol{a} - \boldsymbol{v}_0\|_2^2 + \frac{\lambda}{c}\mathbf{1}^T\rho(\boldsymbol{a}), \boldsymbol{v}_0 = \frac{1}{c}\boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}_0) + \boldsymbol{a}_0 \tag{2.20}$$

$$= \sum_{i=1}^{m}[\frac{1}{2}(\boldsymbol{a}[i] - \boldsymbol{v}_0[i])^2 + \frac{\lambda}{c}(\boldsymbol{a}[i])] \tag{2.21}$$

The global minimizer of Eq.2.19 was given by

$$\boldsymbol{a}_{min} = S_{\frac{\lambda}{c}}(\boldsymbol{v}_0) = S_{\frac{\lambda}{c}}(\frac{1}{c}\boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}_0) + \boldsymbol{a}_0) \tag{2.22}$$

In the k-th iteration, the approximation was calculated by

$$\boldsymbol{a}_{k+1} = S_{\frac{\lambda}{c}}(\boldsymbol{a}_k) = S_{\frac{\lambda}{c}}(\mu \cdot \frac{1}{c}\boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}_k) + \boldsymbol{a}_k) \tag{2.23}$$

with a stepsize of μ. The work [79] proposed a two-step IST (TwIST) structure, where the approximation in the $(k+1)$-th iteration depended on the previous two iterations.

$$\boldsymbol{a}_1 = \Gamma_\lambda(\boldsymbol{x}_0) \tag{2.24}$$

$$\boldsymbol{a}_{k+1} = (1-\alpha)\boldsymbol{a}_{k-1} + (\alpha - \beta)\boldsymbol{a}_k + \beta\Gamma_\lambda(\boldsymbol{x}_k), k \geq 1 \tag{2.25}$$

where $\Gamma_\lambda(\boldsymbol{a}) = S_\lambda(\boldsymbol{a} + \boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{x}))$. A similar approach was proposed in the work [80], where the shrinkage step was operated on the point $\boldsymbol{u}_k$ generated by a linear combination of the previous two iterations $\boldsymbol{a}_{k-1}, \boldsymbol{a}_{k-2}$.

$$\boldsymbol{a}_k = S(\mu\frac{1}{c}\boldsymbol{D}^T(\boldsymbol{D}\boldsymbol{u}_k - \boldsymbol{y}) + \boldsymbol{u}_k) \tag{2.26}$$

$$\boldsymbol{u}_{k+1} = \boldsymbol{a}_k + (\frac{t_k - 1}{t_{k+1}})(\boldsymbol{a}_k - \boldsymbol{a}_{k-1}) \tag{2.27}$$

$$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2} \tag{2.28}$$

In the study [81], a IRLS-based iterative shrinkage algorithm was proposed.

$$\boldsymbol{a}_{k+1} = \boldsymbol{S} \cdot (\boldsymbol{a}_k - \frac{1}{c}\boldsymbol{D}^T(\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}_k)), \quad \boldsymbol{S} = (\frac{\lambda}{c}\boldsymbol{I} + \boldsymbol{W}_k)^{-1}\boldsymbol{W}_k \tag{2.29}$$

where $\boldsymbol{W} = diag(|\boldsymbol{a}_k|)$, and the matrix $\boldsymbol{S}$ played a role of shrinkage operator.

2.2.4 Basis Pursuit

The basis pursuit (BP) was an optimization principle to approximate the example by using a set of optimal atoms with L_1-norm constraint. It was suggested to use the approach named "basis pursuit de-noising" (BPDN) [82] to solve the BP problem. The BPDN was formulated as the solution of

$$BPDN: \qquad \min_{\boldsymbol{a}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 + \lambda\|\boldsymbol{a}\|_1 \tag{2.30}$$

where the λ was a parameter controlled the residual error. It can be found that the BPDN was equivalent to the least absolute shrinkage and selection operator (LASSO) regression [83] in some sense which can be defined as a similar formulation (Eq.2.31)

$$LASSO: \qquad \min_{\boldsymbol{a}} \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{D}\boldsymbol{a}\|_2^2 \leq \varepsilon \;\; s.t. \;\; \|\boldsymbol{a}\|_1 \leq \tau \tag{2.31}$$

where τ was a constant. The BPDN can be reformulated as a linear programming problem, and solved by using a primal-dual logarithmic-barrier interior point algorithm [84]. In the start, an initial basis was constructed by selecting n atoms from the dictionary. Then some atoms used in the current basis were swapped by new atoms and adopted to update the approximation, until the relative changes between the two iterations was small enough.

In order to improve the performance of the BPDN, several approaches have been proposed. In the work [85], a block coordinate relaxation (BCR) based algorithm was developed. The BCR algorithm can be thought as an improvement of the interior point (IP) algorithm to minimize the object function with respect to a block of entries while kept other entries fixed in each iteration. In the BCR,

the strategies of systematic cyclical and optimal descent were applied to select the block of coefficient vector. The work [86] proposed a homotopy-continuation based algorithm to generate the path of solutions with all possible values of λ with a low computational cost. It can fast optimize the parameter of the BPDN. The study [87] applied the spectral projected-gradient (SPG) algorithm [88, 89] to efficiently solve the object function of DPBN. By using the nonmonotone line search scheme [90] and two-point step size gradient method [91], the SPG algorithm had a fast convergence speed than the conventional gradient project method. In the work [92, 93], a fix-point continuation (FPC) algorithm was proposed which applied a gradient descent and shrinkage operator to update the approximation in the each iteration

2.3 Dictionary Learning

In the sparse representation, a set of representative features which is used as the atoms of the dictionary is essential to approximate the example. This dictionary should contain the salient information of the examples. There are two choices to generate the dictionary. The first one is using a parametric function with a group of parameters, for example, the Gabor function [68] or 2D Gabor function [94]. Although the dictionary based on the off-the-shelf signals would be universal to all classes of images, it would not capture the key features for the specific kind of images. Therefore, it is better to learn the dictionary by extracting the representative features from the training data.

Generally, the dictionary learning is an iterative process to reduce the reconstruction error until it is small enough. There are two stages in each iteration: (1) optimizing the sparse approximations with the fixed dictionary (sparse coding stage) and then (2) optimizing the dictionary with the fixed approximation vectors (dictionary updating stage).

The method of optimal direction (MOD) [95] and K-SVD based on the singular value decomposition (SVD) [96] are two of the most popular approaches for the dictionary learning. Besides, it was considered that the dictionary simultaneously used for reconstruction and classification may lose discriminative information, so

some algorithms were proposed to train the discriminative dictionary which can represent the examples in a more discriminative way.

2.3.1 Method of Optimal Direction (MOD)

In order to reduce the reconstruction error of the sparse representation as many as possible in each iteration, the optimal direction was thought to force the derivative of the reconstruction error with respect to the dictionary to 0. So this approach was called *m*ethod of *o*ptimal *d*irection (MOD). The zero-derivation of the reconstruction error with respect to the $\boldsymbol{D}$ can be solved by Eq.2.32

$$\frac{\partial F}{\partial \boldsymbol{D}}(\boldsymbol{D}) = (\boldsymbol{D}\boldsymbol{A} - \boldsymbol{Y})\boldsymbol{A}^T = 0 \tag{2.32}$$

Therefore, the dictionary could be updated by the following equation.

$$\boldsymbol{D} = \boldsymbol{Y}\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1} \tag{2.33}$$

In the work [71], the Lagrange multiplier was introduced and the dictionary was updated by

$$\boldsymbol{D} = \boldsymbol{Y}\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T + \boldsymbol{\Lambda})^{-1} \tag{2.34}$$

where $\boldsymbol{\Lambda} = diag(\lambda), \lambda > 0$. The MOD is a simple algorithm, and can produce a good dictionary to minimize the overall reconstruction of the sparse representation. However, the major disadvantageous of the MOD is the high computational complexity on the matrix inverse, particularly for the dictionary with a large size.

2.3.2 K-SVD

The K-SVD could be seemed as a generative of the K-Means algorithm. According to the Eckart-Young Theorem, a matrix can be approximated by the another one with a lower-rank [97]. And the SVD can decompose the input matrix into several lower-rank matrixes. Therefore, when the k-th atom was being updated, the reconstruction error matrix except the k-th atom $\boldsymbol{E}_k$ can be calculated by Eq.2.35 where $\boldsymbol{a}_T^j$ stands for the j-th column of coefficient matrix and then decomposed into three rank-1 matrixes by using the SVD (Eq.2.36). The first left-singular

vector $\boldsymbol{u}_1$ was used to update the atom, and the multiplication of the fist right-singular vector and the first singular $\boldsymbol{\Delta}_{1,1} \times \boldsymbol{v}_1$ was treated as the corresponding coefficient (Eq.2.37). In order to keep the sparsity, the input reconstruction error matrix should be restricted by choosing the columns that used the updated atom before the decomposition.

$$\boldsymbol{E}_k = \boldsymbol{Y} - \sum_{j \neq k} \boldsymbol{d}_j \boldsymbol{a}_T^j = \boldsymbol{E} + \boldsymbol{d}_k \boldsymbol{a}_T^k \tag{2.35}$$

$$\boldsymbol{E}_k = \boldsymbol{U} \boldsymbol{\Delta} \boldsymbol{V}^T \tag{2.36}$$

$$\boldsymbol{d}_k = \boldsymbol{u}_1, \boldsymbol{a} = \boldsymbol{\Delta}_{1,1} \times \boldsymbol{v}_1 \tag{2.37}$$

2.3.3 Discriminative Dictionary

In the work [98], the dictionary was learned based on the idea that the dictionary $\boldsymbol{D}_i$ associated to the examples of the class S_i should have a better performance to reconstruct this kind of examples than other dictionaries $\boldsymbol{D}_j, j \neq i$, while bad to reconstruct other kinds of examples $S_j, j \neq i$. In order to improve the discriminative power, a discriminative term was introduced based on the soft-max cost function. Assuming that there were total z classes of dictionaries, the discriminative term of the example $\boldsymbol{y}$ belonging to the class S_i was calculated by

$$\begin{aligned} &C_i^{\lambda}(\{R(\boldsymbol{y}, \boldsymbol{D}_t)\}_{t=1}^z) = C_i^{\lambda}(r_1, r_2, ..., r_z) = log(\sum_{t=1}^{z} e^{-\lambda(r_i - r_t)}), \\ &r_t = R(\boldsymbol{y}, \boldsymbol{D}_t) = \|\boldsymbol{y} - \boldsymbol{D}_t \boldsymbol{a}_t\|_2^2, \qquad t = 1, 2, ..., z \end{aligned} \tag{2.38}$$

where λ is a parameter, $\boldsymbol{D}_t$ is the S_t class of the dictionary, $\boldsymbol{a}_t$ and r_t are the approximation vector and reconstruction error of the example obtained by using the t-th class of the dictionary respectively. The value of the discriminative term was close to zero when the class S_i of the dictionary produced the most accurate reconstruction for the example. In other words, the value of r_i was less than other entries in the set $\{r_1, r_2, ..., r_z\}$.

The new dictionary learning problem with the discriminative term (Eq.2.38) was formulated by the following equation

$$\min_{\{\boldsymbol{D}_t\}_{t=1}^z} \sum_{l \in S_i, i=1}^{z} C_i^{\lambda}(R(\boldsymbol{y}_l, \boldsymbol{D}_t)) + \lambda \gamma R(\boldsymbol{y}_l, \boldsymbol{D}_t) \tag{2.39}$$

where $\boldsymbol{y}_l, l \in S_i$ is the l-th of the example and belongs to the class S_i, and the γ is a parameter related to the trade-off between the reconstruction and discrimination. It can improve the discriminative power of the model by decreasing the value of the parameter $\gamma, \gamma \geq 0$.

When the j-th atom of the i-th class of the dictionary $\boldsymbol{d}_{i,j}$ was being updated, firstly a set of examples θ that using the atom in the classes $S_p, p = 1, 2, ..., z$ was constructed.

$$\theta \leftarrow \{l = 1, 2, ..., m | \boldsymbol{a}_{l,i} \neq 0\} \tag{2.40}$$

Then the weights of examples belonging to the set θ for all classes $S_p, p = 1, 2, ..., z$ were calculated. The weight was defined as the derivative of the object function (Eq.2.40) with respect to the i-th class of the reconstruction error.

$$w_l = \frac{\partial C_p^{\lambda}}{\partial r_i}(\{R(\boldsymbol{y}_l, \boldsymbol{D}_t)\}_{t=1}^{z}) + \lambda\gamma \mathbf{1}_i(p), \mathbf{1}_i(p) = \begin{cases} 0 & p = i \\ 1 & p \neq i \end{cases} \tag{2.41}$$

Finally the weighted matrix of the residual error $\boldsymbol{B}$ was constructed by using the selected examples (Eq.2.42), and the atom was updated by the eigenvector of the matrix corresponding to the greatest eigenvalues.

$$\boldsymbol{B} = \sum_{p=1}^{z} \sum_{l \in S_p \cap \omega} w_l (\boldsymbol{r}_l + \boldsymbol{a}_{li}[j]\boldsymbol{d})(\boldsymbol{r}_l + \boldsymbol{a}_{li}[j]\boldsymbol{d})^T, \boldsymbol{r}_l = \boldsymbol{y}_l - \boldsymbol{D}_i \boldsymbol{a}_{l,i} \tag{2.42}$$

The study [99] designed a label consist discriminative dictionary by combing the reconstruction error with the sparse coding error e_{sc} and classification error e_c respectively. The goal of introducing the sparse coding error was to force the examples with the same label had a similar approximation, which was given by

$$e_{sc} = \|\boldsymbol{Q} - \boldsymbol{U}\boldsymbol{A}\|_2^2 \tag{2.43}$$

where $\boldsymbol{Q} \in \mathbb{R}^{k \times m} = [\boldsymbol{q}_1, ..., \boldsymbol{q}_m]$ was the "discriminative" codes associated with the original examples. The entry $q_{i,j}$ was set to 1 if the atom $\boldsymbol{d}_i$ and the example $\boldsymbol{y}_j$ shared the same label, otherwise the entry was set to 0. $\boldsymbol{U}$ was a linear transformation matrix used to transform the approximation $\boldsymbol{a}_i$ to the discriminative sparse codes. The classification error was calculated by

$$e_c = \|\boldsymbol{H} - \boldsymbol{W}\boldsymbol{A}\|_2^2 \tag{2.44}$$

where $\boldsymbol{H} = [\boldsymbol{h}_1, ..., \boldsymbol{h}_m] \in \mathbb{R}^{l \times m}, \boldsymbol{h}_i = [0, 0, ..., 1, ..., 0, 0]^T$ was the labels of the original examples, and the $\boldsymbol{W} \in \mathbb{R}^{l \times k}$ was the model parameters of the classifier. By using the two terms, a new object function was defined

$$\begin{aligned} arg \min_{\boldsymbol{D},\boldsymbol{W},\boldsymbol{U},\boldsymbol{A}} &= \|\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{A}\|_2^2 + \alpha\|\boldsymbol{Q} - \boldsymbol{U}\boldsymbol{A}\|_2^2 + \beta\|\boldsymbol{H} - \boldsymbol{W}\boldsymbol{A}\|_2^2 \\ & s.t. \quad ||\boldsymbol{a}_i||_0 \leq T \end{aligned} \tag{2.45}$$

The Eq.2.45 can be re-formulated and solved by using the K-SVD algorithm [96] to calculate the following equation.

$$\begin{aligned} & arg \min_{\boldsymbol{D},\boldsymbol{W},\boldsymbol{U},\boldsymbol{A}} = \| \begin{pmatrix} \boldsymbol{Y} \\ \sqrt{\alpha}\boldsymbol{Q} \\ \sqrt{\beta}\boldsymbol{H} \end{pmatrix} - \begin{pmatrix} \boldsymbol{D} \\ \sqrt{\alpha}\boldsymbol{A} \\ \sqrt{\beta}\boldsymbol{W} \end{pmatrix} \boldsymbol{A}\|_2^2 \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq T \\ & \rightarrow arg \min_{\widetilde{\boldsymbol{D}}} \|\widetilde{\boldsymbol{Y}} - \widetilde{\boldsymbol{D}}\boldsymbol{A}\|_2^2 \quad s.t. \quad ||\boldsymbol{a}_i||_0 \leq T \end{aligned} \tag{2.46}$$

2.4 Spatial pooling

When the local features instead of global features are calculated on the original images and then sparse coded, it is necessary to summarize the approximation vectors over different local regions into a global descriptor by operating a step of spatial pooling. There are several spatial pooling strategies, for example, the average of the vectors (average pooling), the sums of the vectors (sum pooling) or the maximum of the entries with the same index (maximum pooling). Let $\boldsymbol{z} \in R^k$ and $\boldsymbol{a} \in R^k$ be the approximation vector and global descriptor respectively, and $\{\cdot\}_t$ be the t-th element of the vector, the three pooling strategies of m vectors are given from Eq.2.48 to Eq.2.49 respectively.

$$Average\ pooling: \quad \boldsymbol{z}_t = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{a}_{it}, t = 1, 2, ..., k \tag{2.47}$$

$$Sum\ pooling: \quad \boldsymbol{z}_t = \sum_{i=1}^{m} \boldsymbol{a}_{it}, t = 1, 2, ..., k \tag{2.48}$$

$$Maximum\ pooling: \quad \boldsymbol{z}_t = max(\boldsymbol{a}_{1t}, \boldsymbol{a}_{2t}, ..., \boldsymbol{a}_{mt}), t = 1, 2, ..., k \tag{2.49}$$

Some researchers thought that the spatial pooling would disregard the spatial layout of the local features and lack any sort of meaningful ordering. In order to take advantage of the spatial arrangement of the local features, the work [100]

proposed a "spatial pyramid" based approach. In the beginning, the entire images (or regions of interest) was partitioned into increasing coarser blocks by a multi-level recursive image decomposition. After extracting the local features from the blocks and sparse coding, the "spatial pyramid pooling" was performed to generate the global descriptor by weighted concentrating the sparse approximation of the local features over the blocks at each level of scale.

The concept of spatial pyramid had been adopted in several works. For example, in the work [101], the images were evenly divided into the blocks with 3 different levels respectively: 1×1, 2×2, 3×1. In the work [42], the images were partitioned into $2^l \times 2^l$ blocks in the scales $l = 0, 1, 2$, and the max pooling was performed over different regions and scales.

2.5 Conclusion

In this chapter, we briefly introduce several sparse representation approaches. The calculation of the sparse representation can be divided into two parts: training a good dictionary, and sparse approximate the input vectors according to a given dictionary. If the input of the sparse representation based classification is a set of local features, the step of spatial pooling is essential to summarize the sparse representation of local features into a global descriptor as the input of the classifier.

Chapter 3

Classification of Diffuse lung Disease Patterns

3.1 Introduction

In this chapter, we adopted and optimized the sparse representation approaches to classify the NOR and five kinds of DLD patterns, including CON, GGO, HCM, EMP and NOD. According to our knowledge, there is no work aimed to apply the sparse representation approaches to analyze these six kinds of the pulmonary patterns. Fig. 3.1 illustrates the framework of the proposed method. In the training stage, firstly huge numbers of local features were extracted from the training volumes of interest (VOIs), and used to train an overcomplete dictionary. Secondly, the sparse representation of the local features was calculated according to the dictionary (sparse coding) and the VOI-level descriptors of the training VOIs were generated by the average pooling. Finally, the VOI-level descriptors were used to train a support vector machine (SVM) classifier. In the testing stage, after extracting the local features on the testing VOIs, the dictionary was adopted to calculate the sparse representation of the local features, and the VOI-level descriptors of the testing VOIs were generated. At last, the descriptors were fed into the trained classifier and the results were given.

3.2 Proposed Methods

3.2.1 Extraction of local features

From Fig. 1.6, it can be found that the main finding of the NOD and HCM is the opacities with various shapes, which could be analyzed by the geometrical information. Some other patterns, such as the CON and EMP are mainly featured by the CT values. So, a combination of the measures based on the shape information and CT values could be used for the classification of the DLD patterns. One the other hand, it is thought that the local features can produce a specific semantic interpretation for the image analysis [102, 103]. For example, the line components can be associated with the vessels, and the blob components can be used to identify the nodule candidates. Therefore, we tried to calculate the shape information and CT values based measures on the local regions in our work. In our method, we adopted the local features proposed in the studies [104, 105], which used the eigenvalues of the Hessian matrix to measure the geometrical information.

The calculation of the local features is illustrated in Fig.3.2. At each sampling point on the VOIs, 1. A cubic-shape patch was constructed whose center was located at the sampling point, and four kinds of statistical moments were computed on the patch: mean, standard deviation, skewness and kurtosis. 2. Let the eigenvalues of the Hessian matrix be $\lambda_1, \lambda_2, \lambda_3, \lambda_1 \geq \lambda_2 \geq \lambda_3$, we computed the eigenvalues on each voxel within the patch, and arranged the eigenvalues by the location of the voxels where the eigenvalues were calculated. So three new patches can be generated whose components were λ_1, λ_2 and λ_3 respectively, and the same kinds of statistical moments were calculated on these eigenvalues-based patches. 4. The statistical moments were concentrated into a vector, and a 16-dimensional feature vector was obtained. In the experiments, the step of the sampling points was set to $4 \times 4 \times 4$. The size of the cubic-shape patch was a parameter, which was tuned from $2 \times 2 \times 2$ to $6 \times 6 \times 6$ with a step of $1 \times 1 \times 1$.

3.2.2 Dictionary Learning and Sparse Coding

The K-SVD [96] and OMP [69] are two of the most popular algorithms for the dictionary learning and sparse coding in the sparse representation. By adopting

these two algorithms, we proposed a sparse representation based method that was called SR1 in the thesis. Fig. 3.3 presents the K-SVD algorithm. In the K-SVD, the reconstruction error was iteratively minimized until it was small enough. Two steps were performed in each iteration: (1) computing the $\boldsymbol{a}$ with the fixed $\boldsymbol{D}$ (sparse coding stage); and then (2) updating the $\boldsymbol{D}$ with the fixed $\boldsymbol{a}$ (dictionary updating stage). In the step of the sparse coding, the OMP was used which is recommended in the K-SVD [96]. Fig. 3.4 gives the OMP algorithm. In the step of the dictionary updating, the atoms of the dictionary (columns of $\boldsymbol{D}$) were sequentially modified based on the SVD.

After training the dictionary, the OMP was also applied to calculate the sparse representation of the local features in the SR1, the same as the sparse coding stage of the K-SVD. There were two parameters in the sparse representation, the number of atoms and the sparsity of the representation. We adjusted the number of atoms from 500 to 3000 with an interval of 500, and the sparsity from 2 to 14 with an interval of 2 in the experiments.

3.2.3 Optimization of Sparse Representation

The SR1 using the K-SVD and OMP (see Subsection 3.2.2) can achieve a high classification accuracy, but the runtime of the SR1 was relatively long (see Subsection 3.3.4). Considering that the dictionary learning and sparse coding spent the most time on the training and testing respectively, we also tried to reduce the runtime of these two stages.

