©2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

DOI: 10.1109/TMI.2009.2038575

Quantitative Analysis of Pulmonary Emphysema using Local Binary Patterns

Lauge Sørensen*, Saher B. Shaker, and Marleen de Bruijne

***Abstract*— We aim at improving quantitative measures of emphysema in computed tomography (CT) images of the lungs. Current standard measures, such as the relative area of emphysema (RA), rely on a single intensity threshold on individual pixels, thus ignoring any interrelations between pixels. Texture analysis allows for a much richer representation that also takes the local structure around pixels into account.**

This paper presents a texture classification based system for emphysema quantification in CT images. Measures of emphysema severity are obtained by fusing pixel posterior probabilities output by a classifier. Local binary patterns (LBP) are used as texture features, and joint LBP and intensity histograms are used for characterizing regions of interest (ROI)s. Classification is then performed using a *k* nearest neighbor classifier with a histogram dissimilarity measure as distance.

A 95.2% classification accuracy was achieved on a set of 168 manually annotated ROIs comprising the three classes: normal tissue, centrilobular emphysema, and paraseptal emphysema. The measured emphysema severity was in good agreement with a pulmonary function test (PFT) achieving correlation coefficients of up to $|r| = 0.79$ in 39 subjects. The results were compared to RA and to a Gaussian filter bank, and the texture based measures correlated significantly better with PFT than RA did.

***Index Terms*— Emphysema, local binary patterns, quantitative CT, texture analysis, tissue classification.**

I. Introduction

CHRONIC obstructive pulmonary disease (COPD) is a growing health problem worldwide. In the United States alone, it is the fourth leading cause of morbidity and mortality, and it is estimated to become the fifth most burdening disease worldwide by 2020 [1]. COPD is a chronic lung disease characterized by limitation of airflow. It comprises two components: small airway disease and emphysema, which is characterized by gradual loss of lung tissue. Detection and quantification of emphysema is important, since it is thought to be the main cause of shortness of breath and disability in COPD.

The primary diagnostic tool for COPD is spirometry by which various pulmonary function tests (PFT)s are performed [1]. However, PFTs have a low sensitivity to emphysema and are not capable of detecting early stages of COPD [2]. 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 [3].

In this work, 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*. [4]. These subtypes are the following: centrilobular emphysema (CLE), defined as multiple small low-attenuation areas; paraseptal emphysema (PSE), defined as multiple low-attenuation areas in a single layer along 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), are shown in Fig. 1.

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 emphysema index or density mask [3], which measures the relative amount of lung parenchyma pixels that have attenuation values below a certain threshold. Usually, thresholds in the range -856 to -960 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 craniocaudal distribution of the three emphysema subtypes [5].

One way to objectively characterize the emphysema morphology is to describe the local image structure using texture analysis techniques [6], [7]. Uppaluri *et al*. introduced the idea of classifying emphysema in lung CT images using texture features [8]. Several authors followed this idea and classified regions of interest (ROI)s 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 [9]–[18]. Other examples of texture features used in the lung tissue classification literature are: the

Copyright (c) 2009 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

Asterisk indicates corresponding author.

*Lauge Sørensen is with the Image Group, Department of Computer Science, University of Copenhagen, Denmark.

Saher B. Shaker is with the Department of Cardiology and Respiratory Medicine, Hvidovre University Hospital, Copenhagen, Denmark.

Marleen de Bruijne is with the Image Group, Department of Computer Science, University of Copenhagen, Denmark and with the Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical Informatics, Erasmus MC, Rotterdam, The Netherlands.

gray-level difference method [17], [18]; discrete wavelet frame decomposition using third order B-splines [15]; convolving with partial derivatives of the Gaussian and the Laplacian of the Gaussian [12], [13]; gradient magnitude [16]; and fractal dimension [8], [9], [14]. In some cases, shape, or geometric, measures are also included in conjunction with the texture features [10], [13], [16]. Most works use a mix of rotation invariant and rotation variant texture features, whereas the texture features used in this work are solely rotation invariant.

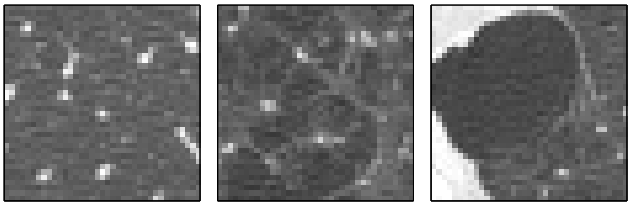
Most of the work on lung texture classification have one or several explicit emphysema classes [8]–[11], [14]–[18]. Multiple emphysema classes are defined by sub-dividing according to disease severity [14], [16] or emphysema morphology [11], [17], [18]. Chabat *et al.* discriminate between CLE and PLE [11] whereas Prasad *et al.* distinguish between different stages of emphysema, ranging from diffuse to bullous emphysema [17], [18]. The work 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 comes from a population of (ex-)smokers, and PLE is known to be more prevalent in subjects with α_1-antitrypsin deficiency than in subjects with smoking-related COPD [2].

A trained classifier can be used for quantification by classifying all pixels in the lung field. In [9], [10], [13], [14], [16]–[18] the full lung is classified either by labeling complete ROIs [9], [13], [14] or by labeling individual pixels [10], [16]–[18]. Xu *et al*. report the percentage of different disease patterns present in a few subjects, but these quantitative measures are not evaluated further [14]. Park *et al*. quantify emphysema by a weighted sum of relative emphysema class areas [16], 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 work presented here appeared in [19] and [20].

The first idea is to use local binary patterns (LBP) originally formulated by Ojala *et al*. [21] as lung texture features. LBP unify structural and statistical information by a histogram of LBP codes that correspond to micro-structures 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 [22] and magnetic resonance image analysis of the brain [23]. In [19], we showed that histogram dissimilarity measures between LBP feature histograms in a *k* nearest neighbor (*k*NN) classifier [24] can discriminate between emphysematous and normal tissue.

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. Further, texture features may be less influenced by inspiration level and noise compared to, e.g., RA, which uses intensity in single pixels. In [20], we showed that this approach agrees well with the outcome of PFTs, achieving a significant correlation. Two fusion schemes are considered in this paper; mean class posterior (MCP) and relative class area (RCA). The second fusion scheme, RCA, is related to the fusion scheme in [16] that uses a weighted sum of relative class areas. The difference is that we consider each relative class area individually.



(a) (b) (c)

Fig. 1. Examples of different lung tissue patterns in CT shown with the window setting $-600/1500$ HU [4]. (a) Normal tissue (NT). (b) Centrilobular emphysema (CLE). (c) Paraseptal emphysema (PSE). The white area in the left part of image (c) is the exterior of the lung.

The proposed system is evaluated in two ways; ROI classification and emphysema quantification on subject level. A data set comprising 2D 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.

II. Methods

The proposed system for emphysema quantification relies on texture classification in local ROIs in the CT images. Three types of texture features are considered, LBP, GFB, and a set of features based on GLCM, GLRLM, and the attenuation histogram. Section II-A describes LBP, and Section II-B describes GFB. Measures on GLCM and GLRLM are the most commonly used features in lung texture classification, and they are therefore not described in detail here. We refer to [6], [25] for a detailed description and to [8], [11], [16]–[18] for examples of applications. Section II-C describes how the texture in the ROIs is characterized by computing distributions of features, or feature histograms, and Section II-D presents a combined measure of histogram dissimilarity between the feature histograms, used to discriminate ROIs with a *k*NN classifier. Finally, Section II-E describes how emphysema is quantified in the CT images by fusing pixel posterior probabilities output by a *k*NN classifier trained on a small set of ROIs.

