

A Comparative Study of Medical Image Classification Using Machine Learning Methods

^[1]P. Vamshi Krishna, ^[2]A. Obulesh, ^[3]Deepika.S, ^[4]A.Mallikarjun Reddy
^[1] Post Graduate Student, ^{[2][3][4]} Assistant Professor
^{[1][2][3][4]} Dept. of CSE, Anurag Group of Institutions, Hyderabad, T.S., India.

Abstract: Interstitial Lung Disease (ILD) are a group of diseases are due to inflammation of lung tissues. Due to unknown cause of the ILDs international multidisciplinary consensus conference, American Thoracic Society and European Respiratory society proposed classification for ILDs. ILD diagnosis involves various stages of questioning and physical examination, testing, x-ray and CT scan. As such, the purpose of this study was to list out the methodologies for classification of ILD disease from medical images and discuss about their metiers and softness. In depth literature survey reveals that there are many methods for classifying ILD disease but very few methodologies uses machine learning issues. In this paper we are discussing about the various lung patterns using different methods like Local Binary Pattern in the process of using the convolutional neural networks. Such that the convolutional neural networks are used in the paper for comparing the various results from the various data sets that are used from the university hospital of Geneva and from Bern University Hospital which consists of HRCT scans and also used the datasets from the publicly available databases of ILD cases used in the “Near-Affine-Invariant Texture learning for lung tissue analysis using isotropic wavelet frames”.

INTRODUCTION

Since ILDs are generally manifested as textural alterations in the lung parenchyma, most of the proposed systems employ texture classification schemes on local regions of interest (ROIs) or volumes of interest (VOIs), depending on the 2D or 3D capabilities of the CT imaging modality employed. By sliding a fixed-scale classifier over pre-segmented lung fields, an ILD quantification map of the entire lung is generated. The latter can be used - either by physicians or CAD systems – to attempt the final diagnosis. The main characteristics of such a system are the chosen feature set and the classification method. The first CAD systems for ILDs proposed classical feature extraction methods to describe 2D texture, such as first order gray level statistics, gray level co-occurrence matrices (GLCM), run-length matrices (RLM) and fractal analysis [41][26]. These features were later merged and referred as the adaptive multiple feature method (AMFM) [65]. AMFM was generally accepted as the state of the art until new systems appeared that utilized more modern texture description techniques and provided a new perspective to the problem. Such systems employed filter banks [81],[20],[65], morphological operations followed by geometric measures [82], wavelet and contourlet transformations [43][8], histograms of oriented gradients [81] and local binary patterns (LBP) [71].

Moreover, some systems exploited the ability of MDCT scanners to achieve almost isotropic 3D sub-millimetre

resolution and expanded some of the already proposed 2D texture feature sets into three dimensions [83],[36],[75],[61]. One of the latest studies on volumetric data proposed the use of multiscale 3D Riesz wavelet frames coupled with wavelet pyramids [11]. The previously presented systems have used hand-crafted features to describe lung tissue, which often fail to adapt to new data or patterns. More recent studies adopted learned schemes for feature extraction which customize the feature set to the training data and have achieved promising results. Most of these use unsupervised techniques, such as bag of features [32][12][68] and sparse representation models [78,45,56]. In these methods, a set of texture atoms or textons is identified by using k-means and k-SVD, on already described local patches. The resulting set of textons constitutes a problem-specific dictionary and every local structure in the image is represented by the closest texton or a linear combination of the entire set. The final global descriptor usually consists of the histogram of textons appearing in the image. Another tool which has been used for extracting learned features is the restricted Boltzmann machine (RBM). RBMs are generative artificial neural networks (ANNs) that are able to capture and reproduce the statistical structure of the input and were employed in [63] for learning multi-scale filters with their responses as the features. Regardless of whether handcrafted or learned features are used, it is also crucial and challenging to choose an appropriate classifier that can optimally handle the properties of the created feature space. Many different approaches can be found in the literature. These use linear

discriminant (LD) [26][20], and Bayesian [65][83], classifiers, k-nearest neighbours (kNN) ANN [20] [71] [61] [12] [82], random forest [49] and support vector machines (SVM) with linear [78],[63] polynomial [81] or radial basis function (RBF) [8][32] kernels. Furthermore, multiple kernel learning classifier (m-MKL) was utilized in [43], while in [45], the minimum reconstruction error served as a classification criterion, after reconstructing the patch using class-specific dictionaries. Some attempts have recently also been made to use deep learning (DL) techniques and especially CNNs, after their impressive performance in large scale colour image classification [14]. Unlike other feature learning methods that build data representation models in an unsupervised manner, CNNs learn features and train an ANN classifier at the same time, by minimizing the classification error. Although the term DL implies the use of many consecutive learning layers, the first attempts on lung CT images adopted shallow architectures.

In [31], a modified RBM was used for both feature extraction and classification of lung tissue, incorporating some features of CNNs. Weight sharing was used among the hidden nodes, which were densely connected to label (output) nodes, while the whole network was trained in a supervised manner, using contrastive divergence and gradient descent. In [64], the authors designed a CNN with one convolutional layer and three dense layers and trained it from scratch. However, the shallow architecture of the network cannot leverage the descriptive ability of deep CNNs. The pre-trained deep CNN of [14] (AlexNet) was used in [64] to classify whole lung slices after fine-tuning with lung CT data. AlexNet was designed to classify natural color images with input size 224×224 , so the authors had to resize the images and artificially generate three.

CNNs are feed-forward ANN inspired by biological processes and designed to recognize patterns directly from pixel images (or other signals), by incorporating both feature extraction and classification. A typical CNN involves four types of layers: convolutional, activation, pooling and fully-connected (or dense) layers. A convolutional layer is characterized by sparse local connectivity and weight sharing. Each neuron of the layer is only connected to a small local area of the input, which resemble the receptive field in the human visual system. Different neurons respond to different local areas of the input, which overlap with each other to obtain a better representation of the image. In addition, the nodes of a convolutional layer are grouped in feature maps sharing the same weights, so the entire procedure becomes equivalent to convolution, with the shared weights being the filters for

each map. Weight sharing drastically reduces the number of parameters of the network and hence increases efficiency and prevents over fitting. Convolutional layers are often followed by a non-linear activation layer, in order to capture more complex properties of the input signal. Pooling layers are also used to subsample the previous layer, by aggregating small rectangular subsets of values. Max or average pooling is usually applied by replacing the input values with the maximum or the average value, respectively. The pooling layers reduce the sensitivity of the output to small input shifts. Finally, one or more dense layers are put in place, each followed by an activation layer, which produce the classification result. The training of CNNs is performed similarly to that of other ANNs, by minimizing a loss function using gradient descent based methods and back propagation of the error. Although the concept of CNNs has existed for decades, training such deep networks with multiple stacked layers was achieved only recently. This is mainly due to their extensive parallelization properties, which have been coupled with massively parallel GPUs, the huge amounts of available data, and several design tricks, such as the rectified linear activation units (ReLU).

