



HAL
open science

An SVM distal lung image classification using texture descriptors

Chesner Désir, Caroline Petitjean, Laurent Heutte, Mathieu Salaün, Luc Thiberville

► **To cite this version:**

Chesner Désir, Caroline Petitjean, Laurent Heutte, Mathieu Salaün, Luc Thiberville. An SVM distal lung image classification using texture descriptors. *Computerized Medical Imaging and Graphics*, 2011, 10.1016/j.compmedimag.2011.11.001. 10.1016/j.compmedimag.2011.11.001 . hal-00673369

HAL Id: hal-00673369

<https://hal.science/hal-00673369>

Submitted on 23 Feb 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An SVM-based distal lung image classification using texture descriptors

Chesner Désir^a, Caroline Petitjean^{a,*}, Laurent Heutte^a, Luc Thiberville^{a,b},
Mathieu Salaün^{a,b}

^a*Université de Rouen, LITIS EA 4108
BP 12, 76801 Saint-Etienne-du-Rouvray, France*

^b*Clinique Pneumologique, CHU de Rouen
1 rue de Germont, 76031 Rouen Cedex, France*

Abstract

A novel imaging technique can now provide microscopic images of the distal lung *in vivo*, for which quantitative analysis tools need to be developed. In this paper, we present an image classification system that is able to discriminate between normal and pathological images. Different feature spaces for discrimination are investigated and evaluated using a Support Vector Machine. Best classification rates reach up to 90% and 95% on non-smoker and smoker groups, respectively. A feature selection process is also implemented, that allows us to gain some insight about these images. Whereas further tests on extended databases are needed, these first results indicate that efficient computer based automated classification of normal vs. pathological images of the distal lung is feasible.

Keywords: Image classification, feature extraction, feature selection, LBP, SIFT, lung, alveoli

1. Introduction

The lungs are the essential respiration organ. They are divided into two anatomic and functional regions: the air conduction system (that includes the trachea, bronchi, and bronchioles) and the gas-exchange region made of alveolar sacs. Whereas the conduction airways can be explored during bronchoscopy, the alveolar region is currently investigated only *in vitro*, using

*Corresponding author: Caroline.Petitjean@univ-rouen.fr

invasive techniques such as open lung biopsies. Recently, a new endoscopic technique, called Fibered Confocal Fluorescence Microscopy (FCFM), has been developed that enables the visualisation of the more distal regions of the lungs *in vivo* [1]. The technique is based on the principle of fluorescence confocal microscopy, where the microscope objective is replaced by a fiberoptic miniprobe, made of thousands of fiber cores. The miniprobe can be introduced into the 2 mm working channel of a flexible bronchoscope to produce *in vivo* endomicroscopic imaging of the human respiratory tract in real time. This very promising technique could replace lung biopsy in the future and might prove to be helpful in a large variety of diseases, including interstitial lung diseases [2].

Images acquired with FCFM represent the alveolar structure (Figures 1 and 2), which can be altered by distal lung pathologies. Medical experts are still examining visually these images to search for typical properties that could be of interest for discriminating pathological subjects from healthy ones. Our aim is thus to provide the clinician with a computer aided-diagnosis tool, so as to help him to analyze these images, by automatically classifying FCFM images into healthy or pathological ones. The most crucial step when designing an automatic image classification system is the choice of relevant features to describe the image. In a preliminary work [7], we have designed a first classification system, so as to discriminate healthy cases from pathological cases, in which images are described by an ad hoc feature vector based on a visual analysis of the images. But several generic, low-level texture descriptions have been proved very efficient for various image description tasks [8]. We thus investigate in this paper those state-of-the-art feature descriptors to describe our FCFM images and compare their results to the ones obtained with the ad hoc feature vector so as to determine which descriptors are best adapted to the classification of FCFM images. Using a Support Vector Machine (SVM) for classification, we show that one low-level feature vector is better adapted to the discrimination between pathological and healthy images. Furthermore, we point out that some insight about these images can be gained by studying, through a feature selection process, the most relevant features.

The remainder of this paper is thus organized as follows: the imaging technique is detailed in Section 2, our classification system is described in Section 3, and results and discussion are provided in Section 4. Conclusion

is presented in Section 5 along with perspectives for this work.