Although the decreasing running time in the dictionary learning seemed uselessness in the clinical workflow, we thought that it can be convenient to update the existing methods for the developers. It is reported that the K-Means algorithm can achieve a competitive performance on the natural image classification to the K-SVD when the same number of atoms were used [106]. Moreover, the K-Means could be seen as a simple version of the K-SVD. Firstly, the dictionary was updated by K times of means operation in the K-Means, or by K times of SVD in the K-SVD. Secondly, in the K-Means, the example was assigned to the closest atom and the coefficient of used atom was set to 1 (the coefficients of other atoms were set to 0). However, in the K-SVD, the sophisticated OMP

algorithm was used to approximate the example. Therefore, the runtime of the K-Means on updating dictionary and approximating example can be shorter than the K-SVD. In additional, the K-Means can be efficiently implemented by using the k-dimensional tree (k-d tree) [107]. Therefore, we tried to adopt the K-Means as a substitute of the K-SVD in the dictionary learning.

In the OMP, the sparse representation of the example was iteratively calculated until the stopping condition was achieved. For each iteration, only one atom was selected and added to the solution support, and then the approximation of the example was solved according to the elements of the current solution support. It could be deduced that the runtime of the OMP can be reduced by selecting enough atoms at one time. We named this method OMP_1 in the thesis. Fig. 3.5 gives the OMP_1 algorithm. After arranging the inner products of the atoms and example by a descending order in the first iteration of OMP, the first sufficient number of atoms were treated as the solution support, and the sparse representation of the example was calculated as the projection onto the linear space spanned by the selected atoms. Although the reconstruction error of the OMP_1 would be larger than the OMP, the performance of this approach can be ensured under a certain sparsity constraint [77].

In order to examine the performances of the sparse representation approaches, we constructed another two sparse representation based methods: SR2 (K-Means+OMP) and SR3 (K-Means+OMP_1). Table 3.1a summarizes the proposed three sparse representation based methods. The parameters of the SR2 and SR3 were same as the SR1: number of atoms and sparsity. The replacement of the dictionary learning approach was evaluated by comparing the SR1 and SR2, because both two methods adopted the OMP for the sparse coding, and the substitution of sparse coding approach was evaluated by comparing the SR2 and SR3, because both two methods adopted K-Means for the dictionary learning. Table 3.1 summarizes the experiments.

3.2.4 Spatial pooling

After calculating the sparse representation of local features, the procedure of the spatial pooling was performed to summarize the sparse representation of the

local features into VOI-level descriptors for each VOI. These descriptors could be thought as the global features of VOIs. We used one of the most popular choices, the average pooling in the work, which could be seen as an average operation on the vectors. Let $\boldsymbol{z} \in R^k$ and $\boldsymbol{a} \in R^k$ be the sparse representation vector and descriptor respectively, and $\{\cdot\}_t$ be the t-th element of the vector, the average pooling of m sparse representation vectors is given by Eq.3.1.

$$Average\ pooling: \quad \boldsymbol{z}_t = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{a}_{it}, t = 1, 2, ..., k \tag{3.1}$$

3.2.5 Classification

In the research, the support vector machine (SVM) was adopted as the classifier. We used a version named LIBSVM [108]. Considering that the sparse representation based classification using the linear kernel can achieve a competitive performance and lower computational cost than the nonlinear kernel [42], we employed the LIBSVM with a linear kernel which is given by Eq.3.2

$$Linear\ kernel: \quad K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j \tag{3.2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are both descriptors. Because the SVM was originally designed as the binary (two-class) classifier, the "one-against-one" approach was applied to extend the binary SVM classifier for the solution of the multi-class tasks in the LIBSVM [108]. There was one parameter in the classification: soft-margin penalty C, which was set to $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$.

3.3 Experiment

3.3.1 Data

We obtained 117 scans from 117 subjects from Tokushima University Hospital in Japan. In the 117 scans, there were 111 scans with the DLDs and 6 scans with the normal tissues. All scans were acquired by Toshiba Aquilion 16-row multi-slice CT when edge-enhanced filtering was not applied. A tube voltage of 120kVp and current of 250mAs were used. The resolution of scans was 512×512, and the in-plane resolution was 0.6mm. The slice thickness was 1mm.

In the previous works [104, 105], the area of abnormal tissues in the VOIs was not considered. We thought that it would be difficult to evaluate the classification of the VOIs with only little area of abnormities. Therefore, we set a low-threshold of 70% for the area of dominant pulmonary patterns in the generation of the VOIs. We used the number of voxels belonging to the specific kind of pulmonary patterns to determine whether this criterion was met or not. Furthermore, because the diagnosis of DLDs mainly depends on the radiologists' individual experiences, we thought that the results made by only one radiologist cannot be used as the "golden standard" for evaluation. So, we requested three radiologists to review the images in the experiments.

The VOIs were constructed according to the following procedures: 1. All scans were reviewed by one radiologist, and a maximum of three axial slices was selected from the top, middle and bottom parts of the lung respectively in each scan. Only one type of pattern dominantly existed on the selected slices, and the radiologist should indicate what the dominant pulmonary texture was and where it existed. 2. Another two radiologists reviewed the results of the first radiologist. Only the slices which were thought to be correct by both two radiologists were saved for the next procedures. 3. The regions of the pulmonary patterns on the selected slices were marked by all three radiologists respectively. 4. The common regions chosen by the radiologists were extracted. 5. The grids with a size of 32×32 were overlaid on the determined regions, and the square-shape patches were generated where the number of voxels belonging to the determined regions was greater than 716 (70% of 32×32). 6. By using the patches (generated in the step 5) as the central-axial slices, the VOIs with a size of $32 \times 32 \times 32$ were constructed, and the label of each VOI was assigned according to the kind of the pulmonary texture dominantly existed on the central-axial slice.

3.3.2 Experimental Setting

In the experiments, we separated the VOIs into two independent sets. One set (1161 VOIs) was adopted as the training set to optimize the parameters of each method and train each method with the optimal parameters. The other set (1049 VOIs) was used as the testing set to evaluate the performance of each method.

The VOIs in the two sets belong to the different patients (subjects). In other words, there were no crossing-subjects between the two sets. The number of VOIs of each type of patterns for training and testing are summarized in Table 4.4.

There were four kinds of parameters in the proposed methods: the size of cube-shape patches, the number of atoms, the sparsity of sparse representation and the parameter related to the classifier. We tuned the values of patch size from $2 \times 2 \times 2$ to $6 \times 6 \times 6$ with a step of $1 \times 1 \times 1$, the number of atoms from 500 to 3000 with an interval of 500 and the sparsity from 2 to 14 with an interval of 2. The parameter of the SVM classifier was set to $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$. A 20-fold cross-validation was used to optimize the parameters in the training stage. In the beginning of the cross-validation, the training data was randomly divided into 20 subsets with a nearly equal size. In each iteration 19 subsets were selected to train a classifier with the fitting parameters, and the remaining set was used to validate the classifier. This procedure was repeated 20 times, so every example in the training set was used to train the classifier 19 times and validate the classifier once.

We simultaneously optimized all kinds of parameters, which means that every possible combination of all kinds of parameters was tried. The combination of parameters which produced the best overall accuracy in the 20-fold cross validation was chosen as the optimal parameters. All methods were operated on the server with a 2.8GHz Intel Core i7 CPU and 24GB RAM.

3.3.3 Baseline methods

We compared the proposed methods with three kinds of published approaches, which were called as SDF [49], CSE [109] and BOW [105] respectively. The parameters of the baseline methods were optimized in the same way as the proposed methods.

1. In the work [49], the pulmonary patterns were determined by six kinds of **s**pecially **d**esigned **f**eatures. So this method was called SDF in the paper. The six features adopted in the SDF were mean and standard deviation of CT values, air density components, nodular components, line components and multilocular

components. The details of the air density components, nodular components, line components and multilocular components had been reported in the work [49]. In order to easily understand, we briefly introduced the calculation of these four components. (1) Air density component: the area whose CT values ranged from -910 HU to -1000 HU. (2) Nodular components and line components: firstly the "white" structures were extracted by performing the morphological white top-hat transform on the original image. Secondly the small noise structures were removed by applying the gray-level thresholding on the detected "white" structures. Finally, the nodule components and line components were determined according to the two conditions: the degree of the circularity and contrast. The degree of the circularity was defined by the fraction of the overlap area of the component with the circle having the same area. And the contrast was defined by the mean value of the five largest pixel values for each detected component on the white top-hat-transformed image. (3) Multilocular components: the candidates were extracted by performing the morphological black top-hat transform. Because the SDF required the 2D regions of interest (ROIs), the central slices in the axial direction of the VOIs were treated as the ROIs in the operation.

The SDF adopted a three-layered artificial neural network (ANN) with back-propagation as the classifier. The training of the ANN was an iterative process. In the beginning, the weights of hidden units were randomly evaluated. Then the resulting output was produced by using the weights and activation function in the hidden layer. Finally the weights were adjusted according to the error between the desired output and resulting output, and used as the training pattern for the next iteration. We used the symmetrical sigmoid given by Eq.3.3 as the activation function.

$$Symmetric\ sigmoid: \quad f(x) = (1 - e^{-x})/(1 + e^{x}) \tag{3.3}$$

There is a parameter in the ANN: the number of hidden units. In the work [49], the number of hidden units was set to 10 empirically. We adjusted the number of hidden units from 5 to 30 with an interval of 5 in the experiments. Additionally, we designed a complementary method of SDF by using the SVM classifier instead of the ANN. This method was called SDF-SVM in the paper. The LIBSVM with

RBF kernel (Eq.3.4) was adopted.

$$RBF\ kernel: \quad K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma||\mathbf{x}_i - \mathbf{x}_j||^2} \tag{3.4}$$

The parameter of the RBF kernel (γ) and the soft-margin penalty (C) were set to be $2^{-10}, 2^{-9}, ..., 2^{1}$ and $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$ respectively.

2. In the work [109], the classification was performed based on the signatures of VOIs. The signature was defined as the centroids and the weight of the clusters (number of voxels in the clusters). The K-means algorithm was used to calculate the centroids of the clusters. The earth mover's distance (EMD) approach was used to measure the similarity between the two signatures, and the nearest neighbor (NN) was employed as the classifier. In order to reduce the computational cost, the canonical signatures for each class were generated by combining and re-clustering the signatures of the training data. Therefore, the VOIs were recognized by comparing the signatures of VOIs with canonical signatures instead of the signatures of training data. Because this technique used **c**anonical **s**ignatures and **e**arth mover's distance (EMD), it was called CSE in the paper. The CSE had only one parameter: the number of clusters. Considering that the larger value was suggested to be avoided, we adjusted the number of clusters from 5 to 60 with a step of 5.

3. The work [105] adopted a model named "bag-of-words" (also named bag-of-features) [110] to generate the global descriptors for classification, so this method was called BOW in the thesis. The main idea of the BOW was to train a code-book (dictionary), and used the histogram of the words (atoms) in the code-book to describe the images. The BOW adopted the K-Means algorithm to cluster the local features, and the centers of the clusters were saved as the words in the code-book. The number of the words were adjusted from 50 to 400 with a step of 50 in the experiments. Because the local features adopted in the BOW were the same as our methods, we adjusted the size of patches from $2 \times 2 \times 2$ to $6 \times 6 \times 6$ with a step of $1 \times 1 \times 1$, the same as proposed methods. The SVM was used as the classifier. Considering that it was reported that χ^2 kernel achieved a better result than other kernels in the work [105], the LIBSVM was employed with the χ^2 kernel. Eq. 3.5 gives the χ^2 kernel, where α is the parameter for the kernel,

and $\mathbf{x}_i, \mathbf{x}_j$ are both histograms with k-bins.

$$RBF\ kernel: \quad H(\mathbf{x}_i, \mathbf{x}_j) = e^{-\alpha \sum_{t=1}^{k} \frac{(\mathbf{x}_{it} - \mathbf{x}_{jt})^2}{\mathbf{x}_{it} + \mathbf{x}_{jt}}} \tag{3.5}$$

The possible value of the soft-margin penalty and α were set to be 2^{-2}, 2^{-1},...,2^{11}, 2^{12} and 2^{-10}, 2^{-0},...,2^{1} respectively.

3.3.4 Results

Table 3.3 gives the overall accuracies of all methods with the optimal parameters. It illustrates that the sparse representation based methods achieved better results than the baseline methods (SR1:96.1%, SR2:95.6% and SR3:96.4% vs. SDF:75.8%, SDF-SVM:73.6%, CSE:65.1% and BOW: 85.5%). The p-values for the proposed methods against the baseline methods were all smaller than 0.0001, which means that there were significant differences between the performances of the proposed methods and baseline methods. The confusion tables of all methods are given from Table 3.4 to Table 3.9.

Table 3.10 shows that the replacement of the K-SVD and OMP by the K-Means and OMP_1 can reduce the runtimes of sparse representation based methods. When the K-SVD was replaced by the K-Means, the runtime of the dictionary learning was decreased by 98.2% (SR1:13520s vs. SR2:241s). On the other hand, when the OMP_1 was substituted for the OMP, the runtime of recognizing one VOI was dropped by 55.2% (SR2:0.29s vs. SR3:0.13s).

3.4 Discussion

The experimental results show that the sparse representation based methods can achieve good results (greater than 95%) for the classification of the pulmonary patterns, and the runtimes of the sparse representation based methods can be decreased by adopting the K-Means and OMP_1.

By using the sparse representation approaches, it can extract the important information of example while removing irrelevant details. Moreover, there would be a similar distribution for the local features extracted from the same kind of examples in the feature space. So the sparse representation of the local features

could be distributed with a low intra-class variation and high inter-class variation in a high dimensional feature space, which is advantageous for the pattern recognition. Experimental results also illustrate that the sparse representation based methods were efficient to classify the pulmonary patterns.

Comparing with the baseline methods, the proposed methods had a noticeably better performance on the classification of the HCM and NOR. For the classification of HCM, the sensitivity and precision by all methods were 99.5% \100.0% (SR1), 99.5% \ 99.0% (SR2), 99.5% \ 99.5% (SR3). And the sensitivity and precision to classify the NOR were 97.1% \ 96.4% (SR1), 97.4% \ 95.0% (SR2), 96.3% \ 96.3% (SR3). Fig. 3.6 and Fig. 3.7 illustrate two example images of the HCM and NOR which were correctly classified by our methods. The HCM has a presence of diffuse decreased lung opacities on the HRCT images ("blacker" than the surrounding normal tissues), slightly similar to the normal tissues in comparison to the increased pulmonary opacities. Although there were many "white" structures in the VOIs, our methods successfully recognized the VOIs.

On the other hand, the classification of the GGO and NOD by using the proposed methods was relatively lower than other patterns, although it were better than the baseline methods. According to the confusion tables of proposed methods, it can be found that all misclassified VOIs of GGO were identified to be EMP or NOD. On the HRCT images, the appearance of the GGO is a hazy increased in the pulmonary attenuation. So this kind of opacity is "whiter" than the normal pulmonary parenchyma, but "blacker" than the soft-tissues such as vessels. The existence of other kinds of opacities would affect the recognition of the GGO. Fig. 3.8 and Fig. 3.9 give the example images of GGO which were classified as the EMP and NOD respectively. For the NOD which appears to be of soft-tissue attenuation ("whiter" than the surrounding normal tissues), the recognition of the NOD would be affected by the distribution and extent of "white" components. In the Fig. 3.10, a large area of soft-tissues outside the lung-field was contained in the VOI. So the pulmonary pattern of this VOI was false classified to be the increased pulmonary opacity (GGO). For the VOI given by Fig. 3.11, the reason of false misclassified would be the relatively lower extent of attenuation in the VOI.

By using the classical sparse representation approaches, the K-SVD and OMP algorithm, the SR1 can achieve high recognition rates for the classification of the pulmonary patterns, but the runtime of the SR1 was relatively long. Considering that the dictionary learning and sparse coding spent the most time on the training and testing respectively, another goal of our work was to reduce the runtime of these two stages. For the optimization of the dictionary learning, we used the K-Means to train the dictionary, which could be seen as a simple version of K-SVD. Fig.3.13 compares the SR1 (using the K-SVD) and SR2 (using the K-Means) when the same parameters were used. It can be found that the SR1 and SR2 had similar classification accuracies (SR1: from 90.2% to 95.2% vs. SR2: from 91.8% to 94.2%), and the recognition rates of both two methods revealed a trend of increase with the raising of the number of atoms. The runtime of dictionary learning of the SR1 was nearly 50 times as long as the SR2 when the same number of atoms were used. It is demonstrated that the replacement of K-SVD by the K-Means will not affect the classification results but can considerably decrease the runtime of the dictionary learning.

Although the runtime of recognizing one VOI was seemed not very long, the CAD systems would be used to analyze the whole lungs of patients in the clinical practice, which can be divided into several hundreds of VOIs. So, the time of recognizing a patient will be several hundred times longer than classifying a VOI, and a small reduction of the runtime in the experiments (classify the individual VOIs) is still meaningful which can lead to a remarkably decrease in the actual practice (recognize the whole lungs of patients).

In order to reduce the runtime of the sparse coding, we applied a simple version of OMP. In the OMP, the sparse representation of the example was iteratively calculated, and only one atom was selected in each iteration. We simplified the OMP by selecting the desired number of atoms at one time (OMP_1). Fig. 3.14 compares the SR2 (using the OMP) and SR3 (using the OMP_1) with the same parameters. The recognition rates of the SR3 were similar to the SR2 when the sparsity was set to 2 and 4 (SR2:95.5% and 95.5% vs. SR1:96.4% and 92.8%). And It shows that the SR3 spent shorter time on classifying the VOI than the SR2 (SR2: from 0.36s to 3.6s vs. SR3: from 0.13s to 0.17s). It is demonstrated that the application of the OMP_1 with high sparsity can optimize the runtime of

the sparse representation based method while keeping the classification accuracy. On the other hand, although the runtime of dictionary learning in the SR2 can be shorter than the SR3 (shown in Table 3.10), in the clinical practice the sparse coding is an important step in the recognition of the examples which should be performed on-line, while the dictionary can be trained off-line. Therefore, we thought that the optimization of the sparse coding is more meaningful than the dictionary learning, and it would be more suitable to apply the OMP_1 in the clinical practice.

We compared our methods with the SDF (including SDF-SVM) due to the two reasons. The first one was that the SDF had been successfully applied to classify the most kinds of pulmonary patterns, including normal tissues and six types of DLD patterns. The second one was that the features extracted from the images were directly used as the input vectors of the classifier without a "sparse coding" step in the SDF. However, the overall accuracy of the SDF was worse in the experiments. We thought that the result of the SDF would be affected by the accuracy of detecting geometrical-based components (nodular, linear and multilocular component), which is still a difficult problem in the CAD especially for the images of severe DLD.

The CSE was slightly similar to the SR2 and SR3. Firstly, the K-Means algorithm was adopted in all three methods. Secondly, the signatures of the VOIs, which were used as the input vectors of the classifier, were generated according to the local features. It could be seemed as a "coding" step but not the sparse coding. Therefore, the CSE was used to compare with our methods. Unfortunately the CSE produced a worst result in the experiments. The reason for the bad performance of the CSE may be that the NN classifier is naive comparing to the SVM.

The bag-of-words is a popular model for the image classification, and the bag-of-words based methods have achieved good results in the previous works. The bag-of-words model could be treated as a special version of the sparse representation, which was implemented with an extremely strict constraint on the sparsity. In the bag-of-words, the input examples are assigned to the closet atom in terms of Euclid distance. In other words, only one atom was used to approximate the example, and the coefficient of the selected atom was fixed at 1. The work [42]

thought that this constraint was too restrictive, so it would produce a large reconstruction error. For the sparse coding strategy, the sparsity constraint was relaxed by allowing a small number of atoms to describe the examples. Although more time would cost, the sparse coding approach can achieve a fine reconstruction. Therefore, it can reserve more important information of the examples, which was advantageous for the classification. On the other hand, the experimental data adopted in our experiments was different with the pervious work [105]. It also would affect the classification of the BOW. Table 3.11 compares the overall accuracy and runtime of the SR3 and BOW. The BOW spent little time both on the dictionary learning and recognizing. However, the SR3 achieved a significantly better overall accuracy. Fig. 3.12 shows two example images of the NOR which were correctly classified by the SR3, but falsely recognized as the NOD by the BOW. The reason of the misclassification may be that the appearance of these two VOIs was similar to the NOD. There were many structures with high CT values (in the presence of "whiter" than the normal pulmonary parenchyma) such as vessels in the VOIs.

There were two main limitations in our work. Firstly, the proposed methods have not been tested on the clinical setting. It is better to evaluate the classification of the DLD patterns on the manually identified VOIs combined with the clinical setting. Secondly, the source codes of our projects have not been optimized, which may increase the runtime and require too much memory in the operation. For example, the calculation of eigenvalues of the Hessian matrix in the feature extraction cost nearly 10GB memory. It would limit the operating platform of our method. Therefore, in future research, we will improve our method for the actual clinical practice on the following aspects. The first one is to evaluate the proposed methods on the clinical setting. The second one is to optimize the source codes of our methods. Besides, considering that the quantitative analysis of the agreements between the results of CAD method and each radiologist can be used to demonstrate the objective of the CAD method [33, 111, 112], we will collect more HRCT scans and measure these agreements. And it is necessary to integrate a lung-field algorithm with our method to automatically extract the lung-field as the pre-processing, and adapt our method according to the experimental results on the clinical setting.

3.5 Conclusion

In this chapter, we proposed and optimized a sparse representation based method to classify the DLD patterns. The method using the K-SVD and OMP can achieve a high classification accuracy (greater than 95%) in the experiments, but spend too much time in the operation. So, in order to optimize this method, we applied the K-Means as a substitute of the K-SVD, and simplified the OMP by selecting a sufficient number of atoms at one time (OMP_1). Experimental results showed that the performances of the sparse representation based methods were significantly better than the baseline methods (SR1: 96.1%, SR2: 95.6%, SR3:96.4% vs. SDF: 75.8%, SDF-SVM: 73.6%, CSE: 65.1% and BOW: 85.5%). Furthermore, when the K-SVD was replaced by the K-Means, the runtime of the dictionary learning was reduced by 98.2% (SR1:13520s vs. SR2:241s). And when the OMP_1 was substituted for the OMP, the runtime of recognizing one VOI was dropped by 55.2% (SR2:0.29s vs. SR3:0.13s). So we concluded that the method using the K-Means and OMP_1 (SR3) was efficient to classify the pulmonary patterns.

