PAPER Special Section on Medical Imaging

A Bag-of-Features Approach to Classify Six Types of Pulmonary Textures on High-Resolution Computed Tomography

Rui XU[†a)], **Yasushi HIRANO**[†], **Rie TACHIBANA**[††], ***and*** **Shoji KIDO**[†], ***Members***

SUMMARY Computer-aided diagnosis (CAD) systems on diffuse lung diseases (DLD) were required to facilitate radiologists to read high-resolution computed tomography (HRCT) scans. An important task on developing such CAD systems was to make computers automatically recognize typical pulmonary textures of DLD on HRCT. In this work, we proposed a bag-of-features based method for the classification of six kinds of DLD patterns which were consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), emphysema (EMP), nodular (NOD) and normal tissue (NOR). In order to successfully apply the bag-of-features based method on this task, we focused to design suitable local features and the classifier. Considering that the pulmonary textures were featured by not only CT values but also shapes, we proposed a set of statistical measures based local features calculated from both CT values and eigenvalues of Hessian matrices. Additionally, we designed a support vector machine (SVM) classifier by optimizing parameters related to both kernels and the soft-margin penalty constant. We collected 117 HRCT scans from 117 subjects for experiments. Three experienced radiologists were asked to review the data and their agreed-regions where typical textures existed were used to generate 3009 3D volume-of-interest (VOIs) with the size of 32×32×32. These VOIs were separated into two sets. One set was used for training and tuning parameters, and the other set was used for evaluation. The overall recognition accuracy for the proposed method was 93.18%. The precisions/sensitivities for each texture were 96.67%/95.08% (CON), 92.55%/94.02% (GGO), 97.67%/99.21% (HCM), 94.74%/93.99% (EMP), 81.48%/86.03%(NOD) and 94.33%/90.74% (NOR). Additionally, experimental results showed that the proposed method performed better than four kinds of baseline methods, including two state-of-the-art methods on classification of DLD textures.

key words: *pulmonary textures classification, diffuse lung diseases, bag-of-features, computer-aided diagnosis (CAD)*

1. Introduction

Diffuse lung diseases (DLD) refer to a heterogeneous group of abnormal conditions or diseases which spread out in large areas of lungs [1]. These diseases typically have similar clinical, radiographic, physiologic, and/or pathologic features. There are two common anatomically imaging modalities, X-ray chest radiography and computed tomography (CT), which are routinely used in the detection and diagnosis of DLD. With the development of techniques on CT, high-resolution computed tomography (HRCT) has played important roles in recent years [2]. Thin slice CT is able to give detailed appearances on pulmonary patterns. This facilitates the diagnosis of DLD, but a huge number of CT slices also requires radiologists to take much more time on reading. Additionally, objective identification of pulmonary textures has not being established in current clinical protocols. Since visual inspection is being carried on according to radiologists' experiences, subjective differences are inevitable. Considering both of the two aspects, a quantitative computer-aided diagnosis (CAD) system is required to give a second opinion to facilitate detection and diagnosis of DLD.

In the past ten years, there have been many researches to develop CAD systems to automatically analyze pulmonary textures on HRCT [3]. Usually, such a CAD system can be divided into two steps. First, a segmentation method is applied to extract the whole lung regions. Then, voxels inside the segmented regions are automatically assigned to labels which refer to the predefined classes of textural patterns. For the second step, it can be generalized as a pattern recognition problem of texture classification. Usually, a region/volume of interest (ROI/VOI) surrounding a under-processing pixel/voxel is constructed and classification methods are applied on it. In this paper, we only focus on how to develop a method to automatically classify pulmonary textures. Segmentation of lungs with severe diffuse disorders in recent years can be referred to the works [24], [25].

Researches on the classification of pulmonary textures can be found in a lot of works [4]–[16]. These methods could be grouped in the aspects of ROIs (VOIs), features, classifiers and textural types. In the majority of these works, 2D square-shaped ROIs are often adopted [4]–[11]. Recently, 3D cubic-shaped VOIs are used because textural patterns can be more discriminative [13]–[16]. An interesting work is given in the work [12], where irregular-shaped 2D ROIs clustered on homogenous textures show that they are more powerful than the 2D square-shaped ROIs. For the classifiers, all of the works adopt mature techniques, such as the Bayesian classifier [4], [13], [14], the artificial neural networks (ANN) [5], [9], the nearest neighbor (NN) classifier [6], [11], [12], [15], [16], and the support vector machine (SVM) [6]–[8], [10]. The types of pulmonary textures are usually determined by the purpose of CAD systems. For the identification of general DLD patterns, the number of textural types is relatively more. For example, five types are considered in [6], [16], six types are considered in [4], [7], [12] and seven types are considered in [5]. The definitions of the pulmonary textures are similar and differences

Manuscript received June 15, 2012.
Manuscript revised September 20, 2012.

[†]The authors are with the Graduate School of Medicine, Yamaguchi University, Ube-shi, 755–8505 Japan.
[††]The author is with the Information Science and Technology Dept., Oshima National College of Maritime Technology, Yamaguchi-ken, 742–2193 Japan.
a) E-mail: xurui@yamaguchi-u.ac.jp

DOI: 10.1587/transinf.E96.D.845

Copyright © 2013 The Institute of Electronics, Information and Communication Engineers

rely on the data or the purpose of clinical application. For CAD systems on a specific purpose, the number of textural type is relatively few. For example, only two types of textures (mild and normal) are considered for the detection of early interstitial lung diseases [9], two kinds of emphysema textures and normal textures are considered for chronic obstructive pulmonary disease (COPD) [10], [11], and a subset of three kinds of DLD textures is considered for interstitial pneumonia [15].

Almost all of the existing methods are dedicated to designing discriminate features for their problems. Feature extraction is a critical problem in pattern recognition. If features are more discriminative, classification becomes easier. Features can be designed by using one or several of the following measures on gray values [9], [15], such as statistical measures on histogram distributions, measures on gray-level co-occurrence matrices (GLCM) and measures on gray-level run-length matrices (GLRLM). More powerful features can be designed if these measures are combined with some other cues on geometrical information, such as fractal features [4], [13], [14] or shapes [7], [8]. There are also works [6], [12], [16], where the filter-bank based features are calculated from several scales and orientations and then sequential forward search is used to select the most discriminate features. Another work designs the compact features for the specific task on hands considering both of the gray-values and geometrical information [5]. There are also works considering invariant-features, such as rotation-invariant local binary patterns (LBP) are adopted to distinguish the normal tissues from two kinds of emphysema for COPD [11].