2. FCFM images

FCFM images represent the alveolar structure, made of elastin fiber (Figures 1-4), with a lateral resolution of $3.5 \mu\text{m}$. The FCFM imaging technique enables to observe the elastin as it is the main endogenous fluorophore at 488 nm in the respiratory tract from the bronchus to the alveoli. The elastin framework appears as a network of (almost) continuous lines in healthy subjects, as shown in Figure 1. In [2], morphometric analysis of the alveolar structures was performed, based on the elastin distribution observed with FCFM. We showed that results were reproducible and concordant to *ex vivo* morphometric analysis based on lung stereology. Our study confirmed that the FCFM observations were reproducible between subjects and between different alveolar areas in the same subject. In the bronchus, the elastin that can be imaged with FCFM is a component of the basement membrane. However, in the alveoli, elastin is the main element of the elastic structure that frames the alveolar ducts and sacs and originates the fluorescent signal. The alveolar walls are not visible during FCFM observation *in vivo*, but only the elastin framework of the interstitial tissue. In this study, the aim is thus to perform an automatic classification of alveolar FCFM images, where the signal only comes from the elastin that belongs to the acinar elastic framework.

Images acquired on smoking subjects differ from the ones acquired on non-smoking ones, notably because of the presence of macrophages, cells which digest cellular debris (cf. Figures 3 and 4). In smoker subjects, the identification of alveolar macrophages using FCFM relies on bronchoalveolar lavage analysis from smokers using fluorescence and conventional microscopy, which confirmed that these fluorescent cells correspond exclusively to alveolar macrophages [2]. These cells are made visible on these images because of tobacco-tar induced fluorescence. We also demonstrated by in-situ microspectrometry experiments that whereas the main endogenous fluorophore is elastin in non-smokers, it is the tobacco-tar in smokers [2].

In this context, a clinical trial is currently being conducted that collects FCFM images in several pathological conditions of the distal lungs, including:

- post radiation therapy fibrosis, silicosis, bronchioloalveolar carcinoma, systemic sclerosis, connective lung (Sharp syndrome), sarcoidosis with

interstitial involvement, alveolar proteinosis, asbestos related interstitial fibrosis, Bleomycin induced lung fibrosis, Furadantine induced lung fibrosis, idiopathic pulmonary fibrosis for non-smoking patients,

- sarcoidosis with interstitial involvement, systemic sclerosis, Amiodarone induced interstitial pneumoniae, smoke induced respiratory bronchiolitis, Langerhans histiocytosis, alveolar proteinosis for smoking patients.

The trial also includes images acquired on healthy smoking and non-smoking volunteers. The images are selected from the *in vivo* FCFM image database by two medical experts, and are labeled as (i) healthy if they are obtained from the healthy volunteer group, and (ii) pathological if they are obtained from a patient with a diagnosed interstitial lung disease and if the image was obtained from a lung segment that appeared abnormal on the chest CT Scan. Note that there is no correlation study between histology and FCFM features intended in this work. The description of the FCFM features is beginning and the description of the FCFM imaging in interstitial lung diseases is not yet established. Therefore, the study is based on images from miscellaneous infiltrative lung diseases patients, as well as healthy volunteers, and is not designed to authorize a lesion/FCFM feature comparison at this point. The database of FCFM images, that includes a total of 133 non-smoker images and 93 smoker images, is used for the elaboration of the automatic image classification system, described in the following section.

3. Image classification method

3.1. Feature extraction

In image classification system, feature extraction is still a critical step. Features are usually adapted to the image content. For instance, blood cell images can be described by frequency analysis, using Fourier-Mellin transform [3]; wooden image texture and color can be described by LBP (Local Binary Patterns) descriptor [4]; protein sub-cellular images can be described by their texture, with cooccurrence matrix using Haralick statistical parameters [5]. The image description can also be local, e.g. by extracting keypoints and describing their neighbourhoods using gradient as local features [6]. Thus, an initial approach to numerically describe FCFM images can be based on a visual analysis of these images. As shown in Figures 1 and 3, the alveolar structure in healthy subjects can be described as contrasted continuous lines and curves. On the other hand, in the pathological subset, the