A. Local binary patterns

LBP were originally proposed by Ojala *et al*. as a gray-scale invariant measure for characterizing local structure in a 3×3 pixel neighborhood [26]. Later, a more general formulation was proposed that further allowed for multi-resolution analysis and rotation invariance [21]. We use the formulation given in [21]. The LBP are obtained by thresholding samples in a local

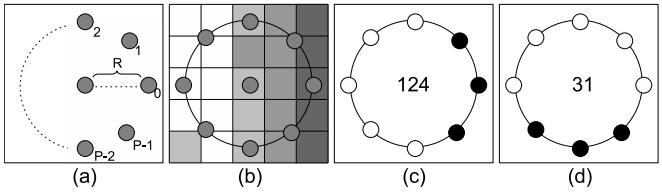


Fig. 2. Illustration of LBP. (a) The filter is defined by two parameters; the circle radius R and the number of samples P on the circle. (b) Local structure is measured w.r.t. a given pixel by placing the center of the circle in the position of that pixel. (c) The samples on the circle are binarized by thresholding with the intensity in the center pixel as threshold value. Black is zero and white is one. The example image shown in (b) has an LBP code of 124. (d) Rotating the example image in (b) ninety degrees clock-wise reduces the LBP code to 31 which is the smallest possible code for this binary pattern. This principle is used to achieve rotation invariance.

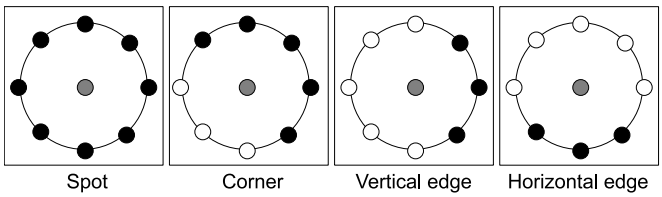


Fig. 3. Various micro-structures measured by LBP. The gray circle indicates the center pixel. Black and white circles are binarized samples; black is zero and white is one.

neighborhood with respect to the center pixel intensity and is given by

$$LBP(\mathbf{x}; R, P) = \sum_{p=0}^{P-1} H(I(\mathbf{x}_p) - I(\mathbf{x}))2^p \tag{1}$$

where I is an image, $\mathbf{x}$ is the center pixel, $\mathbf{x}_p = [-R\sin(2\pi p/P), R\cos(2\pi p/P)]^T + \mathbf{x}$ are P local samples taken at a radius R around $\mathbf{x}$, and $H(\cdot)$ is the Heaviside function. As long as the relative ordering among the gray-scale values in the samples does not change, the output of (1) stays the same; hence, LBP are invariant to any monotonic gray-scale transformation. The application of the LBP filter is illustrated in Fig. 2. Note that, by choosing a fixed sample position on the circle as the "leading bit", in this case the right-most sample, the thresholded samples can be interpreted as bits, and a P bit binary number can be computed.

LBP measure the local structure by assigning unique identifiers, the binary number, to various micro-structures in the image. Thus, LBP capture many structures in one unified framework. In the example in Fig. 2(b), the local structure is a vertical edge with a leftward intensity gradient. Other micro-structures are assigned different LBP codes, e.g., corners and spots as illustrated in Fig. 3. By varying the radius R and the number of samples P, the structures are measured at different scales, and LBP allows for measuring large scale structures without smoothing effects, as is, e.g., the case for Gaussian based filters. We expect emphysematous tissue to contain more edges and homogeneous dark areas compared to normal, healthy tissue. Further, the micro-structures are expected to exist at different scales and frequencies according to the severity of the disease state.

Rotation invariant LBP are achieved by "rotating the circle" until the lowest possible binary number is found

$$LBP^{ri}(\mathbf{x}; R, P) = \min_i (ROR(LBP(\mathbf{x}; R, P), i)) \tag{2}$$

for $i = 0, \ldots, P-1$. $ROR(b, i)$ performs i circular bit-wise right shifts on the P-bit binary number b. When using (2), the horizontal edge and the vertical edge in Fig. 3 are assigned the same LBP code, namely 31. We will use the LBP formulation in (2) in all experiments.

B. Gaussian filter bank

The second type of texture features are computed using a rotation invariant GFB and are based on convolving the image with the Gaussian function

$$G(\mathbf{x}; \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{||\mathbf{x}||_2^2}{2\sigma^2}\right) \tag{3}$$

where σ is the standard deviation, or scale.

Sluimer *et al.* used a similar GFB comprising both rotation variant and invariant filters [12], [13]. Since rotation invariant LBP are used in this work, the GFB we compare with consists of the rotation invariant filters from [12], [13]; Gaussian and Laplacian of the Gaussian, augmented with two more rotation invariant filters; gradient magnitude, which is also used on the original data in [16], and Gaussian curvature.

Letting L_x and L_y denote the first order derivatives of the convolved image $L = I * G(\mathbf{x}; \sigma)$, and L_{xx}, L_{yy} and L_{xy} denote the second order derivatives, the four base filters in the GFB are as follows: the Gaussian function (3) itself, the Laplacian of the Gaussian

$$\nabla^2 G(\mathbf{x}; \sigma) = L_{xx} + L_{yy}, \tag{4}$$

gradient magnitude

$$||\nabla G(\mathbf{x}; \sigma)||_2 = \sqrt{L_x^2 + L_y^2}, \tag{5}$$

and Gaussian curvature

$$K(\mathbf{x}; \sigma) = L_{xx}L_{yy} - L_{xy}^2. \tag{6}$$

C. Feature histograms

Based on the feature values in an ROI, obtained either by computing rotation invariant LBP (2) in all pixels in the ROI or by applying one of the GFB filters (3), (4), (5), or (6), a feature histogram, $f(\text{ROI})$, is computed.

For LBP, the computed LBP codes are directly accumulated into a histogram with the number of bins determined by the number of samples P. In the case of GFB, we employ an adaptive binning principle similar to that of [26]; the total feature distribution across all ROIs in the training set is made approximately uniform. Consequently, densely populated areas in feature space are quantized with a high resolution while sparse areas are quantized with a low resolution. The number of bins is set to $\lfloor \sqrt[3]{N_p} \rfloor$, where N_p is the number of pixels in the ROI.