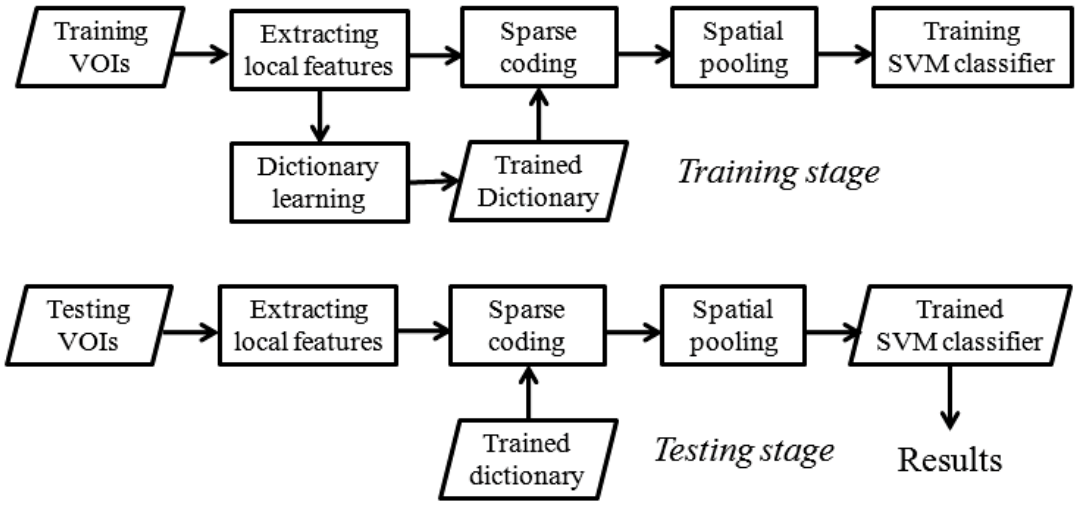


Figure 3.1: The proposed sparse representation based methods

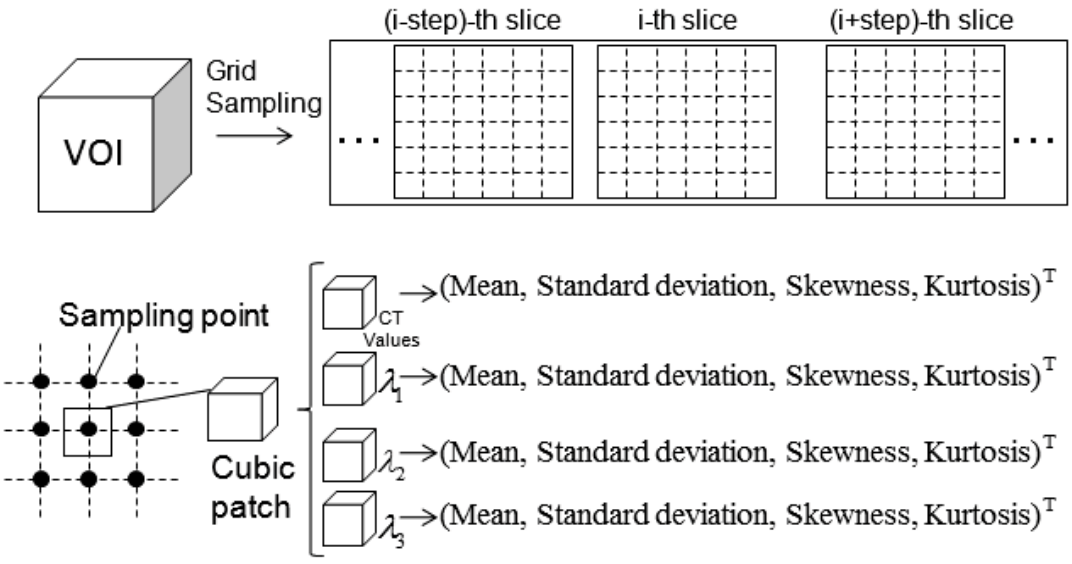


Figure 3.2: Extraction of local features proposed in the work [105]

Input: m examples $\boldsymbol{y} \in \Re^n$ as the columns of the matrix $\boldsymbol{Y} \in \Re^{n \times m}$

Output: The dictionary matrix $\boldsymbol{D} \in \Re^{n \times k}$

1: ***Initialization***: Initial dictionary $\boldsymbol{D}$ with normalized columns by random entries

2: **repeat**

3: ***Sparse Coding Stage***: Defining the sparse representation vector $\boldsymbol{a} \in \Re^k$ as the column of matrix $\boldsymbol{A} \in \Re^{n \times m}$, using any pursuit algorithm to calculate the sparse representation of examples by solving the solution: $\min_{\boldsymbol{a}_i} ||\boldsymbol{y}_i - \boldsymbol{D}\boldsymbol{a}_i||_2^2, 1 \leq i \leq m \;\; subject\ to\ ||\boldsymbol{a}_i||_0 \leq T$

4: ***Dictionary Updating Stage***: Sequentially update the atoms of dictionary, when the i-th atom is updating:

5: Define the group of examples using the atom $\boldsymbol{d}_i$, where the $\boldsymbol{a}_T^i$ represents the i-th row of matrix $\boldsymbol{A} : \omega = \{j | 1 \leq j \leq m, \boldsymbol{a}_T^i(j) \neq 0\}$

6: Compute the reconstruction error matrix $\boldsymbol{E}_i = \boldsymbol{Y} - \boldsymbol{D}\boldsymbol{A} + \boldsymbol{d}_i \boldsymbol{a}_T^i$

7: Restrict the error matrix by choosing the columns corresponding to ω and obtain a new error matrix $\boldsymbol{E}_i^R$

8: Apply singular value decomposition (SVD) to decompose the error matrix and update the atom by the first left-singular vector: $\boldsymbol{E}_i^R = \boldsymbol{U}\boldsymbol{\Delta}\boldsymbol{V}, \boldsymbol{d}_i = \boldsymbol{u}_1$

9: **until** $||\boldsymbol{Y} - \boldsymbol{D}\boldsymbol{A}||_2^2$ is small enough

Figure 3.3: K-SVD algorithm

Input: The example $\boldsymbol{y} \in \mathbb{R}^n$ and the dictionary $\boldsymbol{D} \in \mathbb{R}^{n \times k}, n \ll k$

Output: The sparse representation of example $\boldsymbol{a} \in \mathbb{R}^k$

1: ***Initialization***: Initial residual $\boldsymbol{r}_{(0)} = \boldsymbol{y}$, initial solution support $\boldsymbol{\Omega} = \varnothing$, sparsity constraint T

2: **for** iteration counter $J = 1$, $J < T$, $J++$ **do**

3: Find the atom that has the largest product with the residual $\boldsymbol{r}_{(J-1)}$ and has not been added to the solution support $\boldsymbol{d}_s = \max \boldsymbol{d}_l \cdot \boldsymbol{r}_{(J-1)}, 1 \leq l \leq k, \boldsymbol{d}_l \notin \boldsymbol{\Omega}$

4: Add the selected atom to the solution support $\boldsymbol{\Omega} = \boldsymbol{\Omega} \cup \boldsymbol{d}_s$

5: Compute the sparse representation of example as the projection of example onto the linear space spanned by the elements of solution support and update the residual: $\boldsymbol{a}_{(J)} = (\boldsymbol{\Omega}^T\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^T\boldsymbol{y}, \boldsymbol{r}_{(J)} = \boldsymbol{y} - \boldsymbol{\Omega}\boldsymbol{a}_{(J)}$

6: **if** $||\boldsymbol{r}_{(J)}||_2^2$ is small enough **then**

7: Stop the iteration and output $\boldsymbol{a}_{(J)}$

8: **end if**

9: **end for**

Figure 3.4: Orthogonal-Matching-Pursuit (OMP) algorithm

Input: The example $\boldsymbol{y} \in \mathbb{R}^n$ and the dictionary $\boldsymbol{D} \in \mathbb{R}^{n \times k}, n \ll k$

Output: The sparse representation of example $\boldsymbol{a} \in \mathbb{R}^k$

1: ***Initialization***: Initial solution support $\boldsymbol{\Omega} = \varnothing$, sparsity constraint T

2: ***Computation***:

3: Arrange the inner products of the atoms and example in a descending order, and choose the first T atoms $\boldsymbol{d}_{s1} \cdot \boldsymbol{y} \geq \boldsymbol{d}_{s2} \cdot \boldsymbol{y} \geq ... \geq \boldsymbol{d}_{sT} \cdot \boldsymbol{y} \geq ... \geq \boldsymbol{d}_{sk} \cdot \boldsymbol{y}$

4: Add the selected atoms to the solution support $\boldsymbol{\Omega} = \{\boldsymbol{d}_{s1}, ..., \boldsymbol{d}_{s2}, ..., \boldsymbol{d}_{sT}\}$

5: Compute the sparse representation of example as the projection of example onto the linear space spanned by the elements of solution support: $\boldsymbol{a} = (\boldsymbol{\Omega}^T\boldsymbol{\Omega})^{-1}\boldsymbol{\Omega}^T\boldsymbol{y}$

Figure 3.5: OMP_1 algorithm

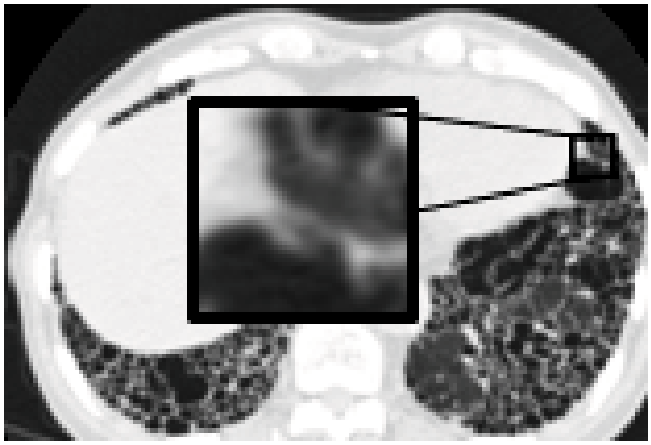


Figure 3.6: Image of good classified by our methods, True:HCM, Estimated:HCM

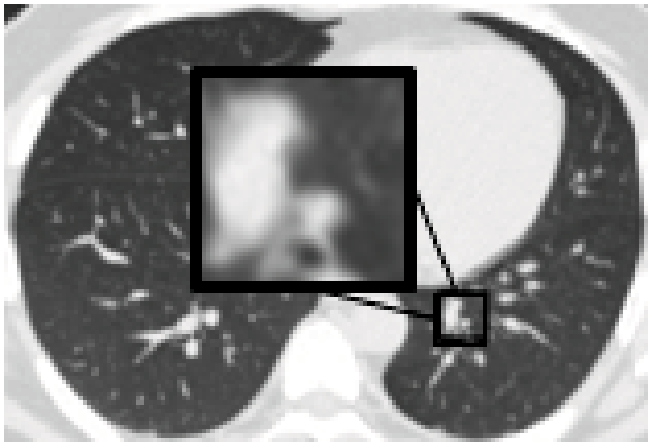


Figure 3.7: Image of good classified by our methods, True:NOR, Estimated:NOR

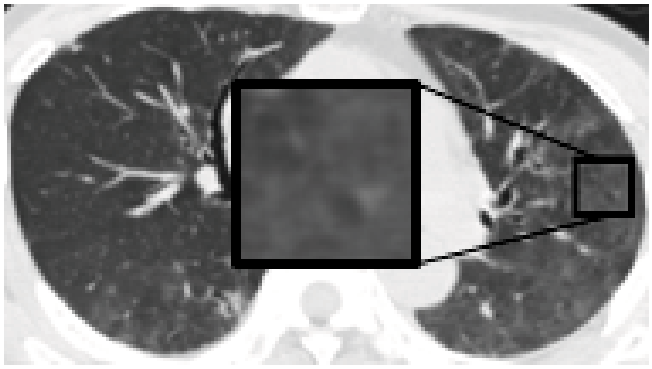


Figure 3.8: Image of bad classified by our methods, True:GGO, Estimated:EMP

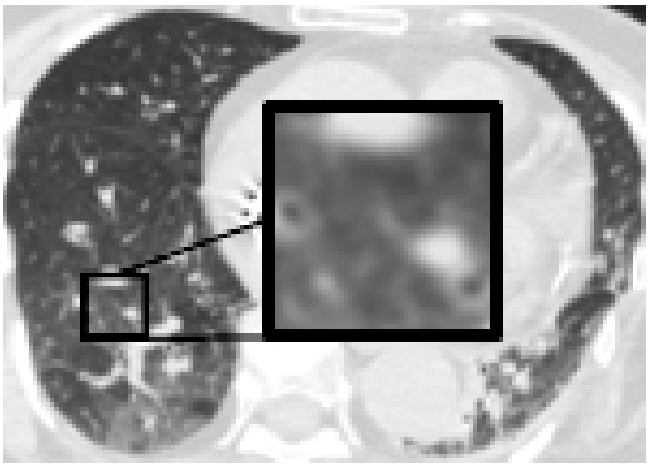


Figure 3.9: Image of bad classified by our methods, True:GGO, Estimated:NOD

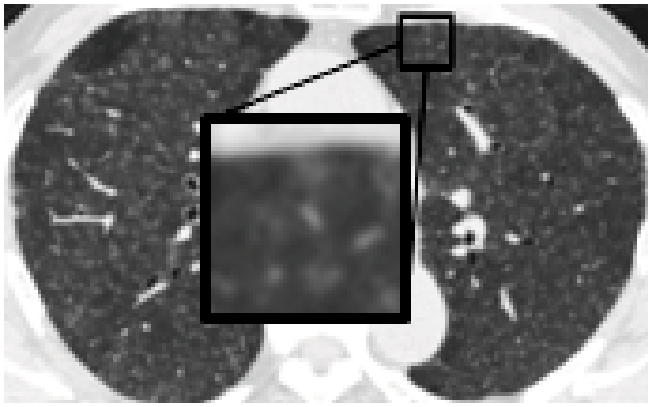


Figure 3.10: Image of bad classified by our methods, True:NOD, Estimated:GGO

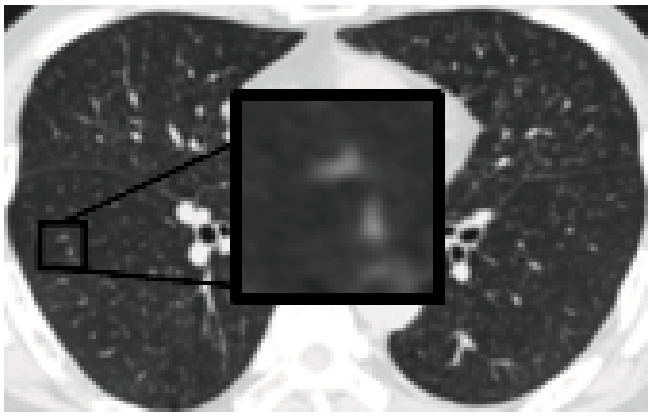


Figure 3.11: Image of bad classified by our methods, True:NOD, Estimated:NOR

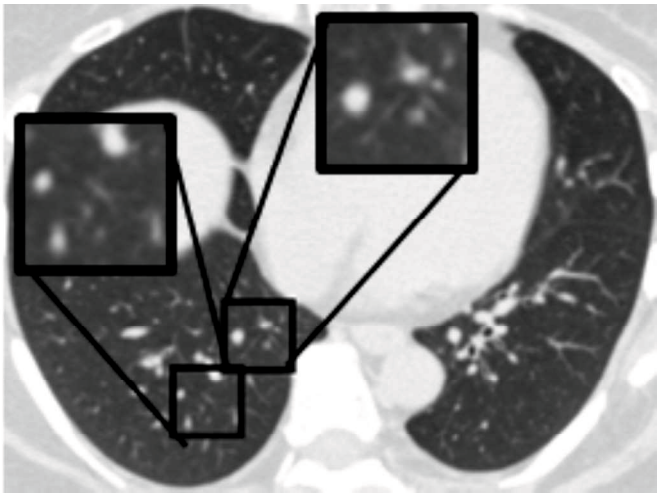
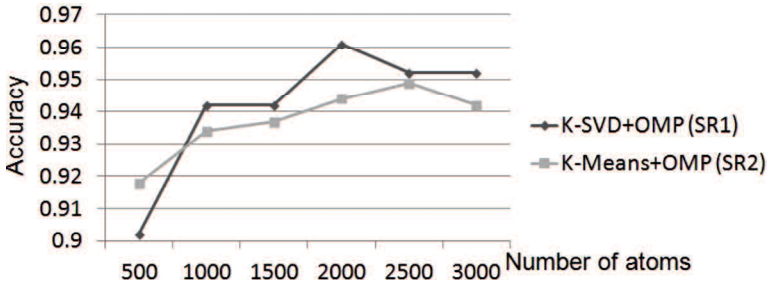
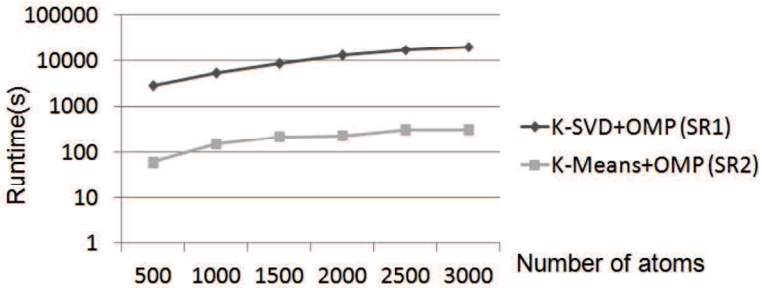


Figure 3.12: Image of bad classified by the BOW, True:NOR, Estimated:NOD

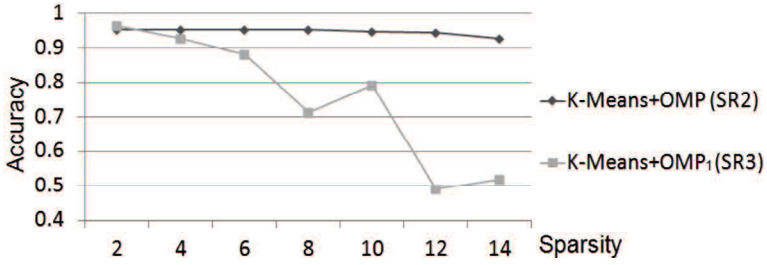


(a) Classification accuracies

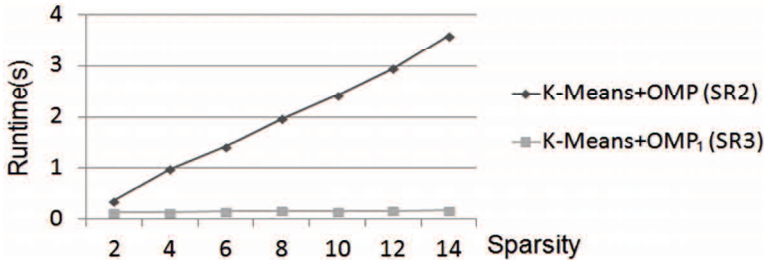


(b) Runtime of the dictionary learning

Figure 3.13: The classification accuracies and runtime of dictionary learning of SR1 (K-SVD+OMP) and SR2 (K-Means+OMP) with the patch size:$4 \times 4 \times 4$ and number of non-zero entries:8



(a) Classification accuracies



(b) Time to recognize one VOI

Figure 3.14: The classification accuracies and runtime to recognize one VOI (except feature calculation) of SR2 (K-Means+OMP) and SR3 (K-Means+OMP_1) with the patch size:$3 \times 3 \times 3$ and the number of atoms:3000.

Table 3.1: Three proposed method to evaluate the sparse representation approaches

(a) The proposed methods

	Dictionary learning	Sparse coding
SR1	K-SVD	OMP
SR2	K-Means	OMP
SR3	K-Means	OMP_1

(b) Evaluation of sparse representation approaches

Evaluation	Comparison of methods
Dictionary learning	SR1 vs. SR2
Sparse coding	SR2 vs. SR3

Table 3.2: Number of VOIs in the training and testing sets

	CON	GGO	HCM	EMP	NOD	NOR	Total
Training set	49	170	221	323	113	285	1161
Testing set	45	160	204	275	92	273	1049

Table 3.3: Comparison on the overall accuracy and statistical difference for the proposed methods and baseline methods

(a) Overall accuracy of all methods

Methods	Overall accuracies
SR1	96.1%
SR2	95.6%
SR3	96.4%
SDF / SDF-SVM	75.8% / 73.6%
CSE	65.1%
BOW	85.5%

(b) P values of statistical differences

	SR1	SR2	SR3	SDF	SDF-SVM	CSE	BOW
SR1 vs.	——	0.57	0.75	<0.0001	<0.0001	<0.0001	<0.0001
SR2 vs.	0.57	——	0.31	<0.0001	<0.0001	<0.0001	<0.0001
SR3 vs.	0.75	0.31	——	<0.0001	<0.0001	<0.0001	<0.0001
SDF vs.	<0.0001	<0.0001	<0.0001	——	0.09	<0.0001	<0.0001
SDF-SVM vs.	<0.0001	<0.0001	<0.0001	0.09	——	<0.0001	<0.0001
CSE vs.	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	——	<0.0001
BOW vs.	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	——

Table 3.4: Confusion table of SR1, overall accuracy:96.1%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	45	0	0	0	0	0	100.0%
GGO	0	142	0	15	3	0	88.8%
HCM	0	0	203	1	0	0	99.5%
EMP	0	1	0	270	0	4	98.2%
NOD	0	3	0	0	83	6	90.2%
NOR	0	0	0	6	2	265	97.1%
Precision	100.0%	97.3%	100.0%	92.5%	94.3%	96.4%	

Table 3.5: Confusion table of SR2, overall accuracy:95.6%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	45	0	0	0	0	0	100.0%
GGO	0	143	0	9	8	0	89.4%
HCM	0	1	203	0	0	0	99.5%
EMP	1	0	2	263	0	9	95.6%
NOD	0	4	0	0	83	5	90.2%
NOR	0	2	0	0	5	266	97.4%
Precision	97.8%	95.3%	99.0%	96.7%	86.5%	95.0%	

Table 3.6: Confusion table of SR3, overall accuracy:96.3%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	45	0	0	0	0	0	100.0%
GGO	0	152	0	7	1	0	95.0%
HCM	0	1	203	0	0	0	99.5%
EMP	0	1	1	268	0	5	97.5%
NOD	0	7	0	0	80	5	87.0%
NOR	0	3	0	5	2	263	96.3%
Precision	100.0%	92.7%	99.5%	95.7%	96.4%	96.3%	

Table 3.7: Confusion table of SDF

(a) Original SDF (using ANN classifier), overall accuracy:75.8%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sentivity
CON	45	0	0	0	0	0	100.0%
GGO	1	157	0	0	0	2	98.1%
HCM	0	14	153	15	8	14	75.0%
EMP	0	0	0	249	7	19	90.5%
NOD	0	31	0	7	41	13	44.6%
NOR	0	80	3	28	12	150	54.9%
Precision	97.8%	55.7%	98.1%	83.3%	60.3%	75.8%	

(b) SDF-SVM,overall accuracy:73.6%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sentivity
CON	45	0	0	0	0	0	100.0%
GGO	0	159	0	0	0	1	99.4%
HCM	0	4	174	7	5	14	85.3%
EMP	0	0	5	188	23	59	68.4%
NOD	0	28	1	0	59	4	64.1%
NOR	0	97	2	5	22	147	53.8%
Precision	100.0%	55.2%	95.6%	94.0%	54.1%	65.3%	

Table 3.8: Confusion table of CSE,overall accuracy:65.1%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	45	0	0	0	0	0	100.0%
GGO	10	94	28	0	28	0	58.8%
HCM	10	29	133	5	27	0	65.2%
EMP	0	0	23	216	16	20	78.5%
NOD	0	2	0	3	61	26	66.3%
NOR	0	17	20	6	96	134	49.1%
Precision	69.2%	66.2%	65.2%	93.9%	26.8%	74.4%	

Table 3.9: Confusion table of BOW,overall accuracy:85.5%

	Estimated Label						
True Label	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	44	0	1	0	0	0	97.8%
GGO	0	160	28	0	28	0	100.0 %
HCM	0	2	201	0	1	0	98.5%
EMP	0	0	7	262	0	6	95.3%
NOD	0	29	0	1	56	6	60.9%
NOR	0	63	0	8	28	174	63.8%
Precision	100.0%	63.0%	84.8%	96.7%	49.6%	93.5%	

Table 3.10: Comparison on runtime of proposed methods

Methods	Time of dictionary learning	Time to recognize one VOI (except feature calculation)
SR1	13520s (using K-SVD)	1.27s (using OMP)
SR2	241s (using K-Means)	0.29s (using OMP)
SR3	350s (using K-Means)	0.13s (using OMP_1)

Table 3.11: Comparison of SR3 and BOW

Methods	Overall accuracy	Time of dictionary learning	Time to recognize one VOI (except feature calculation)
SR3	96.4%	350s (using K-Means)	0.13s
BOW	85.5%	241s (using K-Means)	0.013s

Chapter 4

Classification of Pneumoconiosis Images

4.1 Introduction

According to the Japan Pneumoconiosis Law, the pneumoconiosis is divided into 4 types based on the density and diameter of nodular opacities: Type 1 (no nodules), Type 2 (few small-sized nodules (diameter $\leq$ 10mm)), Type 3-a(numerous small-sized nodules) and Type 3-b (numerous small nodules and presence of large-sized nodules (diameter>10mm)). Fig. 4.1 illustrates the example images of 4 types of the pneumoconiosis.