II. RELATED WORK

In 2012, Krizhevsky et al. [14] won the Image Net Large-Scale Visual Recognition Challenge, convincingly outperforming the competition on a challenging dataset with 1000 classes and 1.2 million images. The proposed deep CNN, also known as AlexNet, consists of five convolutional layers with ReLU activations, some of which are followed by max- pooling layers, and three dense layers with a final 1000-way softmax. The network was trained with stochastic gradient descent (SGD) with a momentum term, maximizing the multinomial logistic regression objective. Deep architectures permit learning of data representations in multiple levels of semantic abstraction, so even high-level visual structures like cars or faces can be recognized in the last layers by combining low-level features of the first, such as edges. Nevertheless, designing a deep CNN for a specific problem is not trivial, since a large number of mutually dependent parameter values and algorithmic choices have to be chosen. Although much research has been conducted in recent years on deep CNNs for colour image classification, very little has been done on the problems of texture recognition and medical image analysis. In this paper, CNN has been proposed for the classification of ILD patterns that exploits the outstanding descriptive capability of deep neural networks. The method has been evaluated on a dataset of

120 cases from two hospitals and the results confirm its superiority compared to the state of the art. To the best of our knowledge, this is the first time a deep CNN has been designed and trained for lung tissue characterization. Finally, we provide empirical rules and principles on the design of CNN architectures for similar texture classification problems.

Lung Texture Analysis in the Literature:

State of the Art A wide range of features have been proposed for characterizing various lung tissue patterns associated with chronic obstructive pulmonary diseases (COPDs) and ILDs [1]:

- 1) gray-level histograms (GLH)[38],[70],[3],[5],[6],[9],[10],[19],[21],[28],[44],[46],[58],[66],[74],[76],[84];
- 2) Mathematical morphology and shape [38],[19],[21],[58];
- 3) gray-level co-occurrence matrices (GLCM) [70],[58],[66],[74],[84],[19],[21],[28],[76];
- 4) run length (RLE) [70],[28],[58],[66],[84],[76]
- 5) filter banks and wavelets [44],[46],[3],[5],[6],[9],[10],[17],[19],[21],[22],[34]
- 6) Others such as fractals and local binary patterns (LBPs) [66],[46],[51],[84],[76]

The heterogeneous feature group composed of GLH, RLE, and GLCM was used in most of the studies [70],[19],[21],[28],[58],[74]. A more comprehensive review of the techniques used for lung tissue categorization in the literature along with qualitative evaluations can be found in [1]. Although being able to well describe the lung tissue patterns, the performance of features derived from GLCM and RLE strongly depends on the underlying parameters (i.e., scales and directions). Most of the texture features used in the literature in lung texture analysis are modelling similar information: the spatial periodicity and scales contained in the images (GLCM, Gabor filters, wavelets, LBPs, and so on). The question is, which one characterizes the patterns best and is the most adaptable to the needs of lung tissue analysis in HRCT imaging. Although not being the most common in the state of the art of computer-aided diagnosis (CAD) in HRCT imaging of the chest, filtering techniques have several desirable properties. First, they are providing continuous responses to transient patterns in images. This is not the case for GLCMs that are based on the sum of occurrences of pixel pairs. GLCMs are as a consequence not able to characterize the similarity between series of contiguous pixels, which are often carrying important information in medical images. Filtering allows us to seek for specific features in the images (i.e., edge or ridge detection) by

modelling the shape of the filters either in the spatial or in the frequency domain. Wavelet transforms (WTs), which are implemented as filter banks, have the desirable property of being multi scale and thus allow covering the frequency domain (scale covariance). Filtering techniques and translation-invariant wavelets offer an over complete feature set able to fit most of the texture functions under the condition that they efficiently derive features from the coefficients, being perfectly complementary to the measures of density using GLH. The specific texture signatures of the lung tissue patterns can hardly be described by deterministic methods as intra class variations are very high due to the influence of factors such as the age of the patient, smoking history, and extend of the disease. Highly flexible texture modelling is required to catch subtle texture signatures of a given lung tissue pattern. In particular, invariance of the texture descriptors to affine properties (i.e., translation, rotation, and scale) is desirable to obtain a system that is able to impartially learn any texture appearance independently of prevailing localizations, orientations, or sizes. The main research contribution, is the development of a near-affine-invariant set of texture features (translation- and rotation-invariant as well as scale covariant) based on the WT. The isotropic properties of poly-harmonic B-spline wavelets [25], the scale-covariant properties of the quincunx sub sampling, and the translation invariance of redundant frame transforms [57],[30]. We investigate the ability of the proposed wavelet-based texture features to discriminate among the classes of healthy and pathological lung tissue types in HRCT images. GLHs model a complementary information corresponding to the density of the structures in CT imaging and are thus used along with WTs.

The primary diagnostic tool for COPD is spirometry by which various pulmonary function tests (PFTs) are performed [39]. However, PFTs have a low sensitivity to emphysema and are not capable of detecting early stages of COPD [37]. Another diagnostic tool that is gaining more and more attention is computed tomography (CT) imaging. CT is a sensitive method for diagnosing emphysema, assessing its severity, and determining its subtype, and both visual and quantitative CT assessment are closely correlated with the pathological extent of emphysema [59]. In this study, we focus on the assessment of emphysema in CT images. Emphysema lesions, or bullae, are visible in CT images as areas of abnormally low attenuation values close to that of air. In CT, emphysema can be classified into three subtypes, or patterns, and we will adopt the naming and definitions used in Webb et al. [77]. These subtypes are the following: cSentrilobular emphysema (CLE), defined as multiple small low-attenuation areas;

paraseptal emphysema (PSE), defined as multiple low-attenuation are as in a single layer long the pleura often surrounded by interlobular septa that is visible as thin white walls; and panlobular emphysema (PLE), defined as a low-attenuation lung with fewer and smaller pulmonary vessels. Examples of CLE and PSE, as well as normal tissue (NT). Common computerized approaches to emphysema quantification in CT are based on the histogram of CT attenuation values, and different quantitative measures of the degree of emphysema can be derived from this histogram. The most common measure is the relative area of emphysema (RA), also referred to as emphyse main dexor density mask [59], which measures the relative amount of lung parenchyma pixels that have attenuation values below a certain threshold. Usually, thresholds in the range to Hounsfield units (HU) are used. Measures based on the attenuation histogram disregard the information present in the morphology of the emphysema subtypes such as shape and size distribution of bullae. This was exemplified in a recent clinical study that reported discrepancies between visual scoring and RA for assessing the cranio caudal distribution of the three emphysema subtypes [73]. One way to objectively characterize the emphysema morphology is to describe the local image structure using texture analysis techniques [13], [54]. Uppaluri et al. introduced the idea of classifying emphysema in lung CT images using texture features [67]. Several authors followed this idea and classified regions of interest (ROIs) of various lung disease patterns using different texture features, mostly measures on gray-level co-occurrence matrices (GLCM), gray-level run-length matrices (GLRLM), and on the attenuation histogram, and different classifiers [65],[7], [33], [35], [28],[52],[53],[60],[80],[84]. Other examples of texture features used in the lung tissue classification literature are: the gray-level difference method [52],[53]; discrete wavelet frame decomposition using third-order B-splines [7]; convolving with partial derivatives of the Gaussian and the Laplacian of the Gaussian [33],[35]; gradient magnitude [80]; and fractal dimension [67],[65],[84]. In some cases, shape, or geometric, measures are also included in conjunction with the texture features [60], [33], [84]. Most studies use a mix of rotation invariant and rotation variant texture features, whereas the texture features used in this study are solely rotation invariant. Most of the studies on lung texture classification have one or several explicit emphysema classes [67],[65],[60], [84],[28], [52],[53],[80]. Multiple emphysema classes are defined by subdividing according to disease severity [84], [80] or emphysema morphology [28], [52], [53]. Chabat et al. discriminate between CLE and PLE [28], whereas