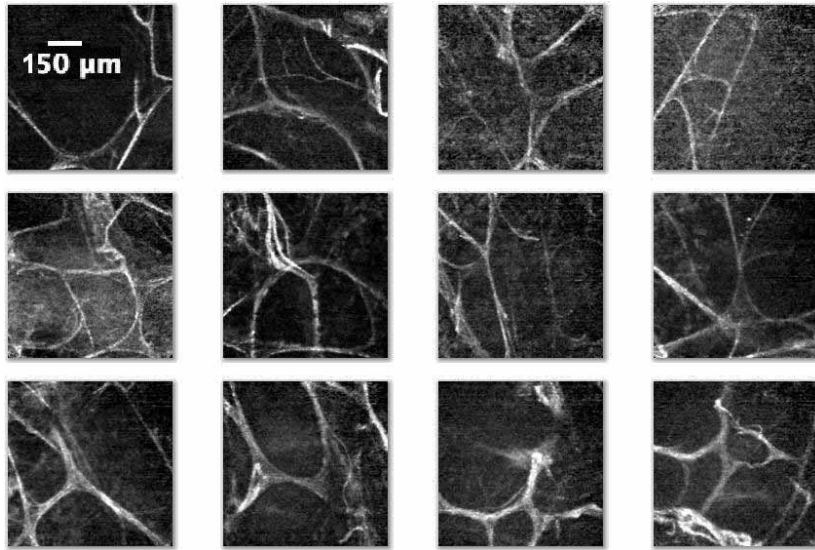


Figure 1: FCFM images of non-smoking healthy subjects

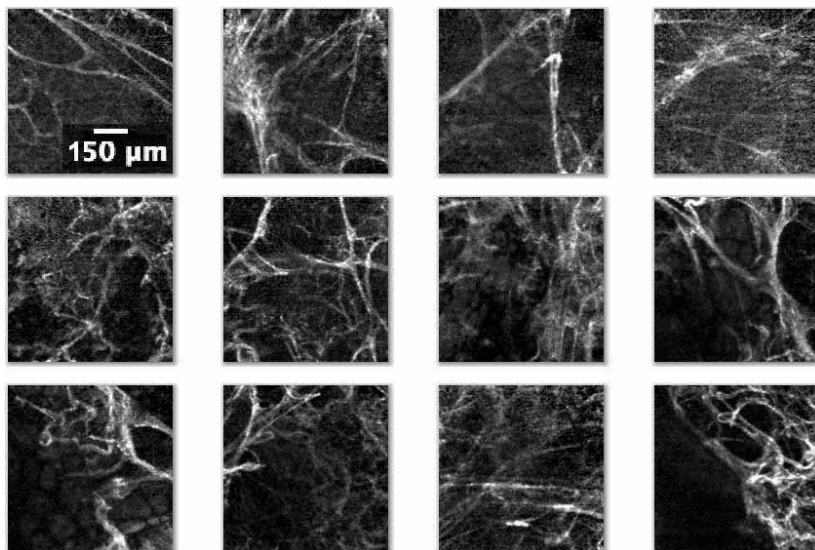


Figure 2: FCFM images of non-smoking pathological subjects

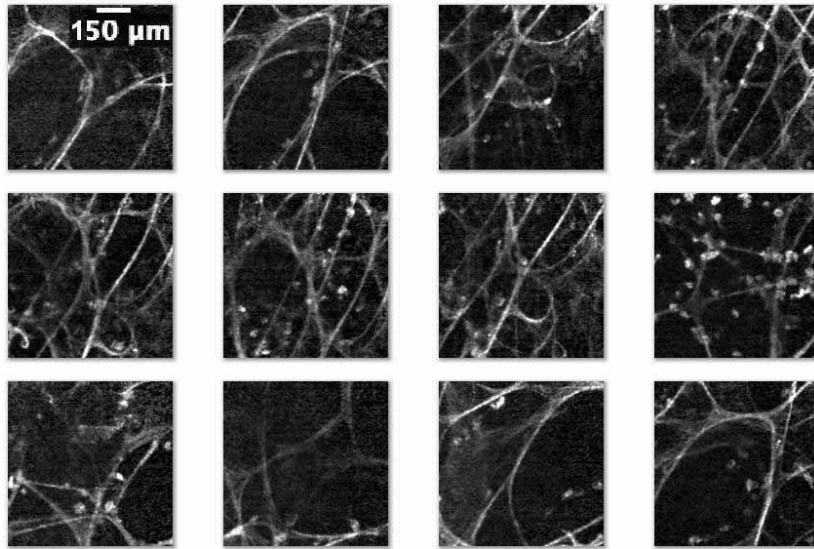


Figure 3: FCFM images of smoking healthy subjects. Notice the presence of macrophages (white spots).

