Texture Based Emphysema Quantification in Lung CT

Lauge Sørensen[1], Saher B. Shaker[2], and Marleen de Bruijne[1,3]

[1] Department of Computer Science, University of Copenhagen, Denmark,
{lauges, marleen}@diku.dk
[2] Department of Cardiology and Respiratory Medicine, Hvidovre University Hospital, Copenhagen, Denmark
[3] Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, The Netherlands

Abstract. In this paper we propose to use texture based pixel classification in lung computed tomography (CT) for measuring emphysema. Two quantitative parameters for emphysema, based on the pixel classification, are suggested; relative class area and mean class posterior.
The approach is evaluated on a group of 39 patients, of whom 20 have been diagnosed with chronic obstructive pulmonary disease, using two different feature groups, local binary patterns and a filter bank based on Gaussian derivatives. The pixel classification based quantitative parameters correlate well with lung function ($r = 0.80$, $p < 10^{-5}$ for the parameter with the highest correlation) and correlate significantly better than the most commonly used CT based emphysema quantification method, namely relative area of low attenuation.

1 Introduction

Chronic obstructive pulmonary disease (COPD) is a major cause of death and a growing health problem worldwide. In the United States it is the fourth leading cause of morbidity and mortality and it is estimated to be ranked the fifth most burdening disease worldwide by 2020 [1]. COPD is a chronic lung disease characterized by limitation of airflow in the airway and it comprises two components: Chronic bronchitis, which is an inflammation of the small airways, and emphysema, which is characterized by gradual loss of lung tissue.

The primary diagnostic tools for COPD are lung function tests (LFT). Another diagnostic tool that is gaining more and more attention is computed tomography (CT) imaging. CT is a sensitive method for diagnosing emphysema and both visual and quantitative CT are closely correlated with the pathological extent of emphysema [2]. This makes CT suitable for both early detection and study of COPD as well as for monitoring the effect of different treatments.

We focus on the assessment of emphysema, which is thought to be the main cause of shortness of breath and disability in COPD. Emphysema is usually classified into three subtypes, or patterns, and we will adopt the naming and definitions from Webb et al. [3]. These subtypes are: Centrilobular emphysema (CLE) defined as multiple, small, spotty lucencies, that may have thin walls,

paraseptal emphysema (PSE) defined as multiple, lucencies in a single layer along the pleura, commonly with thin walls visible, and panlobular emphysema (PLE) defined as a lucent lung with small pulmonary vessels.

In CT emphysema lesions, or bullae, are visible as areas of abnormally low attenuation values, close to that of air. Different objective quantitative measures of emphysema can be derived from the histogram of CT attenuation values. The most common measure is the emphysema index or relative area of low attenuation (RA) [2], which measures the amount of lung parenchyma pixels that have values below a certain threshold relative to the total amount of lung parenchyma pixels. Common for the quantitative methods based on the attenuation histogram is that they ignore the possibly valuable information inherent in the emphysema disease patterns, such as subtype, shape, and size distribution.

One way to objectively analyze the properties of the disease patterns is to use texture analysis [4]. Several publications exist on characterizing emphysema and other disease patterns in regions of interest (ROI) in lung CT images using texture features [5–9]. In [5, 7, 8] the entire lung is labelled by subdividing the lung into adjacent ROIs followed by a classification step that assigns the same label to all pixels within a ROI. In [5, 7] the labelled result is evaluated by comparing the agreement between the output of the classification and that of human expert readings of the same ROIs. [8] Reports the percentage of different disease patterns present.