Prasad et al. distinguish between different stages of emphysema, ranging from diffuse to bullous emphysema [52], [53]. The study described in this paper has two emphysema classes defined based on morphology, namely, CLE and PSE. PLE is not considered since only 2 out of 39 subjects had PLE as leading pattern in the data used in the experiments. The data come from a population of (ex-) smokers, and PLE is known to be more prevalent in subjects with -antitrypsin deficiency than in subjects with smoking-related COPD [37]. A trained classifier can be used for quantification by classifying all pixels in the lung field. In [65], [60], [33], [84], [80,52,53] the full lung is classified either by labelling complete ROIs [65], [33], [84] or by labelling individual pixels [60], [80], [52], [53]. Xu et al. report the percentage of different disease patterns present in a few subjects, but these quantitative measures are not evaluated further [84]. Park et al. quantify emphysema by a weighted sum of relative emphysema class areas [80], and it is to our knowledge the only emphysema based quantitative study on a group of subjects in the lung CT texture analysis literature. This paper proposes two new ideas in the area of lung texture analysis in CT images. The specific application is emphysema quantification, but the ideas are also applicable to other lung disease patterns. Preliminary versions of the study presented here appeared in and [47], [48]. The first idea is to use local binary patterns (LBPs) originally formulated by Ojala et al. [72] as lung texture features. LBP unify structural and statistical information by a histogram of LBP codes that correspond to microstructures in the image at different scales. LBP have shown promising results in various applications in computer vision and have successfully been applied in a small number of other medical image analysis tasks, e.g., in mammographic mass detection [16] and magnetic resonance image analysis of the brain [24].

The second idea is to fuse the posterior probabilities obtained from a classification of all pixels in the lung field into quantitative measures of emphysema severity. Texture-based classification allows for quantification of different emphysema subtypes, which may be important in phenotyping emphysema for increased understanding of COPD. Furthermore, texture features may be less influenced by inspiration level and noise compared to, e.g., RA, which uses intensity in single pixels. In [48], this approach agrees well with the outcome of PFTs, achieving a significant correlation. Two fusion schemes are considered; mean class posterior (MCP) and relative class area (RCA). The second fusion scheme, RCA, is related to the fusion scheme in [80] that uses a weighted sum of RCAs. The difference is that we consider each RCA

individually. The proposed system is evaluated in two ways; ROI classification and emphysema quantification on subject level. A dataset comprising 2-D high-resolution CT (HRCT) slices with manually annotated ROIs is used for these purposes. The LBP features are compared to two other sets of features, one based on a Gaussian filter bank (GFB) and one comprising measures on GLCM, GLRLM, and the attenuation histogram.

III. DATASETS

The dataset used for training and evaluating the proposed method was made using two databases of ILD CT scans from two different Swiss university hospitals: The first is the publicly available multimedia database of ILDs from the University Hospital of Geneva [2], which consists of 109 HRCT scans of different ILD cases with 512×512 pixels per slice. Manual annotations for 17 different lung patterns are also provided, along with clinical parameters from patients with histological proven diagnoses of ILDs. The second database was provided by the Bern University Hospital, "Inselspital", and consists of 26 HRCT scans of ILD cases with resolution 512×512 . The scans were produced by different CT scanners with slightly different pixel spacing so a pre-processing step was applied which rescaled all scans to match a specific spacing value (i.e. 0.4 mm). However, the use of different reconstruction kernels by the scanners, still remains an open issue that complicates the problem even further. The image intensity values were cropped within the window $[-1000, 200]$ in HU and mapped to $[0, 1]$. Experienced radiologists from the "Inselspital" annotated (or re-annotated) both databases by manually drawing polygons around the six most relevant ILD patterns, namely GGO, reticulation, consolidation, micro nodules, honeycombing and a combination of GGO and reticulation. Healthy tissue was also added, leading to 7 classes. The annotation focused on typical instances of the considered ILD patterns, excluding ambiguous tissue areas that even experienced radiologists find difficult to classify. Hence, tissue outside the polygons may belong to any pattern, including that considered. Moreover, the annotators tried to avoid the Broncho vascular tree which (in a complete CAD system) should be segmented and removed, before applying the fixed-scale classifier. Annotation of the lung fields was also performed for all scans.

The considered classes appeared in the annotations of 94 out of the 109 scans of the Geneva database, to which the 26 cases from "Inselspital" were added, giving a total of 120 cases. On the basis of the ground truth polygons of these cases, this extracted in total 14696 non-overlapping

image patches of size 32×32 , unequally distributed across the 7 classes. Each pattern provides the number of ground truth polygons, the average and standard deviation of their area, the number of cases in which it was annotated and the number of extracted patches. The healthy pattern was only annotated in 8 cases, which however proved to be enough, since its texture does not present large deviations. It has to be noted that one case may contain multiple types of pathologies, so the sum of cases is larger than 120. The patches are entirely included in the lung field and have an overlap with the ground truth polygons of at least 80%. For each class, 150 patches were randomly selected for the test and 150 for the validation set. The choice of 150 was made based on the patch number of the rarest class (i.e. honeycombing) leaving about 50% of the patches for training. On the remaining patches, data augmentation was employed in order to maximize the number of training samples and equalize, at the same time, the samples' distribution across the classes. Data augmentation has often been employed in image classification, in order to increase the amount of training data and prevent over-fitting [14]. To this end, 15 label-preserving transformations were used, such as flip and rotation, as well as the combinations of the two. For each class, the necessary number of augmented samples was randomly selected, so all classes would reach the training set size of the rarest class, i.e. 5008, leading to 35056 equally distributed training patches. The CNN decides on the optimal architecture and configuration of a CNN, one should first comprehend the nature of the problem considered - in this case - the classification of ILD patterns. Unlike arbitrary objects in color images, which involve complex, high-level structures with specific orientation, ILD patterns in CT images are characterized by local textural features. Although texture is an intuitively easy concept for humans to perceive, formulating a formal definition is not trivial, which is the reason for the many available definitions in the literature [55]. Here, we define texture as a stochastic repetition of a few structures (textons) with relatively small size, compared to the whole region. Image convolution highlights small structures that resemble the convolution kernel throughout an image region, and in this way the analysis of filter bank responses has been successfully used in many texture analysis applications. This encourages the use of CNNs to recognize texture by identifying the optimal problem-specific kernels; however some key aspects stemming from our definition of texture have to be considered: (i) The total receptive field of each convolutional neuron with respect to the input (i.e. the total area of the original input "seen" by a convolutional neuron) should not be larger than the characteristic local structures of texture, otherwise non-

local information will be captured, which is irrelevant to the specific texture, (ii) since texture is characterized by fine grained low-level features, no pooling should be carried out between the convolutional layers, in order to prevent loss of information, (iii) each feature map outputted by the last convolutional layer should result in one single feature after pooling, in order to gain some invariance to spatial transformations like flip and rotation. Unlike color pictures that usually have high-level geometrical structure (e.g. the sky is up), a texture patch should still be a valid sample of the same class when flipped or rotated. Architecture: On the basis of these principles, we designed the network presented in Fig. 1.