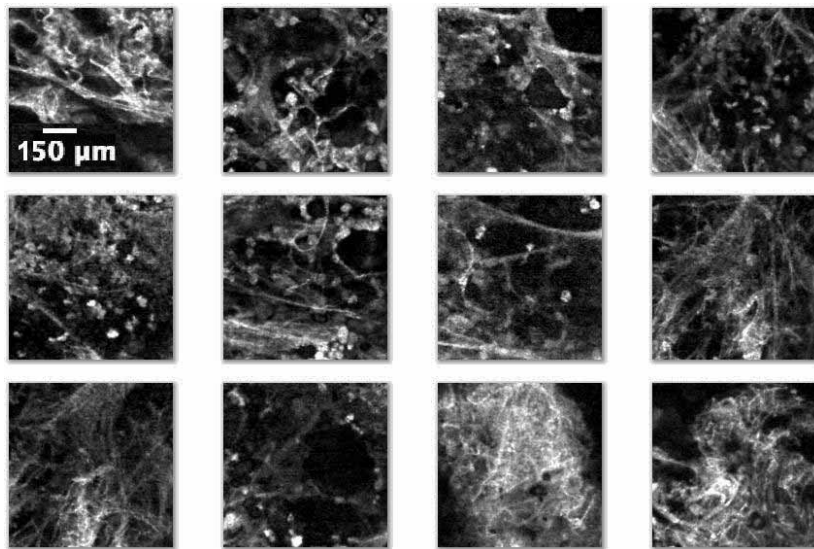


Figure 4: FCFM images of smoking pathological subjects. Notice the presence of macrophages (white spots).

disorganization of the meshing is illustrated by the numerous irregularities and the tangle of the fibered structures (Figures 2 and 4). Differences are mostly visible for the structure shape, the image texture and the contrast, implying that numerical features could therefore be chosen among the ones that best describe the visual differences from these three points of view.

The structure contrast can be characterized by studying first order statistics of the distribution of gray levels, and computing pixel densities. Because structures also show local differences, local parameters are computed on sub-windows of the image. Subwindows are obtained by dividing by 4 the image height and width (Figure 5). Features characterizing the image contrast are (i) first order statistics on global and local histogram: mean, variance, skewness, kurtosis, entropy, (ii) global and local pixel densities obtained on binarized images using Otsu thresholding, (iii) the sum of the image gradient values, obtained using Prewitt operator. We could suppose that pathological images will have high values for densities, because of the tangle of the fibered structure. The complexity of the structure shape can be characterized by studying the image skeleton. After skeletonization [9] obtained on the binary image, the number of junction points is computed. One can suppose that on clearly organized, healthy images, this number will be small, contrary to pathological images where the meshing mess will induce a higher number of points. At last, the image texture can be characterized by Haralick parameters computed from gray level cooccurrence matrices (GLCM) [5]. The GLCM provide the joint distribution of gray-level intensities between two image points. These two points are located according to several configurations, along 4 directions (0° , 45° , 90° , 135°) in our case. Energy, contrast, homogeneity and correlation are then computed on the GLCM.

This first characterization of FCFM images leads to a 120-feature vector called ad hoc feature vector, as shown in Table 1.

On the other hand, FCFM images can also be described using low-level, generic feature vectors, thus making little assumption about the image content. An image texture can be classically characterized with:

- Haralick parameters [5], computed from GLCM, as explained above. In order to design a feature vector based exclusively on Haralick parameters, we have included almost all of the features originally proposed by Haralick, obtained on a large number of GLCM. Thus, the following

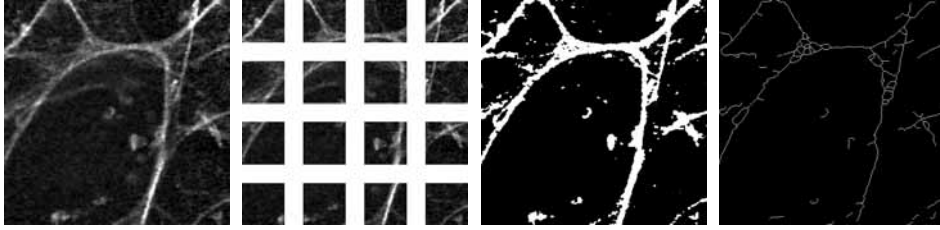


Figure 5: From left to right: original FCFM image, 16 subwindows, binarized image, skeleton on binarized image