The current diagnosis criterion of the pneumoconiosis is made based on the CXR images, but the HRCT image is thought to be more sensitive than CXR images for the diagnosis. For example, the superimposition of the large-sized nodules on the CXR images would lead to a misdiagnosis for the cases of Type 3-b. Therefore, we designed a method to classify the four categories of the pneumoconiosis on HRCT images. According to our knowledge, there is no work to classify the pneumoconiosis on HRCT images.

Although the performance of the bag-of-features method on the classification of the six kinds of pulmonary patterns would be limited (see Chapter 3), it achieved a good result when few number of pulmonary patterns were classified [113, 114]. Considering that the classification of the pneumoconiosis can be seemed as the classification of two pulmonary patterns: NOR and NOD, and the

bag-of-feature method can save the runtime, we adopted this method to classify the pneumoconiosis. In additional, we design a novel nodule filter to detect the pneumoconiotic nodules. Comparing with the conventional filters, more pneumoconiotic nodules can be extracted by using the proposed filter.

4.1.1 Nodule Extraction Method

The pneumoconiosis is diagnosed according to the size and density of the nodular opacities. So, it is essential to detect pneumoconiotic nodules for the classification. Researchers have proposed some methods to detect the nodules on the HRCT images. Considering that the intensity of nodules is higher than its surrounding normal tissues, many intensity-based methods have been developed, such as local maximum density algorithm [115], local maximal gradient strength [116] and morphological approaches [117]. However, most pneumoconiotic nodules are low-contrast to vessels in intensity, and many nodules are connected to vessels and airways. It is difficult to select suitable thresholds.

Researchers also designed several model-based methods to extract pulmonary nodules. Osman [118] designed a three-dimensional nodule template to detect nodule candidates by convolving regions of interest (ROIs) with the proposed template. Farag [119] used four different types of deformable templates to describe typical geometry and gray-level distribution of pulmonary nodules. Felzenszwalb [120] proposed an object detection method based on mixtures of multiscale deformable part models. Tbellotti [121] combined an active contour model with region growing method to extract pulmonary nodules on CT images.

Besides the model-based methods, the methods based on the eigenvalues of Hessian matrix are another popular choice to detect the pulmonary nodules. It is reported that the conditions of eigenvalues of the Hessian matrix can be used to indicate the shape of objects [122]. Table 4.1 summarizes the conditions of eigenvalues of the structures with three typical kinds of shapes: nodules, vessels and cortex. Researchers have proposed some nodule filters [122, 123] to enhance the nodule candidates at first, and then extract the nodules by thresholding the response of the filter.

The conventional methods devote to improve the accuracy of nodule extraction, in other words, the ability to distinguish nodules from other objects. So, the large-sized pneumoconiotic nodules can be well detected, but most small-sized pneumoconiotic nodules with a irregular shape may be failed detected. It would affect the classification of Type 1,2 and 3-a because the diagnosis of these three types of cases is according to the density of small-sized nodules. In order to improve the performance of the classification, we designed a novel filter to detect nodule candidates.

4.2 Proposed Methods

Considering that the pneumoconiosis are diagnosed according to the size (Type 3-b and other types) and density of nodules (Type 1,2,and 3-a), we recognized the cases of type 3-b at first, and then classified the cases of type 1,2 and 3-a. Fig. 4.2 illustrates the framework of proposed method. Firstly, the large-sized nodules were extracted on all data, and the cases of Type 3-b and non-3-b were classified. Secondly, the small-sized nodules were extracted from the cases of type non-3-b; Thirdly, the local features were calculated on the patches located on the centers of detected nodules. Finally, the bag-of-features was used to generate the image-level descriptors as the input vectors of a SVM classifier, and the cases of type 1,2 and 3-a were classified.

4.2.1 Nodule Extraction

In the research, we designed two filters based on the eigenvalues of the Hessian matrix to detect nodule candidates as many as possible at first, and then carefully eliminated false positives. Let the eigenvalues be λ_1, λ_2 and $\lambda_3, \lambda_3 \leq \lambda_2 \leq \lambda_1$, considering that the three eigenvalues are approximately equal in the nodules' center ($\lambda_1 \approx \lambda_2 \approx \lambda_3$), we adopted an monotonically increasing function to represent the difference between the eigenvalues λ_i and λ_j,

$$f(\lambda_i, \lambda_j) = \frac{2(\lambda_i/\lambda_j)}{1 + (\lambda_i/\lambda_j)^2} \tag{4.1}$$

When $\lambda_i = \lambda_j$, the function comes to its maximum. Our nodule filter was defined as the product of differences between (λ_1, λ_2) and (λ_2, λ_3)

$$f_{blob} = f(\lambda_1, \lambda_2) \times f(\lambda_2, \lambda_3) = \begin{cases} \frac{4(\lambda_1/\lambda_3)}{1+(\frac{\lambda_1}{\lambda_2})^2+(\frac{\lambda_2}{\lambda_3})^2+(\frac{\lambda_1}{\lambda_3})^2}, & \lambda_3 \leq \lambda_2 \leq \lambda_1 < 0 \\ 0, & otherwise \end{cases} \tag{4.2}$$

Fig. 4.3a gives the response of our nodule filter. The less the difference between eigenvalues, the greater will be response. The proposed nodule filter can produce the maximum response on the voxels located in the centers of nodules. In other words, the response reached the maximum where the eigenvalues meet the condition $\lambda_1 \approx \lambda_2 \approx \lambda_3$. And the response of the filter descends with the increased difference between the eigenvalues.

In order to enhance the nodules with various scales, the nodule filter was convolved with Gaussian kernels. By adjusting the standard deviation of Gaussian convolution, the nodules with specific range of sizes can be enhanced. Fig. 4.3b illustrates the color-code result of nodule enhancement by the proposed nodule filter. The values of the standard variation were fixed to 0.6, 1, 1.5 and 2 voxels. When the response increases, the color will gradually change from blue by yellow to red. Fig. 4.3c shows the result of nodule extraction after thresholding. This threshold is called nodular threshold in this thesis. It can be found that a large number of nodule candidates can be extracted, and our nodule filter had a good sensitivity on detecting small-sized nodules. Fig. 4.3d shows the result of enhancing the large-sized nodules when the scales were selected as 4 and 5 voxels.

It is inevitable that there were numerous false positives in the result of extracting small-sized nodules, for example, the vascular intersections whose condition of eigenvalues is similar with nodules. We designed a vessel filter to remove the false positives meanwhile preserving nodules as many as possible. Since the condition of eigenvalues in vessels' centers is $\lambda_3 \approx \lambda_2 \ll \lambda_1 \approx 0$ [122], the vessel filter was defined as:

$$f_{line} = \begin{cases} exp[-(\frac{\lambda_1}{\lambda_2})^{1/2}] \times \frac{2(\frac{\lambda_2}{\lambda_3})}{1+(\frac{\lambda_2}{\lambda_3})^2}, & \lambda_3 \leq \lambda_2 < 0, \lambda_3 \leq \lambda_2 \leq \lambda_1 \\ 0, & otherwise \end{cases} \tag{4.3}$$

Fig. 4.3e shows the response of the proposed vessel filter. The proposed filter can produce the maximum on the voxels located in the center line of the vessels.

After enhancing the small-sized nodules and vessels, the false positives were removed as follows. 1.Thresholding the response of nodule filter and vessel filter respectively. 2.Removing the objects which exist in the both of two results. The threshold of proposed vessel filter is called vascular threshold in this thesis. By using the proposed filters, the result of extracting small nodules was determined by the two parameters: nodular threshold and vascular threshold. Fig. 4.3f illustrates the result of subtracting false positives from Fig. 4.3c with a vascular threshold of 0.7. The removed false positives are given in Fig. 4.3g. In the experiments, the nodular threshold was adjusted from 0.35 to 0.55 with an interval of 0.05, and the vascular threshold was adjusted from 0.6 to 0.8 with an interval of 0.05. Considering that the scales of most false positives were less than the large-sized nodules, the detected large-sized nodules were directly adopted in the experiments without removing false positives.

4.2.2 Calculation of Local Features

After extracting the small-sized nodules, we calculated a set of 21 features based on three kinds of measures on the patches whose centers located on the centers of the detected nodules. The first kind of local feature was the number of voxels belonging to nodules within the patches. Considering that it may be several parts of nodules rather than entire nodules existed in the patches, the number of voxels was more suitable to indicate the density of nodules in the local regions. The other two kinds of local features were the measures based on the GLCM [22] and GLRLM [21] respectively. From Fig. 4.1, it is clearly that the density of nodules can affect the textures of pulmonary tissues. The more nodules exist, the greater extent the tissues changes. Therefore, we adopted the statistical texture features to discriminate various pneumoconiotic patterns. Table 4.2 and Table 4.3 summarize the measures based on the GLCM and GLRLM respectively. The size of patches was a parameter in the calculation of local features. We thought that the relatively large-sized patches would be better to indicate the information of nodular intensity. So it was adjusted from $15 \times 15 \times 15$ to $40 \times 40 \times 40$ with an interval of 5 voxels in the experiments.

4.2.3 Bag-of-features Based Classification

We used the bag-of-features to calculate the image-level descriptors as the input vectors of a SVM classifier. Fig. 4.4 illustrates the framework of bag-of-features based classification. After calculating huge number of local features from the input data, the local features were aggregated into limit number of the clusters, and the centers of clusters were saved as the words of a code-book (atoms of the dictionary). By using the code-book, each local feature was assigned to the closet atom in the term of Euclidean distance, and the histograms indicated the frequencies of the words were generated. These histograms were used as the input vectors of the classifier.

In the experiment, we adopted the K-Means algorithm to aggregate the local features and calculate the centers of the clusters. The centers of the clusters were saved as the atoms of the dictionary. The number of clusters in the K-means was a parameter of the bag-of-features, which was adjusted from 50 to 200 with an interval of 25.

For the SVM classifier, we used a version named LIBSVM with RBF (Radial basis function) kernel [108]. The RBF kernel is defined as:

$$RBF\ kernel: \quad G(h, h^{'}) = exp(-\gamma||h - h^{'}||^2) \tag{4.4}$$

where h and $h^{'}$ are both the N-bin histogram. The SVM classifier has two parameters, the soft-margin penalty C and the parameter related to the RBF kernel γ. The values of the C and γ were set to be $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$ and $2^{-10}, 2^{-9}, ..., 2^{1}$ respectively.

4.3 Experiment

4.3.1 Data

We collected 175 HRCT scans from 112 different subjects from Kochi University to evaluate the performance of the proposed method. All scans were aquired by GE Lightspeed VCT when edge-enhanced filtering was not applied. The resolution of HRCT scans was 512×512. The slice thickness was 1mm, and the in-plane resolution was 0.625mm. A tube voltage of 140kV and a tube current of 250mA

were used. The scans were reviewed by experienced radiologists and divided into 4 types.

4.3.2 Experimental Setting

We separated the HRCT scans into two independent sets by splitting each type of pneumoconiosis nearly in half. The training set (90 cases) was adopted to optimize the parameters and train the algorithms with the optimized parameters, and the test set (85 cases) was used to evaluate performance of the method. The number of scans of each type of pneumoconiosis for training and test are summarized in Table 4.4. There was no crossing-subject scan between the two sets.

There were six parameters in the proposed method: nodular threshold, vascular threshold, size of patch, number of clusters, the soft-margin penalty C and the parameter related to the RBF kernel γ. When the cases of type 3-b was recognized, we set the nodular threshold and vascular threshold to 0.01 and 0.99 respectively. In the classification of type 1,2 and 3-a, all parameters were simultaneously adjusted. For the nodular threshold and vascular threshold, we tried the values from 0.35 to 0.55 and 0.6 to 0.8 respectively. The intervals of two thresholds were both 0.05. The path size was adjusted from 15×15×15 to 40×40×40 with a step of 5×5×5. On the choice of number of clusters, we tried the number from 50 to 200 with an interval of 25. The values of the C and γ were set to be $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$ and $2^{-10}, 2^{-9}, ..., 2^{1}$ respectively.

We used leave-one-out to optimize parameters in the training stage. The combination of possible values which produced the best classification accuracy was chosen as the optimal parameter. Then each method was trained with the optimal parameters on training set, and finally evaluated on the testing set.

4.3.3 Baseline methods

Sato et al. [122] and Li et al. [123] have proposed two kinds of state-of-the-art nodule filters based on the eigenvalues of the Hessian matrix respectively. These

two filters are given by Eq.4.5 and Eq.4.6 respectively.

$$Sato's\ filter: \quad S_{blob} = \begin{cases} |\lambda_3|(\frac{\lambda_2}{\lambda_3})^{\gamma}(\frac{\lambda_1}{\lambda_2})^{\gamma}, & \lambda_3 \leq \lambda_2 \leq \lambda_1 < 0 \\ 0, & otherwise \end{cases} \tag{4.5}$$

$$Li's\ filter: \quad L_{blob} = \begin{cases} |\lambda_1|^2/|\lambda_3|, & \lambda_3 \leq \lambda_2 \leq \lambda_1 < 0 \\ 0, & otherwise \end{cases} \tag{4.6}$$

where γ is a parameter of Sato's filter. The paraments of the baseline filters were optimized in a same way as our method.

4.3.4 Results

The overall accuracies of all three methods are given in Table 4.5. The proposed method achieved a recognition rate of 90.6%, better than the baseline filters(Sato's filter:80.0%, Li's filter:76.5%). Table 4.9 shows that the p values of the statistical difference were less than 0.05, which means that there were significantly statistical differences between the methods. The confusion tables of the methods are summarized in Table 4.6 (our method), Table 4.7 (Sato's filter) and Table 4.8 (Li's filter) respectively.

4.4 Discussion

According to the study [123], the quality of the nodule filter can be evaluated from the two aspects: sensitivity and specificity. The sensitivity means that the nodule filter should give a strong response to blob-like objects, and the specificity indicates that the nodule filter should not produce strong response to other objects, such as line-like objects. Because the pneumoconiosis is featured by large numbers of nodular opacities, a filter with good sensitivity is required to detect pneumoconiotic nodules as many as possible for the classification. However, the conventional filters aimed to improve the discriminatory power(specificity) rather than detecting more nodule candidates(sensitivity).

The Sato's filter can be simplified as $-\lambda_1$ when γ is fixed to 1, one of the preferred values. This filter had a good performance on discriminating nodules from other tissues, but the response of tiny nodules was as low as the false positives. Fig. 4.5a and Fig. 4.5b illustrate the results of nodule enhancement by

using the Sato's filter and the extracted nodules by thresholding of the result of nodule enhancement. It can be found that numerous nodules were removed after thresholding. Comparing with the Sato's filter, the Li's filter had a better specificity but worse sensitivity on detecting small-sized nodules. Fig. 4.5c and Fig. 4.5d show the result of nodule enhancement and its thresholding by Li's filter respectively. The number of nodules was further decreased after thresholding.

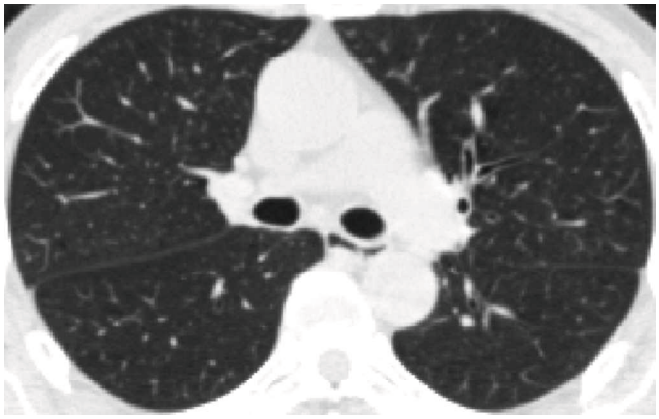
We compared the two kinds of the baseline filters in Fig. 4.5f. It is clearly that the response of Sato's filter descended slower than Li's filter. So the small-sized nodules can be given a stronger response whose magnitude of eigenvalues was relative small, and greater number of nodules can be kept after thresholding. It can be deduced that the slower descending response was advantageous to improve the sensitivity of nodule filter.

In order to extract more pneumoconiotic nodules, we proposed a novel nodule filter in the study. Because all eigenvalues are approximately equal in the nodules' center, we adopted the product of differences between (λ_1, λ_2) and (λ_2, λ_3) to indicate the shape of objects. Although this product was also considered in the Sato's filter, it was simplified and not used in the computation. We adopted an monotonically increasing function to define the proposed nodule filter which can avoid the simplification of the eigenvalues. Fig. 4.5g compares the response between the proposed filter and Sato's filter. It shows that the decreasing of response of proposed filter was slower than the Sato's filter. So more nodules can be extracted, as be shown in Fig. 4.5e. Experimental results demonstrate that increasing extracted nodules was advantageous for the classification of the Type 2 and Type 3-a which are featured by the density of the small-sized nodules. The sensitivities of Type 2 and Type 3-a by using all filters were 91.7% \ 92.9% (ours), 25.0% \ 85.7% (Sato's) and 41.7% \ 78.6% (Li's) respectively.

4.5 Conclusion

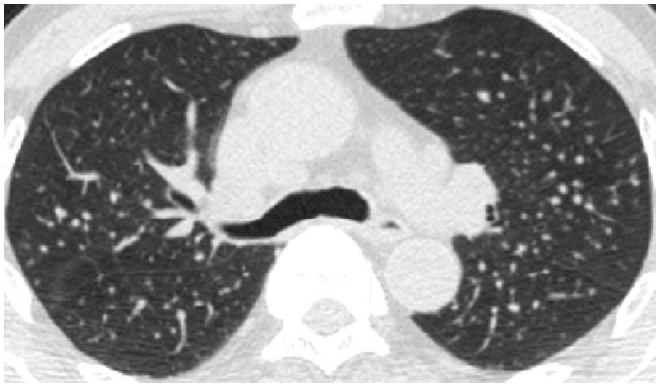
In this chapter, we proposed a method to classify the four kinds of the pneumoconiosis on the HRCT images. Considering that main finding of the pneumoconiosis is the nodular opacities, but the performance of the conventional nodule extraction approaches would be limited for the detection of the small-sized and

irregular nodules, we designed two eigenvalues based filters to detect the nodules and then remove the false positives while keeping the nodules as many as possible. A total of 21 kinds of measures was calculated on the patches located on the centers of the detected nodules. The image-level descriptor of each image was generated by using the bag-of-features and employed as the input vector of the SVM classifier with RBF kernel. The overall accuracy by using our filter was 90.6%, better than the two kinds of baseline filters (Sato's filter: 80.0%, Li' filter: 76.5%). We thought that the proposed method would be useful for the CAD of the pneumoconiosis.

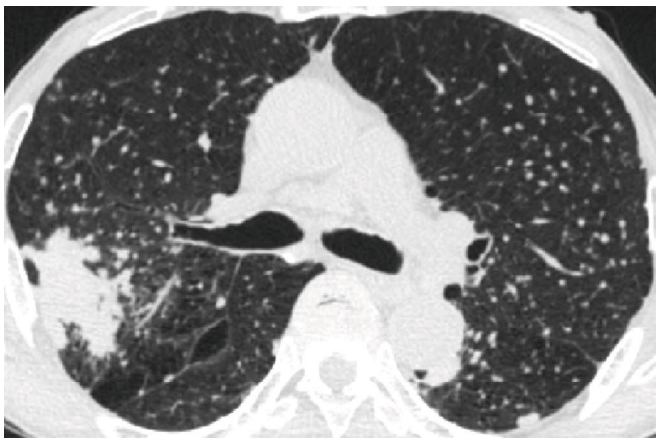


(a) Type 1 (absence of nodule)

(b) Type 2 (few small nodules)



(c) Type 3-a (numerous small nodules)



(d) Type 3-b (numerous small nodules and presence of larger nodules)

Figure 4.1: HRCT images of pneumoconiosis

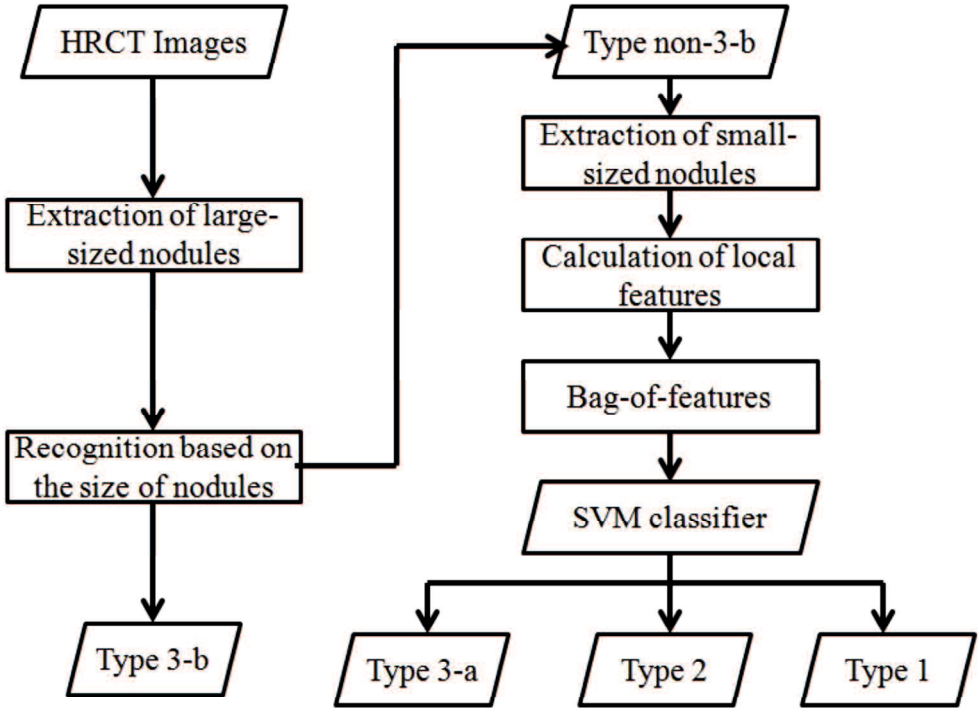
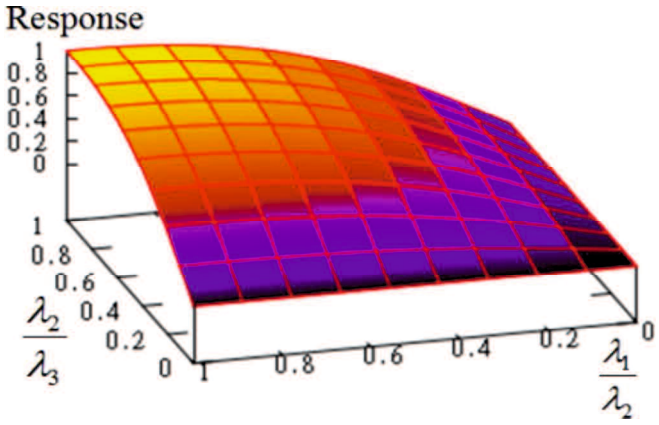
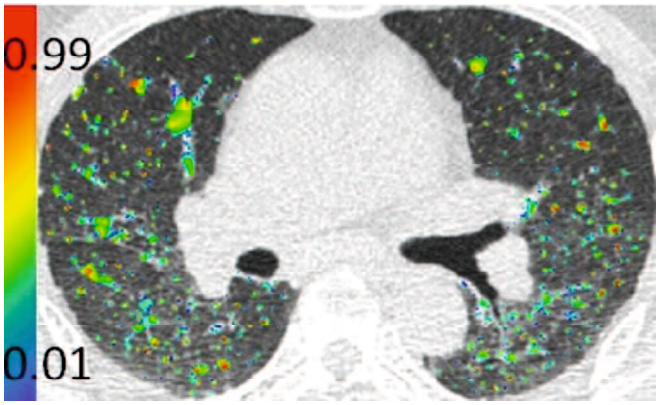