The input of the network is a 32×32 image patch, which is convolved by a series of 5 convolutional layers. The size of the kernels in each layer was chosen to be minimal, i.e. 2×2. The use of small kernels that lead to very deep networks was proposed in the VGG-net [42], which was ranked at the top of ILSVRC 2014 challenge by employing 3×3 kernels and up to 16 convolutional layers. Here, we go one step further by shrinking the kernel size even more to involve more non-linear activations, while keeping the total receptive field small enough (6×6) to capture only the relevant local structure of texture. Each layer has a number of kernels proportional to the receptive field of its neurons, so it can handle the increasing complexity of the described structures. The size of the rectangular receptive field is 2×2 for the first layer and is increased by 1 in each dimension, for each layer added, leading to an area of (L+1)² for the Lth layer. Hence, the number of kernels we use for the Lth layer is k(L+1)², where the parameter k depends on the complexity of the input data and was set to 4 after relevant experiments. An average pooling layer follows, with size equal to the output of the last convolutional layer (i.e. 27×27). The resulting features, which are equal to the number of feature maps of the last layer i.e. f = 36k, are fed to a series of 3 dense layers with sizes 6f, 2f and 7, since 7 is the number of classes considered. The use of large dense layers accelerated convergence, while the problem of overfitting was solved by adding a dropout layer before each dense layer. Dropout can be seen as a form of bagging; it randomly sets a fraction of units to 0, at each training update, and thus prevents hidden units from relying on specific inputs [29].

1) Activations: It is well-known that the choice of the activation function significantly affects the speed of convergence. The use of the ReLU function $f(x) = \max(0, x)$ has been proven to speed up the training process many times compared to the classic sigmoid alternative. In this study, we also noticed that convolutional activations have a strong influence on the descriptive ability of the network. Driven by this

observation and after experimenting with different rectified activations, we propose the use of LeakyReLU [15], a variant of ReLU, for activating every convolutional layer. Unlike ReLU, which totally suppresses negative values, leaky ReLU assigns a non-zero slope, thus allowing a small gradient when the unit is not active (Eq. 1)

$$f_x = \begin{cases} x, & x > 0 \\ \alpha x & \text{else} \end{cases} \dots\dots(1)$$

where α is a manually set coefficient. LeakyReLU was proposed as a solution to the “dying ReLU” problem, i.e. the tendency of ReLU to keep a neuron constantly inactive as may happen after a large gradient update. Although a very low negative slope coefficient (i.e. $\alpha = 0.01$) was originally proposed, here we increase its value to 0.3, which considerably improves performance. Similar observations have also been reported in other studies [18]. A very leaky ReLU seems to be more resilient to overfitting when applied to convolutional layers, although the exact mechanism causing this behavior has to be further studied. For the dense part of the network, the standard ReLU activation was used for the first two layers and softmax on the last layer, to squash the 7dimensional output into a categorical probability distribution.

2) Training method: The training of an ANN can be viewed as a combination of two components, a loss function or training objective, and an optimization algorithm that minimizes this function. In this study, we use the Adam optimizer [23] to minimize the categorical cross entropy. The cross entropy represents the dissimilarity of the approximated output distribution (after softmax) from the true distribution of labels. Adam is a first-order gradient-based algorithm, designed for the optimization of stochastic objective functions with adaptive weight updates based on lower-order moments. Three parameters are associated with Adam: one is the learning rate and the other two are exponential decay rates for the moving averages of the gradient and the squared gradient. After relevant experiments, we left the parameters to their default values namely, learning rate equal to 0.001 and the rest 0.9 and 0.999, respectively. The initialization of the convolutional layers was performed using orthogonal matrices multiplied with a scaling parameter equal to 1.1, while a uniform distribution was utilized for the dense layers, scaled by a factor proportional to the square root of the layer’s number of inputs [40]. The weight updates are performed in mini-batches and the number of samples per batch was set to 128. The training ends when the network does not significantly improve its performance on the validation set for a predefined number of epochs. This

number is set to 200 and the performance is assessed in terms of average f-score (F avg) over the different classes (Eq. 2). An improvement is considered significant if the relative increase in performance is at least 0.5%.

$$F_{avg} = \frac{2}{7} \sum_{c=1}^7 \frac{recall_c * precision_c}{recall_c + precision_c} \dots\dots\dots (2)$$

$$recall_c = \frac{\text{samples correctly classified as } c}{\text{samples of class } c} \dots\dots (3)$$

$$Precision_c = \frac{\text{samples correctly classified as } c}{\text{samples classified as class } c}$$

$$Accuracy = \frac{\text{correctly classified samples}}{\text{total number of samples}}$$

3.1 ILD

The dataset used is part of a publicly available database of ILD cases [4] containing HRCT images with a slice thickness of 1 mm. 1448 hand-drawn regions of interest (ROIs) were annotated in 2-D HRCT slices in a collaborative fashion by two radiologists with 15 and 20 years of experience at the University Hospitals of Geneva (HUG). The in-plane resolution of the images is 512×512 pixels with an inter slice gap of 10 mm. A complete description of the database can be found in [4]. HRCT image series of 85 patients are used to evaluate the performance of the proposed approach. The hand-drawn ROIs are sub divided into 32×32 blocks for evaluating the methods. For blocks to be part of one of the tissue classes, at least 75% of the pixels need to be in the annotated region. The distributions and visual aspects of the five lung tissue classes are detailed in Table III. A total of 17848 blocks were used for the evaluation. The diagnosis of each case was confirmed either by pathology (biopsy and bronchoalveolar washing) or by a laboratory/specific test. In some cases with an early stage of ILD, healthy tissue was annotated in normal parts of the lungs to increase the amount of healthy annotated tissue, as finding HRCT scans of healthy patients was difficult.

In order to estimate the generalization performance of the classification of 32×32 blocks, a leave-one-patient-out cross validation (LOPO CV) is used. LOPO CV splits training and testing sets based on patients and has several advantages when compared to other validation methods [50]. First, when compared to the classical leave-one-out

(LOO) CV, it ensures that all ROIs belonging to the same patient are contained in the same fold and thus do not allow to train and test with identical patients. This situation corresponds to the clinical routine where the CAD system is trained using the entire database and unseen ROIs from an unknown patient are classified. Second, LOPO has the advantage of LOO where global experience is perfectly reproducible when compared to N-folds-fold CV because no random draw is carried out with LOO to create the folds. This is desirable when searching for optimal parameters where the variation of the evaluation conditions introduced by a random draw of the folds in N-folds-fold CV can lead to an inappropriate choice of parameters, especially when the number of patients is fairly low for some classes. At last, the computational cost is affordable with N-folds equals to the number of cases N-cases. The number of classifiers to train is equal to N-cases.