As noted previously, LBP are invariant to any monotonic gray-scale transformation of the image. This is, however, not always desirable when dealing with CT images, where values

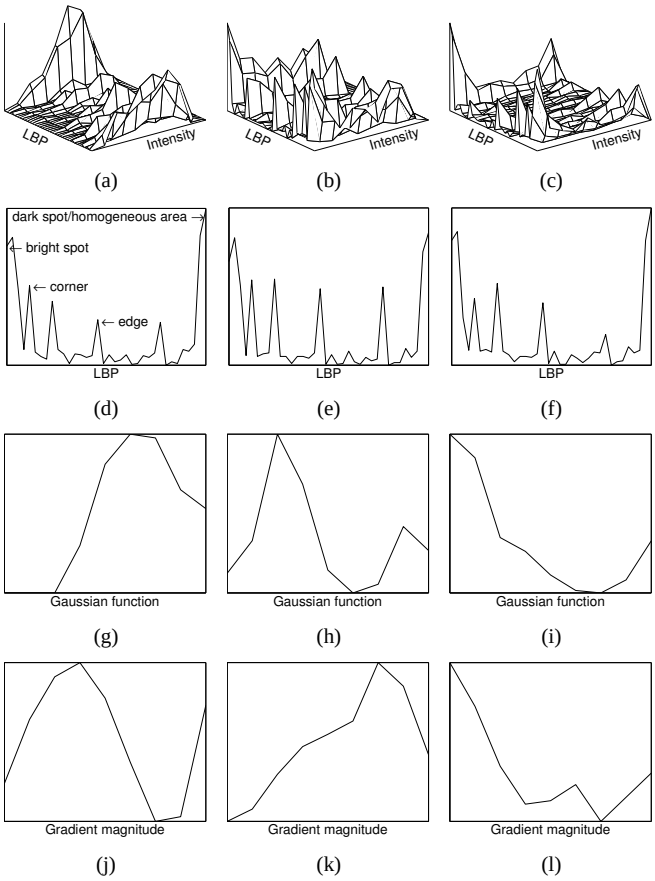


Fig. 4. Examples of feature histograms. (a, d, g, j) Are computed from the NT ROI in Fig. 1(a), (b, e, h, k) from the CLE ROI in Fig. 1(b), and (c, f, i, l) from the PSE ROI in Fig. 1(c). (a-c) Joint LBP and intensity histograms for $R = 1$ and $P = 8$. (d-f) LBP histograms for $R = 1$ and $P = 8$. (g-i) Gaussian function filter response histograms for $\sigma = 0.5$. (j-l) Gradient magnitude filter response histograms for $\sigma = 1$.

are measurements of a physical property of the tissue displayed [12]. Therefore, we include intensity information in the feature histogram by forming the joint histogram between the LBP codes and the intensities in the center pixels. The intensities are binned using the same adaptive principle as used for the GFB filter values [26].

Examples of feature histograms computed from the three different ROIs in Fig. 1 are shown in Fig. 4. Only few bins contain the mass in the LBP histograms, and these bins correspond to different micro-structures such as edges, corners, and spots, as indicated with arrows in Fig. 4(d). One of the discriminating bins when comparing the NT ROI to the CLE ROI is the edge bin as expected, see Fig. 4(d) and Fig. 4(e). The joint LBP and intensity histogram captures information about at which intensities the different micro-structures reside, thus improving discrimination of the NT ROI from the CLE ROI when, e.g., looking at the edge bin in Fig. 4(a) and Fig. 4(b).

D. Classifier

The feature histograms are used to classify ROIs or center pixels of ROIs. For this purpose, we use the *k*NN classifier [24] with the distance between two ROIs being a combined histogram dissimilarity between feature histograms. *k*NN is the natural classifier of choice when working in a distance representation of objects, and it is also the classifier employed in the LBP literature [21], [26]. Further, *k*NN is a non-parametric classifier and therefore able to handle multi-modal class distributions in feature space, which might be the case in lung texture classification. For example, an emphysema class containing samples from different disease stages might contain patterns of varying bullae sizes and varying number of edges, giving rise to a multi-modal class distribution.

Each ROI is represented by a set of feature histograms. When using LBP, the set comprises feature histograms that are measured with different radii for multi-scale analysis. When using GFB, the set comprises feature histograms measured with different filters at different scales. Dissimilarities between ROIs are expressed as dissimilarities between feature histogram sets

$$D_{set}(\text{ROI}_j, \text{ROI}_k) = \sum_{i}^{N_f} D(f_i(\text{ROI}_j), f_i(\text{ROI}_k))$$

where N_f is the number of histograms, $D(\cdot,\cdot)$ is a histogram dissimilarity measure, and $f_i(\cdot)$ are individual feature histograms. In this paper, we use negated histogram intersection [27] as histogram dissimilarity measure

$$D(H, K) = -\sum_{i=1}^{N_b} \min(H_i, K_i)$$

where H and K are histograms each with N_b bins. In this work, all feature histograms are normalized to sum to one, thus $D(\cdot,\cdot) \in [-1, 0]$.

We use a posterior probability estimator for the *k*NN classifier that includes distances to prototypes in the estimation. A principle similar to [28] is employed; the estimation is based on the distance to the n'th nearest prototype of each class where n is the number of prototypes of the majority class within the *k* nearest neighbors of $\mathbf{x}$. The posterior probability of class ω_i given pixel $\mathbf{x}$ is therefore given by

$$P(\omega_i|\mathbf{x}) = \frac{|D_{set}(\text{ROI}(\mathbf{x}), \text{ROI}_n^{\omega_i})|}{\sum_{j=1}^{N_c} |D_{set}(\text{ROI}(\mathbf{x}), \text{ROI}_n^{\omega_j})|} \tag{7}$$

where N_c is the number of classes, $\text{ROI}_n^{\omega_i}$ is the n'th nearest prototype of class ω_i, and $\text{ROI}(\mathbf{x})$ is the ROI centered on $\mathbf{x}$.

E. Emphysema quantification

Prior to classification of the lung field, the lung parenchyma pixels are segmented in the HRCT slice using a combination of thresholding and connected component analysis. Manual editing was needed afterwards in one third of the cases and required simple outlining of a few of the larger airways. In principle, automated methods such as [29], [30] could be used here instead. We denote the obtained segmentation S. Each segmented lung parenchyma pixel is classified by classifying the ROI centered on the pixel.

It should be noted that pixels that are not part of the lung segmentation S are not classified, but they can still contribute

to the classification. For example, part of the exterior of the lung is in the local neighborhood when classifying a pixel at the border of the lung. In this way, all potentially relevant structural information is incorporated, such as proximity to the border of the lung or to the large vessels and airways.

The pixel probabilities are fused to obtain one measure for the complete lung field that can be used for emphysema quantification. There are several ways of doing this, e.g., averaging, voting, or the maximum rule [31]. In this work, we evaluate averaging of soft and hard classification results. The considered quantitative measures for emphysema are the mean class posterior (MCP_{ω_i}) and the relative class area (RCA_{ω_i}). MCP_{ω_i} is given by

$$\mathrm{MCP}_{\omega_i} = \frac{1}{|S|} \sum_{\mathbf{x}_j \in S} P(\omega_i|\mathbf{x}_j) \tag{8}$$

where $|S|$ is the number of lung parenchyma pixels in segmentation S and $P(\omega_i|\mathbf{x}_j)$ is obtained using (7). RCA_{ω_i} is given by

$$\mathrm{RCA}_{\omega_i} = \frac{1}{|S|} \sum_{\mathbf{x}_j \in S} \delta(\arg\max_c P(\omega_c|\mathbf{x}_j) - i) \tag{9}$$

where $\delta(\cdot)$ denotes the Kronecker delta function.

III. Experiments and Results

A. Data

The data comes from an exploratory study carried out at the Department of Respiratory Medicine, Gentofte University Hospital [32] and consist of CT images of the thorax acquired using General Electric equipment (LightSpeed 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, 1.25 mm slice thickness, 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 (FEV_1) were measured [33]. FEV_1 is adjusted for age, sex, and height by dividing with a predicted value according to these three parameters, thereby obtaining FEV_1%pred.