In this paper we propose a quantitative measure for emphysema, based on a pattern classification approach that utilizes local texture information. Compared to RA, a pattern classification approach allows for more than the two classes healthy and emphysematous, making it possible to quantify different subtypes of emphysema, which may be related to prognosis of the patient. Further, texture may be less influenced by inspiration level, compared to using intensity alone. We perform full lung classification by computing the posterior class probability for each pixel in the lung based on the local neighborhood around the pixel. Two ways of deriving a quantitative measure for emphysema, from the posterior, are investigated and evaluated. The first approach is to perform a hard classification and compute the relative area of the classes, the second approach is to compute the average posterior probability of each class in the lung. Thus, we obtain two measures per class. To our knowledge the use of the mean class posterior probability for quantifying emphysema is novel. We experiment with using two different kinds of features for this purpose, local binary patterns [10] and a filter bank based on Gaussian derivatives. These two feature groups were previously tested and evaluated on a set of hand picked ROIs, achieving an accuracy of 95.2% and 94.0% respectively in discriminating between normal tissue (NT), CLE, and PSE [9]. The experimental results reported in [9] are based on a subset of the CT images used in the experiments in this paper.

2 Methods

In the following the lung pixel classification system is described. Sections 2.1 and 2.2 describe the features we use for characterizing the lung texture, namely local binary patterns and a filter bank based on Gaussian derivatives, Section 2.3 describes the classification framework, and Section 2.4 describes how the classification framework is used for pixel classification and how the classification result is turned into a quantitative measure for emphysema.

2.1 Local Binary Patterns

The first group of features is based on the local binary patterns (LBP) proposed by Ojala et al. [10]. LBP measures the local structure at a given pixel by thresholding P samples on a circle of radius R around the pixel using the intensity in the pixel as threshold. The resulting thresholded samples are interpreted as a binary number, that provides a unique code for each kind of local structure or pattern. The operator is highly non-linear and detects microstructures in the image at different resolutions governed by the parameter R, for example spots, edges, corners, etc. Applying the LBP operator to an ROI results in an LBP code image. Based on this an LBP histogram is formed by accumulating the LBP codes directly into a histogram. We use the rotation invariant formulation of LBP, see [9, 10] for more details. LBP are by design gray-scale invariant, and this is not a desired property when dealing with CT images, where values are measurements of a physical property of the tissue displayed. Therefore the distribution of the intensities is included, by forming the joint histogram between the LBP and the intensities in the center pixels.

2.2 Gaussian function and its derivatives

The second group of features is based on the Gaussian function $G(\mathbf{x};\sigma)$ and combinations of derivatives of $G(\mathbf{x};\sigma)$. We use $\mathbf{x} = [x, y]^T$ to denote the pixel position. By varying the standard deviation σ of the function in a discrete manner we obtain a whole bank of filters that can be applied to the image by convolution. The Gaussian function itself is included to make the filter bank sensitive to offsets in absolute intensity. The filters that we use are similar to those used in [7], except that the filters we use are all rotation invariant. The filter bank comprise the following filters: $G(\mathbf{x};\sigma)$; the Laplacian of the Gaussian $\nabla^2 G(\mathbf{x};\sigma)$; the gradient magnitude $|\nabla G(\mathbf{x};\sigma)|_2$; the Gaussian curvature $K(\mathbf{x};\sigma) = \partial^2 G(\mathbf{x},\sigma)/\partial x^2 + \partial^2 G(\mathbf{x},\sigma)/\partial y^2 - 2\partial^2 G(\mathbf{x},\sigma)/\partial xy$. Feature histograms are obtained by convolving the ROIs with each filter and making histograms of each filter response.

2.3 Classification framework

Classification is done using the kNN classifier [11] with combined histogram similarities as distance measure. We use histogram intersection as the similarity

Model	k	Window size	Features	Feature specific
Model 1	1	31×31 pixels	Joint LBP and intensity histogram, see Section 2.1	One resolution: R = 1, P = 8
Model 2	1	31×31 pixels	Histograms of filter responses of the Gaussian derivative based filters, see Section 2.2	Two filters: $G(\mathbf{x}, \sigma = 0.5)$ $\lvert\nabla G(\mathbf{x}, \sigma = 1)\rvert_2$