Recently there have been researches to apply new developed techniques in computer vision to design more discriminative features for the classification of pulmonary textures, taking examples of a bag-of-features technique [10] and a signature-matching technique [16] reported recently. Both of the two methods are originally used in image retrieval, i.e. [19] and [26] respectively. They are similar in methodology since both of them require clustering methods (i.e. *k*-means) to assign a large number of pre-extracted local features into some clusters. According to the numbers of features assigned to each cluster, a histogram is calculated and treated as a discriminative feature for classification. The difference is that the clustering is only operated once in the training stage for the bag-of-features method. However, the clustering is operated on each sample everytime for the signature-matching method. The bag-of-feature method is firstly applied in the work [10] to classify three kinds of pulmonary patterns (normal tissues, centrilobular emphysema and paraseptal emphysema) for COPD. The signature-matching method is applied to classify five kinds of pulmonary textures of DLD in [16].

In this paper, we try to develop a classification method for the general DLD textures, including the following six kinds of textures, consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), nodular (NOD), emphysema (EMP), and normal tissue (NOR). Their examples are given in Fig. 1. We tried both of the works [10], [16], however obtained unsatisfied results. The bag-of-feature based method is more efficient in practice than the signature-matching method, since it only needs the clustering once in the training stage. In this paper, we adopt the bag-of-features based method to resolve our problem. We improve several aspects in order to make it work properly for our task. Comparing with the work [10], the major differences rely in the following five aspects. First, our aim to classify pulmonary textures is for DLD rather than COPD. According to our knowledge, there is no work to apply the bag-of-features based method to recognize the general DLD textures. Second, we classify six kinds of textures, much more than the work [10]. Third, we apply 3D VOIs instead of 2D ROIs in order to get more discriminative features for classification. Fourth, we propose a new kind of local features considering both shapes and CT values according to textural characteristics. Fifth, we design a powerful SVM classifier by optimizing both kernels and soft-margin-related constant.

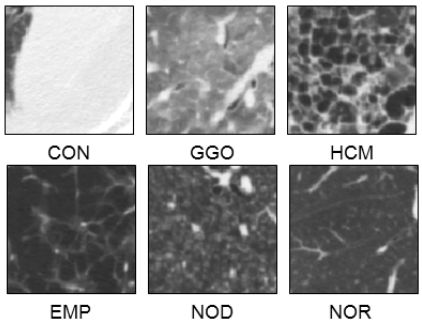


Fig. 1 Examples of six kinds of pulmonary textures to be classified in HRCT. They are consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), nodular (NOD), emphysema (EMP), and normal tissue (NOR).

A preliminary version of this work was published [17]. The major differences against the preliminary version rely on two aspects. First, we adopt a different experimental setting. Here, we separate samples into the training and testing sets. We train and optimize parameters of methods on the training set, and then test the optimally trained method on the testing set. The way that training and testing samples are completely blind to each other is more reasonable than the 20-fold cross validation test adopted in the previous version. Second, we compare the proposed method with other two kinds of the state-of-the-art methods [5], [16] in this paper.

2. The Proposed Method

We proposed a bag-of-features based method to classify the six typical pulmonary textures on HRCT. A bag-of-features framework can be divided into three parts, local features calculation, training of a code-book, and training of a classifier. Compared to the work [10], the major differences in methodology rely on the local features and the classifier. In

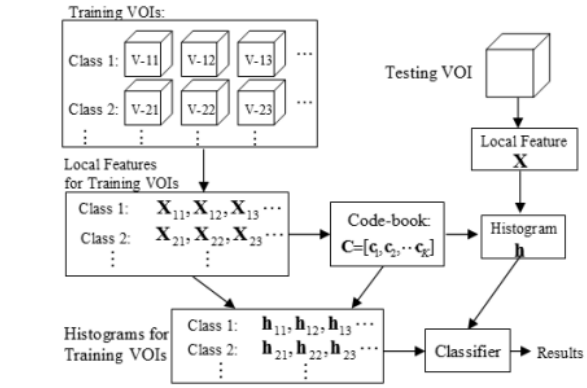
this section, we will describe the proposed method in details. In order to make it more readable, we firstly introduce the bag-of-feature framework. Then, we describe the proposed local features and classifier's training in the following two subsection.

2.1 Bag-of-Features Approach

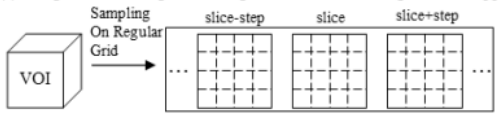
The bag-of-features framework can be expressed by Fig. 2 (a). We separately describe the training and testing stages of this framework. In the training stage, there are two kinds of training, the training of a code-book and a classifier. The input of training is a set of VOIs, which is denoted by V_{ij} to represent the j-th VOI of the i-th class. For each VOI, a large number of local features are calculated. Here, $\mathbf{X}_{ij}$ means the set of local features calculated from the j-th VOI of the i-th class. The features of different VOIs for all classes are mixed together, and then a set of clusters' centers is calculated by a clustering method. These centers are called the codebook denoted as $\mathbf{C} = [\mathbf{c_1}, \mathbf{c_2}, \ldots, \mathbf{c_k}]$, where k is the number of the clusters. It should be noted that the input of the clustering can be local features of all of training VOIs or just a subset of the training VOIs. Usually, this can not change results much. When the code-book is obtained, local features of each training VOI can be assigned into their nearest cluster's centers. The number of local features assigned to each cluster gives a histogram for each VOI to show the frequency of clusters' centers that local features belong to. The histogram is denoted by $\mathbf{h}_{ij}$ for the training VOI V_{ij}. Finally, all of the histograms $\mathbf{h}_{ij}$ are treated as discriminative features to train a classifier. In the testing stage, local features are firstly calculated for a testing VOI. Then, the pre-calculated codebook is used to assign local features to each cluster in order to calculate the histogram $\mathbf{h}$. Finally, the histogram is fed into the trained classifier to get a recognized result. It should be noticed that the codebook is only trained once in the training stage.

The main idea of bag-of-features approach is to represent the VOIs by a histogram whose bins are the elements of the code-book trained from local features. Such a codebook can be seen as an intermediate layer to interpret VOIs. Since local features are calculated from limited local regions, each of them only reflects information of partial objects. Although local features are huge, they are usually clustered into limited number of centers in high-dimensional feature spaces. Just as words are basic elements of sentences, such centers of clusters can also be considered as the basic elements to represent the underlying objects. Codebook construction can be generalized as a vector quantization problem. Some clustering methods, such as k-means, are usually adopted to resolve this problem. The number of clusters (k) is the parameter determined by experiments.

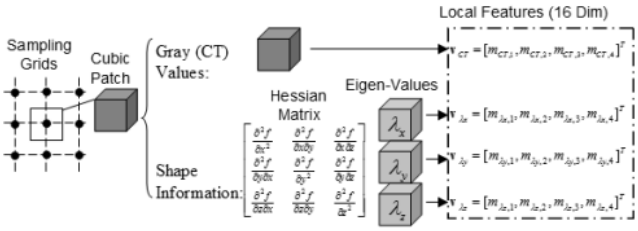
Local features can be calculated at the regularly sampling grids or at the pre-determined key-point positions. According to the literature [18], there seems to be no evidences to show which one is better. In this work, we prefer to adopt the way of regularly sampling grids, since it is easier and



(a). Diagram of training and testing Procedures in the bag of words approach.