Table 1: Ad hoc feature set

	Features	Number
Contrast	Global histogram statistics	5
	Local histogram statistics	80
	Density	1
	Local densities	16
	Sum of image gradient	1
Shape	Number of junction points in skeleton	1
Texture	Haralick parameters	16
Total		120

features are computed: energy, contrast, correlation, variance, inverse different moment, entropy, sum average, sum entropy, sum variance, difference entropy, difference variance, and two information measures of correlation. The only discarded feature is the maximum correlation coefficient, which is too computationally expensive. To these 13 parameters we added dissimilarity, a measure of homogeneity [10]. All these 14 parameters are computed over 10 co-occurrence matrices, obtained via the following classical translation vectors: $[0\ 1]$, $[-1\ 1]$, $[-1\ 0]$, $[-1\ -1]$, $[0\ 2]$, $[-1\ 2]$, $[-1\ -2]$, $[-2\ 1]$, $[-2\ 0]$, $[-2\ -1]$.

- the SIFT (Scale Invariant Feature Transform) algorithm, shown to be one of the most efficient for image description [8]. SIFT consists in extracting robust keypoints in the image, and describing the neighborhood of each of them with a histogram of oriented gradients [11]. The final feature vector represent a 128-bin histogram of the gradient values

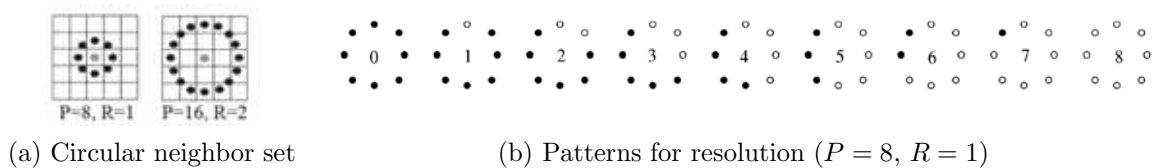


Figure 6: Local binary patterns (Figure (b) is from [4], with kind permission from Springer Science+Business Media)

in each keypoint neighborhood.

- the LBP (Local Binary Patterns) operator [12], that consists in computing the distribution of binary patterns in the circular neighborhood of each pixel. These patterns are obtained by thresholding neighboring pixels compared to the central pixel. This neighborhood is characterized by a radius R and a number of neighbors P , as shown on Figure 6(a). The possible patterns obtained for resolution ($P = 8, R = 1$) are provided in Figure 6(b): for example pattern #4 detects edges, whereas pattern #8 detects flat areas. In some preliminary experiments [13], we have shown that a better parameterization of LBP for FCFM images is to concatenate the LBP values for neighborhood set of values ($P = 8, R = 1$), ($P = 16, R = 2$) and ($P = 24, R = 3$).

Table 2 summarizes all the feature vectors tested in this study and their respective length.

Table 2: Feature vectors used to characterize FCFM images

Feature vector name	Number of features
Ad hoc description	120
Haralick parameters on GLCM	140
LBP	54
SIFT	128

3.2. Classifier

On the previously cited features an SVM classifier is implemented [14]. SVM is one of the best performing and most used classification algorithm. The support vector machine classifier is a binary classifier algorithm that

looks for an optimal hyperplane as a decision function in a high-dimensional space. The associated kernel function is classically chosen as a cubic polynomial kernel. In order to improve the prediction performance of the classifier, and to provide faster and more cost-effective decision, variable selection [15] can be used. It can also provide a better understanding of which visual features discriminate the data. Support Vector Machine with Recursive Feature Elimination (SVM-RFE) is one way to perform variable selection [16]. The goal is to find a subset of size r among d variables ($r < d$) which maximizes the performance of the predictor. The method is based on a sequential backward selection. One feature at a time is removed until r features are left. The removed variables are the ones that minimize the variation of the margin.

4. Experimental results

4.1. Experimental protocol

The SVM classifier and SVM-RFE based feature selection [16] are implemented using the SVM and Kernel Methods Matlab Toolbox [17]. The number of images for the smoker and the non-smoker groups is provided in Table 3. Because of the relatively small number of non-smoker and smoker images, a 10-fold cross-validation process is used. The system performance is assessed with the mean recognition rates \pm standard deviation on the 10 folds.