Table 1. Parameter settings for the two models used in the pixel classification experiments. k is the number of neighbors used in the kNN classifier. The unit of R and σ is pixels and the unit of P is samples.

measure between histogram H and histogram K

$$L_{hist}(H, K) = 1 - \sum_{b=1}^{N_b} \min(H_b, K_b), \tag{1}$$

where H_b denotes bin b of histogram H, N_b is the number of histogram bins and the histograms are assumed normalized to sum to one. In the case of measuring combined histogram similarity based on different histograms, e.g. different Gaussian based filters, the similarities are computed individually for each feature histogram using (1) and summed afterwards

$$L(\mathbf{x}, \omega_i, m) = \sum_{n}^{N_f} L_{hist}(f_n(\mathbf{x}), M_{m,n}^{\omega_i}), \tag{2}$$

where $M_{m,n}^{\omega_i}$ is the n'th feature histogram of prototype m from class ω_i, $f_n(\mathbf{x})$ denotes some function that extracts the local neighborhood around pixel $\mathbf{x}$ in the current image and computes the n'th feature histogram, and N_f is the number of feature histograms used in the combination. The histograms of the intensities and the filter responses are constructed using non-linear binning, where the binning is found by employing two rules on the total distribution of the ROIs in the training set: The total distribution should be approximately uniform and the number of bins is $\lfloor \sqrt[3]{N_p} \rfloor$, where N_p is the number of pixels in the ROI.

In [9] the accuracy of both feature groups is estimated as an average of a number of leave-one-patient-out experiments, in which the optimal filters and parameters are selected on the training set and can vary in between the experiments. Table 1 summarizes the parameter settings for the two feature groups that we will be using in this paper, which are those that were most often selected in the experiments in [9].

2.4 Quantification by pixel classification

Prior to classification, the lung parenchyma is extracted using a combination of thresholding, connected component analysis, and manual editing. The posterior probability of class ω_i given pixel $\mathbf{x}$ is computed based on the combined histogram similarity (2) with the closest prototype histograms of each class and is given by

$$P(\omega_i|\mathbf{x}) = \frac{\min_m L(\mathbf{x}, \omega_i, m)}{\sum_{j=1}^{3} \min_m L(\mathbf{x}, \omega_j, m)}. \tag{3}$$

The size of the local neighborhood is the same as the window size used in the model, i.e. 31×31 pixels, see Table 1, and the pixel being classified is in the center of this window. The classification result is a posterior class probability for each pixel in the lung.

A hard classification can be obtained by using the maximum a posteriori (MAP) rule in each pixel [11]. It should be noted that pixels that are not part of the lung segmentations are not classified, but they can still contribute to the classification, e.g. 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 potential relevant structural information is included, like being at the border of the lung or near large vessels and airways.

The quantitative measures for emphysema that we propose are the relative class area (RCA) and the mean class posterior (MCP). RCA is defined by the relative amount of pixels with a given class label, obtained using the MAP rule, divided by the total number of lung pixels N_l

$$\mathrm{RCA}_{\omega_i} = \frac{1}{N_l} \sum_{j}^{N_l} \delta(\arg\max_c P(\omega_c|\mathbf{x}_j) - i), \tag{4}$$

where δ denotes the Kronecker delta function. MCP is given by averaging the posterior class probability of a given class, obtained using (3), across all pixels in the lung

$$\mathrm{MCP}_{\omega_i} = \frac{1}{N_l} \sum_{j}^{N_l} P(\omega_i|\mathbf{x}_j). \tag{5}$$