The 39 subjects were divided into three groups: 9 healthy lifelong non-smokers (referred to as never-smokers), 10 smokers without COPD (referred to as healthy smokers), and 20 smokers 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 [1] as follows: no COPD, defined as $\mathrm{FEV}_1/\mathrm{FVC} \geq 0.7$ and $\mathrm{FEV}_1\%\mathrm{pred} \geq 80\%$; moderate to severe COPD, defined as $\mathrm{FEV}_1/\mathrm{FVC} < 0.7$ and $30\% \leq \mathrm{FEV}_1\%\mathrm{pred} < 80\%$. Of the 39 subjects, 19 were women and 20 were men. Table I summarizes the characteristics of the three groups.

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.

TABLE I

Group characteristics reported as mean values, with standard deviation in parentheses and range in square brackets. *n* is the number of subjects in a group.

Group	Age	FEV_1	FEV_1%pred	$\mathrm{FEV}_1/\mathrm{FVC}$
Never-	59	3.15	103	80
smokers	(9)	(0.77)	(9)	(4)
$n = 9$	[47-73]	[2.02-4.08]	[93-121]	[76-89]
Healthy	58	2.90	101	78
smokers	(10)	(0.47)	(8)	(5)
$n = 10$	[47-73]	[1.95-3.60]	[85-113]	[68-78]
COPD	64	1.62	57	54
smokers	(8)	(0.57)	(12)	(7)
$n = 20$	[49-80]	[0.94-2.73]	[37-76]	[42-67]

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 center pixel of the emphysematous area. Square ROIs of a given width centered on the marked pixel were subsequently extracted. PLE was excluded due to under-representation in the data, only two subjects had PLE as leading pattern. Therefore, we are dealing with the three classes $\omega_i \in \{\mathrm{NT}, \mathrm{CLE}, \mathrm{PSE}\}$ in all the experiments.

B. Feature and parameter selection

When using the GFB, feature selection is applied using the sequential forward selection algorithm [24] for deciding which filters at which scales to include. When using LBP, several combinations of radii for multi-resolution analysis are evaluated. In both approaches, different *k*'s in the *k*NN classifier as well as different ROI sizes are evaluated during training. In all cases, parameters and feature sets are optimized based on validation classification accuracy.

C. Classification of ROIs

Classification performance is evaluated by leave-one-subejct-out error estimation on the set of manually annotated ROIs. Six different approaches are evaluated and compared.

The ROIs are represented as points in a feature space, with all features standardized to unit variance, in the first two approaches, and Euclidean distance in the feature space is used in the *k*NN classifier.

1) GFB1: The feature vector consists of the first four central moments computed from histograms of GFB filter responses. Standard histograms are used instead of applying the adaptive binning approach described in Section II-C, and the four filters described in Section II-B are used, resulting in a $16 \times$ scales dimensional feature vector. This set of features resembles the features used in [12], [13].

2) Intensity, Co-occurrence, and Run-length (ICR): The feature vector consists of the following features: the first four central moments of the intensity histogram; the gray-level co-occurrence matrix (GLCM) based measures contrast, correlation, energy, entropy, and homogeneity [6], [25]; and the gray-level run-length matrix (GLRLM) based measures short run emphasis, long run emphasis, gray-level nonuniformity, run-length nonuniformity, and run percentage [6], [25]. The resulting feature vector is 14 dimensional. This set of features resembles the features used in [8], [9], [11], [14], [16]–[18].

The remaining four approaches all use the methods described in Sections II-C and II-D with different feature histograms.

3) INT: Intensity histograms.

4) GFB2: GFB filter response histograms.

5) LBP1: Basic rotation invariant LBP histograms.

6) LBP2: Joint 2D LBP and intensity histograms.

In each leave-out trial, all ROIs from one subject are held out and used for testing. The remaining subjects are separated into a training set and a validation set. In this separation, balanced class distributions are ensured by placing half the subjects representing one class in the training set and the rest in the validation set. The optimal parameter setting is learned using the training and validation sets and can differ for each test subject. Subsequently, the ROIs in the test set are classified using the optimal parameter setting and all the ROIs in the training set and validation set as prototypes in the *k*NN classifier.

In GFB1 and GFB2, the following scales are used for all filters: $\sigma = \{0.5, 1, 2, 4, 8\}$ pixels. In ICR, GLCM and GLRLM are computed using the orientations $\{0, 45, 90, 135\}^\circ$ and the lengths $\{1, 2, \ldots, 5\}$ pixels, and the following binnings of intensity values are evaluated: $\{16, 32, 64\}$ number of bins. The GLCMs are symmetric and mean GLCM measures across orientation and length are used [11], [16]. GLRLM are computed using the Gray Level Run Length Matrix Toolbox [34], and mean GLRLM measures across orientation are used. In LBP1 and LBP2, the following radii and corresponding number of samples are used: $R = \{1, 2\}$ pixels and $P = \{8, 16\}$ samples. Common parameters considered for all six approaches are as follows: ROI size $= \{31 \times 31, 41 \times 41, 51 \times 51\}$ pixels and number of neighbors in the *k*NN classifier $k = \{1, 2, \ldots, 10\}$.

The estimated classification accuracies of the six approaches are summarized in Table II. LBP2 performs best, achieving a classification accuracy of 95.2%. However, it is not significantly different ($p = 0.72$), according to a McNemar's test [35], from the second best approach, GFB2, which achieves an accuracy of 94.0%. As expected, including intensity is important. This is seen in the performance gain between LBP1 and LBP2. In fact, intensity alone performs better than LBP alone, as seen when comparing INT to LBP1. LBP2 performs significantly better than the four approaches GFB1, ICR, INT, and LBP1 ($p < 0.05$). We will focus on the two best performing approaches, GFB2 and LBP2, in the remaining part of Section III.

The confusion matrices in Table III show that LBP2 and GFB2 generally agree on the class labels. Further, GFB2 never mistakes the emphysema classes, and LBP2 only labels a PSE pattern as CLE once. The agreement between the two approaches is further investigated in Section III-D.

TABLE II

ROI CLASSIFICATION ACCURACY AND *p*-VALUE FOR DIFFERENCE WITH LBP2 ACCORDING TO MCNEMAR'S TEST

Approach	Accuracy	*p*-value
GFB1	61.3	$< 10^{-4}$
ICR	89.3	0.016
INT	87.5	0.004
GFB2	94.0	0.724
LBP1	79.2	$< 10^{-4}$
LBP2	95.2	-

TABLE III

CONFUSION MATRICES SHOWING THE TRUE LABEL (ROWS) VS. LABEL ASSIGNED BY THE *k*NN CLASSIFIER (COLUMNS) FOR THE TWO BEST PERFORMING APPROACHES.

LBP2

	NT	CLE	PSE
NT	55	0	4
CLE	1	49	0
PSE	2	1	56

GFB2

	NT	CLE	PSE
NT	55	0	4
CLE	2	48	0
PSE	4	0	55

The parameter settings and filters that were most often selected in the leave-one-subject-out error estimation, for LBP2 and GFB2, are shown in Table IV. The tendency is small scale features, small ROIs, and small *k*.

D. Parenchyma classification

In this section, results of applying the trained classifiers to all pixels within the lung fields are compared for LBP2 and GFB2.