(b). Local features calculated at regular sampling grids in each VOI.



(c). Local features based on gray (CT) values and shape information.

Fig. 2 The proposed bag-of-features based method.

faster. Figure 2 (b) illustrates the regular sampling grids used in this work. In the preliminary work [17], we adopted the VOI's size and the grid step to be 32×32×16 and 4×4×4 respectively, and obtained $9 \times 9 \times 5 = 405$ sampling points. We consider that it would be more reasonable to use an iso-sized VOI, so we set the VOI size to be 32×32×32 in this work. Since we use the same grid step, we can get can get 9×9×9=729 sampling points on each VOI. Local features are calculated inside local regions, called patches, whose centers are located at these sampling points. The patch size is the parameter determined by experiments, either.

In our opinion, the bag-of-features approach is a common framework. According to different recognition tasks, the three basic parts, local feature calculation, code-book construction and classifier should be adjusted in order to make it achieve better results. In the proposed method, we use k-means method to train the code-book, just the same as the work [10]. However, we make the improvements on the calculation of local features and the training of classifiers in order to make it work properly for our task. In the next two subsection, we will describe these two parts.

2.2 Local Features

CT values of a 2D squared patch are rearranged as a vector to be treated as local features for 2D ROIs in [10], It is reported that such local features achieved good results for the recognition of normal tissues and two kinds of emphy-

sema patterns. We tried this method, but the results were not satisfied. Therefore, we reconsider about our problem and design a new kind of local feature which is suitable to our task.

The categories of pulmonary textures in our work are much more. Some textural patterns, such as NOR, CON and EMP, are mainly characterized by CT values. However, other patterns, such as NOD, are mainly featured by 3D shape information. Since 3D shape information can not be visualized from 2D slices, 3D VOIs are adopted in this work. Therefore, local features are calculated from 3D cubic-shaped patches instead of 2D square-shaped patches. Combining both of the CT and shape information, we proposed a new kind of local features, whose calculation can be represented by Fig. 2 (c). At each sampling point, we can construct a cubic-shaped patch whose center is that sampling point. Each patch is a 3D volume, whose size is a parameter. Supposing CT values in a patch at a position $\mathbf{x} = [x, y, z]^T$ can be represented by $f(\mathbf{x})$. We calculate four kinds of statistical measures according to Eq. (1) as the features reflecting the information of CT values, denoted as $\mathbf{v}_{CT} = [m_{CT,1}, m_{CT,2}, m_{CT,3}, m_{CT,4}]^T$. These four kinds of statistical measures are mean (m_1), standard deviation (m_2), skewness (m_3) and kurtosis (m_4).

$$\begin{cases} m_1 = \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_i) \\ m_2 = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(f(\mathbf{x}_i) - m_1)^2} \\ m_3 = \frac{1}{N}\sum_{i=1}^{N}(\frac{f(\mathbf{x}_i)-m_1}{m_2})^3 \\ m_4 = \frac{1}{N}\sum_{i=1}^{N}(\frac{f(\mathbf{x}_i)-m_1}{m_2})^4 - 3 \end{cases} \tag{1}$$

where N is the number of voxels in the patch.

Additionally, we use the eigen-values of Hessian matrix to calculate another kind of features to represent the information of shapes. A Hessian matrix at a position $\mathbf{x} = [x, y, z]^T$ can be calculated by Eq. (2).

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix} \tag{2}$$

The 3×3 matrix $\mathbf{H}(\mathbf{x})$ has 3 eigen-values, which are $\lambda_1, \lambda_2, \lambda_3$ listed in a descending order. For each position of a patch volume, we can calculate these three kinds of eigen-values. Arranging them by their positions' order, we can get three volumes whose component are the eigen-values. Calculating the same four kinds of statistical measure according to Eq. (1) respectively on the three volumes, we can obtain 3 vectors as the features reflecting information of shapes. The 3 vectors are denoted as $\mathbf{v}_{\lambda 1} = [m_{\lambda 1,1}, m_{\lambda 1,2}, m_{\lambda 1,3}, m_{\lambda 1,4}]^T$, $\mathbf{v}_{\lambda 2} = [m_{\lambda 2,1}, m_{\lambda 2,2}, m_{\lambda 2,3}, m_{\lambda 2,4}]^T$ and $\mathbf{v}_{\lambda 3} = [m_{\lambda 3,1}, m_{\lambda 3,2}, m_{\lambda 3,3}, m_{\lambda 3,4}]^T$ respectively. Catenating $\mathbf{v}_{CT}$, $\mathbf{v}_{\lambda 1}$, $\mathbf{v}_{\lambda 2}$ and $\mathbf{v}_{\lambda 3}$, we get a 16 dimensional feature, denoted as $\mathbf{v} = [\mathbf{v}_{CT}^T, \mathbf{v}_{\lambda 1}^T, \mathbf{v}_{\lambda 2}^T, \mathbf{v}_{\lambda 3}^T]^T$, which is able to reflect the information of both CT values and shapes.

For the proposed local features, the only parameter is the size of cubic-shaped patches. This parameter should be optimized in the training stage. No matter how the size of patches varies, the compactness of such local features does not change. In another words, the local feature is always a 16 dimensional vector. Additionally, these statistical measures are invariant to translation and rotation.

2.3 Classifier

We choose the SVM classifier in the proposed method. The SVM classifier finds a hyperplane which separates two-class data with maximal margin [27]. In order to deal with the outliers in training samples, a soft-margin constraint term is usually added into the original cost function of SVM. This soft-margin term is penalized by a constant C, which is the parameter of the maximal soft-margin classifier. Additionally, the training samples are usually not linear. In order to make the classifier to be robust to nonlinear data, kernel methods are introduced to make it to be a kernelized maximal soft-margin classifier. An usual choice of kernels is the Gaussian kernel, which is adopted in the work [10]. According to the recent report [22] in computer vision, some other kernels are superior to the Gaussian kernel only if sample-vectors are histograms. In the proposed method, the histograms of VOIs are treated as discriminative features for the classifier training, so we try both the Gaussian kernel and the kernels recommended by the work [22]. The kernel definitions are given by Eq. (3).

$$\begin{cases} G(\mathbf{h}, \mathbf{h}') = \exp(-\gamma \|\mathbf{h} - \mathbf{h}'\|^2) \\ K_1(\mathbf{h}, \mathbf{h}') = \sum_{i=1}^{M} \min(h_i, h_i') \\ K_2(\mathbf{h}, \mathbf{h}') = \sum_{i=1}^{M} \sqrt{h_i h_i'} \\ K_3(\mathbf{h}, \mathbf{h}') = \exp(-\alpha \sum_{i=1}^{M} \frac{(h_i - h_i')^2}{h_i + h_i'}) \end{cases} \tag{3}$$

where $\mathbf{h} = [h_1, \ldots, h_M]^T$ is the histogram with M-bins. $G(\mathbf{h}, \mathbf{h}')$ is the *Gaussian* kernel, $K_1(\mathbf{h}, \mathbf{h}')$ is the *histogram intersection* kernel, $K_2(\mathbf{h}, \mathbf{h}')$ is the *Bhattacharyya* kernel, $K_3(\mathbf{h}, \mathbf{h}')$ is the χ^2-kernel. γ is the parameter for the gaussian kernel and α is the parameter for the χ^2-kernel.