3 Experiments and Results

3.1 Data

The data used for the experiments consists of a set of thin-slice CT images of the thorax. CT was performed using GE equipment (LightSpeed QX/i; GE Medical Systems, Milwaukee, WI, USA) with four detector rows, 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 milliampere (mA). The slices were reconstructed using a high spatial resolution (bone) algorithm. A population of 39 individuals, 9 healthy non-smokers, 10 smokers without COPD, and 20 smokers diagnosed with moderate or severe COPD according to LFT [1] were scanned in the upper, middle, and lower lung, resulting in a total of 117 CT slices. Visual assessment of the leading pattern, either NT, CLE, PSE, or PLE, and severity, ranging from 0 to 5, in each of these slices was done individually by an experienced chest radiologist and a CT experienced pulmonologist. In cases of disagreement, consensus readings were obtained. 216 non-overlapping ROIs were annotated in the slices representing the three classes: NT (107 observations, of which 48 were near the lung border or hilum area), CLE (50 observations), and PSE (59 observations). PLE was excluded due to underrepresentation in the data

	Model 2 NT	Model 2 CLE	Model 2 PSE
Model 1 NT	49.6	2.1	3.9
Model 1 CLE	2.4	18.8	3.5
Model 1 PSE	2.3	0.9	16.6

Table 2. Confusion matrix showing the percentage of pixel labels that the two models agree/disagree on.

set as only 2 out of the 20 individuals diagnosed with COPD had PLE as the leading pattern. These 216 ROIs were used as prototypes in the kNN classifier.

The 39 individuals also underwent LFT, performed according to the European Respiratory Society recommendations, prior to the CT scanning of the lungs. One widely used LFT is forced vital capacity in one second (FEV_1) which is the amount of air in liters that you can forcibly blow out in one second. FEV_1 can be adjusted for age, sex, and height by dividing by a predicted value according to these three parameters, thereby obtaining FEV_1%pred, and it is this LFT that we will use in the evaluation.

3.2 Lung pixel classification

Each of the 39 individuals were in turn measured using our proposed approach by classifying the pixels in each of the three CT slices using either Model 1 or Model 2, while leaving the prototypes coming from that individual out of the kNN classifier. The classification result is then used as a quantitative measure for emphysema by applying (4) or (5) to the posterior. Figure 3.2 shows the CT slices from two different patients, along with obtained pixel classifications and NT pixel posterior, when using Model 1. For comparison the RA below -910 HU (RA910) is shown in Figure 3.2 bottom-left.

3.3 Comparison of Model 1 and Model 2

The confusion matrix in Table 2 reveals that the two models generally are in good agreement; in 85% of the pixels, the two models agree on the class label. The highest level of disagreement is in the cases where Model 1 labels a pixel as NT or CLE whereas Model 2 labels that pixel as PSE, which happens in 3.9% and 3.5% of the cases. Correlating the class posteriors also shows a high degree of agreement between the two models, with $r = 0.93$ for NT, $r = 0.93$ for CLE, and $r = 0.91$ for PSE.

Two specific cases of disagreement are shown in Figure 2. In the first case Model 1 labelled half the pixels as CLE whereas Model 2 labelled these pixels mostly NT and partly CLE. Interestingly RA910 does not label one single pixel as emphysematous. Perhaps the models have picked up on a texture pattern and really found some emphysema, which was not possible using a simple threshold of -910 HU. Increasing the threshold to -860 HU reveals some low attenuation pixels partly in the areas labelled CLE by Model 1. Thus it seems that there is