Only one parameter setting is considered for each representation based on the most frequent parameters in Table IV. For LBP2, we use $\{LBP^{ri}(\mathbf{x}; R = 1, P = 8), k = 1, \text{ROI size} = 31 \times 31\}$, and for GFB2, we use $\{\{G(\mathbf{x}; \sigma = 0.5), ||\nabla G(\mathbf{x}; \sigma = 1)||_2\}, k = 1, \text{ROI size} = 31 \times 31\}$. The set of annotated ROIs serve as prototypes in the *k*NN classifier. When classifying the HRCT slices from a particular subject, all the ROI prototypes coming from that same subject are left out in the *k*NN classifier.

Fig. 5 shows examples of the resulting posterior class probabilities assigned by the classifiers in a never-smoker HRCT slice and a COPD smoker HRCT slice. The never-smoker has many high NT probability pixels assigned by both LBP2 and GFB2 as seen in Fig. 5(c) and Fig. 5(d), whereas the COPD smoker has many high CLE probability pixels and some high PSE probability pixels, see Figs. 5(i), 5(j), 5(m), and 5(n). For the shown COPD smoker, the consensus reading of the leading pattern is CLE in all three slices. The LBP2 posterior seems more localized than the GFB2 posterior. See, e.g., the low NT posterior area in the anterior part of the left lung in the slice in Figs. 5(e) and 5(f) and the high CLE posterior area in the same positions in Figs. 5(i) and 5(j).

TABLE IV

MOST FREQUENTLY SELECTED PARAMETERS AND FILTER COMBINATIONS FOR LBP2 AND GFB2 IN THE LEAVE-ONE-SUBJECT-OUT EXPERIMENTS IN SECTION III-C. ONLY PARAMETERS AND FILTERS SELECTED IN AT LEAST 20% OF THE LEAVE-OUT TRIALS ARE SHOWN. OTHER GFB2 FILTERS THAT ARE SELECTED TOGETHER WITH THE REPORTED GFB2 FILTER COMBINATIONS IN LESS THAN 20% OF THE INDIVIDUAL EXPERIMENTS ARE NOT SHOWN.

LBP2					**GFB2**						
$LBP^{ri}(\mathbf{x}; R=1, P=8)$	$LBP^{ri}(\mathbf{x}; R=2, P=16)$	$\{LBP^{ri}(\mathbf{x}; R=1, P=8), LBP^{ri}(\mathbf{x}; R=2, P=16)\}$	$k=1$	ROI size $=31$	$\{G(\mathbf{x};\sigma=0.5), \lVert\nabla G(\mathbf{x};\sigma=1)\rVert_2\}$	$\{G(\mathbf{x};\sigma=0.5), \lVert\nabla G(\mathbf{x};\sigma=2)\rVert_2\}$	$\{G(\mathbf{x};\sigma=0.5), \lVert\nabla G(\mathbf{x};\sigma=1)\rVert_2, \lVert\nabla G(\mathbf{x};\sigma=2)\rVert_2\}$	$\{G(\mathbf{x};\sigma=0.5), \lVert\nabla G(\mathbf{x};\sigma=1)\rVert_2, \lVert\nabla G(\mathbf{x};\sigma=4)\rVert_2\}$	$k=1$	$k=3$	ROI size $=31$
56%	20%	24%	96%	96%	92%	28%	24%	28%	64%	20%	96%

TABLE V

CONFUSION MATRIX BETWEEN GFB2 AND LBP2 ACROSS ALL SUBJECTS FOR ALL LUNG PARENCHYMA PIXELS. THE NUMBERS REPORTED ARE IN PERCENTAGE OF TOTAL NUMBER OF LUNG PARENCHYMA PIXELS.

	LBP2 NT	**LBP2 CLE**	**LBP2 PSE**
GFB2 NT	48.2	1.0	4.1
GFB2 CLE	2.9	16.8	2.5
GFB2 PSE	2.9	0.4	21.3

Correlating the class posteriors shows a high degree of agreement between LBP2 and GFB2; $r = 0.93$ ($p < 10^{-4}$) when correlating $P(\text{NT}|\mathbf{x})$ of the two classifiers, $r = 0.94$ ($p < 10^{-4}$) in the case of $P(\text{CLE}|\mathbf{x})$, and $r = 0.91$ ($p < 10^{-4}$) in the case of $P(\text{PSE}|\mathbf{x})$. Further, class label agreements between LBP2 and GFB2 in each lung parenchyma pixel are shown in the confusion matrix in Table V. This result is based on a hard classification obtained by applying the maximum a posteriori rule in each pixel. The two classifiers generally are in good agreement; in 86.3% of the pixels, the two classifiers agree on the class label.

E. Emphysema quantification

In this section, we evaluate the value of fusing pixel posterior probabilities, computed using the proposed classification system, into a single measure for emphysema.

The full lung classification results of Section III-D are turned into quantitative measures of emphysema using MCP_{ω_i} according to (8) and using RCA_{ω_i} according to (9). These measures are computed across the three HRCT slices representing a subject. We evaluate the obtained measures by correlating with FEV_1%pred, which is one of the standard PFTs for diagnosing subjects with COPD [1]. The common CT based measure RA is also included in the evaluation, in this case using a threshold of -910 HU [3], [5], [32] (RA_{910}). The results are shown in Table VI where the NT based measures achieve correlation coefficients ranging from $r = 0.75$ to $r = 0.77$. For comparison, RA_{910} correlates significantly worse with FEV_1%pred than the NT based measures do ($p < 0.05$) according to a Hotelling/Williams test [36]. In the Hotelling/Williams test, we correct for the difference in signs.

TABLE VI

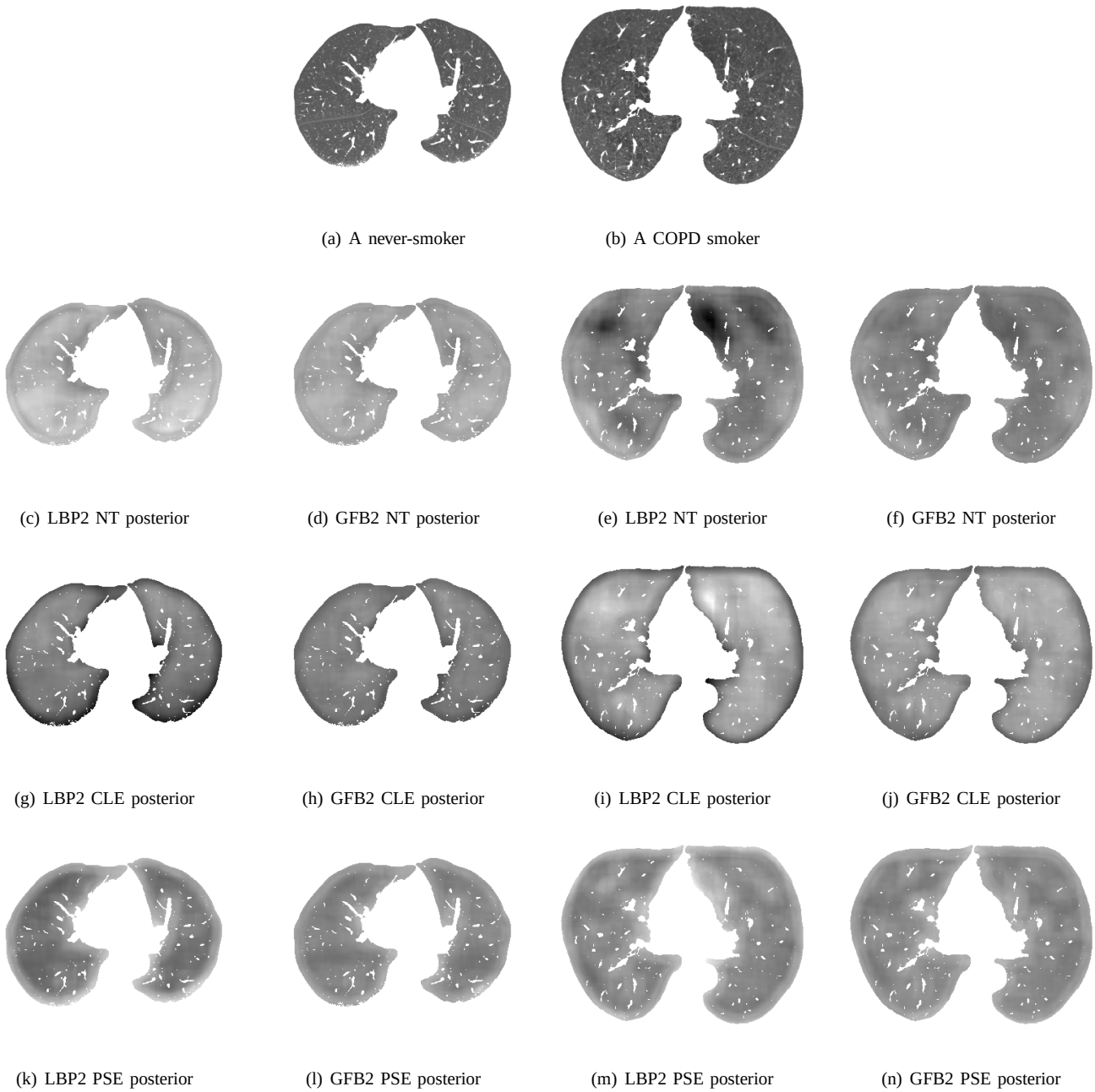
CORRELATION OF CT BASED EMPHYSEMA MEASURES WITH FEV_1%PRED. THE p-VALUES OF THE CORRELATIONS ARE SHOWN IN PARENTHESES.