The original SVM classifier is designed for the two classes classification problems. Here, we choose one-against-one approach [28] to make it to be used for the multi-classes classification problems. For the training data from the i-th and the j-th classes, we first resolve the two-classes classification problem. Then a voting strategy is used to assign the data to the class which gets the maximal votes. The details can be referred to the work [21].

It can be seen that our SVM classifier includes three kinds of parameters, which are the soft-margin penalty C, the types of kernels, and the parameters related to kernels. These parameters should be optimized in the training stage.

3. Four Baseline Methods

The propose method is compared with four kinds of baseline methods, which are called CTV-2D, CTV-3D, CanSigEMD

and SpeDesFea in this paper.

CTV-2D is the method proposed in [10]. Although it is originally not applied for the recognition of general DLD textures, since it is also a bag-of-features based method we compare it with the proposed method. The major differences between the two methods rely in the following three aspects. First, 2D ROIs are used in CTV-2D, so patches are 2D squares. Second, local features are a vector whose components are CT values of all pixels in 2D square-shaped patches. Third, only the Gaussian kernel is considered for the SVM classifier based on the original work. Since CT values and 2D ROIs are considered in this method, we call it CTV-2D.

CTV-3D is an improved version for CTV-2D. It is a bag-of-feature based method, either. We improve this method in the following two aspects. First, 3D VOIs are adopted in stead of 2D ROIs, so patches are 3D cubic volumes, the same as the proposed method. The other improvement is about the SVM classifier. We try all the four kinds of kernels shown by Eq. (3). Intensities of all voxels inside 3D patches are arranged as a vector as the local feature for CTV-3D, so its local feature is actually the same as the CTV-2D method. We call this improved version as CTV-3D, since 3D VOIs are used and only CT values are considered for local features.

CanSigEMD is a signature-matching based method implemented according to the work [16], which is a state-of-the-art recognition method on DLD textural patterns. From the aspect that k-means is operated on local features to calculate clusters' centers and local features are assigned into each cluster center to calculate histograms as discriminative features, CanSigEMD is somewhat similar as the bag-of-features based methods, however they are very different in methodology. The major difference is that a signature-matching based method requires to do clustering on each VOI (sample) in every time to obtain different clusters' centers; however, the bag-of-features based methods only do k-means once in the training stage to get the same clusters' centers. In another words, the clustering in CanSigEMD is adaptive to each VOI, and it should be performed in both of training and testing stage for each VOI. Therefore, both of histograms and clusters' centers should be considered in classification for CanSigEMD. The combination of histograms and clusters' centers is called a signature. According to [16], signatures are fed into a nearest neighbor (NN) classifier in CanSigEMD. Earth mover's distance (EMD) is considered to measure the distance of two signatures [26]. In order to save computing cost of comparing the signature of a test VOI against all signatures of training VOIs, signatures of the same class are re-clustered by k-means to generate one canonical signature for this class in [16]. Therefore only the canonical signature of each class should be compared to the signature of a test VOI in the NN classifier. In CanSigEMD, intensities of each voxel in a VOI is treated as the 1-dimensional local features. Since it uses canonical signatures and EMD for classification, we call it CanSigEMD in this paper.

SpeDesFea is a well-known method for DLD texture classification, originally proposed in [5]. According to our knowledge, this work is successfully applied to classify the most number of categories (7 categories) for DLD textures in current literatures. Additionally, this work inspires us to start our research in DLD textural classification. Based on these two reasons, we treat it as the state-of-the-art work and compare it with the proposed method. Although there are some other works which can be seen as the state-of-the-art methods, such as the work in [12]. Since irregular-shaped 2D ROIs are adopted in [12], we can not directly compare it with the proposed work. SpeDesFea operates on 2D ROIs. It adopts six kinds of specially designed features, three of which are related to the gray values and the others are related to geometric patterns. The gray values measures are the mean and the standard deviation of CT values, and also the fraction of the area with density components. The geometric measures are the mean of CT values for the extracted nodular, line and multilocular components. In the original method, an three-layered artificial neural networks (ANN) with a back-propagation algorithm is adopted as the classifier. Here, we use the SVM classifier instead of ANN. Since both SVM and ANN are mature techniques of classifier, this change will not essentially affect results of the original work. We can only use the Gaussian kernel for the SVM classifier because the six-measure-features are not histograms. Since it uses specially designed features, we call it SepDesFea in this paper.

4. Experiments

4.1 Data

We collect 117 scans from 117 different subjects from Tokushima University Hospital, Tokushima, Japan. All scans are acquired from 16-row multi-slice CT (Aquilion, Toshiba Co.) based on non-edge-enhanced kernel. A tube voltage of 120 kVp and current of 250 mAs are used. The image data are reconstructed to 512×512 matrices. Slice thickness is 1 mm, and the in-plane resolution is about 0.6 mm, with differences depended on patients' sizes. In the 117 scans, there are 6 scans with normal tissues in lungs and 111 scans with pulmonary disorders. The use of the data in this study has been approved by the local ethics committee.

The regions of the six types of patterns are marked by three experienced radiologists according to the following procedure. First, one radiologist is asked to review all of the scans. From each scan, the maximum of three axial slices with typical textural patterns are selected from the top, middle and bottom parts of lungs. In order to avoid the confusion of multi-textures existence on the same slice, this radiologist only can select the slice where one kind of textural patterns dominantly exists and also should indicate what the pattern is. Then the rest two radiologists are asked to review the results of the first radiologist, and only the agreed slices are saved for the next process. Second, the three radiologists are asked to separately mark the regions where the identified

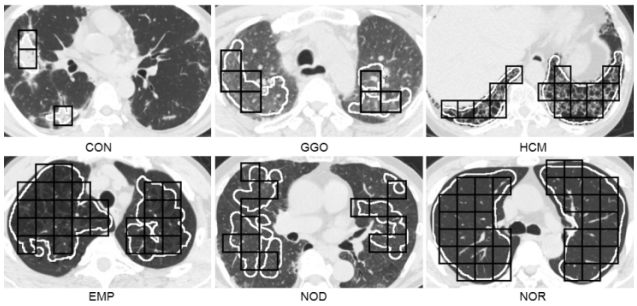


Fig. 3 The regions with typical textures agreed by three experienced radiologists and examples of the cental-axial slices of 3D-VOIs with the size of $32 \times 32 \times 32$. The regions agreed by three experienced radiologists are marked by white contours. The central-axial slices of 3D-VOIs are marked by black boxes. We accepted VOIs including regions outside the white contours and lungs in order to increase more variations to develop a robust recognition method. The six kinds of textures are consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), nodular (NOD), emphysema (EMP), and normal tissue (NOR).

textures are located by using a drawing tool on the selected slices. Finally, the regions commonly agreed by the three radiologists are extracted as where the six kinds of textures are located. These regions are saved as slice-masks where non-zero intensities of the voxels indicate the textures' types.