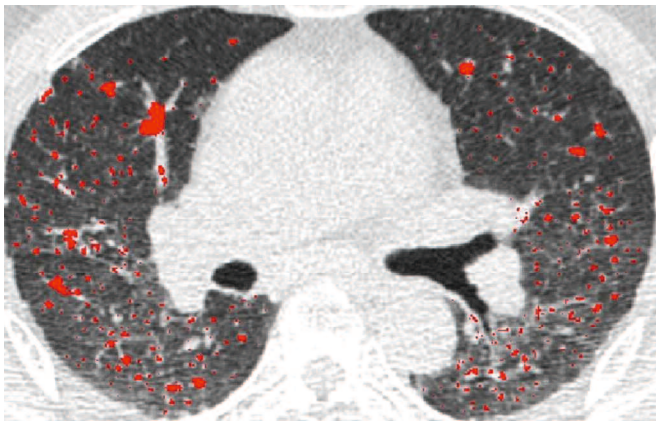
Figure 4.2: Overall scheme of proposed CAD method



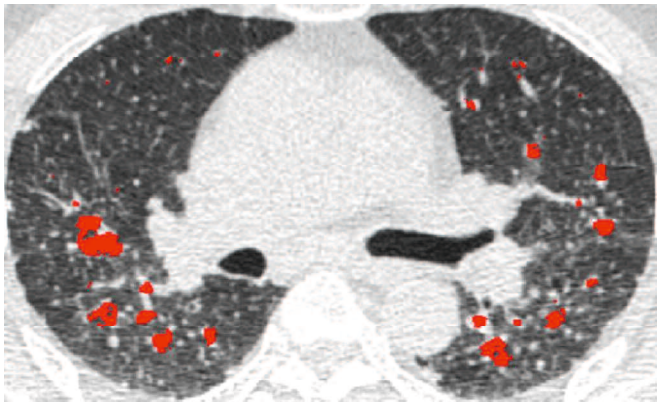
(a) Response of nodule filter



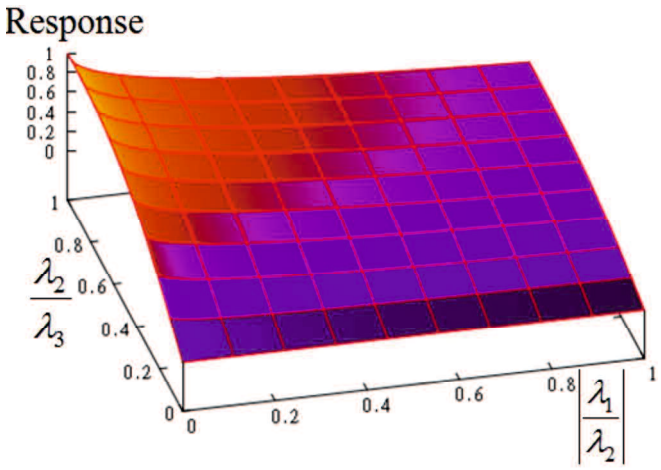
(b) Multiscale enhancement of small nodules with scales of 0.6, 1, 1.5, and 2 voxels



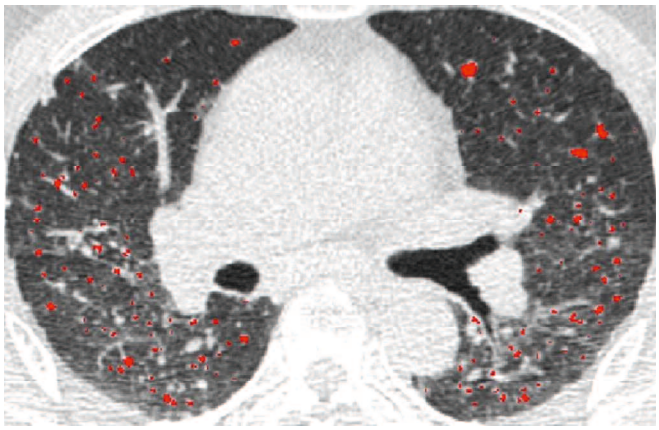
(c) Thresholding of the nodule enhancement



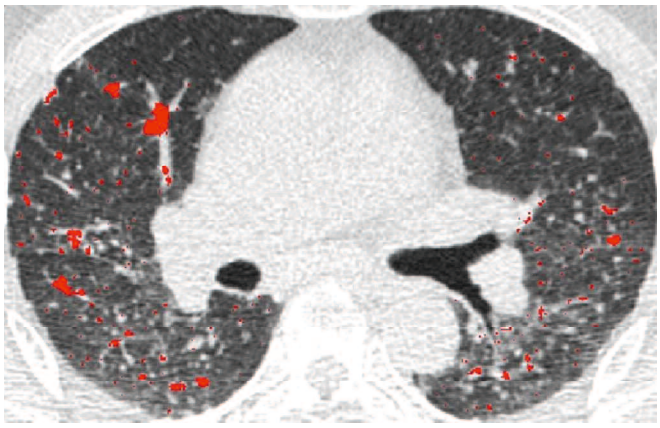
(d) Multiscale enhancement of large-sized nodules with scales of 4 and 5 voxels



(e) Response of our vessel filter

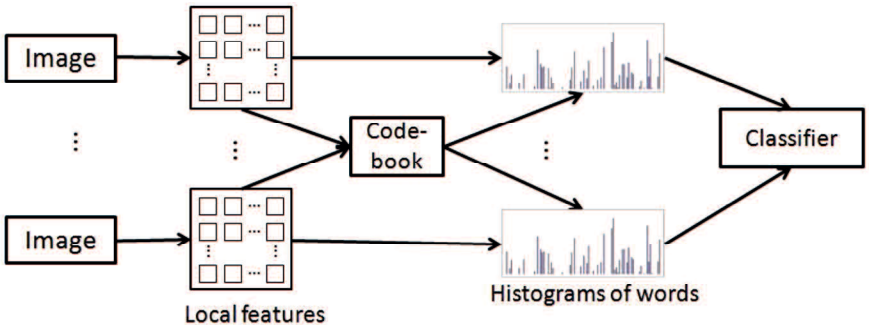


(f) Result of subtracting false positives

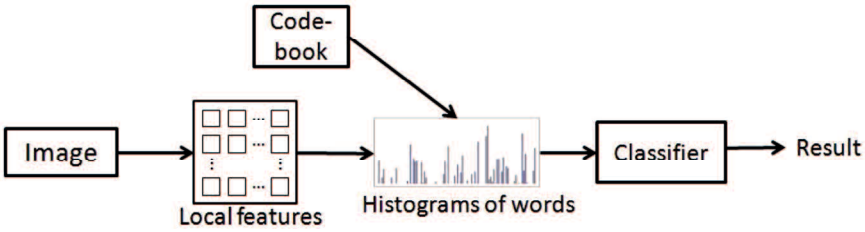


(g) Removed false positives

Figure 4.3: Result of nodule extraction by our filter

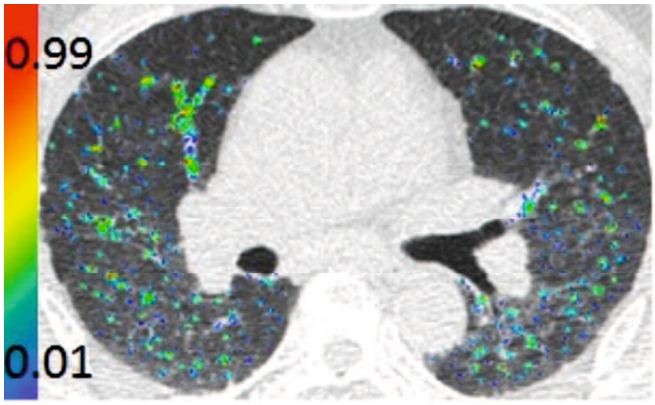


(a) Training

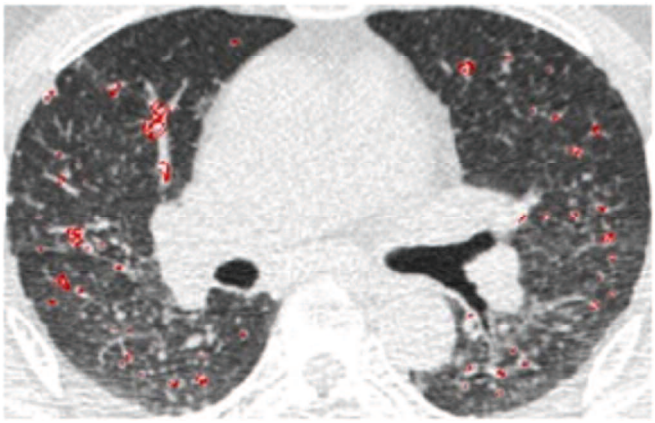


(b) Test

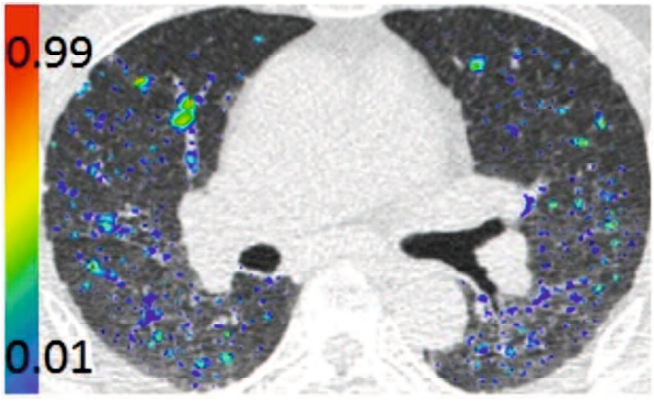
Figure 4.4: Bag-of-features



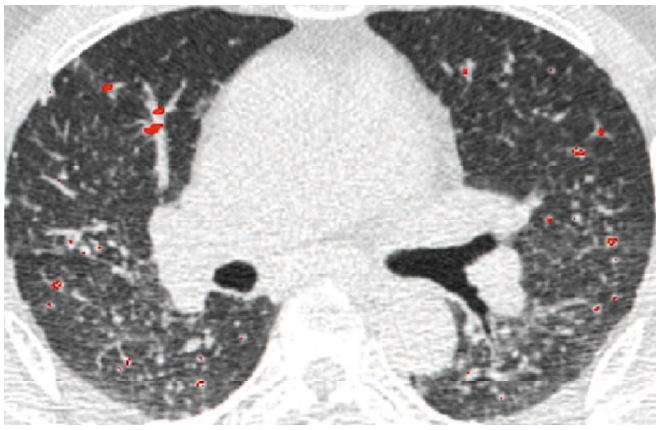
(a) Enhancement of small nodules by Sato's filter



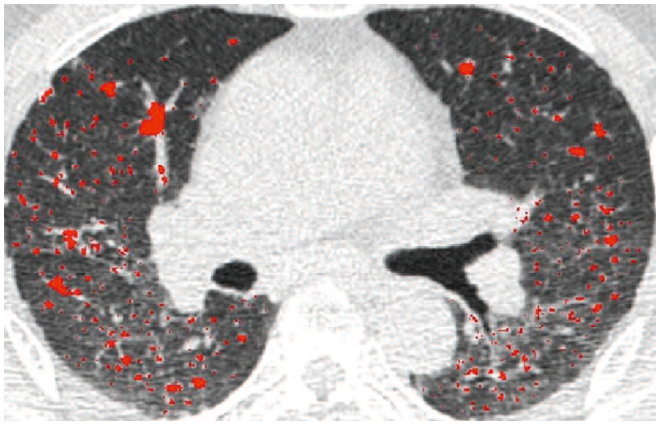
(b) Thresholding of (a) with a nodular threshold of 0.5



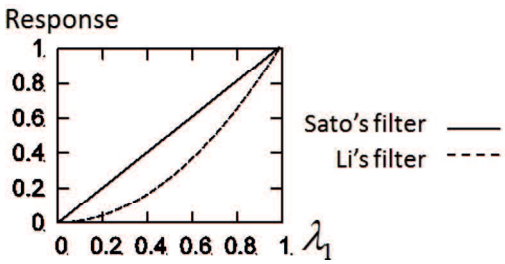
(c) Enhancement of small nodules by Li's filter



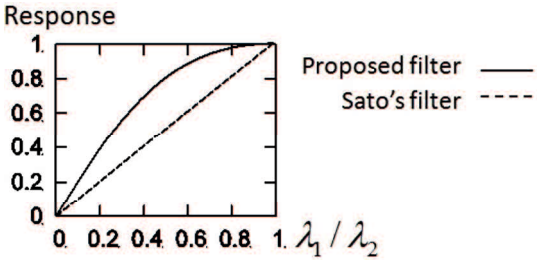
(d) Thresholding of (c) with a nodular threshold of 0.5



(e) Result of nodule extraction by proposed filter with a nodular threshold of 0.5



(f) Comparison of the response between Sato's and Li's filter



(g) Comparison of response between proposed and Sato's filter

Figure 4.5: Comparison of three filters

Table 4.1: Eigenvalue condition of local structures, $\lambda_1, \lambda_2, \lambda_3$, represent the eigenvalues of Hessian matrix [122]

Structures	Eigenvalue condition
Nodule	$\lambda_3 \approx \lambda_2 \approx \lambda_1 \ll 0$
Vessel	$\lambda_3 \approx \lambda_2 \ll \lambda_1 \approx 0$
Cortex	$\lambda_3 \ll \lambda_2 \approx \lambda_1 \approx 0$

Table 4.2: The features based on gray level co-occurrence matrix (GLCM)

Entropy
Energy
Contrast
Homogeneity
Mean
Variance
Correlation
Maximum Probability
Inverse Difference
Cluster Tendency

Table 4.3: The features based on run-length matrix (GLRLM)

Short Run Emphasis
Long Run Emphasis
Low Gray-level Run Emphasis
High Gray-level Run Emphasis
Short Run Low Gray-level Run Emphasis
Short Run High Gray-level Run Emphasis
Long Run Low Gray-level Run Emphasis
Long Run High Gray-level Run Maximum Emphasis
Gray-Level Non-uniformity
Run Length Non-uniformity

Table 4.4: Number of scans for the four types of pneumoconiosis cases

	Type 1	Type 2	Type 3-a	Type 3-b	Total
Training set	23	12	17	38	90
Test set	23	12	14	36	85

Table 4.5: Comparison of three methods

Method	Accuracy
Proposed	90.6%
Sato's	80.0%
Li's	76.5%

Table 4.6: Confusion table of proposed method, overall accuracy:90.6%

	Estimated Types				
True Types	Type 1	Type 2	Type 3-a	Type 3-b	Sensitivity
Type 1	23	0	0	0	100%
Type 2	0	11	1	0	91.7%
Type 3-a	0	1	13	0	92.9%
Type 3-b	0	3	3	30	83.3%
Precision	100.0%	73.3%	76.5%	100.0%	

Table 4.7: Confusion table of Sato's filter, overall accuracy:80.0%

	Estimated Types				
True Types	Type 1	Type 2	Type 3-a	Type 3-b	Sensitivity
Type 1	23	0	0	0	100.0%
Type 2	1	3	8	0	25.0%
Type 3-a	0	2	12	0	85.7%
Type 3-b	0	1	5	30	83.3%
Precision	95.8%	50.0%	48.0%	100.0%	

Table 4.8: Confusion table of Li's filter, overall accuracy:76.5%

	Estimated Types				
True Types	Type 1	Type 2	Type 3-a	Type 3-b	Sensitivity
Type 1	19	0	4	0	82.6%
Type 2	0	5	7	0	41.7%
Type 3-a	2	1	11	0	78.6%
Type 3-b	0	0	6	30	83.3%
Precision	90.5%	83.3%	39.3%	100.0%	

Table 4.9: Statistical differences by MCNEMAR's test

Methods	P value
Proposed vs Sato's	0.046
Proposed vs Li's	0.016

Chapter 5

CAD Tools

5.1 Introduction

In our research, we designed two methods to classify the DLD patterns and pneumoconiosis on the HRCT images respectively. The Insight Segmentation and Registration Toolkit (ITK) was used to develop the methods. The ITK is a powerful toolkit to process the medical images, but it cannot be used to visualized operate and demonstrate the medical images. Considering that it would be convenient for the radiologists to use the CAD methods with an interactive graphical user interface (GUI), the visualized tool is required to operate the CAD applications in the clinical practice. In our laboratory, a visualized CAD platform named MARIMO (Medical Analyzer of Radiology for Images of Multi-Organs) had been developed by Shikata [124], and we implemented the two proposed applications as the plug-ins of the MARIMO.

5.1.1 Medical Analyzer of Radiology for Images of Multi-Organs (MARIMO)

The MARIMO is a visualized platform developed for the radiologists' clinical practice. There are three main characteristics for the MARIMO. (1) The MARIMO can be operated on several operating systems, such as Windows, Linux and MacOS. (2) The MARIMO provides the ability of multi-threading programming,

which can reduce the runtime of the medical image processing. (3) Most applications are implemented as the plug-ins, so it can be easily updated by adding or replacing the desired plug-ins. So far, many plug-ins have been developed to read, process and save the 3-dimension HRCT images. In the practice, the source codes of projects were firstly compiled as the dynamic link library (DLL) files, and then placed in the plug-in subdirectory of the MARIMO. When the MARIMO start, the plug-ins in the plug-in subdirectory were loaded and shown on the toolbar. Fig. 5.1a and Fig.5.1b show the startup screen of the MARIMO without and with plug-ins respectively. Fig. 5.2 gives the image of MARIMO with a three views (axial, sagittal and coronal direction) of a HRCT image.

5.2 CAD of DLDs by MARIMO

In the research, we designed two algorithms to classify the DLD patterns and pneumoconiosis respectively. In order to visualized operate the proposed algorithms, these applications were implemented as the plug-ins of the MARIMO. Furthermore, considering that only the regions located within the lung-field should be analyzed, if the mask image of the lung-field was not provided, it is necessary to automatically extract the lung-field as a preprocessing step of the classification.

5.2.1 Lung Segmentation

We integrated a simple lung-field segmentation algorithm based on the intensity-threshold to extract the lung-field. Fig. 5.3a gives the framework of extracting the lung-field. Firstly, the second largest structure was extracted by using the connect component labeling on the original image after thresholding (the largest structure was usually the background). This structure was composed of the lung and bronchus. Fig. 5.3b and Fig. 5.3c give the images of extracted structure and its volume rendering. Secondly, the region-growing approach with the manually set seed was used to extract the bronchus. The image of the seed point is shown in the Fig. 5.3d. Fig. 5.3e and Fig.5.3f show the images of the bronchus and its volume rendering respectively. Thirdly, the bronchus was removed from the structure extracted in the first step, and an initial result of the lung-field segmentation was

obtained. Fig. 5.3g gives the result of this step. It can be found that there were numerous holes in the image. Finally, the morphological dilate and then erode operation were implemented to fill the holes. Fig. 5.3h and Fig. 5.3i demonstrate the final result of the lung-field segmentation and its volume rendering.

The threshold of lung-field and bronchus is important for the extraction of lung structures and bronchus in the above method. However, in the clinical, it is difficult to find suitable thresholds for the lungs, especially for the lungs suffered from the severe DLDs, so this method would fail to extract the lung-field. In order to improve the performance of our applications, we also integrated a sophisticated algorithm based on the chamfer distance transform. This method spent little runtime and need no to input any parameters. Although the lung can be successfully extracted by this algorithm, but the main disadvantage is that many other structures would be extracted. Fig. 5.4 gives the results of the lung extraction and its volume rendering based on the chamfer distance transform.

5.2.2 Classification of DLD patterns and Pneumoconiosis

For the classification of DLD patterns, besides the sparse represent based method proposed in this thesis, we also integrated another two state-of-the-art methods [105, 124] which were developed by our laboratory in the previous research. These three methods were named "sparse representation", "bag-of-features" and "shape-texture" method respectively. The user can select any one method to classify the DLD patterns. In the classification, firstly we divided the lung-field into several 3D cubic-shape patches. The size of patches can be adjusted by the user. Then the VOIs with a size of $32 \times 32 \times 32$ were generated on the centers of the patches, and classified by the selected method. The recognized pulmonary pattern of each VOI was indicated by a non-zero value, which was used to evaluate the pulmonary pattern of the voxels within corresponding patch. Finally, the classification result of voxels on the input HRCT image was saved as a image-mask and output. Fig. 5.5a shows the dialogue of classifying the DLD patterns. The classification method and size of patches can be set on the dialogue. The classification of the CON, GGO, HCM and EMP by using the MARIMO are illustrated from

Fig. 5.5b to Fig. 5.5e respectively. After classification, the percentage of each abnormal patterns can be summarized and reported, as be shown in the Fig. 5.6.

In order to improve the performance on the classification of DLD patterns in the clinical application, it is important to often update the classification model. The conventional training process can be divided into two stages. At first, the radiologists collect the HRCT scans, and mark the normal or abnormal regions on the scans. Then the programmers generate the VOIs according to the data, and train the classification model. Unfortunately, the radiologists may spend a long time for a new model. So, we thought it would be convenient for the radiologists if they can train the classification model by themselves.