The data come from an exploratory study carried out at the Department of Respiratory Medicine, Gentoft University Hospital [69] and consist of CT images of the thorax acquired using General Electric (GE) equipment (Light Speed QX/i; GE Medical Systems, Milwaukee, WI, USA) with four detector rows. A total of 117 HRCT slices were acquired by scanning 39 subjects in the upper, middle, and lower lung. The CT scanning was performed using the following parameters: in-plane resolution 0.78 0.78 mm, slice thickness 1.25 mm, tube voltage 140 kV, and tube current 200 mAs. The slices were reconstructed using a high-spatial-resolution (bone) algorithm.

Prior to CT imaging, the subjects underwent PFTs, and both the forced vital capacity (FVC) and the forced expiratory volume in one second were measured [62]. is adjusted for age, sex, and height by dividing with a predicted value according to these three parameters, thereby obtaining % . The 39 subjects were divided into three groups: nine healthy life long non-smokers (referred to as never smokers), ten smokers without COPD (referred to as healthy smokers), and 20smokers diagnosed with moderate or severe COPD (referred To as COPD smokers). The COPD diagnosis was based on the recorded PFTs and done according to the Global Initiative for Chronic Obstructive Lung Disease criteria [39] as follows: no COPD, defined as FEV1 /FVC >= 0.7 and FEV1 % pred >= 80%; moderate to severe COPD, defined as FEV1 /FVC < 0.7 and 30% <= FEV1% pred < 80%. Of the 39 subjects, 19 were women and 20 were men. An experienced chest radiologist and a CT experienced pulmonologist each assessed the leading pattern, either NT, CLE, PSE, or PLE, in each of the 117 slices. Overall, the observers agreed in 53% of the slices, and they agreed on the emphysema class in 60% of slices where both decided

on an emphysema pattern. 168 non overlapping ROIs were annotated manually in 25 of the subjects, representing the three classes: NT (59 observations), CLE (50 observations), and PSE (59 observations). The NT ROIs were annotated in never smokers, and the CLE and PSE ROIs were annotated in healthy smokers and COPD smokers within the area(s) of the leading emphysema pattern by approximately marking the centre pixel of the emphysematous area. Square ROIs of a given width centred on the marked pixel were subsequently extracted. PLE was excluded due to underrepresentation in the data, only two subjects had PLE as leading pattern. Therefore, we are dealing with the three classes in all the experiments.

3.2 Multi-Detector CT (MDCT)

A pilot clinical case sample was acquired consisting of 30 MDCT scans corresponding to 5 normal patients and 25

patients diagnosed with IP secondary to connective tissue diseases, radio logically manifested with ground glass, reticular and honeycombing patterns (Fig. 1). MDCT scans were obtained with a Multi-slice (16x) CT (Light Speed, GE), in the Department of Radiology at the University Hospital of Patras, Greece. Acquisition parameters of tube voltage, tube current and slice thickness were 140 kVp, 300 mA and 1.25 mm, respectively. The image matrix size was 512x512 pixels with average pixel size of 0.89 mm. The MDCT scans were used to extract VOIs for training the classifiers employed for IP pattern identification and characterization. These sets consisted of 1173 cubic VOIs of variable size, defined by an expert radiologist, exploiting a home developed graphical user interface, representing patterns corresponding to reticular (458), ground glass opacities (195) honeycombing (249) and normal LP (271).

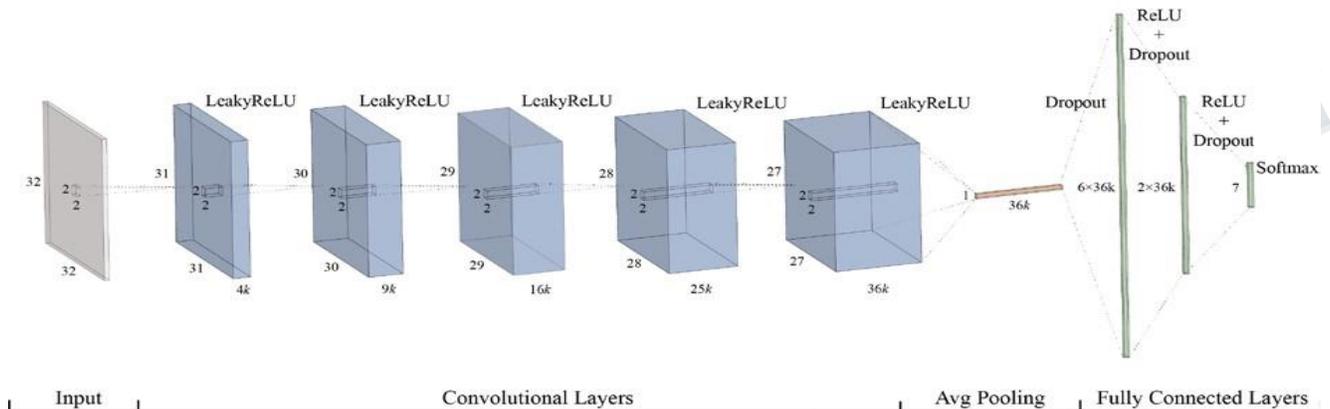


Figure 1 Architecture of CNN for lung pattern classification. Value of k is set to 4

IV. EXPERIMENTS

The ILD classifications are based on the training, validation and testing methods. The training is provided for the sample data sets or on the training data sets. The validation is used for the tuning of the parameters. The test and the test results are used for assessing the performance of the system. F-Score is used as the primary measure over the different classes (Eq 2). The accuracy is computed using the correctly classified samples over total number of samples (Eq 3). Due to the use of different patterns, datasets the performances are not comparable. But the datasets may affect the performance of the methods and not relative performance rank.

Table 1

Theano [27] framework was implemented while for AlexNet and VGGNet we used Caffe [79]. For lung pattern classification, the methods which did not use the convolutional neural network used python and MATHLAB.

The results from various comparison are:

- This uses 3 parts for obtaining results
- 1. Tuning of hyper-parameters
- 2. Comparison with state of the art
- 3. Analysis of the system performance

Tuning of hyper-parameters show choices for the architecture and training procedure.

Dropout fraction	Pooling type	Pooling percentage	Kernel number multiplier (k)	Number of kernels for L_{th} layer	Number of conv layers	Kernel size	Input scale factor	Activation function	Testing F_{avg}	# Epochs \times Epoch time
0	Avg	100%	4	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.7908	90 X 11s
0.5	Max	100%	4	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.8105	69 X 11s
0.5	Avg	50%	4	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.7895	249 X 11s
0.5	Avg	25%	4	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.7452	286 X 12s
0.5	Avg	100%	4	17	5	2 X 2	1	LReLU(0.3)	0.8446	300 X 12s
0.5	Avg	100%	4	36	5	2 X 2	1	LReLU(0.3)	0.8508	386 X 32s
0.5	Avg	100%	3	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.8266	427 X 7s
0.5	Avg	100%	5	$k(L+1)^2$	5	2 X 2	1	LReLU(0.3)	0.8425	362 X 14s
0.5	Avg	100%	4	$k(L+1)^2$	7	2 X 2	1	LReLU(0.3)	0.8432	295 X 23s
0.5	Avg	100%	4	$k(L+1)^2$	6	2 X 2	1	LReLU(0.3)	0.8559	215 X 18s
0.5	Avg	100%	4	$k(L+1)^2$	4	2 X 2	1	LReLU(0.3)	0.8443	372 X 6s
0.5	Avg	100%	4	$k(L+1)^2$	7	2 X 2	1	LReLU(0.3)	0.8432	295 X 23s

Table 1 provides classification of the performance of the different configurations on network's architecture and the training time required. The k multiplier is used for identification of the number of kernels for which 4 is used as optimal choice for performance and efficiency.