Measure		***r***
LBP2	MCP_{NT}	0.77 ($p < 10^{-4}$)
	MCP_{CLE}	-0.74 ($p < 10^{-4}$)
	MCP_{PSE}	-0.40 ($p = 0.011$)
	RCA_{NT}	0.77 ($p < 10^{-4}$)
	RCA_{CLE}	-0.78 ($p < 10^{-4}$)
	RCA_{PSE}	-0.66 ($p < 10^{-4}$)
GFB2	MCP_{NT}	0.76 ($p < 10^{-4}$)
	MCP_{CLE}	-0.79 ($p < 10^{-4}$)
	MCP_{PSE}	-0.28 ($p = 0.088$)
	RCA_{NT}	0.75 ($p < 10^{-4}$)
	RCA_{CLE}	-0.74 ($p < 10^{-4}$)
	RCA_{PSE}	-0.66 ($p < 10^{-4}$)
RA_{910}		-0.62 ($p < 10^{-4}$)

All measures, except LBP2 and GFB2 based MCP_{PSE}, separate the group of COPD smokers from the combined group of never-smokers and healthy smokers according to a rank sum test ($p < 0.05$). The separation can also be seen for LBP2 based MCP_{NT} in Fig. 6(a). The figure also shows that the individual features of the joint LBP and intensity feature histogram measure different properties of the parenchyma at subject level. Using intensity alone, i.e., parenchyma density, results in the picture shown in Fig. 6(b), and using LBP alone, i.e., parenchyma micro-structures, results in the picture shown in Fig. 6(c).

IV. DISCUSSION AND CONCLUSION

The proposed classification system using LBP2 achieves an ROI classification accuracy of 95.2%, see Table II, with an NT sensitivity and specificity of 97.3% and 93.2% respectively. This is better than using GFB or ICR, and is within the $75-100\%$ range of NT sensitivities and specificities reported in the literature [9], [11]–[16]. The experiments revealed that



(a) A never-smoker (b) A COPD smoker

(c) LBP2 NT posterior (d) GFB2 NT posterior (e) LBP2 NT posterior (f) GFB2 NT posterior

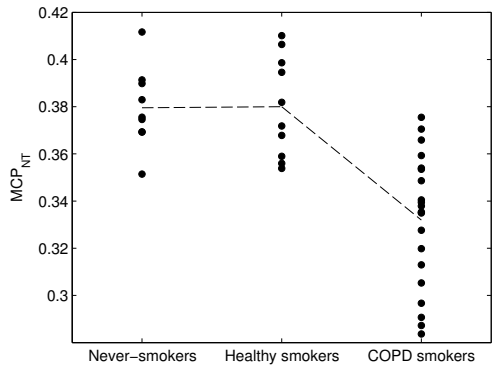
(g) LBP2 CLE posterior (h) GFB2 CLE posterior (i) LBP2 CLE posterior (j) GFB2 CLE posterior

(k) LBP2 PSE posterior (l) GFB2 PSE posterior (m) LBP2 PSE posterior (n) GFB2 PSE posterior

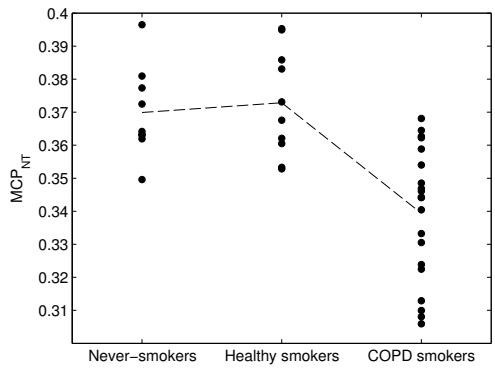
Fig. 5. An HRCT slice from a never-smoker and from a COPD smoker together with posterior probabilities computed in each lung parenchyma pixel. White is high probability and black is low probability. (a, b) Original HRCT slices shown with the window setting $-600/1500$ HU [4]. (c, g, k) LBP2 based posteriors for the never-smoker. (d, h, l) GFB2 based posteriors for the never-smoker. (e, i, m) LBP2 based posteriors for the COPD smoker. (f, j, n) GFB2 based posteriors for the COPD smoker.

using LBP in isolation does not work well in the presented application. This was to be expected, since LBP by design are invariant to monotonic intensity transformations and therefore discard the density information contained in the CT image intensities. Including intensity information via the joint LBP and intensity histogram combines complementary information in the form of micro-structures and densities. Hereby it is measured at which densities the different micro-structures reside which improves discrimination considerably. This is illustrated when comparing Figs. 4(a) and 4(b) to Figs. 4(d) and 4(e) where the differences between the joint histograms 4(a) and 4(b) are much more obvious than between the LBP histograms 4(d) and 4(e). Other feature sets, like GFB, also include intensity and hereby also mix structure and density information. However, obtaining a similar representation to LBP2 in GFB, i.e., joint histograms of structure and density, would require histograms of much higher dimensionality with one dimension for each type of micro-structure such as edge, blob, etc., potentially leading to problems with overfitting in cases of limited number of training samples.