The VOIs used in experiments are generated from the slice-masks. 32×32 grids are overlaid on them and square-shaped patches whose centers cover non-zero pixels are selected. The selected patches are treated as the central-axial slices of $32 \times 32 \times 32$ VOIs. Figure 3 gives examples to show the regions agreed by the three radiologists and the generated VOIs. It can be seen that we accept VOIs including regions outside the white contours (regions agreed by the three radiologists). Even there are VOIs including regions outside lungs. This is because that we hope to collect VOIs with variation as much as possible in order to develop a robust algorithm to recognize pulmonary textures. Finally, we obtain 3009 3D VOIs for the six kinds of textures.

4.2 Experimental Setting and Tuning Parameters

In this paper, the experimental setting is totally different from our preliminary work [17]. The differences rely on both of the complete separation of training and testing data sets and the ways of tuning parameters. In the new experimental setting, the 3009 VOIs are separated into two sets by splitting each kind of texture nearly in half. One set is treated as the training data set, which is firstly used for tuning parameters of each method and then training each method by the optimized parameters. The other set is treated as the testing data set to evaluate their performances. Table 1 gives the numbers of VOIs of each type of textures for the training and testing data sets. It should be noticed that VOIs in the two sets belong to different subjects which means there are no crossing-subjects in the two sets. In parameters' optimization, 20-fold cross-validation (CV) tests are used. We randomly divide training samples into 20 groups, each of which has similar relative ratios on samples of each class. In the preliminary work [17], we optimized parameters by a fix-one-then-turning-one method, which meant that we only adjust one parameter in each time while setting the others to be fixed values. For example, we initially adjust the first parameter by setting the others to be any certain values; when we fix the optimal value of the first parameter we use it and adjust the second one. This process is operated in turns until all parameters are optimized. However, this fix-one-then-tuning-one method can not test all combination of the parameters and it could miss the optimal values of parameters. Therefore, we give up this method and try the combinations of all possible values of parameters in the 20-fold CV tests to adjust parameters for each method. The combination which gives the best overall recognition rate is chosen as the optimized parameters. By using the optimized parameters, the training set data is used again to train each method and then each method is evaluated on the testing data set.

Table 1 Numbers of 3D-VOIs for the six types of textures in the training and testing set. There are no cross-subjects VOIs in the two sets. The six kinds of textures are consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), nodular (NOD), emphysema (EMP), and normal tissue (NOR).

	Training Set	Testing Set
CON	63	61
GGO	260	251
HCM	272	254
EMP	354	383
NOD	181	179
NOR	384	367
Total	1514	1495

We change the experimental setting according to the following two aspects. First, the completely separation of VOIs into two sets will make the training and testing sets totally blind with each other. This is more similar as practical applications, so the evaluation of the performances for each method is more reasonable. Second, since we try all combination of possible values of parameters for each method, it is able to make each method to achieve its best performance.

Two kinds of baseline methods, CTV-2D and SepDes-Fea, require 2D ROIs. In order to make the generated 3D VOIs to be used for them, the central slices in the axial direction are chosen as the 2D ROIs to evaluate these two methods.

The details of the training for each method are described as follows. For CTV-3D and the proposed methods, the parameters are the same, including size of cubic-patches, the clusters' number and the parameters related to SVM classifier. Here, the possible values for the size of cubic-patches are from $2 \times 2 \times 2$ to $8 \times 8 \times 8$ with a step of $1 \times 1 \times 1$, and possible cluster's numbers are from 50 to 400 with a step of 50. For SVM classifiers, we adjust parameters related to kernels and soft-margin constant. We set possible values for the soft-margin constant to be $2^0, 2^1, \ldots, 2^{10}$. Only the gaussian kernel and the χ^2-kernel have parameters, which are γ and α. We set their possible values to be $2^{-10}, 2^{-9}, \ldots, 2^1$. Parameters of CTV-2D is similar to

Table 2 Optimized parameters for each method. CTV-2D is the method proposed in [10]. CTV-3D is an improved version of CTV-2D. CanSigEMD is the method proposed in [16], and SpeDesFea is implemented according to [5].

Methods	Optimized Parameters
CTV-2D	squared-patch size = 3×3, clusters' number = 350, SVM: *Gaussian*-kernel, $\lambda = 0.031, C = 2.0$
CTV-3D	cubic-patch size = $4 \times 4 \times 4$, clusters' number = 350, SVM: χ^2-kernel, $\alpha = 2.0, C = 4.0$
CanSigEMD	clusters' number = 20
SpeDesFea	SVM: *Gaussian*-kernel, $\lambda = 0.125, C = 1024$
Proposed Method	cubic-patch size = $3 \times 3 \times 3$, clusters' number = 200, SVM: χ^2-kernel, $\alpha = 0.5, C = 4.0$

CTV-3D and the proposed method. The differences are that squared-patch sizes are from 2×2 to 8×8 and the Gaussian-kernel is only used for SVM classifier. It should be noticed that there are two kinds of training in the three bag-of-features based methods, k-means clustering and training of SVM classifiers. Since k-means training takes too much time in the whole 20-fold CV tests, we only train it once by random selecting 50 VOIs for each class before 20-fold CV tests start. According to our experiences, this way can not change results much. CanSigEMD has only one parameter, the clusters' number. According to the work [16], the value should be not too large, so we adjust it from 5 to 50 with a step of 5. For SpeDesFea, we adjust the parameters related to the Gaussian-kernel based SVM classifier. Possible values of γ and the soft-margin constant are set to be the same as the proposed method. The optimized parameters for each method are listed by Table 2.

4.3 Results

Using the optimized parameters, shown by Table 2, we evaluate the performance of each method on the testing data set. We compare the proposed method with the four baseline methods by using three kinds of measures, which are the overall accuracy ($\frac{\text{number of corrected classifications}}{\text{total number}}$), precision ($\frac{\text{true positive}}{\text{true positive + false positive}}$) and sensitivity ($\frac{\text{true positive}}{\text{true positive + false negative}}$). The confusion tables, precisions and sensitivities for the five methods are given by Table 3. The comparison of overall accuracy is given by Table 4(a). Additionally, we do the McNemar's tests to compare the statistical differences between the proposed method and the four kinds of baseline methods. The p values for the four baseline methods against the proposed method are given by Table 4(b). All of the p values are less than 0.0001, which means the statistical differences are significantly high.