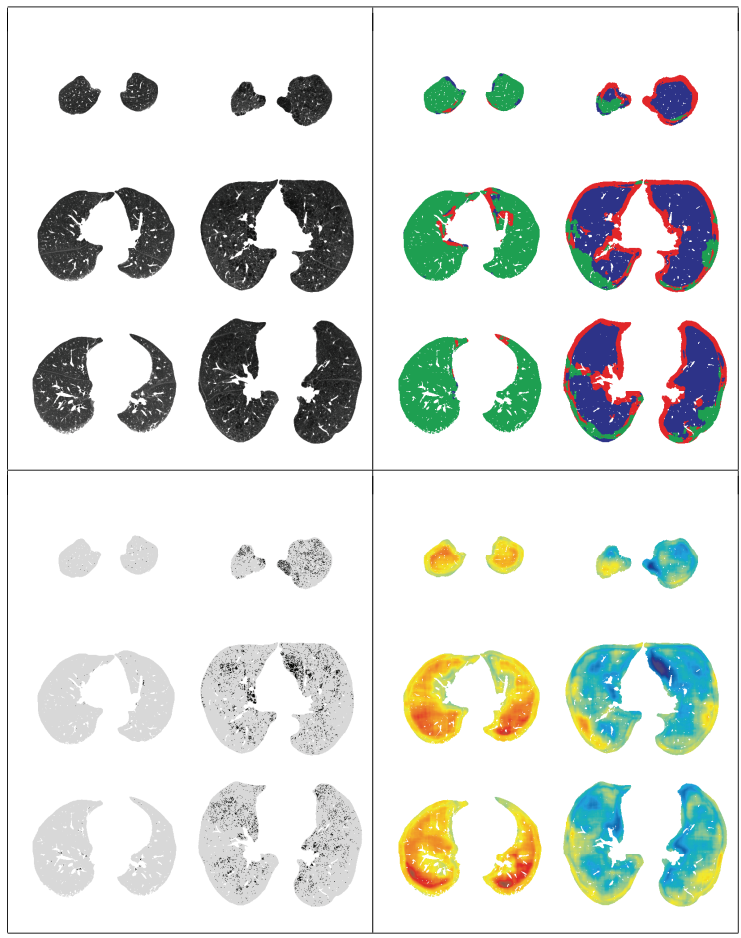

Fig. 1. Example images from two patients. The images are organized as follows: First column of each sub-box is a healthy non-smoker and second column is a patient diagnosed with moderate COPD according to LFT [1]. First row of each sub-box corresponds to the upper scan, second row to the middle scan, and third row to the bottom scan. **Top-left:** The segmented images used in the classification. The images are shown with the window setting -600/1500 HU. **Top-right:** The pixel classification result obtained when applying Model 1 to the lung segmentations shown top-left. Green corresponds to NT, blue to CLE, and red to PSE. **Bottom-left:** A threshold of -910 HU applied to the lung segmentation shown top-left. The areas below the threshold are indicated in black and the lung segmentations are indicated in light gray. **Bottom-right:** Posteriori NT probabilities. Dark red means that a NT prototype ROI is very similar to the given pixel's neighborhood in histogram feature space and corresponds to a NT probability of 0.48 when using (3). Dark blue means that all NT prototype ROIs are dissimilar and corresponds to a NT probability of 0.16. Refer to the electronic version for colors.

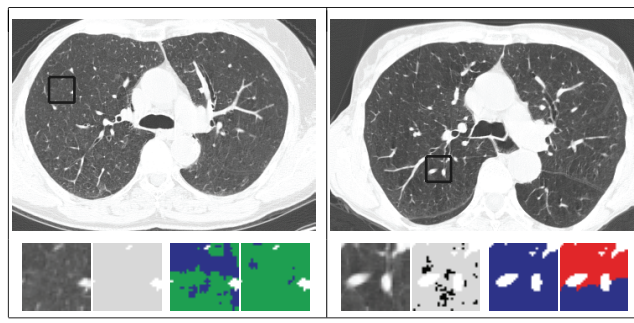

Fig. 2. Visual inspection of the classification results obtained using the two models in Table 1. **Left:** A non-smoker. **Right:** A smoker diagnosed with COPD. **Top:** The CT slice overlayed with a square indicating where the particular case is taken from. **Bottom:** Zoom in on the cases, from left to right; the original image, a threshold of -910 HU applied to the image with the areas below the threshold indicated in black and the lung segmentations indicated in light gray, the classification result obtained using Model 1, and the classification result obtained using Model 2. Green corresponds to NT, blue to CLE, and red to PSE. Refer to the electronic version for colors.

some emphysematous pattern present that Model 1 picks up. In the second case in the right part of Figure 2 Model 1 has labelled all lung pixels as CLE whereas Model 2 has labelled some CLE and some PSE. The ROI in the second case is from within the lung and thus by definition it should not be PSE. An explanation of why Model 2 labels many of the pixels PSE could be the large vessels seen within the ROI. Since we are using histograms, the spatial information in the ROI is ignored, and thus emphysematous regions with large vessels can share similarities with emphysema at the boundary in histogram feature space.