In the research, we designed some functions to train the model for the classification of the DLD patterns. The new training process composed of following three steps. Firstly, the radiologists selected some slices from the HRCT images and marked the interested regions on the selected slices. Fig. 5.7 illustrates the plug-ins to select the slices and draw interest regions on the HRCT images (these plug-in were developed in the work [124]). Secondly, the regions marked by different radiologists were summarized, and the common regions agreed by all radiologists were generated. Fig. 5.8 illustrate the plug-in for the summarization of regions. Finally, the MARIMO constructed the VOIs according to the common regions, and trained the classification model by the VOIs. Fig. 5.9a gives the images of the plug-in for the training classification model. There were two ways to construct the VOIs. The first one was to divide the selected slices into the patches with a size of 32×32. If the center of the patch located in the regions marked by the radiologists, the VOI was generated by treating this patch as the center axial-slice of the VOI with a size of $32 \times 32 \times 32$. The second method for the generation of VOIs was to move a grid with a size of 32×32 on the selected slices. When the area of marked regions in the current grid was larger than the input threshold and these regions had not been used to generate the VOIs, a patch was extracted on the current grid and then the VOI was generated according to this patch. The range of parameters of CAD methods can be adjusted by the users. Fig. 5.9b gives the screen of setting parameters to train the sparse representation method. We provide a default range of 500-3000 with an interval of 500 for number of atoms, 2-4 with an interval of 2 for the sparsity and $2 \times 2 \times 2$ - $6 \times 6 \times 6$

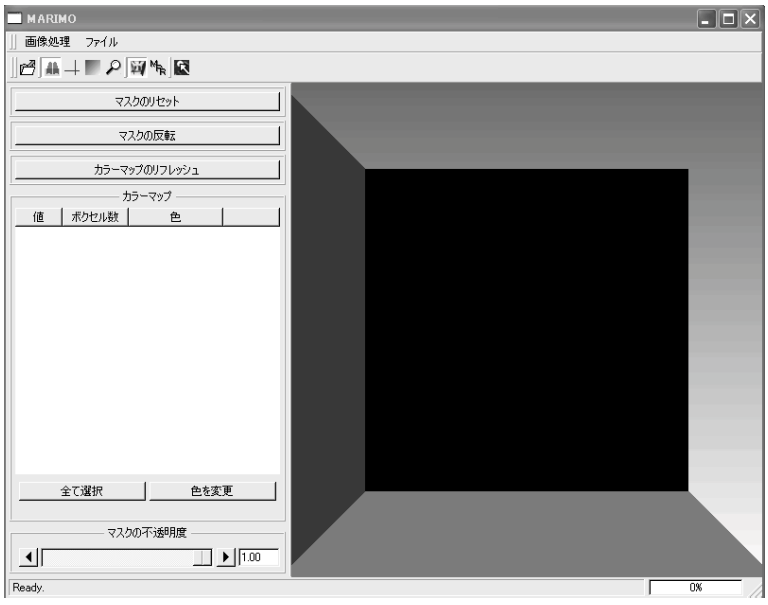
with an interval of $1 \times 1 \times 1$ for the size of patch. In the bag-of-features method, the default range of number of atoms and size of patches were set to be 50-400 and 2-6 respectively. Fig. 5.9c shows the screen of setting parameters to train the bag-of-features method. For the SVM classifier adopted in all three methods, the possible values of the soft-margin constant were set to be $2^{-2}, 2^{-1}, ..., 2^{11}, 2^{12}$, and the gamma related to the χ^2 kernel in the bag-of-features and RBF kernel in the shape-textures were set to be $2^{-12}, 2^{-1}, ..., 2^{1}, 2^{2}$. These parameters can be optimized by two kinds of evaluation approaches: n-fold cross validation or leave-one-out. After selecting the classification model and setting the range of parameters, the MARIMO optimized the parameters on the input data. The combination of parameters which produced the best overall accuracy were treated as the optimal parameters. Then the MARIMO trained the selected method with the optimal parameters on the input data again, and the classification model was output.

For the classification of the pneumoconiosis, there are three kinds of outputs after the classification: the dialogue box (Fig. 5.10a) and text file (Fig. 5.10b) that used to report the recognized class of the pneumoconiosis, and the mask of the lung indicated the detected nodules (Fig. 5.10c).

5.3 Conclusion

In this chapter, we introduced a visualized CAD platform named MARIMO, and our methods were implemented as the plug-ins of the MARIMO. Considering that only the regions within the lung-filed should be analyzed, the lung-field segmentation was integrated with the proposed methods to automatically extract the lung-field as the preprocessing. We thought that the application of our methods by using MARIMO was useful for the radiologists in the clinical practice.

(a) MARIMO without plug-ins



(b) MARIMO with plug-ins

Figure 5.1: The startup screen of the MARIMO

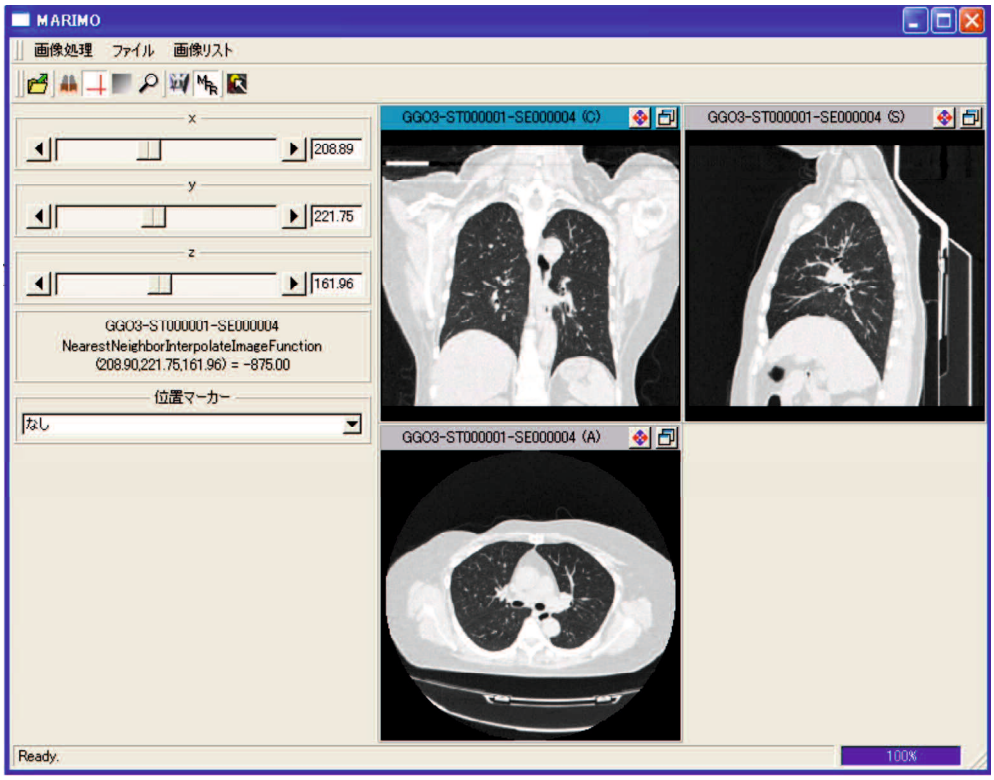
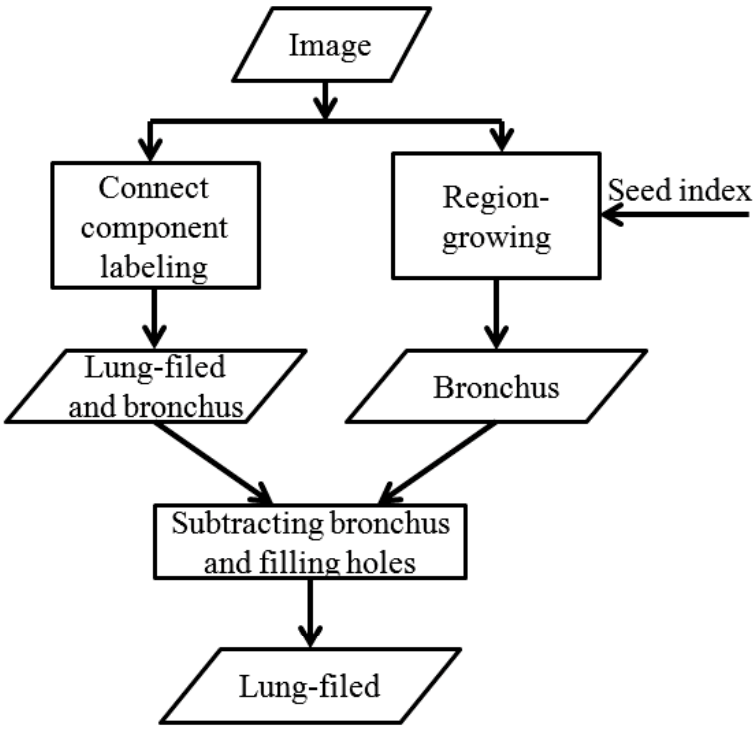
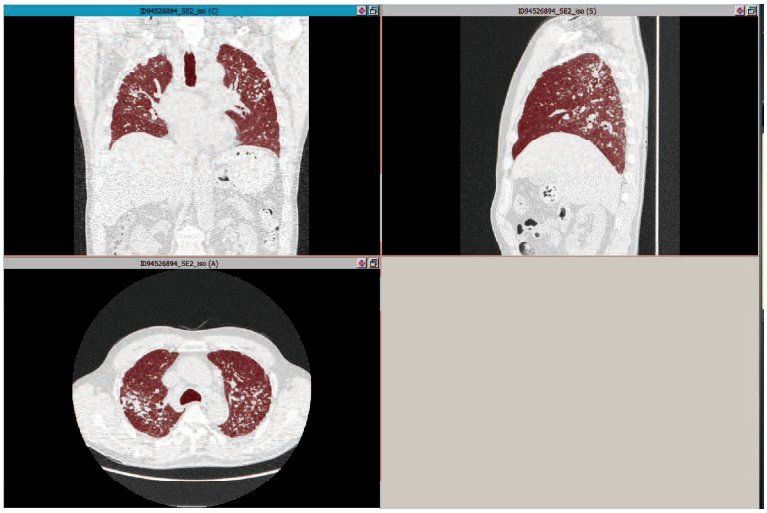


Figure 5.2: The three views of a HRCT image (axial, sagittal and coronal direction)

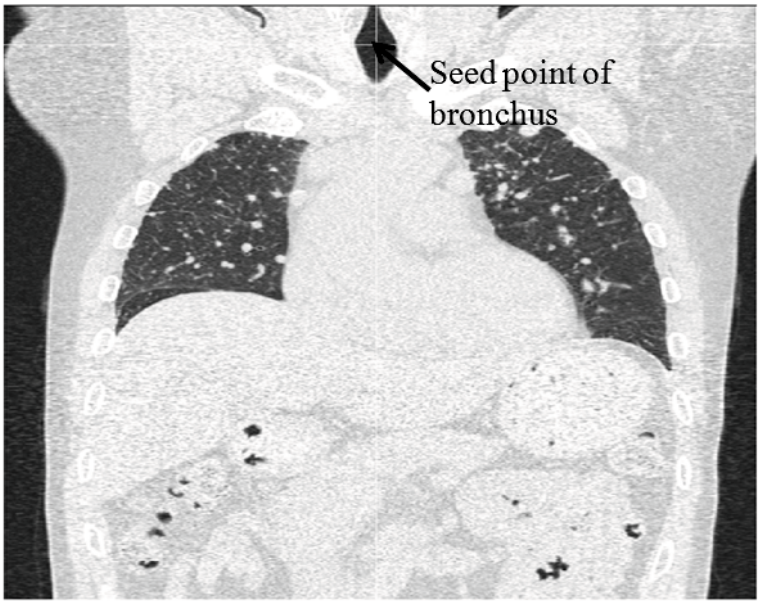


(a) The framework of the lung-field segmentation

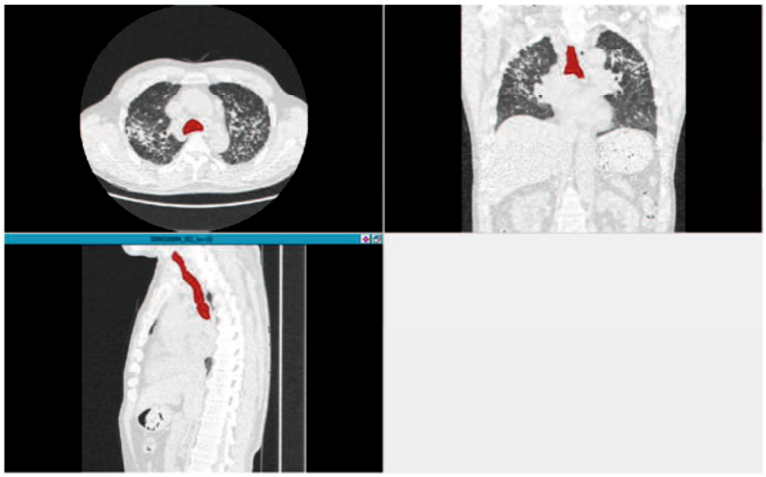


(b) Image of the 2nd largest object composed of the lung-field and bronchus

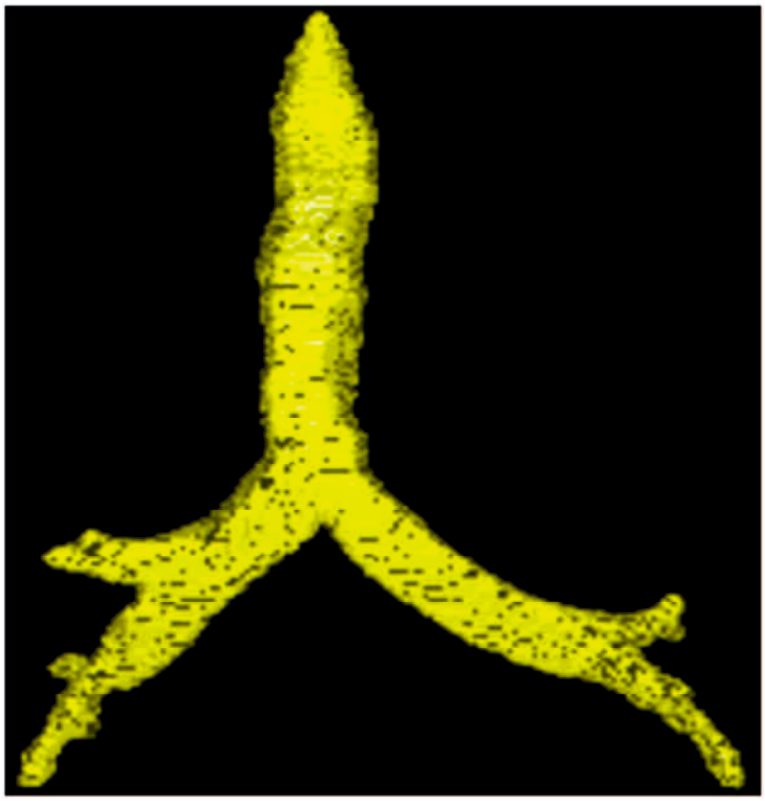
(c) Volume rendering of the object composed of the lung-field and bronchus



(d) Seed point for extracting the bronchus



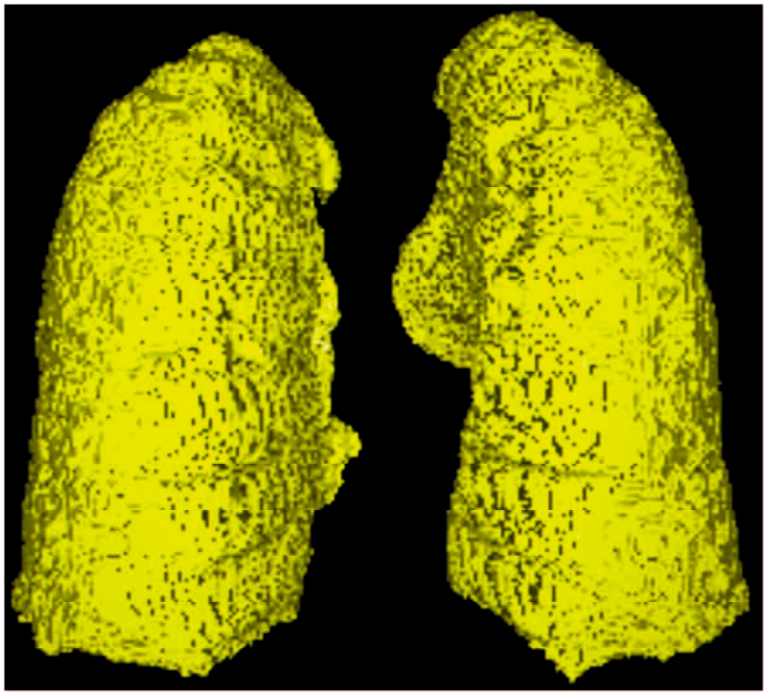
(e) Image of the extracted bronchus



(f) Volume rendering of the extracted bronchus

(g) Lung-field without filling holes

(h) Final result of the lung-field segmentation



(i) Volume rendering of lung-field

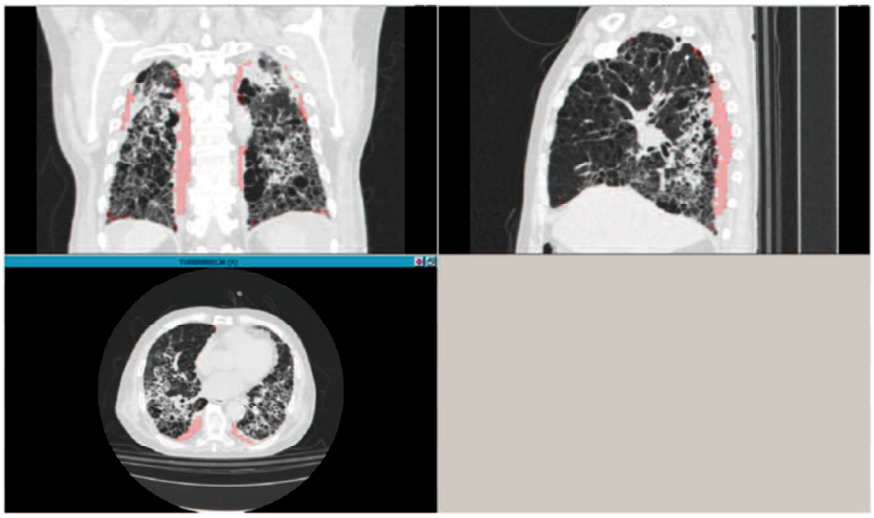
Figure 5.3: Lung-field segmentation

(a) Lung extraction

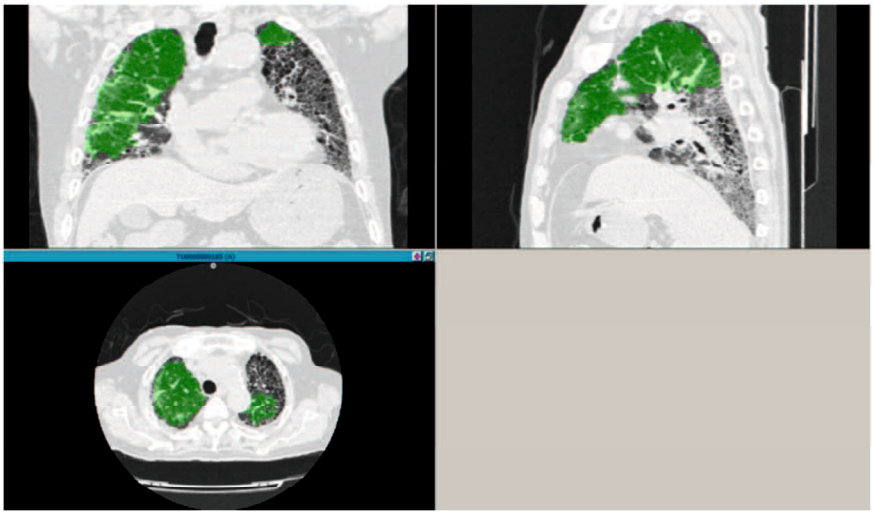
(b) Volume rendering of lung-field

Figure 5.4: Lung extraction based on chamfer distance transform

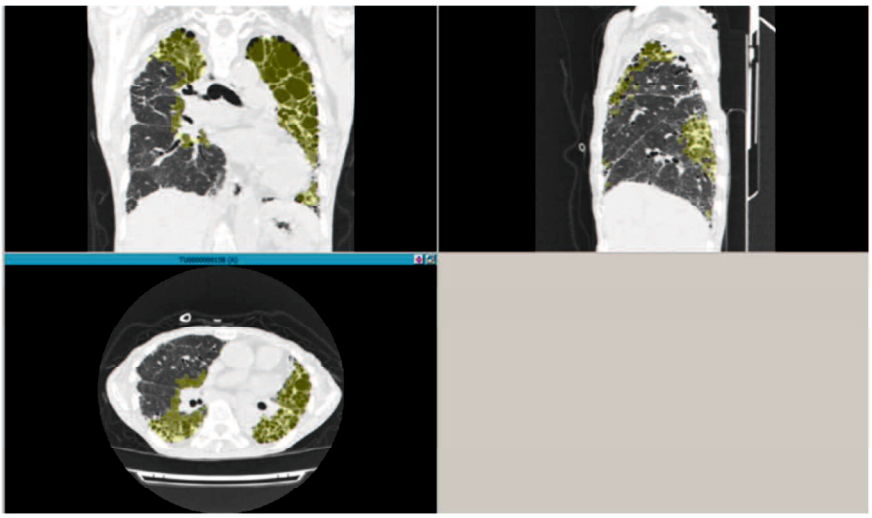
(a) Selection of the methods to classify the DLD patterns