4.4 Discussion

According to these results, it can be seen that the performance of the proposed method is better than the four baseline methods. CTV-3D is the baseline method which is closest to the performances of our method. Additionally, it is the most similar to our method in methodology, either, only with the difference on local features. Local features of the proposed method make use of information of both CT values and shapes; however, CTV-3D only consider the information of CT values. This makes the proposed method more powerful to classify the six kinds of textures. According to the Table 3(b) and the Table 3(e), it can be seen that CTV-3D is good at classifying textures of CON, GGO, HCM and EMP; however it can not differentiate NOD and NOR very well. NOD is featured by tiny and spread focal-opacities, dominated by shape information. CT values is not discriminative enough to distinguish NOD from NOR. Since we consider both of the information of CT values and shapes, the proposed method is able to distinguish the two kinds of textures better.

CTV-2D also achieves relatively better results in the four base-line methods. Although its performance on CON is similar to CTV-3D and the proposed method, the performances on the other five textures are worse. CTV-2D is different from the proposed method in three aspects, which are 2D-ROIs, CT values based local features and Gaussian kernel based SVM classifier. The information only on CT values and the 2D-ROIs make CTV-2D can not obtain enough powers to distinguish the DLD patterns. Additionally, only the Gaussian-kernel based SVM classifier leads to a little worse performance, either. According to our experiences, other kernels, especially the χ^2-kernel, sometimes can increase 1-2% on the recognition rate.

Unfortunately, CanSigEMD gives the worst results in the four baseline methods. We consider there may be three reasons. First, the six kinds of texture classification problem is more difficult than the problem in the work [16], where only five kinds of textures are considered. Especially, the addition of the texture NOD makes the problem much harder, since it is very easy to be confused with EMP and NOR. Second, as we mentioned in several times, it can be difficult to get discriminative features by only considering the information of CT values. This can be demonstrated by the confusion table of this method given by Table 3(c). It can be seen that some textures featured by CT values (i.e. CON and EMP) achieve relatively better results. The third reason is that the nearest neighbor classifier is relative naive to handle this difficult problem compared with the SVM classifier. Additionally, CanSigEMD needs to do clustering on each ROI on both of the training and testing stages, however, the clustering is only carried on the training stage in our method. Therefore, our method is more efficient than CanSigEMD in practice.

SpeDesFea is implemented according to the state-of-the-art work [5], which is reported that 7 kinds of DLD textures are recognized well. However, it does not give satisfied results in our evaluation. In the aspect of methodology, it is quite different from our method. Although we use a SVM classifier to replace the ANN classifier suggested by the original work [5], this should not change performances much since both of them are mature techniques. We consider that the following two reasons may explain why it performs worse in our experiments. The first one is related to features. In SepDesFea, six kinds of compact measures

Table 3 Comparison on confusion tables, precisions and sensitivities for the proposed method and the four baseline methods. CTV-2D is the method proposed in [10]. CTV-3D is an improved version of CTV-2D. CanSigEMD is the method proposed in [16], and SpeDesFea is implemented according to [5]. The six kinds of textures are consolidation (CON), ground-glass opacity (GGO), honeycombing (HCM), nodular (NOD), emphysema (EMP), and normal tissue (NOR).

(a) Confusion table of CTV-2D

	Estimated Labels						
True Labels	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	58	0	3	0	0	0	95.08%
GGO	1	214	9	0	5	22	85.26%
HCM	0	0	244	10	0	0	96.06%
EMP	0	2	5	343	27	6	89.56%
NOD	0	3	1	19	141	15	78.77%
NOR	0	0	0	15	49	303	82.56%
Precision	98.31%	97.72%	93.13%	88.63%	63.51%	87.57%	

(b) Confusion table of CTV-3D

	Estimated Labels						
True Labels	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	58	1	2	0	0	0	95.08%
GGO	2	236	2	1	3	7	94.02%
HCM	0	0	251	3	0	0	98.82%
EMP	0	11	2	357	8	5	93.21%
NOD	0	6	0	31	132	10	73.74%
NOR	0	3	0	5	47	312	85.01%
Precision	96.67%	91.83%	97.67%	89.92%	69.47%	93.41%	

(c) Confusion table of CanSigEMD

	Estimated Labels						
True Labels	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	56	0	5	0	0	0	91.80%
GGO	10	137	43	0	59	2	54.58%
HCM	22	39	145	10	38	0	57.09%
EMP	0	0	20	307	31	25	80.16%
NOD	0	1	5	28	87	58	48.60%
NOR	0	0	18	132	51	166	45.23%
Precision	63.64%	77.40%	61.44%	64.36%	32.71%	66.14%	

(d) Confusion table of SpeDesFea

	Estimated Labels						
True Labels	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	58	2	1	0	0	0	95.08%
GGO	1	211	7	1	6	25	84.06%
HCM	2	7	207	24	8	6	81.50%
EMP	0	7	5	326	21	24	85.12%
NOD	0	7	12	35	98	27	54.75%
NOR	0	15	8	38	34	272	74.11%
Precision	95.08%	84.73%	86.25%	76.89%	58.68%	76.84%	

(e) Confusion table of the proposed method

	Estimated Labels						
True Labels	CON	GGO	HCM	EMP	NOD	NOR	Sensitivity
CON	58	3	0	0	0	0	95.08%
GGO	2	236	4	0	4	5	94.02%
HCM	0	0	252	2	0	0	99.21%
EMP	0	8	2	360	7	6	93.99%
NOD	0	5	0	11	154	9	86.03%
NOR	0	3	0	7	24	333	90.74%
Precision	96.67%	92.55%	97.67%	94.74%	81.48%	94.33%	

Table 4 Comparison on recognition accuracy and statistical differences for the proposed method and the four baseline methods. CTV-2D is the method proposed in [10]. CTV-3D is an improved version of CTV-2D. CanSigEMD is the method proposed in [16], and SpeDesFea is implemented according to [5].

(a) Recognition accuracy

Methods	Accuracy
CTV-2D	87.15%
CTV-3D	90.03%
CanSigEMD	60.07%
SpeDesFea	78.39%
Proposed	93.18%

(b) Statistical differences by McNemar's tests

Methods	P Values
Proposed vs. CTV-2D	< 0.0001
Proposed vs. CTV-3D	< 0.0001
Proposed vs. CanSigEMD	< 0.0001
Proposed vs. SpeDesFea	< 0.0001

are calculated as features, which are three measures on CT values and and three measures on geometrical information. According to the work [5], the CT values based measures can be straightly calculated; while the geometrical measures should be calculated on extracted nodular, line and multi-locular components. The extraction of these components is not a trivial work. Not only the image noise but also the complex opacities can lead the extraction results unreasonable. The extraction based method may be work well on images with mild DLD textures, however, we doubt its performances on images with severe diffuse lung diseases. In our method, the local features are calculated directly from the original images. Since it does not depend on any pre-extracting results, it is more robust for the recognition of DLD patterns. The second reason is related to VOIs. A condition that the processed textures should take more than 70% area of 2D-ROIs is considered in the work [5]. In our work, we consider to choose a VOIs' central slice whose center should be covered by the processed textures because we hope to add more variations in the training and testing in order to develop a robust method. Therefore, the ROIs considered in this work are much harder.