Table 3: Number of FCFM images for the smoker and non-smoker databases

Classes \ Database	Non-smoker	Smoker
Healthy	31	60
Pathological	102	33
Total	133	93

4.2. Classification results

From the recognition rates reported in Table 4, one can see that texture-based feature vectors outperform the ad hoc feature set. In particular, the LBP operator seems to provide an adequate description for discriminating healthy and pathological FCFM images. Note that results obtained on the smoker group are always better than those obtained on the non-smoker group,

but not with the SIFT based feature vector. As a matter of fact, the SIFT algorithm looks for some salient points in the image. In the smoker images, these points can be hidden behind the macrophages, thus preventing a better discrimination of these images (Figure 7).

Since the LBP-based feature vector offers the best performance, it is selected for further analysis. The confusion matrix for the LBP vector, provided in Table 5, shows that the system performs well on healthy images for both groups, since more than 96% of healthy images have been well classified. The false alarm rate, i.e healthy patients mistaken for pathological ones, is quite similar for both groups. However, non detection, i.e. pathological subjects mistaken for healthy ones, is more at risk than false alarm. This non detection rate, which reaches 15.83% for the non-smoker group, leaves room for improvement.

In the next section, the results of a feature selection process applied on the LBP feature vector are presented, to refine the interpretation of FCFM image classification.

Table 4: Recognition rates obtained with the different feature vectors

	Non-smoker	Smoker
Ad hoc description	77.92± 9.27%	82.50± 12.61%
Haralick parameters on GLCM	80.17± 17.26%	90.00± 10.41%
LBP	90.08± 8.47%	95.00± 6.67%
SIFT	81.08± 14.11%	81.67± 10.41%

Table 5: Confusion matrices of pathological (P) vs. healthy (H) classification for the LBP based feature vector

		Non-smoker		Smoker	
		Predicted labels		Predicted labels	
		P	H	P	H
True labels	P	84.17±16.01%	15.83±16.01%	93.33±11.06%	6.67±11.06%
	H	4±4.90%	96.00±4.90%	3.33±10.00%	96.67±10.00%

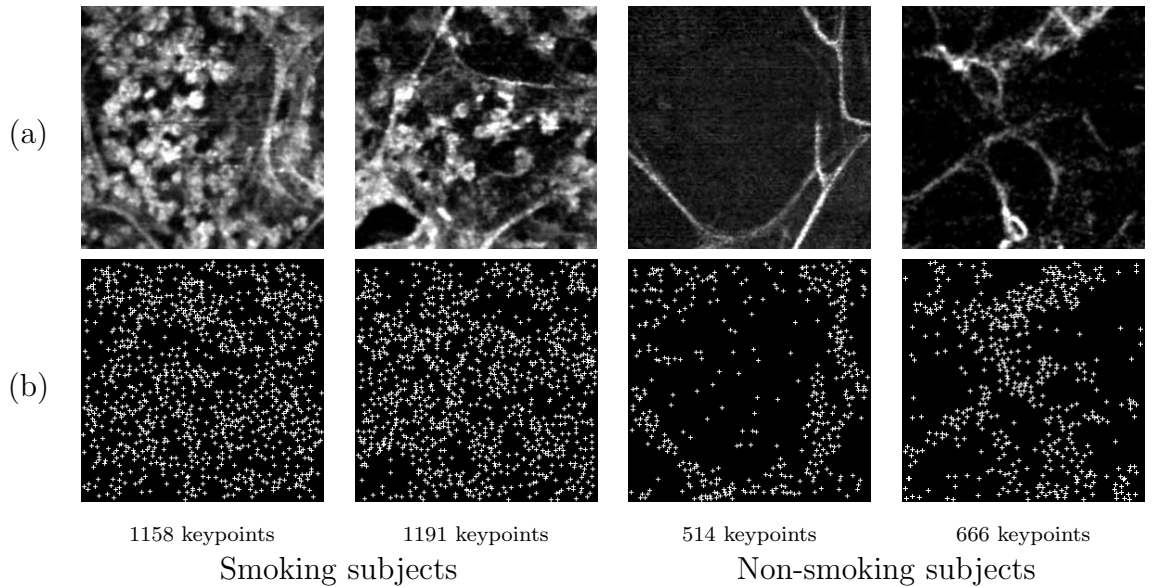


Figure 7: (a) Original FCFM images and (b) their corresponding relevant SIFT keypoints. Note how the presence of macrophage prevent from correctly highlighting the underlying network.