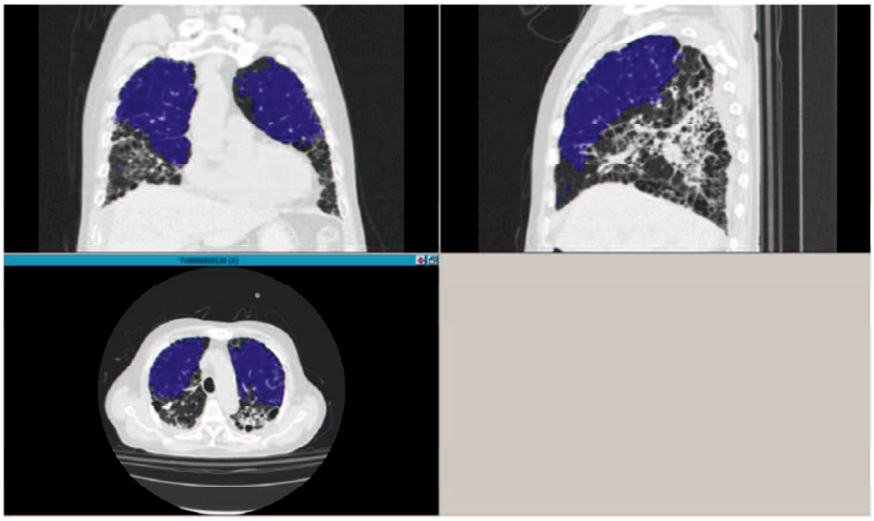
(b) Classified CON



(c) Classified GGO



(d) Classified HCM



(e) Classified EMP

Figure 5.5: Implementation of the DLD classifier

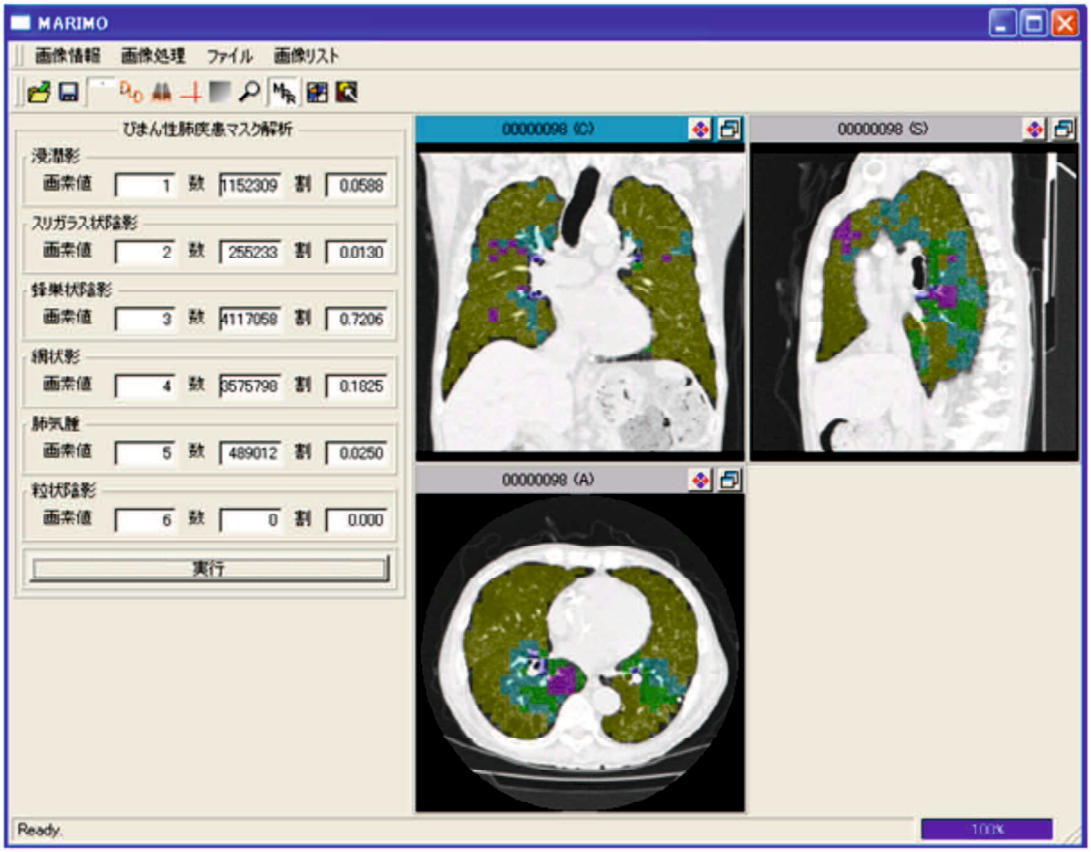
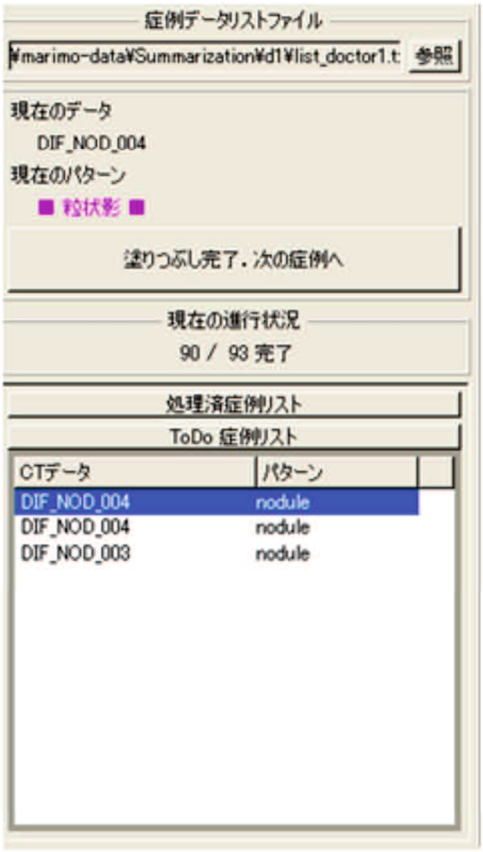
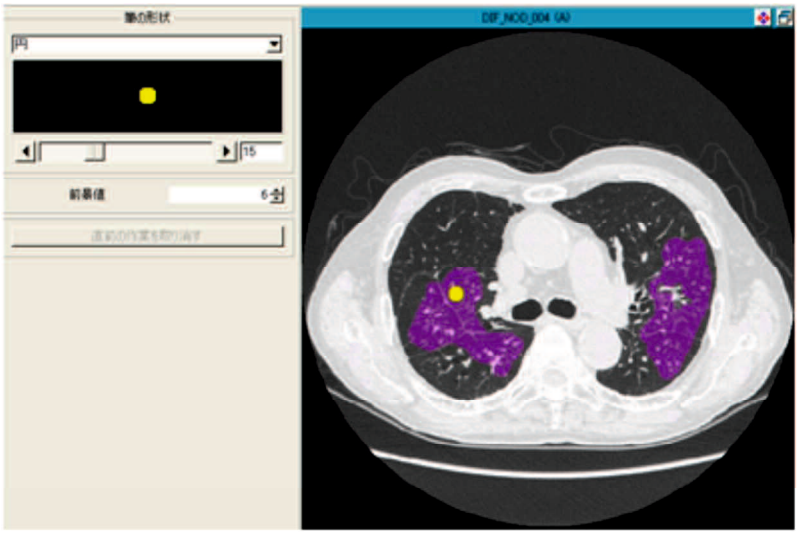


Figure 5.6: Report of abnormal pulmonary patterns



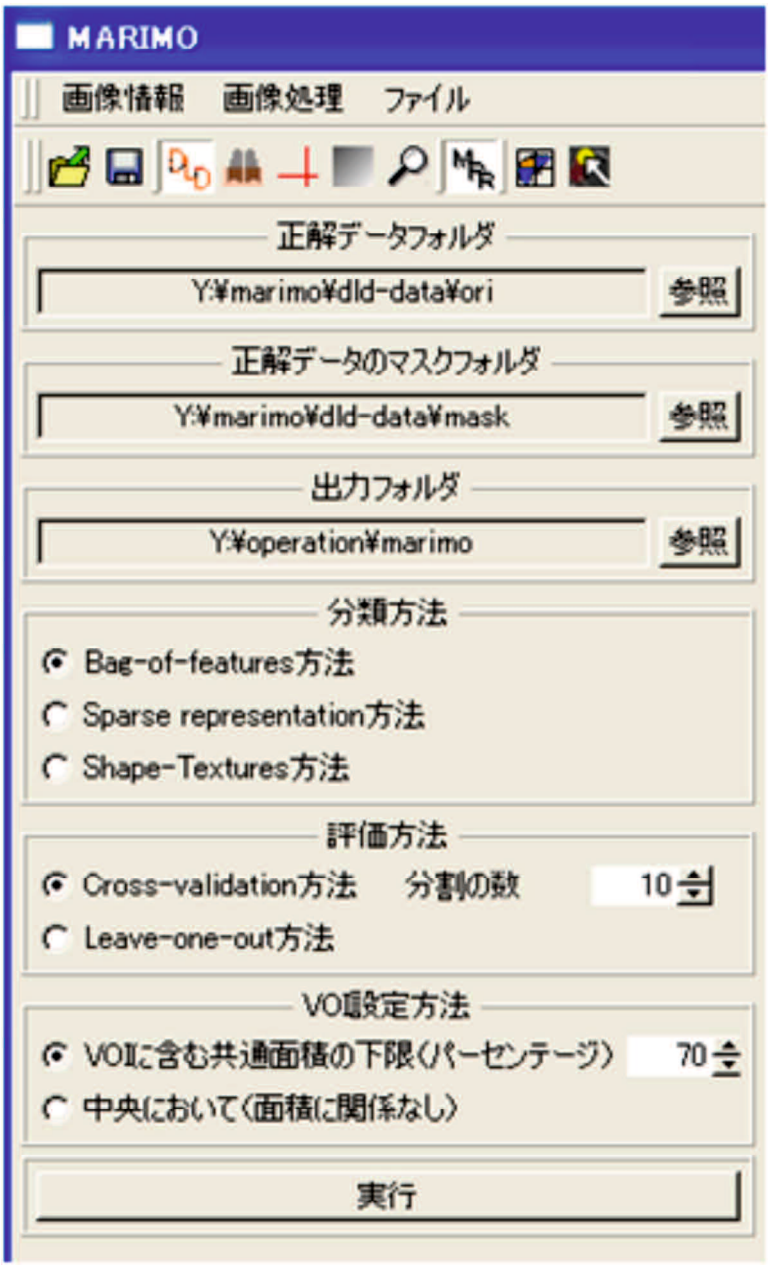
(a) Selection of slice



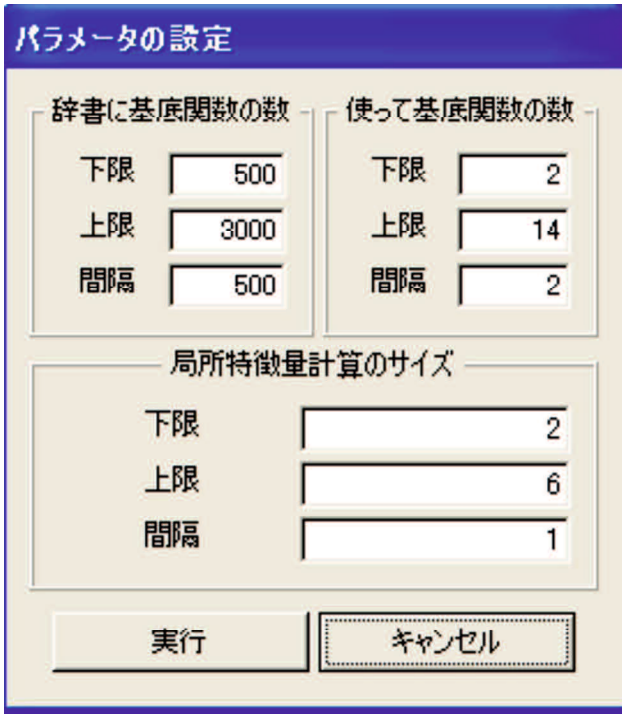
(b) Drawing interest regions

Figure 5.7: Preparation of interest regions by radiologists

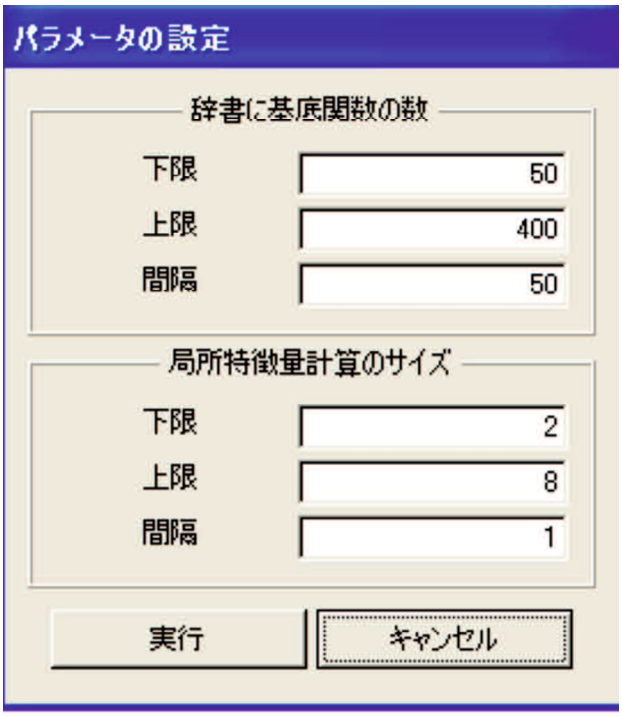
Figure 5.8: Summarization of regions marked by radiologists



(a) Start screen of the training filter for classification of DLD patterns

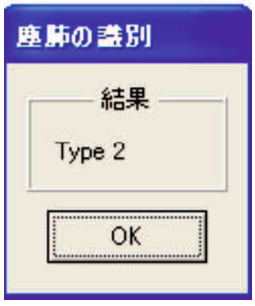


(b) Setting parameters of the sparse representation method

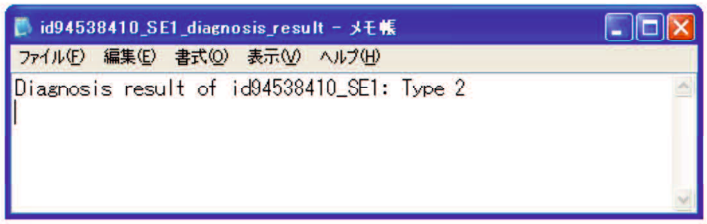


(c) Setting parameters of the bag-of-features method

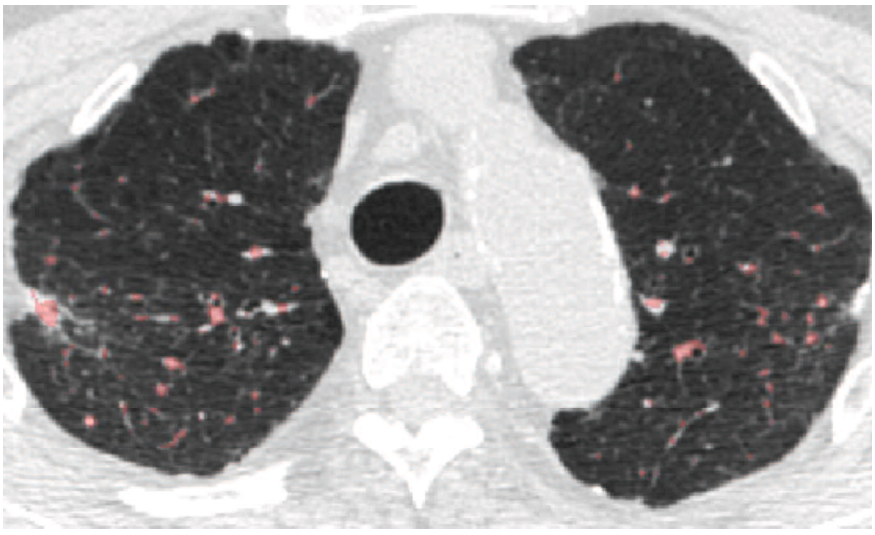
Figure 5.9: Implementation of the DLD training filter



(a) The dialogue as the output of classification of Pneumoconiosis



(b) The txt file as the output of classification of the DLD patterns



(c) The detected of nodules as the output of CAD of Pneumoconiosis

Figure 5.10: Implementation of the pneumoconiosis classifier

Chapter 6

Conclusion

In the thesis, we presented two methods to classify the DLD patterns and pneumoconiosis on the HRCT images, and implemented the proposed methods as the plug-ins of a visualized CAD platform named MARIMO developed in our laboratory. In order to improve the performance on the classification, we introduced a concept of the sparse representation to calculate the image features. The main idea of the sparse representation is to select few number of key features to approximate the examples. So it is important to construct a set of representative features which can represent the salient information of examples, and approximate the examples in a suitable way. Experimental results show that the our methods achieved good performance and would be useful in the clinical practice.

In the Chapter 3, we used and optimized the sparse representation approaches to classify the normal tissues and five kinds of DLD patterns, including consolidation, ground-glass opacity, honeycombing, emphysema and nodular. The K-SVD and OMP are two of the most popular approaches for the dictionary learning and sparse coding respectively, and the method using these two algorithms achieved a good classification accuracy. However, it need too much time in the operation. Considering that it is report that the application of the K-Means can obtain a competitive result on the classification of natural image to the K-SVD, we used the K-Means as a substitute of the K-SVD. On the other hand, we simplified the OMP by selecting a sufficient number of atoms at one time (OMP_1). The performance of the OMP_1 can be ensured under a certain sparsity constraint although the reconstruction error of OMP_1 is larger than the OMP. We designed three

methods for evaluation: SR1 (K-SVD+OMP), SR2 (K-Means+OMP) and SR3 (K-Means + OMP_1). The overall accuracies of all three methods were greater than 95%, which means that the replacement of the K-SVD and OMP by the K-Means and OMP_1 had not influence on the recognition of pulmonary patterns. Furthermore, the runtimes of the dictionary learning and recognition were reduced by 98.2% and 55.2% respectively when the K-Means and OMP_1 were employed. Therefore, we conclude that the method using K-Means and OMP_1 was efficient to classify the DLD patterns.

In the Chapter 4, we proposed a method to classify the four kinds of the pneumoconiosis on the HRCT images. Because the diagnosis of the pneumoconiosis is according to the size and density of nodules, we designed a two-step classification strategy. Firstly, we classified the Type 3-b and other kinds of data based on the size of nodules. Then, we classified the Type 1, 2 and 3-a by using the bag-of-features method which could be seemed as a special version of the sparse representation on the sparsity of the approximation. Considering that the nodular opacities are the main finding of pneumoconiosis, but the performance of conventional methods for the nodule extraction is limited. We designed two eigenvalues based filters to detect the nodules and then remove the false positives while keeping the nodules as many as possible. The recognition rate of our method was 90.6%, better than the baseline methods. It demonstrates the proposed method would be useful for the CAD of the pneumoconiosis.

The above two algorithms had good performance on the classification, but the operation of these algorithms was based on the command-line interface. It would be inconvenient for the radiologists in the clinical practice. In order to facilitate the radiologists' diagnosis, the proposed methods were visualized implemented through a graphic use interface in the Chapter 6. We designed several plug-ins to classify the DLD patterns and pneumoconiosis, and train the model for the classification of DLD patterns. When the MARIMO started, the plug-ins were shown in the menu. After reading the image, user can select CAD method to analyze the image. Finally, the results were output. We thought that the MARIMO with the proposed plug-ins would have a great potential for the CAD of the DLDs on the HRCT images.

In future research, we will improve the proposed methods from the two aspects. The first one is to optimize our methods based on the experimental results in the actual clinical setting, and integrate an advanced lung-field segmentation algorithm. The second one is to classify the pneumoconiosis according to the international classification of HRCT for occupational and environmental respiratory diseases (ICOERD)[125].

List of Figures

List of Tables

Bibliography

[1] E. R. S. American Thoracic Society, "American thoracic society/european respiratory society international multidisciplinary consensus classification of the idiopathic interstitial pneumonias," *American Journal of Respiratory and Critical Care Medicine*, vol. 165, pp. 277–304, 2002. 1

[2] G. Raghu and K. K. Brown, "Interstitial lung disease: clinical evaluation and keys to an accurate diagnosis," *Clinics in Chest Medicine*, vol. 25, no. 3, pp. 409–419, 2004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0272523104000607 1

[3] O. J. Dempsey, K. M. Kerr, H. Remmen, and A. R. Denison, "How to investigate a patient with suspected interstitial lung disease," *BMJ*, vol. 340, 6 2010. 1, 2, 10, 101

[4] K. Honma, K. Chiyotani, and K. Kimura, "Silicosis, mixed dust pneumoconiosis, and lung cancer," *American journal of industrial medicine*, vol. 32, no. 6, pp. 595–599, 12 1997. 2

[5] M. Katabami, H. Dosaka-Akita, K. Honma, Y. Saitoh, K. Kimura, Y. Uchida, H. Mikami, Y. Ohsaki, Y. Kawakami, and K. Kikuchi, "Pneumoconiosis-related lung cancers: preferential occurrence from diffuse interstitial fibrosis-type pneumoconiosis," *Am J Respir Crit Care Med*, vol. 162, no. 1, pp. 295–300, 2000. [Online]. Available: http://www.biomedsearch.com/nih/Pneumoconiosis-related-lung-cancers-preferential/1090
2

[6] *The report of occputational diseases in 2010 in China*, National Health and Family Planning Commission of the Peopples's Repulic of China, 4 2011. 2

[7] *The report of occputational diseases in 2011 in China*, National Health and Family Planning Commission of the Peopples's Repulic of China, 9 2012. 2

[8] *The report of occputational diseases in China 2012*, National Health and Family Planning Commission of the Peopples's Repulic of China, 9 2013. 2

[9] S. J. Bourke, "Interstitial lung disease: progress and problems," *Postgraduate Medical Journal*, vol. 82, no. 970, pp. 494–499, 2006. [Online]. Available: http://pmj.bmj.com/content/82/970/494.abstract 2

[10] S. O. C. C. BRITISH THORACIC SOCIETY, "The Diagnosis, Assessment and Treatment of Diffuse Parenchymal Lung Disease in Adults," *Thorax*, vol. 54, no. Supplement 1, pp. S1–S28, Apr. 1999. [Online]. Available: http://thorax.bmj.com/cgi/content/long/54/suppl_1/S1 2

[11] T. M. Maher, "Diffuse parenchymal lung disease," *Medicine*, vol. 40, no. 6, pp. 314 – 321, 2012, ¡ce:title¿Respiratory Disorders: Part 3 of 3¡/ce:title¿. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S135730391200059X 2

[12] M. Gulati, "Diagnostic assessment of patients with interstitial lung disease," *Primary Care Respiratory Journal*, vol. 20, no. 2, pp. 120–127, June 2011. [Online]. Available: http://www.thepcrj.org/journ/view_article.php?article_id=801 2, 14, 103

[13] D. P. Naidich, "Pulmonary parenchymal high-resolution ct: to be or not to be," *Radiology*, vol. 171, no. 1, pp. 22–24, 1989, pMID: 2648472. [Online]. Available: http://pubs.rsna.org/doi/abs/10.1148/radiology.171.1.2648472 2

[14] I. H. K. David M Hansell, "The role of high-resolution computed tomography in the work-up of interstitial lung disease," *Thorax*, vol. 46, no. 2, pp. 77–84, February 1991. 2, 11, 101

[15] Z. Tatjana and S. Busayarat, *Theory and Applications of CT Imaging and Analysis*. InTech, 2011, ch. Computer-aided Analysis and Interpretation of HRCT Images of the Lung. [Online]. Available: http://www.intechopen.com/books/theory-and-applications-of-ct-imaging-and-analysis/co 2

[16] A. Savranlar, R. Altin, K. Mahmutyazicioglu, H. Ozdemir, L. Kart, T. Ozer, and S. Gündogdu, "Comparison of chest radiography and high-resolution computed tomography findings in early and low-grade coal worker's pneumoconiosis," *Eur J Radiol*, vol. 51, no. 2, pp. 175–80, 2004. [Online]. Available: http://www.biomedsearch.com/nih/Comparison-chest-radiography-high-resolution/152465 2

[17] S. Padley, D. Hansell, C. Flower, and P. Jennings, "Comparative accuracy of high resolution computed tomography and chest radiography in the diagnosis of chronic diffuse infiltrative lung disease," *Clinical Radiology*, vol. 44, no. 4, pp. 222 – 226, 1991. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0009926005801837 3

[18] W. R. W. R. Webb, D. P. Naidich, and N. L. Muller, *High-resolution CT of the lung*, 3rd ed. Philadelphia ; London : Lippincott Williams & Wilkins, 2001, previous ed.: 1996. 3

[19] C. V. Zwirewich, J. R. Mayo, and N. Müller, "Low-dose high-resolution ct of lung parenchyma." *Radiology*, vol. 180, no. 2, pp. 413–417, 1991. [Online]. Available: http://pubs.rsna.org/doi/abs/10.1148/radiology.180.2.2068303 3

[20] V. B. Antunes, G. d. S. P. Meirelles, D. Jasinowodolinski, C. A. d. C. Pereira, C. G. Y. Verrastro, F. A.-o. G. Torlai, and G. D'Ippolito, "Observer agreement in the diagnosis of interstitial lung diseases based on hrct scans," *Jornal Brasileiro de Pneumologia*, vol. 36, pp. 29 – 36, 02 2010. [Online]. Available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-37132010000100007&nrm=iso 4