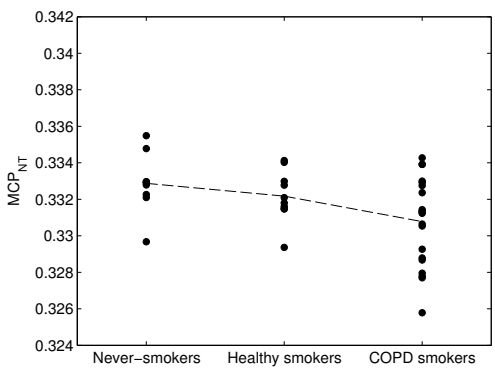
CT based computerized quantitative measures of emphysema are often evaluated by correlating the obtained CT



(a) LBP and intensity jointly, i.e., LBP2 based.



(b) Intensity alone, i.e., INT based.



(c) LBP alone, i.e., LBP1 based.

Fig. 6. $\mathrm{MCP_{NT}}$ for all 39 subjects divided on the three groups; never-smokers, healthy smokers, and COPD smokers. The dashed lines are connecting the means of the three groups.

measures with other markers for disease, such as PFTs or plasma biomarkers, in the clinical literature. A few examples of such studies are: [32], [37]–[40]. In this paper, we have performed a similar evaluation and correlated the proposed quantitative measures with another marker for emphysema in Section III-E, namely $\mathrm{FEV_1}$%pred, and in general the correlations were strong, up to $|r| = 0.79$. It is known that PFTs are noisy measurements [41], and that they are affected by other phenomena than emphysema, e.g., inflammation in the airways. Still, some degree of agreement between PFTs and CT based emphysema measurements is expected. In the future, we intend to further evaluate the proposed CT based emphysema measures by utilizing more meta data. For example, measures derived from blood samples, other PFTs, or health status questionnaires [42] could be used.

Park *et al*. previously performed emphysema quantification based on a hard classification of the lung parenchyma pixels. A weighted sum of the relative areas of mild end severe emphysema was used, and they reported a correlation of -0.47 with $\mathrm{FEV_1}$%pred [16]. Based on the results of our experiments, nothing conclusive can be stated about fusion of soft versus fusion of hard classifications, i.e., (8) versus (9). Both methods work well, and as seen in Table VI, all the NT based measurements correlate significantly with $\mathrm{FEV_1}$%pred with correlation coefficients in the range $r = 0.75$ to $r = 0.77$. It should be noted that in [16], the correlation between RA, using a threshold of -950 HU, and $\mathrm{FEV_1}$%pred is -0.42. In our data, agreement between CT and PFT generally seems to be better with the correlation between $\mathrm{RA_{910}}$ and $\mathrm{FEV_1}$%pred being -0.62. This fact, as well as the fact that we are dealing with a broader range of subjects compared to [16], could explain the difference in correlation level.

Different features may capture different information in the CT images, and though the per pixel posterior probabilities of LBP2 and GFB2 are highly correlated, there may still be something to gain by combining the output of the two classifiers. This was tested by combining the pixel posteriors $P(\omega_i|\mathbf{x})$ of LBP2 and GFB2 with the sum rule and the maximum rule respectively [43], followed by posterior fusion of the combined pixel posteriors. These results did not show any significant improvement in correlation with $\mathrm{FEV_1}$%pred, which indicates that the two classifiers indeed capture similar information in the CT images.

In this work, we have done no preprocessing of the HRCT slices prior to computing feature histograms. Instead, we have relied on the filters to perform the necessary processing, e.g., noise smoothing in the GFB by selecting the appropriate σ or picking up certain micro-structures from the noisy background in LBP by selecting the appropriate radius. It is up to the training procedure to pick these settings.

Further, all available information has been taken into account in the feature histogram estimation by including all pixels in the ROIs instead of first excluding, e.g., the vessels [11], [14] or the airways [14]. Thus, also pixels outside the lung fields and pixels from non-parenchyma structures within the lungs contribute. This can be seen as an implicit way of encoding context information in the feature histograms, with the position being “near the border of the lung” or “near the hilar area”. It may lead to slight overestimation of PSE at the border, see Fig. 5, but in practice, this should not be a real problem as long as the prototype set contains samples at the border representing all classes. In our data, NT and CLE ROIs were mainly annotated in the central parts of the lungs. Nevertheless, the proposed classification system is capable of discriminating between normal tissue and emphysematous tissue within the lung, as seen in the confusion matrices for LBP2 and GFB2 in Table III.

Fig. 6(a) reveals that very similar measures are obtained for the never-smokers and the healthy smokers. One might expect the healthy smokers measures to be slightly lower than the never-smokers on the $\mathrm{MCP_{NT}}$ scale due to early

stages of emphysema not yet detectable by PFTs. Basing the measurements only on intensity feature histograms results in the healthy smokers having an even larger probability of NT as seen in Fig. 6(b), indicating a difference in density for the two groups. This corresponds well with recent results indicating that lung parenchyma is more dense in healthy smokers than in never-smokers possibly due to smoke induced inflammation [40], [44]. On the other hand, basing the measurements solely on LBP feature histograms results in a slight drift downwards as seen in Fig. 6(c), suggesting that there may be structural differences that can be captured at an early stage by LBP. As described in Section II-A, LBP are gray-scale invariant and therefore not affected by parenchymal density changes. This also implies that the proposed classification system should be less sensitive to inspiration level as compared to, e.g., RA_{910}.

Basing the discrimination of ROIs on dissimilarities between sets of feature histograms, using a combined histogram dissimilarity directly as distance in a *k*NN classifier, works well in this setting. Both LBP2 and GFB2 achieve good ROI classification accuracies and high correlations with FEV_1%pred. Using full feature histograms differs from the common approach of using measures derived from feature histograms, such as moments of filter response histograms or GLCM measures, as features in a feature space [8]–[18]. Looking at Fig. 4, taking only the first four moments of the GFB histograms could potentially discard valuable information about the shape of the histograms such as the presence of multiple mods. A previous comparative study of texture features for classification reported similar findings on two standard texture data sets [26]. In this work, we wanted to exploit the full feature histograms and therefore used the *k*NN classification framework with histogram dissimilarity as distance, and LBP and GFB were shown to work very well in this setting. It remains of course a possibility that in a different classification scheme, relying on features rather than on dissimilarity measures, a different feature set would perform as good as or even better than LBP. Alternatively, histogram dissimilarities could be applied within the dissimilarity-based classification schemes proposed by Pekalska *et al.* [45].

The experiments carried out in this work are all done on HRCT slices, but the general framework could easily be extended to 3D. However, no true extension of rotation invariant LBP to 3D exists. Two approximative extensions of LBP to 3D are presented in [46], with the specific application being temporal texture data. The first approach forms a helical path in the temporal direction. This idea could be applied in volumetric CT by, e.g., forming helical paths in various directions and combining the resulting LBPs. The second approach in [46] computes 2D LBPs in three orthogonal planes and combines these.

In conclusion, we propose to use texture measures such as LBP for quantitative analysis of pulmonary emphysema in CT images of the lung. ROI classification experiments showed good classification performance, with an accuracy of 95.2%, and quantitative measures of emphysema derived by fusing posterior probabilities achieved high correlation with PFT, up to $|r| = 0.79$ ($p < 10^{-4}$). Overall, LBP seem to perform slightly better than a rotation invariant GFB, although the difference was not significant in our experiments. MCP_{NT} correlated significantly better with pulmonary function than the most common standard CT measure, RA, which suggests that texture based measures may be better indicators of the degree of emphysema. In addition, LBP seem to pick up certain micro-structures that are more frequent in smokers, including smokers who still have good lung function, than in people who never smoked. This structural information improves discrimination in our experiments and may also improve sensitivity to early changes in lung tissue integrity.