4.3. Feature selection

The SVM-RFE provides a rank for each feature of the LBP-based feature set, thus allowing the most discriminative ones to be identified. To fairly assess the effect of feature selection, the 10 folds have been divided into a learning set (6 folds), used to rank the features with the SVM-RFE, a validation set (3 folds) to determine the number of features to be obtained, and a test set (last fold) is finally used to evaluate a classical SVM in the selected subspace. Note that since variance on the results is quite high (cf. Table 4) due to the size of the database, the features retained for each validation fold are not always the same. The total selected feature space is thus the union of the feature sets selected for each fold during the validation process.

To determine the adequate number of features that must be provided to the predictor, we have computed a statistical test based on the variance of the test error on the validation folds. This test is used to select, among the list of ranked features provided by the SVM-RFE on the learning folds, those features that do not perturb this variance significantly. Let X_i be the mean recognition rate obtained with the first i ranked features (from the best one

to the i th one) on the validation folds. Given a sample $X = \{X_1, \dots, X_n\}$ of n observations ($n = 54$ for our LBP-based feature set) assumed to be *i.i.d.*, and an expected variance s on this sample, the following statistics is computed:

$$T = (n - 1) \cdot \frac{S^{*2}}{s^2}$$

where $S^{*2} = \frac{1}{n-1} \cdot \sum_{i=1}^n [X_i - E(X)]^2$ is the unbiased estimation of the real variance and $E(X)$ the expectation of X .

T follows a χ^2 distribution with $(n - 1)$ degrees of freedom. For a determined level of confidence $(1 - \alpha)$, χ^2 statistical tables provide the degree of significance of the difference between the real value of the variance S^{*2} of our sample and the expected variance s . In our experiments, the level of confidence has been fixed to $1 - \alpha = 95\%$ and two values for the expected variance have been tested, namely $s = 5\%$ and $s = 2.5\%$.

Table 6 shows the results obtained on the test folds (mean \pm stdev) with a classical SVM: in the original feature space (no RFE) and in the selected subspace obtained for $s = 5\%$ and $s = 2.5\%$. We can observe that as expected, no or few features are discarded for $s = 2.5\%$, whereas for $s = 5\%$ the number of features drops down to 40 for non-smoker images, and 25 for smoker images, without decreasing significantly recognition accuracy, at least on the smoker group. These results confirm that a part of the features may be used to discriminate between healthy and pathological images. Investigating those retained features allows us to gain some insight about the usefulness of these features. We have observed that the most discriminating features are mostly the binary patterns that represent edges in the images whereas the discarded ones rather represent flat areas. This phenomenon is illustrated in Figure 8 for one resolution level (namely $(P = 24, R = 3)$). For each pixel of the original image, the presence of an edge pattern is indicated by a white pixel in Figure 8(b), and the presence of a flat area pattern is similarly indicated in Figure 8(c). Note how edge patterns allow for better representation of both classes (pathological and healthy) and thus, better discrimination.

Table 6: Recognition rates obtained with the LBP feature vectors using SVM-RFE (in brackets the number of retained features)

		Non-smoker	Smoker
LBP	<i>no RFE</i>	$90.08 \pm 8.47\%$ (54)	$95.00 \pm 6.67\%$ (54)
	2.5%	$90.08 \pm 8.47\%$ (54)	$91.67 \pm 10.54\%$ (31)
	5%	$84.75 \pm 12.79\%$ (40)	$94.17 \pm 6.51\%$ (25)