[21] D. hui Xu, A. S. Kurani, J. D. Furst, and D. S. Raicu, "Run-length encoding for volumetric texture," in *The 4th IASTED International Conference on Visualization, Imaging, and Image Processing*, 2004, pp. 534–539. 4, 60

[22] A. S. Kurani, D. hui Xu, J. Furst, and D. S. Raicu, "Co-occurrence matrices for volumetric data," in *7th IASTED International Conference on Computer Graphics and Imaging*, Aug 2004, pp. 447–452. 4, 60

[23] J. Wang, F. Li, K. Doi, and Q. Li, "A novel scheme for detection of diffuse lung disease in mdct by use of statistical texture features," in *Proc.SPIE*, ser. Medical Imaging 2009: Computer-Aided Diagnosis, vol. 7260, Lake Buena Vista, February 2009, p. 726039. 4

[24] S. C. Park, J. Tan, X. Wang, D. Lederman, J. K. Leader, S. H. Kim, and B. Zheng, "Computer-aided detection of early interstitial lung diseases using low-dose ct images," *Physics in Medicine and Biology*, vol. 56, no. 4, p. 1139, 2011. 4

[25] P. Korfiatis, A. Karahaliou, A. Kazantzi, C. Kalogeropoulou, and L. Costaridou, "Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector ct," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 3, pp. 675–680, May 2010. 4

[26] R. UPPALURI, T. MITSA, M. SONKA, E. A. HOFFMAN, and G. M. LENNAN, "Quantification of pulmonary emphysema from lung computed tomography images," *American journal of respiratory and critical care medicine*, vol. 156, no. 1, pp. 248–254, July 1997. 4

[27] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. Hoffman, "Mdct-based 3-d texture classification of emphysema and early smoking related lung pathologies," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 4, pp. 464–475, April 2006. 4

[28] Y. Xu, E. J. van Beek, Y. Hwanjo, J. Guo, G. McLennan, and E. A. Hoffman, "Computer-aided classification of interstitial lung diseases via

mdct: 3d adaptive multiple feature method (3d amfm)," *Academic Radiology*, vol. 13, no. 8, pp. 969 – 978, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1076633206002716 4

[29] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51 – 59, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0031320395000674 5

[30] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, Jul 2002. 5

[31] B. B. Mandelbrot, *The Fractal Geometry of Nature*. Henry Holt and Company, 1983. 5

[32] R. Uppaluri, E. A. Hoffman, M. Sonka, P. G. Hartley, G. W. Hunninghake, and G. Mclennan, "Computer recognition of regional lung disease patterns," *American Journal of Respiratory and Critical Care Medicine*, vol. 160, no. 2, pp. 648–654, August 1999. 5

[33] S. O. Park, J. B. Seo, N. Kim, S. H. Park, Y. K. Lee, B. Park, Y. S. Sung, Y. Lee, J. Lee, and S.-H. Kang, "Feasibility of automated quantification of regional disease patterns depicted on high-resolution computed tomography in patients with various diffuse lung diseases," *Korean Journal of Radiology*, vol. 10, no. 5, pp. 455–463, 10 2009. 5, 41

[34] Y. Song, W. Cai, Y. Zhou, and D. Feng, "Feature-based image patch approximation for lung tissue classification," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 4, pp. 797–808, April 2013. 5

[35] L. Sorensen, S. Shaker, and M. De Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 559–569, 2010. 5

[36] A. Foncubierta-Rodrguez, A. Depeursinge, and H. Mller, "Using multiscale visual words for lung texture classification and retrieval," in *Medical Content-Based Retrieval for Clinical Decision Support*, ser. Lecture Notes in Computer Science, H. Mller, H. Greenspan, and T. Syeda-Mahmood, Eds. Springer Berlin Heidelberg, 2012, vol. 7075, pp. 69–79. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-28460-1_7 5

[37] A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser, and H. Muller, "Lung tissue classification using wavelet frames," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, Aug 2007, pp. 6259–6262. 5

[38] A. Shamsheyeva and A. Sowmya, "The anisotropic gaussian kernel for svm classification of hrct images of the lung," in *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, Dec 2004, pp. 439–444. 5

[39] A. Depeursinge, A. Foncubierta-Rodriguez, D. Ville, and H. Mller, "Lung texture classification using locally-oriented riesz components," in *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2011*, ser. Lecture Notes in Computer Science, G. Fichtinger, A. Martel, and T. Peters, Eds. Springer Berlin Heidelberg, 2011, vol. 6893, pp. 231–238. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-23626-6_29 5

[40] Z. Tu, "Auto-context and its application to high-level vision tasks," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, June 2008, pp. 1–8. 6

[41] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1031–1044, 2010. 6

[42] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, June 2009, pp. 1794–1801. 6, 27, 32, 40

[43] B. Wohlberg, "Noise sensitivity of sparse signal representations: reconstruction error bounds for the inverse problem," *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3053–3060, 2003. 6

[44] M. Liu, L. Lu, X. Ye, S. Yu, and M. Salganicoff, "Sparse classification for computer aided diagnosis using learned dictionaries," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011*, ser. Lecture Notes in Computer Science, G. Fichtinger, A. Martel, and T. Peters, Eds. Springer Berlin Heidelberg, 2011, vol. 6893, pp. 41–48. 6

[45] K. Vo and A. Sowmya, "Multiscale sparse representation of high-resolution computed tomography (hrct) lung images for diffuse lung disease classification," in *2011 18th IEEE International Conference on Image Processing*, 2011, pp. 441–444. 6

[46] M. Zhang, X. Zhou, S. Goshima, H. Chen, C. Muramatsu, T. Hara, R. Yokoyama, M. Kanematsu, and H. Fujita, "An application to pulmonary emphysema classification based on model of texton learning by sparse representation," in *Proc.SPIE*, 2012, pp. 831 534–831 534–7. 6

[47] M. Giger, N. Karssemeijer, and S. Armato, "Guest editorial computer-aided diagnosis in medical imaging," *Medical Imaging, IEEE Transactions on*, vol. 20, no. 12, pp. 1205–1208, Dec 2001. 6

[48] B. van Ginneken, C. M. Schaefer-Prokop, and M. Prokop, "Computer-aided diagnosis: How to move from the laboratory to the clinic," *Radiology*, vol. 261, no. 3, pp. 719–732, 2011. 7

[49] Y. Uchiyama, S. Katsuragawa, H. Abe, J. Shiraishi, F. Li, Q. Li, C. Zhang, K. Suzuki, and K. Doi, "Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography," *Medical Physics*, vol. 30, no. 9, pp. 2440–2454, 2003. 7, 34, 35

[50] H. SHIKATA and S. KIDO, "A pattern classification system for abnormal attenuations caused by diffuse lung diseases in chest x-ray 3d ct images by means of texture analysis and svm," *The IEICE transactions on*

information and systems (Japanese edtion), vol. 91, no. 7, pp. 1895–1903, jul 2008. [Online]. Available: http://ci.nii.ac.jp/naid/80019582082/en/ 7

[51] R. Kruger, W. B. Thompson, and A. F. Turner, "Computer diagnosis of pneumoconiosis," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. SMC-4, no. 1, pp. 40–49, Jan 1974. 7

[52] A. F. Turner, R. P. Kruger, and W. B. Thompson, "Automated computer screening of chest radiographs for pneumoconiosis," *Invest Radiol*, vol. 11, no. 4, pp. 258–266, 1976. [Online]. Available: http://www.biomedsearch.com/nih/Automated-computer-screening-chest-radiographs/783 7

[53] P. Yu, H. Xu, Y. Zhu, C. Yang, X. Sun, and J. Zhao, "An automatic computer-aided detection scheme for pneumoconiosis on digital chest radiographs," *J. Digital Imaging*, vol. 24, no. 3, pp. 382–393, 2011. 7

[54] V. Murray, M. S. Pattichis, H. Davis, E. S. Barriga, and P. Soliz, "Multiscale am-fm analysis of pneumoconiosis x-ray images," in *ICIP*. IEEE, 2009, pp. 4201–4204. 7

[55] E. Okumura, I. Kawashita, and T. Ishida, "Computerized analysis of pneumoconiosis in digital chest radiography: effect of artificial neural network trained with power spectra," *J Digit Imaging*, vol. 24, no. 6, pp. 1126–1132, 2011. [Online]. Available: http://www.biomedsearch.com/nih/Computerized-Analysis-Pneumoconiosis-in-Digital/21 7

[56] B. Zhu, H. Chen, B. Chen, Y. Xu, and K. Zhang, "Support vector machine model for diagnosing pneumoconiosis based on wavelet texture features of digital chest radiographs," *Journal of Digital Imaging*, vol. 27, no. 1, pp. 90–97, 2014. [Online]. Available: http://dx.doi.org/10.1007/s10278-013-9620-9 7

[57] E. Okumura, I. Kawashita, and T. Ishida, "Development of cad based on ann analysis of power spectra for pneumoconiosis in

chest radiographs: effect of three new enhancement methods," *Radiological Physics and Technology*, pp. 1–11, 2014. [Online]. Available: http://dx.doi.org/10.1007/s12194-013-0255-9 7

[58] R. Smithuis, "Chest x-ray - lung disease four-pattern approach," adiology Department of the Rijnland Hospital, Tech. Rep., February 2014. [Online]. Available: http://www.radiologyassistant.nl/en/p50d95b0ab4b90/chest-x-ray-lung-disease.html 12, 101

[59] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004. [Online]. Available: http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94 16

[60] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, June 2005, pp. 886–893 vol. 1. 16

[61] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, Apr. 1995. [Online]. Available: http://dx.doi.org/10.1137/S0097539792240406 16

[62] D. L. Donoho, "For most large underdetermined systems of equations, the minimal l1-norm near-solution approximates the sparsest near-solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006. [Online]. Available: http://dx.doi.org/10.1002/cpa.20132 16

[63] I. Gorodnitsky and B. Rao, "Sparse signal reconstruction from limited data using focuss: a re-weighted minimum norm algorithm," *Signal Processing, IEEE Transactions on*, vol. 45, no. 3, pp. 600–616, Mar 1997. 17

[64] B. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *Signal Processing, IEEE Transactions on*, vol. 47, no. 1, pp. 187–200, Jan 1999. 17

[65] B. Rao, K. Engan, S. Cotter, J. Palmer, and K. Kreutz-Delgado, "Subset selection in noise based on diversity measure minimization," *Signal Processing, IEEE Transactions on*, vol. 51, no. 3, pp. 760–770, March 2003. 17

[66] S. Chen and D. Donoho, "Basis pursuit," in *Signals, Systems and Computers, 1994. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on*, vol. 1, Oct 1994, pp. 41–44 vol.1. 17

[67] ——, "Basis pursuit," Department of Statistics, Standdord University, Tech. Rep., 1994. 17

[68] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, Dec 1993. 17, 22

[69] Y. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, vol. 1, 1993, pp. 40–44. 17, 29

[70] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, pp. 407–499, 2004. 17

[71] H. Lee, A. Battle, R. Raina, and A. Y. Ng, "Efficient sparse coding algorithms," in *In NIPS*. NIPS, 2007, pp. 801–808. 18, 23

[72] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 3360–3367. 18

[73] L. Karlovitz, "Construction of nearest points in the lp, p even, and l norms. i," *Journal of Approximation Theory*, vol. 3, no. 2, pp. 123 – 127, 1970. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0021904570900195 18

[74] C. L. Lawson, "Contributions to the theory of linear least maximum approximation," Ph.D. dissertation, University of California, Los Angeles, 1961. 19

[75] J. Murray and K. Kreutz-Delgado, "Sparse image coding using learned overcomplete dictionaries," in *Machine Learning for Signal Processing, 2004. Proceedings of the 2004 14th IEEE Signal Processing Society Workshop*, Sept 2004, pp. 579–588. 19

[76] D. Donoho, "De-noising by soft-thresholding," *Information Theory, IEEE Transactions on*, vol. 41, no. 3, pp. 613–627, May 1995. 19

[77] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, 1st ed. Springer Publishing Company, Incorporated, 2010. 19, 31

[78] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004. [Online]. Available: http://dx.doi.org/10.1002/cpa.20042 19

[79] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007. 20

[80] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Img. Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009. [Online]. Available: http://dx.doi.org/10.1137/080716542 20

[81] T. Adeyemi and M. Davies, "Sparse representations of images using overcomplete complex wavelets," in *Statistical Signal Processing, 2005 IEEE/SP 13th Workshop on*, July 2005, pp. 865–870. 21

[82] S. S. Chen, D. L. Donoho, Michael, and A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, pp. 33–61, 1998. 21

[83] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, pp. 267–288, 1994. 21

[84] M. H. Wright, "The interior-point revolution in optimization: history, recent developments, and lasting consequences," *Bull. Amer. Math. Soc. (N.S)*, vol. 42, pp. 39–56, 2005. 21

[85] S. Sardy, A. G. Bruce, and P. Tseng, "Block coordinate relaxation methods for nonparametric signal denoising with wavelet dictionaries," *Journal of Computational and Graphical Statistics*, vol. 9, pp. 361–379, 2000. 21

[86] D. Malioutov, M. Cetin, and A. Willsky, "Homotopy continuation for sparse signal representation," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, vol. 5, March 2005, pp. v/733–v/736 Vol. 5. 22

[87] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008. [Online]. Available: http://link.aip.org/link/?SCE/31/890 22

[88] E. G. Birgin, J. E. M. Martinez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM Journal on Optimization*, pp. 1196–1211, 2000. 22

[89] E. G. Birgin, J. M. Martínez, and M. Raydan, "Algorithm 813: Spg - software for convex-constrained optimization," *ACM Trans. Math. Softw.*, vol. 27, no. 3, pp. 340–349, Sep. 2001. [Online]. Available: http://doi.acm.org/10.1145/502800.502803 22

[90] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for newton's method," *SIAM J. Numer. Anal.*, vol. 23, no. 4, pp. 707–716, Aug. 1986. [Online]. Available: http://dx.doi.org/10.1137/0723046 22

[91] J. BARZILAI and J. M. BORWEIN, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988. [Online]. Available: http://imajna.oxfordjournals.org/content/8/1/141.abstract 22

[92] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l1-minimization: Methodology and convergence," *SIAM Journal on Optimization*, vol. 19, no. 3, pp. 1107–1130, 2008. 22

[93] E. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation applied to compressed sensing: implementation and numerical experiments," *J. Comput. Math.*, vol. 28, no. 2, pp. 170–194, 2010. 22

[94] T. S. Lee, "Image representation using 2d gabor wavelets," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, pp. 959–971, 1996. 22

[95] K. Engan, S. Aase, and J. Hakon Husoy, "Method of optimal directions for frame design," in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, vol. 5, 1999, pp. 2443–2446 vol.5. 22

[96] M. Aharon, M.Elad, and A.Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov 2006. 22, 26, 29, 30

[97] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, no. 3, pp. 211–218, 1936. [Online]. Available: http://dx.doi.org/10.1007/BF02288367 23

[98] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Discriminative learned dictionaries for local image analysis," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, June 2008, pp. 1–8. 24

[99] Z. Jiang, Z. Lin, and L. S. Davis, "Learning a discriminative dictionary for sparse coding via label consistent k-svd," in *Computer Vision and Pattern*

Recognition (CVPR), 2011 IEEE Conference on, 2011/06// 2011, pp. 1697 – 1704. 25

[100] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, 2006, pp. 2169–2178. 26

[101] X. Zhou, K. Yu, T. Zhang, and T. S. Huang, "Image classification using super-vector coding of local image descriptors," in *Proceedings of the 11th European Conference on Computer Vision: Part V*, ser. ECCV'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 141–154. [Online]. Available: http://dl.acm.org/citation.cfm?id=1888150.1888162 27

[102] Y.-L. Boureau, F. Bach, Y. LeCun, and J. Ponce, "Learning mid-level features for recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 2559–2566. 29

[103] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Found. Trends. Comput. Graph. Vis.*, vol. 3, no. 3, pp. 177–280, Jul. 2008. [Online]. Available: http://dx.doi.org/10.1561/0600000017 29

[104] R. Xu, Y. Hirano, R. Tachibana, and S. Kido, "Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, ser. Lecture Notes in Computer Science, vol. 6893. Springer Berlin Heidelberg, 2011, pp. 193–190. 29, 33

[105] ——, "A bag-of-features approach to classify six types of pulmonary textures on high-resolution computed tomography," *IEICE TRANSACTIONS on Information and Systems*, vol. E96-D, no. 4, pp. 845–855, April 2013. 29, 33, 34, 36, 41, 43, 78, 101

[106] J. Zepeda, E. Kijak, and C. Guillemot, "Sift-based local image description using sparse representations," in *Multimedia Signal Processing, 2009. MMSP '09. IEEE International Workshop on*, 2009, pp. 1–6. 30

[107] T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 881–892, 2002. 31

[108] C. Chang and C. Lin, "Libsvm: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1–27, May 2011. 32, 61

[109] V. A. Zavaletta, B. J. Bartholmai, and R. A. Robb, "High resolution multidetector ct-aided tissue analysis and quantification of lung fibrosis," *Academic radiology*, vol. 14, no. 7, pp. 772–787, July 2007. 34, 36

[110] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Workshop on Statistical Learning in Computer Vision, ECCV*, 2004, pp. 1–22. 36

[111] S. O. Park, J. B. Seo, N. Kim, Y. K. Lee, J. Lee, and D. S. Kim, "Comparison of usual interstitial pneumonia and nonspecific interstitial pneumonia: quantification of disease severity and discrimination between two diseases on hrct using a texture-based automated system." *Korean J Radiol*, vol. 12, no. 3, pp. 297–307, 2011. [Online]. Available: http://www.biomedsearch.com/nih/Comparison-Usual-Interstitial-Pneumonia-Nonspecific 41

[112] R. Yoon, J. Seo, N. Kim, H. Lee, S. Lee, Y. Lee, J. Song, J. Song, and D. Kim, "Quantitative assessment of change in regional disease patterns on serial hrct of fibrotic interstitial pneumonia with texture-based automated quantification system," *European Radiology*, vol. 23, no. 3, pp. 692–701, 2013. [Online]. Available: http://dx.doi.org/10.1007/s00330-012-2634-8 41

[113] M. Gangeh, L. Srensen, S. Shaker, M. Kamel, M. Bruijne, and M. Loog, "A texton-based approach for the classification of lung parenchyma in ct images," in *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2010*, ser. Lecture Notes in Computer Science, T. Jiang, N. Navab, J. Pluim, and M. Viergever, Eds. Springer

Berlin Heidelberg, 2010, vol. 6363, pp. 595–602. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-15711-0_74 56

[114] N. Kato, M. Fukui, and T. Isozaki, "Bag-of-features approach for improvement of lung tissue classification in diffuse lung disease," in *Proc.SPIE*, ser. Medical Imaging 2009: Computer-Aided Diagnosis, vol. 7260, Lake Buena Vista, February 2009, pp. 72 600C–72 600C–10. [Online]. Available: +http://dx.doi.org/10.1117/12.810976 56

[115] B. Zhao, G. Gamsu, M. S. Ginsberg, L. Jiang, and L. H. Schwartz, "Automatic detection of small lung nodules on ct utilizing a local density maximum algorithm," *J Appl Clin Med Phys*, vol. 4, no. 3, pp. 248–60, 2003. [Online]. Available: http://www.biomedsearch.com/nih/Automatic-detection-small-lung-nodules/12841796.htr 57

[116] P. Tao, F. Griess, Y. Lvov, M. Mineyev, B. Zhao, D. Levin, and L. Kaufman, "Characterization of small nodules by automatic segmentation of x-ray computed tomography images." *Journal of Computer Assisted Tomography*, vol. 28, no. 3, pp. 372–377, 2004. [Online]. Available: http://www.biomedsearch.com/nih/Characterization-small-nodules-by-automatic/151005 57

[117] K. T. Bae, J. S. Kim, Y. H. Na, K. G. Kim, and J. H. Kim, "Pulmonary nodules: Automated detection on ct images with morphologic matching algorithm–preliminary results," *Radiology*, vol. 236, no. 1, pp. 286–293, 2005. [Online]. Available: http://radiology.rsnajnls.org/cgi/content/abstract/236/1/286 57

[118] O. Osman, S. Ozekes, and O. N. Ucan, "Lung nodule diagnosis using 3d template matching," *Comput. Biol. Med.*, vol. 37, no. 8, pp. 1167–1172, Aug. 2007. [Online]. Available: http://dx.doi.org/10.1016/j.compbiomed.2006.10.007 57

[119] A. Farag, A. El-Baz, G. Gimelarb, R. Falk, and S. Hushek, "Automatic detection and recognition of lung abnormalities in helical ct images

using deformable templates," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2004*, ser. Lecture Notes in Computer Science, C. Barillot, D. Haynor, and P. Hellier, Eds. Springer Berlin Heidelberg, 2004, vol. 3217, pp. 856–864. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-30136-3_104 57

[120] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010. 57

[121] R. Bellotti, F. De Carlo, G. Gargano, S. Tangaro, D. Cascio, E. Catanzariti, P. Cerello, S. C. Cheran, P. Delogu, I. De Mitri, C. Fulcheri, D. Grosso, A. Retico, S. Squarcia, E. Tommasi, and B. Golosio, "A cad system for nodule detection in low-dose lung cts based on region growing and a new active contour model," *Medical Physics*, vol. 34, no. 12, pp. 4901–10, 2007. [Online]. Available: http://www.biomedsearch.com/nih/CAD-system-nodule-detection-in/18196815.html 57

[122] Y. Sato, C. fredrik Westin, A. Bhalerao, S. Nakajima, N. Shiraga, S. Yoshida, Z. Kikinis, S. Tamura, and R. Kikinis, "Tissue classification based on 3d local intensity structures for volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 6, p. 2000, 2000. 57, 59, 62, 73, 103

[123] Q. Li, S. Sone, and K. Doi, "Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans," *Medical Physics*, vol. 30, no. 8, pp. 2040–2051, Aug 2003. 57, 62, 63

[124] H. SHIKATA and S. KIDO, "A pattern classification system for abnormal attenuations caused by diffuse lung diseases in chest x-ray 3d ct images by means of texture analysis and svm," *The IEICE transactions on information and systems (Japanese edetion)*, vol. 91, no. 7, pp. 1895–1903,

jul 2008. [Online]. Available: http://ci.nii.ac.jp/naid/80019582082/en/ 76, 78, 79

[125] Y. Kusaka, K. G. Hering, and J. E. Parker, *International Classification of HRCT for Occupational and Environmental Respiratory Diseases*. Berlin: Springer, 2005. 100

Acknowledgements

First and foremost, I would like to show my greatest appreciation to Professor Shoji Kido, my advisor, for his excellent supervision, constructive suggestion, great patience and support during my research period. I could not have imagined having a better advisor for my study.

I received generous support from Dr. Rui Xu for his generous help and invaluable guidance both on the research and everyday life life.

I have greatly benefited from Associate Professor Yasushi Hirano and Dr. Rie Tachibana. Their instructions have been of great help to my study.

I would like to express my gratitude to all the members of the laboratory of Computer-aided Diagnosis and Biomedical Imaging Research Biomedical Engineering, Yamaguchi University.

Finally, thanks to my parents and my wife. Their love and encouragement always support me to overcome difficulties in my work.