ACKNOWLEDGMENT

This work is financially supported by the Danish Council for Strategic Research under the Programme Commission for Nanoscience and Technology, Biotechnology and IT (NABIIT), the Netherlands Organization for Scientific Research (NWO), and AstraZeneca, Lund, Sweden.

Fig. 1, part of Fig. 2, part of Fig. 3, part of Table II, and part of the introduction have been reused from [19] with kind permission of Springer Science+Business Media.

REFERENCES

[1] 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," *Am. J. Respir. Crit. Care Med.*, vol. 176, no. 6, pp. 532–555, Sep 2007.

[2] J. W. Gurney, "Pathophysiology of obstructive airways disease." *Radiol. Clin. North Am.*, vol. 36, no. 1, pp. 15–27, Jan 1998.

[3] 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.

[4] W. R. Webb, N. Müller, and D. Naidich, *High-Resolution CT of the Lung, Third Edition*, J.-R. John, Ed. Lippincott Williams & Wilkins, 2001.

[5] 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.

[6] 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.

[7] M. Tuceryan and A. K. Jain, *The Handbook of Pattern Recognition and Computer Vision (2nd Edition)*. World Scientific Publishing, 1998, ch. Texture Analysis, pp. 207–248.

[8] R. Uppaluri, T. Mitsa, M. Sonka, E. A. Hoffman, and G. McLennan, "Quantification of pulmonary emphysema from lung computed tomography images," *Am. J. Respir. Crit. Care Med.*, vol. 156, no. 1, pp. 248–254, Jul 1997.

[9] 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.*, vol. 160, no. 2, pp. 648–654, Aug 1999.

[10] O. Friman, M. Borga, M. Lundberg, U. Tylén, and H. Knutsson, "Recognizing emphysema - a neural network approach," in *ICPR (1)*, Mar 2002, pp. 512–515.

[11] 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, Sep 2003.

[12] 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.

[13] 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.

[14] Y. Xu, M. Sonka, G. McLennan, J. Guo, and E. A. Hoffman, "MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies." *IEEE Trans. Med. Imag.*, vol. 25, no. 4, pp. 464–475, Apr 2006.

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

[16] 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, June 2008.

[17] 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.

[18] 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.

[19] 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, D. N. Metaxas, L. Axel, G. Fichtinger, and G. Székely, Eds., vol. 5241. Springer, Sep 2008, pp. 934–941.

[20] L. Sørensen, S. Shaker, and M. de Bruijne, "Texture based emphysema quantification in lung CT," in *Proc. of The First International Workshop on Pulmonary Image Analysis*, 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.

[21] 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.

[22] 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, N. Ayache, S. Ourselin, and A. J. Maeder, Eds., vol. 4791. Springer, Nov 2007, pp. 286–293.

[23] D. Unay, A. Ekin, M. Cetin, R. Jasinschi, and A. Ercil, "Robustness of local binary patterns in brain MR image analysis," in *Proc. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBS 2007*, Aug 2007, pp. 2098–2101.

[24] A. Jain, R. Duin, and J. Mao, "Statistical pattern recognition: a review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[25] R. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.

[26] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, Jan 1996.

[27] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comput. Vision*, vol. 7, no. 1, pp. 11–32, Nov 1991.

[28] R. P. W. Duin and D. M. J. Tax, "Classifier conditional posterior probabilities," in *SSPR/SPR*, ser. Lecture Notes in Computer Science, A. Amin, D. Dori, P. Pudil, and H. Freeman, Eds., vol. 1451. Springer, Aug 1998, pp. 611–619.

[29] S. Hu, E. Hoffman, and J. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, Jun 2001.

[30] P. Lo, J. Sporring, J. J. H. Pedersen, and M. de Bruijne, "Airway tree extraction with locally optimal paths," in *MICCAI (1)*, ser. Lecture Notes in Computer Science, G.-Z. Yang, D. J. Hawkes, D. Rueckert, J. A. Noble, and C. J. Taylor, Eds., vol. 5762. Springer, Sep 2009, pp. 51–58.

[31] M. Loog and B. Van Ginneken, "Static posterior probability fusion for signal detection: applications in the detection of interstitial diseases in chest radiographs," in *ICPR (1)*, Aug 2004, pp. 644–647.

[32] 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," *The Clinical Respiratory Journal*, vol. 2 (1), pp. 17–25, 2008.

[33] 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.

[34] X. Wei, "Gray level run length matrix toolbox v1.0," Software, Beijing Aeronautical Technology Research Center, 2007. [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/17482

[35] T. G. Dietterich, "Approximate statistical test for comparing supervised classification learning algorithms," *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, Oct 1998.

[36] J. V. Sickle, "Analyzing correlations between stream and watershed attributes," *J. AM. WATER RESOUR. AS.*, vol. 39, no. 3, pp. 717–726, Jun 2003, errata: 41(3) 741–741.

[37] M. Kinsella, N. L. Müller, R. T. Abboud, N. J. Morrison, and A. Dy-Buncio, "Quantitation of emphysema by computed tomography using a "density mask" program and correlation with pulmonary function tests." *Chest*, vol. 97, no. 2, pp. 315–321, Feb 1990.

[38] G. A. Gould, A. T. Redpath, M. Ryan, P. M. Warren, J. J. Best, D. C. Flenley, and W. MacNee, "Lung CT density correlates with measurements of airflow limitation and the diffusing capacity." *Eur. Respir. J.*, vol. 4, no. 2, pp. 141–146, Feb 1991.

[39] M. Schwaiblmair, T. Beinert, M. Seemann, J. Behr, M. Reiser, and C. Vogelmeier, "Relations between cardiopulmonary exercise testing and quantitative high-resolution computed tomography associated in patients with alpha-1-antitrypsin deficiency." *Eur. J. Me.d Res.*, vol. 3, no. 11, pp. 527–532, Nov 1998.

[40] K. Soejima, K. Yamaguchi, E. Kohda, K. Takeshita, Y. Ito, H. Mastubara, T. Oguma, T. Inoue, Y. Okubo, K. Amakawa, H. Tateno, and T. Shiomi, "Longitudinal follow-up study of smoking-induced lung density changes by high-resolution computed tomography," *Am. J. Respir. Crit. Care Med.*, vol. 161, no. 4, pp. 1264–1273, Apr 2000.

[41] A. Dirksen, N. H. Holstein-Rathlou, F. Madsen, L. T. Skovgaard, C. S. Ulrik, T. Heckscher, and A. Kok-Jensen, "Long-range correlations of serial FEV1 measurements in emphysematous patients and normal subjects," *J. Appl. Physiol.*, vol. 85, no. 1, pp. 259–265, Jul 1998.

[42] P. W. Jones, "Health status measurement in chronic obstructive pulmonary disease," *Thorax*, vol. 56, no. 11, pp. 880–887, Nov 2001.

[43] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, March 1998.

[44] H. Ashraf, P. Lo, S. Shaker, M. de Bruijne, A. Dirksen, P. Tønnesen, M. Dahlbäck, and J. H. Pedersen, "Change in smoking habits affects lung density by CT," in *American Thoracic Society International Conference*, 2009.

[45] E. Pekalska and R. P. W. Duin, "Dissimilarity representations allow for building good classifiers," *Pattern Recog. Lett.*, vol. 23, no. 8, pp. 943–956, Jun 2002.

[46] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun 2007.