Optimizer	Loss Function	F_{avg}	Accuracy	Epoch
SGD	Cross-entropy	0.834	0.8428	333
AdaGrad	Cross-entropy	0.8219	0.8228	257
Adam	MSE	0.8499	0.8523	155
Adam	Cross-entropy	0.8547	0.8561	386

Table 2 depicts the usage of different optimizers and loss function for training of the CNN.

Comparison with the state of the art compares CNN with other methods that are best using different classifiers and features. Every author and parameter are tuned using trial and error procedure. The below table 3 shows the comparison

Table 3 Comparison of the proposed with state-of-the-art methods using handcrafted features

Method	Features	Classifier	F_{avg}	Accuracy
Gangeh	Intensity	SVMRF	0.7127	0.7152
Sorensen	LBP + histogram	kNN	0.7322	0.7333
Anthimopoulos	Local DCT + histogram	RF	0.7786	0.7809
PROPOSED METHOD USING DEEP CONVOLUTIONAL NEURAL NETWORK	CNN		0.8547	0.8561

REFERENCES

- [1] A. Depeursinge, "Affine-invariant texture analysis and retrieval of 3D medical images with clinical context integration," Ph.D. dissertation, Serv. Med. Informat. Univ. of Geneva, 2010.
- [2] A. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, P. A. Poletti, and H. Müller, "Building a reference multimedia database for interstitial lung diseases," *Comput. Med. Imaging Graph.*, 36 (3), pp. 227–238, 2012.
- [3] A. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Müller, "3D case-based retrieval for interstitial lung diseases," in *Medical Content-Based Retrieval for Clinical Decision Support (Lecture Notes in Computer Science Series)*. New York: Springer-Verlag, 2010, pp. 39–48.
- [4] A. Depeursinge, A. Vargas, A. Platon, A. Geissbuhler, P.-A. Poletti, and H. Müller, "Building a reference multimedia database for interstitial lung diseases," *Comput. Med. Imag. Graph.*, vol. 36, no. 3, pp. 227–238, 2012.
- [5] A. Depeursinge, D. Racoceanu, J. Iavindrasana, G. Cohen, A. Platon, P.-A. Poletti, and H. Müller, "Fusing visual and clinical information for lung tissue classification in high-resolution computed tomography," *Artif. Int. Med.*, vol. 50, no. 1, pp. 13–21
- [6] A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser, and H. Müller, "Lung tissue classification using Wavelet frames," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2007, pp. 6259–6262
- [7] A. Depeursinge, D. Sage, A. Hidki, A. Platon, P.-A. Poletti, M. Unser, and H. Müller, "Lung tissue classification using wavelet frames," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 2007*, Aug. 2007, pp. 6259–6262.
- [8] A. Depeursinge, D. Van De Ville, A. Platon, A. Geissbuhler, P. A. Poletti, and H. Müller, "Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 665–675, 2012.
- [9] A. Depeursinge, D. Van De Ville, M. Unser, and H. Müller, "Lung tissue analysis using isotropic polyharmonic B-spline wavelets," in *Proc. MICCAI Workshop Pulmonary Image Anal.*, New York, 2008, pp. 125–134.
- [10] A. Depeursinge, J. Iavindrasana, A. Hidki, G. Cohen, A. Geissbuhler, A. Platon, P.-A. Poletti, and H. Müller, "Comparative performance analysis of state-of-the-art classification algorithms applied to lung tissue categorization," *J. Digit. Imag.*, vol. 23, no. 1, pp. 18–30, 2010.
- [11] A. Depeursinge, P. Pad, A. S. Chin, A. N. Leung, D. L. Rubin, H. Müller, and M. Unser, "Optimized steerable wavelets for texture analysis of lung tissue in 3-D CT: Classification of usual interstitial pneumonia," in *12th International Symposium on Biomedical Imaging*, 2015, pp. 403–406.
- [12] A. Foncubierta-Rodríguez et al., "Using Multiscale Visual Words for Lung Texture Classification and Retrieval", *Medical Content-Based Retrieval for Clinical Decision Support, Lecture Notes in Computer Science*, vol. 7075, pp 69-79, 2012
- [13] A. H. Mir, M. Hanmandlu, and S. N. Tandon, "Texture analysis of CT images," *IEEE Eng. Med. Biol. Mag.*, vol. 14, no. 6, pp. 781–786, Nov./ Dec. 1995.
- [14] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst.*, p. 9, 2012.
- [15] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," *Icml-2013*, vol. 28, 2013.
- [16] A. Oliver, X. Lladó, J. Freixenet, and J. Martí, "False positive reduction in mammographic mass detection using local binary patterns," in *MICCAI (1)*, ser. (Lecture Notes in Computer Science 4791), N. Ayache, S. Ourselin, and A. J. Maeder, Eds. New York: Springer-Verlag, Nov. 2007, pp. 286–293.
- [17] A. Tolouee, H. Abrishami-Moghaddam, R. Garnavi, M. Forouzanfar, and M. Giti, "Texture analysis in lung HRCT images," in *Proc. Digital Imaging Comput.: Tech. Appl.*, 2008, pp. 305–311.