Experimental results show that the proposed method performs best. The reason relies on the successfully application of the bag-of-features framework on this DLD textural classification problem. The power of the bag-of-features

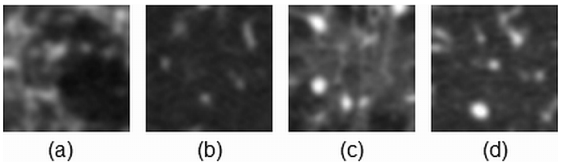


Fig. 4 Some examples of VOIs misclassified among NOD, EMP and NOR. (a). An example of NOD is misclassified to be EMP. (b). An example of NOD is misclassified to be NOR. (c). An example of EMP is misclassified to be NOD. (d). An example of NOR is misclassified to be NOD.

method is that it interprets a VOI of DLD textures by a histogram whose bins are the clusters of a set of local features. This new way has a underlying statistical advantage to interpret a VOI, so it is more robust than conventional methods which directly calculate a feature vector on a VOI. The superior characteristics of the bag-of-feature method is the first aspect to ensure the proposed method works best. Second, the proposed local features and adjustment of parameters related to the SVM classifier make the bag-of-features method achieve the best performances. Since the proposed local features make use of CT values and shapes information both of which are the typical characters of DLD textures, they are able to give more discriminative histogram to facilitate the classification. The adjustment of different types of kernels and the soft-margin penalty constant can make the SVM classifier to achieve its best power. Therefore, the proposed method is superior to the other four kinds of methods in experiments.

The sensitivity and precision are beyond 90% for the textures of CON, GGO, HCM, EMP and NOR by using the proposed method. However, the performance on NOD is relatively worse. Especially, the misclassification are mainly among NOD, EMP and NOR in the proposed method. Figure 4 gives some examples of VOIs on this kind of misclassification. Figure 4 (a) and Fig. 4 (b) are examples of NOD misclassified to be EMP and NOR respectively. It can be seen that both of them are not typical VOIs of NOD. A typical VOI of NOD has the character of spreading tiny nodular opacities, shown by Fig. 3. Since an large emphysema-like region exists in the right-bottom part of Fig. 4 (a), this could be why it is misclassified to be EMP. For the VOI shown by Fig. 4 (b), since there are only few nodular opacities, it is misclassified to be NOR. Although we use a strict protocol to generate typical VOIs of different textures only in regions agreed by three experienced radiologists, because there are lots of variations in DLD textures it could be difficult to judge where the typical textures exist sometimes, especially for the most difficult texture of NOD. Therefore, a post-check of generated VOIs by radiologists may be required in future research to ensure that all generated VOIs are typical ones. Figure 4 (c) and Fig. 4 (d) are the examples of EMP and NOR which are misclassified to be NOD respectively. It can be seen that the crossing-points of vessels can be viewed clearly on these two cases. This could be the reason why they are misclassified to be NOD, so the elimination of vessels inside lungs could be required in future research.

Finally, it should be noticed that overall recognition rate reported in this paper is a little lower than what we reported in the preliminary work [17] because the new experimental setting makes the training and testing set totally blind with each other. However, this can not change the conclusion that the proposed method achieves good results in the recognition of the six kinds of textures.

5. Conclusion

We proposed a bag-of-features approach to automatically classify six kinds of typical pulmonary textures on HRCT for DLD. Some textures, such as CON, NOR and EMP, are mainly characterized by CT values; while for other patterns, such as NOD and HCM, are mainly featured by shape information. Considering the two characters, we proposed a new kind of local features for our bag-of-features approach. The features were mainly based on four kinds of statistical measures, which were mean, standard deviation, skewness and kurtosis. Because they were calculated from both the original CT values and eigenvalues of Hessian matrices, they were able to combine both of the information on CT values and shapes. Therefore, the new local features were discriminative to distinguish the six kinds of patterns. The proposed method was trained (1514 ROIs) and tested (1495 ROIs) on two completely-separated data sets. The experiments showed that the overall recognition accuracy for the proposed method was 93.18%. The values of precision/sensitivity for each texture were 96.67%/95.08% (CON), 92.55%/94.02% (GGO), 97.67%/99.21% (HCM), 94.74%/93.99% (EMP), 81.48%/86.03%(NOD) and 94.33%/90.74% (NOR). We also implemented four kinds of baseline methods, including two implemented version of the state-of-the-art methods [5], [16], in order to compare with the proposed method. According to experimental results, our method was superior to the four methods. Additionally, the McNemar's tests showed the proposed method was statistically different from the four baseline methods, either. Therefore, we concluded that the proposed method achieved good results for the classification of the six kinds of textures and this technique would be useful for a CAD system on diffuse lung diseases.

Although this method performed well, we noticed that the misclassification were mainly among NOD, EMP and NOR. We will try to resolve this problem in our future work. A post-check on VOIs for typical textures, a pre-processing to eliminate vessels and the adoption of more powerful local features (i.e. SIFT [23]) would be good ideas to resolve this problem. Additionally, reticular with GGO is a typical textures of DLD, either. Since we only collect fewer cases for this texture, we does not consider it in this paper. We will collect more data, and consider the texture of reticular with GGO in future. Finally, using irregular-shaped VOIs (or ROIs), such as what is suggested in [12], would be considered in our future research.

Acknowledgement

- This research was supported by MEXT KAKENHI Grant Number 21103008.
- We greatly appreciate the help of Professor Junji Ueno with the Tokushima University to collect the data for this research.