3.4 Relation to lung function

We evaluate the quantitative measures obtained from the pixel classification by correlating them with two other measures for emphysema, namely FEV_1%pred, representing the classical objective way of measuring COPD by LFT, and an emphysema score (ES) computed by summing the visually assessed emphysema severity across the three slices. ES represents the subjective way of measuring emphysema by human visual assessment. For reference, we also compute RA910 for each of the patients and correlate that with FEV_1%pred and ES. The Correlation results are reported in Table 3, where the correlations with FEV_1%pred are computed using the Pearson correlation coefficient r and the correlations with ES are computed using the Kendall tau correlation coefficient τ.

RCA correlate well with FEV_1%pred and generally also MCP. All texture based quantitative parameters for emphysema show significant correlation with

Measure	FEV_1%pred	ES	Separation
Model 1 RCA_{NT}	0.80 ($< 10^{-5}$)	-0.56 (0.006)	$< 10^{-5}$
Model 1 RCA_{CLE}	-0.77 ($< 10^{-5}$)	0.55 (0.006)	$< 10^{-5}$
Model 1 RCA_{PSE}	-0.77 ($< 10^{-5}$)	0.49 (0.067)	$< 10^{-5}$
Model 1 MCP_{NT}	0.73 ($< 10^{-5}$)	-0.56 ($< 10^{-5}$)	$< 10^{-5}$
Model 1 MCP_{CLE}	-0.58 (0.0001)	0.48 (0.0001)	0.0005
Model 1 MCP_{PSE}	-0.73 ($< 10^{-5}$)	0.50 ($< 10^{-5}$)	$< 10^{-5}$
Model 2 RCA_{NT}	0.79 ($< 10^{-5}$)	-0.56 (0.004)	$< 10^{-5}$
Model 2 RCA_{CLE}	-0.76 ($< 10^{-5}$)	0.53 (0.013)	$< 10^{-5}$
Model 2 RCA_{PSE}	-0.74 ($< 10^{-5}$)	0.46 (0.158)	$< 10^{-5}$
Model 2 MCP_{NT}	0.73 ($< 10^{-5}$)	-0.54 ($< 10^{-5}$)	$< 10^{-5}$
Model 2 MCP_{CLE}	-0.63 ($< 10^{-5}$)	0.48 (0.0001)	0.0001
Model 2 MCP_{PSE}	-0.69 ($< 10^{-5}$)	0.49 (0.0001)	$< 10^{-5}$
RA910	-0.62 ($< 10^{-5}$)	0.61 ($< 10^{-5}$)	$< 10^{-5}$
FEV_1%pred	-	-0.44 ($< 10^{-5}$)	$< 10^{-5}$

Table 3. Correlations with FEV_1%pred and ES as well as ability to separate patient groups according the a rank sum test. p-values of the correlations are shown in parenthesis next to the correlation coefficients.

FEV_1%pred and all except one, MCP_{CLE} of Model 1, have a higher correlation with FEV_1%pred than RA910. RCA_{NT} measured using Model 1 achieves the highest correlation of $r = 0.80$, which is significantly better than the correlation of $r = -0.62$ for RA910 ($p = 0.006$) according to a Hotelling/Williams test [12]. Note that we in the Hotelling/Williams test inverted the signs of the RA910 measures so that the two correlations being compared have the same sign. Looking at the correlations with ES, RA910 achieves the highest correlation coefficient ($\tau = 0.61$). Only two correlation coefficients are not significantly different from zero, and that is RCA_{PSE} using both models. All the texture based quantitative parameters for emphysema, as well as FEV_1%pred and RA910, can separate non-smokers/healthy smokers from smokers diagnosed with COPD according to a rank sum test ($p \leq 10^{-4}$).