- [18] B. Xu, N. Wan B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolution Network," *ICML Deep Learn.*, pp. 1–5, 2015.
- [19] C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, and A. M. Aisen, "Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision," in *Proc. 10th Nat. Conf. Artif. Intell.*, Orlando, FL, 1999, pp. 760–767.
- [20] C. Sluimer, P. F. van Waes, M. a Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Med. Phys.*, vol. 30, no. 12, pp. 3081–3090, 2003.
- [21] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, and L. S. Broderick, "ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases," *Comput. Vis. Image Understand.*, vol. 75, no. 1–2, pp. 111–132, 1999.
- [22] C.-T. Liu, P.-L. Tai, A. Y.-J. Chen, C.-H. Peng, T. Lee, and J.-S. Wang, "A content-based CT lung retrieval system for assisting differential diagnosis images collection," in *Proc. 2nd Int. Conf. Multimedia Expo.*, Tokyo, Japan, 2001, pp. 172–177.
- [23] D. P. Kingma and J. L. Ba, "Adam: a Method for Stochastic Optimization," *Int. Conf. Learn. Represent.*, pp. 1–13, 2015.
- [24] D. Unay, A. Ekin, M. Cetin, R. Jasinschi, and A. Ercil, "Robustness of local binary patterns in brain MR image analysis," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 2007*, Aug. , pp. 2098–2101.
- [25] D. VanDeVille, T. Blu, and M. Unser, "Isotropic polyharmonic B-Splines: Scaling functions and wavelets," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1798–1813, Nov. 2005.
- [26] Delorme, Stefan, Mark-Aleksi Keller-Reichenbecher, Ivan Zuna, Wolfgang Schlegel, and Gerhard Van Kaick. "Usual interstitial pneumonia: quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis." *Investigative radiology* 32, no. 9 (1997): 566-574.
- [27] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley, and Y. Bengio, "Theano: new features and speed improvements," *arXiv Prepr.*, pp. 1–10, 2012.
- [28] F. Chabat, G.-Z. Yang, and D. M. Hansell, "Obstructive lung diseases: Texture classification for differentiation at CT," *Radiology*, vol. 228, no. 3, pp. 871–877, 2003.
- [29] G. Hinton, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
- [30] G. van de Wouwer, P. Scheunders, and D. V. Dyck, "Rotation-invariant texture characterization using isotropic wavelet frames," in *presented at the 14th Int. Conf. Pattern Recognit. Brisbane, Qld., Australia*, 1998.
- [31] G. van Tulder and M. de Bruijne, "Learning Features for Tissue Classification with the Classification Restricted Boltzmann Machine," in *Medical Computer Vision: Algorithms for Big Data*, 2014, pp. 47–58.
- [32] Gangeh, Mehrdad J., Lauge Sørensen, Saher B. Shaker, Mohamed S. Kamel, Marleen De Bruijne, and Marco Loog. "A texton-based approach for the classification of lung parenchyma in CT images." In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pp. 595-602. Springer Berlin Heidelberg, 2010
- [33] I. C. Sluimer, M. Prokop, I. Hartmann, and B. van 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, Jul. 2006.
- [34] I. C. Sluimer, P. F. van Waes, M. A. Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Med. Phys.*, vol. 30, no. 12, pp. 3081–3090, 2003.
- [35] I. C. Sluimer, P. F. van Waes, M. A. Viergever, and B. van Ginneken, "Computer-aided diagnosis in high resolution CT of the lungs," *Med. Phys.*, vol. 30, no. 12, pp. 3081–3090, Dec. 2003.
- [36] I. Mariolis, P. Korfiatis, L. Costaridou, C. Kalogeropoulou, D. Daoussis, and T. Petsas, "Investigation of 3D textural features' discriminating ability in diffuse

- lung disease quantification in MDCT,” 2010 IEEE Int. Conf. Imaging Syst. Tech. IST 2010 - Proc., pp. 135–138, 2010.
- [37] J. W. Gurney, “Pathophysiology of obstructive airways disease,” *Radiol. Clin. North Amer.*, vol. 36, no. 1, pp. 15–27, Jan. 1998
- [38] K. Doi, “Current status and future potential of computer-aided diagnosis in medical imaging,” *Brit. J. Radiol.*, vol. 78, pp. 3–19, 2005.
- [39] K. F. Rabe, S. Hurd, A. Anzueto, P. J. Barnes, S. A. Buist, P. Calverley, Y. Fukuchi, C. Jenkins, R. Rodriguez-Roisin, C. van Weel, and J. Zielinski, “Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary,” *Amer. J. Respir. Crit. Care Med.*, vol. 176, no. 6, pp. 532–555, Sep. 2007
- [40] K. He, X. Zhang, S. Ren, and J. Sun, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification.”
- [41] K. R. Heitmann, H. Kauczor, P. Mildenerger, T. Uthmann, J. Perl, and M. Thelen, “Automatic detection of ground glass opacities on lung HRCT using multiple neural networks,” *Eur. Radiol.*, vol. 7, no. 9, pp. 1463–1472, 1997.
- [42] K. Simonyan and A. Zisserman “Very Deep Convolutional Networks for Large-Scale Image Recognition”, in International Conference on Learning Representations 2015.
- [43] K. T. Vo and A. Sowmya, “Multiple kernel learning for classification of diffuse lung disease using HRCT lung images,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2010, pp. 3085–3088, 2010.
- [44] K. Vo and A. Sowmya, “Directional multi-scale modeling of highresolution computed tomography (HRCT) lung images for diffuse lung disease classification,” in *Proc. 13th Int. Conf. Comput. Anal. Images Patterns*, vol. 5702. New York: Springer-Verlag, 2009, pp. 663–671.
- [45] K.T. Vo et al., "Multiscale sparse representation of HRCT lung images for diffuse lung disease classification," In Int'l Conf. on Image Processing (ICIP), pp.441-444, 2012
- [46] L. Sørensen, S. B. Shaker, and M. D. Bruijne, “Quantitative analysis of pulmonary emphysema using local binary patterns,” *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 559–569, Feb. 2010.
- [47] L. Sørensen, S. B. Shaker, and M. de Bruijne, “Texture classification in lung CT using local binary patterns,” in *MICCAI (1)*, ser. (Lecture Notes in Computer Science 5241), D. N. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds. New York: Springer-Verlag, Sep. 2008, pp. 934–941.
- [48] L. Sørensen, S. Shaker, and M. de Bruijne, “Texture based emphysema quantification in lung CT,” in *Proc. 1st Int. Workshop Pulm. Image Anal.*, M. Brown, M. de Bruijne, B. van Ginneken, A. Kiraly, J. Kuhnigk, C. Lorenz, K. Mori, and J. Reinhardt, Eds., Sep. 2008, pp. 5–14.
- [49] M. Anthimopoulos, S. Christodoulidis, a Christe, and S. Mougiakakou, “Classification of interstitial lung disease patterns using local DCT features and random forest,” 2014 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., pp. 6040–6043, 2014.
- [50] M. Dundar, G. Fung, L. Bogoni, M. Macari, A. Megibow, and B. Rao, “A methodology for training and validating a CAD system and potential pitfalls,” in *Proc. 18th Int. Congr. Exhib. Comput. Assist. Radiol. Surg.*, 2004, vol. 1268, pp. 1010–1014.
- [51] M. Gangeh, L. Sørensen, S. Shaker, M. Kamel, M. D. Bruijne, and M. Loog, “A texton-based approach for the classification of lung parenchyma in CT images,” in *Proc. Med. Image Comput. Comput. Assisted Intervention (Lecture Notes in Computer Science Series)*, vol. 6363. Berlin, Germany: Springer-Verlag, 2010, pp. 595–602
- [52] M. Prasad, A. Sowmya, and I. Koch, “Designing relevant features for continuous data sets using ICA,” *Int. J. Comput. Intell. Appl.*, vol. 7, no. 4, pp. 447–468, Dec. 2008.
- [53] M. Prasad, A. Sowmya, and P. Wilson, “Multi-level classification of emphysema in HRCT lung images,” *Pattern Anal. Appl.*, vol. 12, no. 1, pp. 9–20, Jan. 2009.
- [54] M. Tuceryan and A. K. Jain, “Texture analysis,” in *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed. Singapore: World Scientific, 1998, pp. 207–248