References

[1] W.M. Thurlbeck, R.R. Miller, N.L. Müller, and E.C. Rosenow, Diffuse Diseases of the Lung: A Team Approach, B.C. Decker, 1991.
[2] W. Webb, N.L. Müller, and D. Naidich, High Resolution CT of the Lung, 4th edition, Lippincott Williams & Wilkins, 2008.
[3] I. Sluimer, A. Schilham, M. Prokop, and B.V. Ginneken, "Computer analysis of computed tomography scans of the lung: A survey," IEEE Trans. Med. Imaging, vol.25, no.4, pp.385–405, 2006.
[4] R. Uppaluri, E.A. Heitmman, M. Sonka, P.G. Hartley, G.W. Hunninghake, and G. Mclennan, "Computer recognition of regional lung disease patterns," Am. J. Respir. Crit. Care Med., vol.160, no.2, pp.648–654, 1999.
[5] 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," Med. Phys., vol.30, no.9, pp.2440–2454, 2003.
[6] I.C. Sluimer, P.F.V. Waes, M.A. Viergever, and B.V. Ginneken, "Computer-aided diagnosis in high-resolution CT of the lungs," Med. Phys., vol.30, no.12, pp.3081–3090, 2003.
[7] S.O. Park, J.B. Seo, N. Kim, S.H. Park, Y.K. Lee, B.W. 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 J. Radiol., vol.10, no.5, pp.455–463, 2009.
[8] N. Kim, J.B. Seo, Y. Lee, J.G. Lee, S.S. Kim, and S.H. Kang, "Development of an antomatic classification system for differentiation of obstructive lung disease using HRCT," J. Digital Imaging, vol.22, no.2, pp.136–148, 2009.
[9] S.C. Park, J. Tan, X.W. Wang, D. Lederman, J.K. Leader, S.H. Kim, and B. Zheng, "Computer-aided detection of early intersitial lung diseases using low-dose CT images," Phys. Med. Biol., vol.56, no.4, pp.1139–1153, 2011.
[10] M.J. Gangeh, L. Sørensen, S.B. Shaker, M.S. Kamel, M.D. Bruijne, and M. Loog, "A texton-based approach for the classification of lung parenchyma in CT images," Proc. MICCAI'10, LNCS, vol.6363, pp.595–602, 2010.
[11] L. Sørensen, S.B. Shaker, and M. de Bruijne, "Quantitative analysis of pulmonary emphysema using local binary patterns," IEEE Trans. Med. Imaging, vol.29, no.2, pp.559–69, 2010.
[12] I.C. Sluimer, M. Prokop, L. Hartmann, and B.V. Ginneken, "Automated classification of hyperlucency, fibrosis, ground glass, solid, and focal lesions in high-resolution CT of the lung," Med. Phys., vol.33, no.7, pp.2610–2620, 2006.
[13] Y. Xu, M. Sonka, G. McLennan, J.F. Guo, and E.A. Hoffman, "MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies," IEEE Trans. Med. Imaging, vol.25, no.4, pp.464–475, 2006.
[14] Y. Xu, J.R.V. Beek, Y. Hwanjo, J.F. Guo, G. McLennan, and E.A. Hoffman, "Computer-aided classification of intersititial lung diseases Via MDCT: 3D adaptive multiple feature method (3D AMFM)," Acad. Radiol., vol.13, no.8, pp.969–978, 2006.
[15] P.D. Korfiatis, A.N. Karahaliou, A.D. Kazantzi, C. Kalogeropoulou, and L.L. Costaridou, "Texture-based identification and characterization of interstitial pneumonia patterns in lung multidetector CT," IEEE Trans. Inf. Technol. Biomed., vol.14, no.3, pp.675–680, 2010.
[16] V.A. Zavaletta, B.J. Bartholmai, and R.A. Robb, "High resolution multi-detector CT aided tissue analysis and quantification of lung fibrosis," Acad. Radiol., vol.14, no.7, pp.772–787, 2007.
[17] 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," Proc. MICCAI'11, LNCS, vol.6893, pp.183–190, 2011.
[18] T. Tuytelaars and K. Mikolaczyk, "Local invarriant feature detectors: A survey," Foundations and Trands in Computer Graphics and Vision, vol.3, no.3, pp.177–280, 2007.
[19] G. Csurka, C.R. Dance, L. Fan, and C. Bray, "Visual categorization with bags of keypoints," Proc. ECCV'04 Workshop on Statistical Learning in Computer Vision, pp.1–22, 2004.
[20] C.D. Manning and H. Schfutze, Foundation of Statistical Natural Language Processing, The MIT Press, 2001.
[21] C.C. Chang and C.J. Lin, "LIBSVM: A library for support vector machines," ACM Trans. Intell. Sys. Technol., vol.2, no.3, pp.1–27, 2011.
[22] C.H. Lampert, "Kernel methods in computer vision," Foundations and Trends in Computer Graphics and Vision, vol.4, no.3, pp.193–285, 2009.
[23] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol.60, no.2, pp.91–110, 2004.
[24] J.H. Wang, F. Li, and Q. Li, "Automated segmentation of lungs with severe interstitial lung disease in CT," Med. Phys., vol.36, no.10, pp.4592–4599, 2009.
[25] P. Korfiatis, C. Kalogeropoulou, A. Kazantzi, S. Skiadopoulos, and L. Costaridou, "Texture classification-based segmentation of lung affacted by interstitial pneumonia in high-resolution CT," Med. Phys., vol.35, no.12, pp.5290–5301, 2008.
[26] Y. Rubner, C. Tomasi, and L.J. Guibas, "The Earth mover's distance as a metric for image retrieval," Int. J. Comput. Vis., vol.40, no.2, pp.99–121, 2000.
[27] V. Vapnik, Statistical Learning Theory, Wiley, 1998.
[28] C.W. Hsu and C.J. Lin, "A comparison of methods for multi-class support vector machines," IEEE Trans. Neural Netw., vol.13, no.2, pp.415–425, 2002.

Rui Xu received the B.S. and M.S. degrees in 2001 and 2004, respectively, from School of Electronic and Information, South China University of Technology. He received the Ph.D. degree in 2007 from the Graduate School of Science and Engineering, Ritsumeikan University, Japan. He worked in Digital Technology Research Center, Sanyo Electric Co., Ltd., Japan, from 2008 to 2010. Since April 2010, he has been a research staff in Graduate School of Medicine, Yamaguchi University, Japan. His research fields are medical image processing and pattern recognition. He is a member of JAMIT (Japanese Society of Medical Imaging Technology) and IEEE.

Yasushi Hirano received the B.S. and M.S. degrees in electronic informatics, and Ph.D degree in information engineering from Nagoya University in 1995, 1997 and 1999, respectively. His major fields of study are pattern recognition, image processing, parallel processing and inter-university authentication infrastructure. He was a Research Fellow of the Japan Society for the Promotion of Science (PD) in 1999–2000, an assistant professor at Department of information engineering, Nagoya University in 2000–2002, assistant professor at Information Technology Center, Nagoya University in 2002–2004, and an associate professor at Information Technology Center, Nagoya University in 2004–2010. He has been an associate professor at Graduate School of Medicine, Yamaguchi University since 2010. He is a member of IEEE, JAMIT (Japanese Society of Medical Imaging Technology), and JSMBE (Japanese Society for Medical and Biological Engineering).

Rie Tachibana received her M.S. and Ph.D. degrees in System Engineering from Yamaguchi University, Japan, in 2003 and 2006, respectively. She is currently a lecturer in the Department of Information Science and Technology, Oshima National College of Maritime Technology, Japan. Her research interests include medical image processing and computer-aided diagnosis.

Shoji Kido received his M.D degree from Osaka University in 1988. He received his Ph.D. degrees in Medicine and Information Science from Osaka University in 1992 and 1999, respectively. He was a staff radiologist in Yao Municipal Hospital, Nishinomiya Municipal Hospital, and Osaka Medical Center for Cancer and Cardiovascular Diseases from 1992 to 1999. He was a Professor in the Department of Computer Science and Systems Engineering at Yamaguchi University from 1999 to 2004. He is currently a Professor in the Department of Applied Medical Engineering Science, Graduate School of Medicine, Yamaguchi University.