4 Discussion and Conclusion

It is not surprising that the texture based emphysema parameters perform differently than RA910, since the two approaches are very different in the amount of information they utilize. RA910 is based on a single threshold, -910 HU, and makes a decision for each pixel based only on the information in that particular pixel. The texture based emphysema parameters on the other hand, base the decision on all pixels in a local neighborhood and thus incorporate much more information. Further, the decision is not based on a specific threshold parameter, but is based on the distribution of the attenuation values as well as measurements of local structure. A consequence of this is that the proposed texture based approaches are expected to be less sensitive to differences in inspiration level. It is known that RA is sensitive to changes in inspiration level, since inspiration level influences the lung density, and thereby the CT attenuation values [13].

The insensitivity to inspiration level is something that is not investigated in this paper, but could be interesting to evaluate in the future on longitudinal data.

There is a tendency that RCA correlate better with FEV_1%pred than MCP, however the difference is not significant in all cases. This tendency could be due to uncertainty at the boundaries, evident in the bottom-right part of Figure 3.2, causing boundary effects as seen in the top-right part of Figure 3.2.

To conclude, we have proposed new parameters for quantifying emphysema in lung CT using texture based pixel classification. The proposed measures generally correlate well with lung function and the highest correlation, $r = 0.80$, is achieved by the relative normal tissue area.

Acknowledgements. This work is partly funded by the Danish Council for Strategic Research (NABIIT), the Netherlands Organisation for Scientific Research (NWO), and AstraZeneca, Lund, Sweden.

References

1. Rabe et al.: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med **176**(6) (Sep 2007) 532–555
2. Mller et al.: "density mask". an objective method to quantitate emphysema using computed tomography. Chest **94**(4) (Oct 1988) 782–787
3. Webb et al.: High-Resolution CT of the Lung, Third Edition. Lippincott Williams & Wilkins (2001)
4. Tuceryan, M., Jain, A.K.: Texture analysis. In: The Handbook of Pattern Recognition and Computer Vision (2nd Edition). World Scientific Publishing (1998) 207–248
5. Uppaluri et al.: Computer recognition of regional lung disease patterns. Am J Respir Crit Care Med **160**(2) (Aug 1999) 648–654
6. Chabat et al.: Obstructive lung diseases: texture classification for differentiation at CT. Radiology **228**(3) (Sep 2003) 871–877
7. Sluimer et al.: Automated classification of hyperlucency, fibrosis, ground glass, solid and focal lesions in high resolution ct of the lung. Medical Physics **33**(7) (2006) 2610–2620
8. Xu et al.: MDCT-based 3-D texture classification of emphysema and early smoking related lung pathologies. IEEE Trans Med Imaging **25**(4) (Apr 2006) 464–475
9. Srensen, L., Shaker, S.B., de Bruijne, M.: Texture classification in lung ct using local binary patterns. MICCAI **11**(Pt. 1) (2008) 934–941
10. Ojala et al.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell **24**(7) (2002) 971–987
11. Jain et al.: Statistical pattern recognition: a review. IEEE Trans Pattern Anal Mach Intell **22**(1) (Jan. 2000) 4–37
12. Sickle, J.V.: Analyzing correlations between stream and watershed attributes. Journal of the American Water Resources Association **39**(3) (June 2003) 717–726 Errata: 41(3) 741–741.
13. Shaker et al.: Volume adjustment of lung density by computed tomography scans in patients with emphysema. Acta Radiol **45**(4) (Jul 2004) 417–423