- [55] M. Tuceryan and A. K. Jain, *The Handbook of Pattern Recognition and Computer Vision*, C. H. Chen, L. F. Pau, and P. S. Wang, Eds. Singapore: World Scientific, 1998.
- [56] M. Zhang et al. "Pulmonary Emphysema Classification based on an Improved Texton Learning Model by Sparse Representation", *Proc. SPIE*, vol. 8670, 2013.
- [57] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 4, no. 11, pp. 1549–1560, Nov. 1995.
- [58] N. Kim, J. B. Seo, Y. Lee, J. G. Lee, S. S. Kim, and S. H. Kang, "Development of an automatic classification system for differentiation of obstructive lung disease using HRCT," *J. Digit. Imag.*, vol. 22, no. 2, pp. 136–148, 2009.
- [59] N. L. Müller, C. A. Staples, R. R. Miller, and R. T. Abboud, "Density mask". An objective method to quantitate emphysema using computed tomography," *Chest*, vol. 94, no. 4, pp. 782–787, Oct. 1988.
- [60] O. Friman, M. Borga, M. Lundberg, U. Tylén, and H. Knutsson, "Recognizing emphysema—A neural network approach," in *Proc. ICPR(1)*, Mar. 2002, pp. 512–515.
- [61] P. D. Korfiatis, A. N. Karahaliou, A. D. Kazantzi, C. Kalogeropoulou, and L. I. Costaridou, "Texture-based identification and characterization of interstitial pneumonia patterns in lung multi-detector CT," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 675–680, 2010.
- [62] P. Quanjer, G. Tammeling, J. Cotes, O. Pedersen, R. Peslin, and J. Yernault, "Lung volumes and forced ventilatory flows. Report working party standardization of lung function tests, european community for steel and coal. Official statement of the european respiratory society," *Eur. Respir. J. Suppl.*, vol. 16, pp. 5–40, Mar. 1993.
- [63] Q. Li, W. Cai, and D. D. Feng, "Lung image patch classification with automatic feature learning," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 6079–6082, 2013.
- [64] Q. Li, W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen, "Medical image classification with convolutional neural network," in *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*, 2014, vol. 2014, no. December, pp. 844–848.
- [65] R. Uppaluri, E. A. Hoffman, M. Sonka, P. G. Hartley, G. W. Hunninghake, and G. McLennan, "Computer recognition of regional lung disease patterns," *Am. J. Respir. Crit. Care Med.*, 160 (2), pp. 648–654, 1999.
- [66] R. Uppaluri, E. A. Hoffman, M. Sonka, G. W. Hunninghake, and G. McLennan, "Interstitial lung disease: A quantitative study using the adaptive multiple feature method," *Amer. J. Respirat. Crit. Care Med.*, vol. 159, no. 2, pp. 519–525, 1999.
- [67] R. Uppaluri, T. Mitsa, M. Sonka, E. A. Hoffman, and G. McLennan, "Quantification of pulmonary emphysema from lung computed tomography images," *Amer. J. Respir. Crit. Care Med.*, vol. 156, no. 1, pp. 248–254, Jul. 1997.
- [68] R. Xu et al., "Classification of diffuse lung disease patterns on high-resolution computed tomography by a bag of words approach," *Med Image Comput Comput Assist Interv.*, vol. 14(Pt 3), pp. 183–90, 2011.
- [69] S. B. Shaker, K. A. von Wachenfeldt, S. Larsson, I. Mile, S. Persdotter, M. Dahlbäck, P. Broberg, B. Stoel, K. S. Bach, M. Hestad, T. E. Fehniger, and A. Dirksen, "Identification of patients with chronic obstructive pulmonary disease (COPD) by measurement of plasma biomarkers," *Clin. Respir. J.*, vol. 2, no. 1, pp. 17–25, 2008.
- [70] S. Delorme, M.-A. Keller-Reichenbecher, I. Zuna, W. Schlegel, and G. V. Kaick, "Usual interstitial pneumonia: Quantitative assessment of high-resolution computed tomography findings by computer-assisted texture-based image analysis," *Investigat. Radiol.*, vol. 32, no. 9, pp. 566–574, 1997.
- [71] Sørensen, Lauge, Saher B. Shaker, and Marleen De Bruijne. "Quantitative analysis of pulmonary emphysema using local binary patterns." *Medical Imaging, IEEE Transactions on* 29, no. 2 (2010): 559–569.
- [72] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
-

- [73] T. Stavngaard, S. B. Shaker, K. S. Bach, B. C. Stoel, and A. Dirksen, "Quantitative assessment of regional emphysema distribution in patients with chronic obstructive pulmonary disease (COPD)," *Acta Radiol.*, vol. 47, no. 9, pp. 914–921, Nov. 2006
- [74] T. Zrimec and J. S. J. Wong, "Improving computer aided disease detection using knowledge of disease appearance," in *Proc. MEDINFO, 2007*, vol. 129, pp. 1324–1328.
- [75] V. A. Zavaletta, B. J. Bartholmai, and R. a. Robb, "High Resolution Multidetector CT-Aided Tissue Analysis and Quantification of Lung Fibrosis," *Acad. Radiol.*, vol. 14, no. 7, pp. 772–787, 2007.
- [76] V. A. Zavaletta, B. J. Bartholmai, and R. A. Robb, "Nonlinear histogram binning for quantitative analysis of lung tissue fibrosis in high-resolution CT data," in *Medical Imaging 2007: Physiology, Function, and Structure from Medical Images*, vol. 6511, A. Manduca and X. P. Hu, Eds. Bellingham, WA: SPIE, 2007, p. 65111Q.
- [77] W. R. Webb, N. Müller, and D. Naidich, *High-Resolution CT of the Lung*, J.-R. John, Ed., 3rd ed. Baltimore, MD: Lippincott Williams & Wilkins, 2001
- [78] W. Zhao et al., "Classification of diffuse lung diseases patterns by a sparse representation based method on HRCT images," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS 2013*, pp.5457-5460, 2013.
- [79] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional Architecture for Fast Feature Embedding," 2014.
- [80] Y. S. Park, J. B. Seo, N. Kim, E. J. Chae, Y. M. Oh, S. D. Lee, Y. Lee, and S.-H. Kang, "Texture-based quantification of pulmonary emphysema on high-resolution computed tomography: Comparison with density-based quantification and correlation with pulmonary function test," *Invest Radiol.*, vol. 43, no. 6, pp. 395–402, Jun. 2008.
- [81] Y. Song, W. Cai, Y. Zhou, and D. D. Feng, "Feature-based image patch approximation for lung tissue classification," *IEEE Trans. Med. Imaging*, vol. 32, no. 4, pp. 797–808, 2013.
- [82] Y. Uchiyama, S. Katsuragawa, H. Abe, J. Shiraishi, F. Li, Q. Li, C.-T. 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.
- [83] Y. Xu, E. J. R. 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)," *Acad. Radiol.*, vol. 13, no. 8, pp. 969–978, 2006.
- [84] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. A. Hoffman, "MDCTbased 3D texture classification of emphysema and early smoking related lung pathologies," *IEEE Trans. Med. Imag.*, vol. 25, no. 4, pp. 464–475, Apr. 2006.