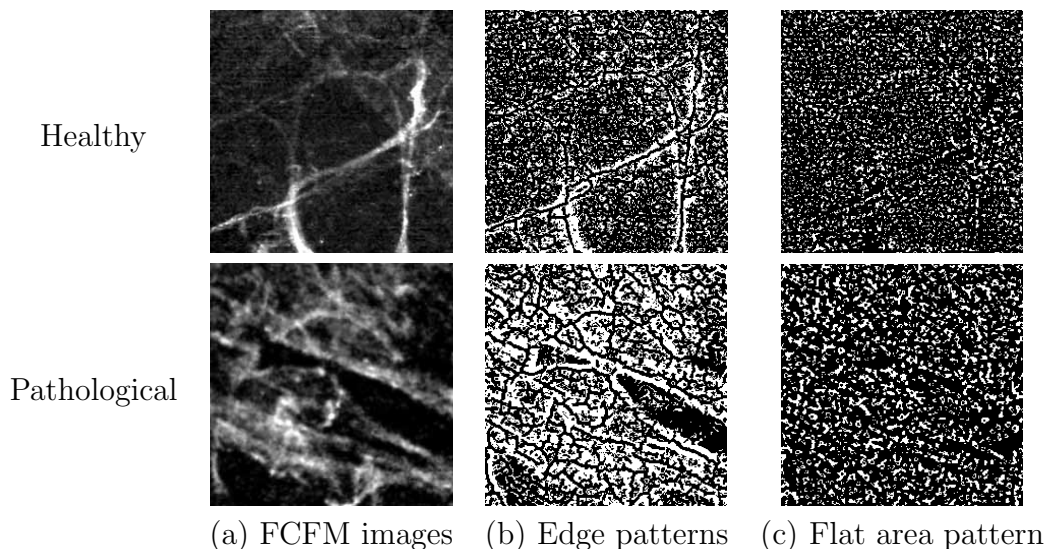


Figure 8: Original FCFM images with a representation of the presence of flat area and edge LBP patterns

5. Conclusion

The present work deals with the classification of a new category of images of the distal lung. The images were acquired using fibered confocal fluorescence microscopy, a technique that enables the observation of *in vivo* alveolar structures. Such images are not well described so far and are difficult to discriminate by pathologists and respiratory physicians. In this paper, we have proposed several possible descriptions for these images, one based on their visual analysis, and the other ones texture-based. These are incorporated into a classification system, so as to aid the clinician to discriminate between healthy and pathological subjects. Best performance is obtained using the LBP based description of the images, on both non-smoker and smoker groups. A feature selection process has allowed to extract the most discrimi-

native features, which mostly represent the images edges, without decreasing significantly classification accuracy. This result could guide the clinician to focus on the image parts that are the most discriminant. Other potential applications include the automated recognition of the area to biopsy during endoscopy.

However, the current database should be extended to confirm these first results. Because the clinical trial is ongoing, this will be feasible in the near future. In order to improve the reliability of the system, the non detection rate could be reduced by introducing for instance a reject rule. Classification methods could also give information about which part of the alveolar structure might be more altered by pathologies. A future goal will also be to discriminate between different interstitial lung diseases (abestosis, systemic sclerosis, fibrosis, sarcoidosis).

References

- [1] L. Thiberville, S. Moreno-Swirc, T. Vercauteren, E. Peltier, C. Cave, G. Bourg-Heckly, In vivo imaging of the bronchial wall microstructure using fibered confocal fluorescence microscopy, *American Journal of Respiratory and Critical Care Medicine* 175 (2007) pp. 178–187.
- [2] L. Thiberville, M. Salaün, S Lachkar, S. Moreno-Swirc, C. Vever-Bizet, G. Bourg-Heckly, Human in vivo fluorescence microimaging of the alveolar ducts and sacs during bronchoscopy, *Eur Respir J.* 33(5) (2009) 974–85.
- [3] J. Dahmen, J. Hektor, R. Perrey, H. Ney, Automatic classification of red blood cells using gaussian mixture densities, in: *Bildverarbeitung für die Medizin*, 2000, pp. 331–335.
- [4] T. Ojala, M. Pietikäinen, T. Mäenpää, Gray scale and rotation invariant texture classification with local binary patterns, *ECCV* (2000) 404–420.
- [5] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *Systems, Man and Cybernetics* 3(6) (1973) 610–621.
- [6] K. Mikolajczyk, C. Schmid, Scale and affine invariant interest point detectors, *Int. J. Comput. Vision* 60 (1) (2004) 63–86.

- [7] C. Petitjean, J. Benoist, L. Thiberville, M. Salaiün, L. Heutte, Classification of in-vivo endomicroscopic images of the alveolar respiratory system, in: Proceedings of the IAPR Conference on Machine Vision Applications, Yokohama, Japan, 2009, pp. 471–474.
- [8] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 (10) (2005) 1615–1630.
- [9] G. S. Dibajaa, E. Thiel, Skeletonization algorithm running on path-based distance maps, *Image and Vision Computing* 14:47-57.
- [10] W. Pratt, *Digital Image Processing*, 2nd Edition, John Wiley & Sons, 1991.
- [11] D. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* 60 (2) (2004) 91–110.
- [12] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (7) (2002) 971987.
- [13] A. Saint-Réquier, B. Lelandais, C. Petitjean, C. Désir, L. Heutte, Characterization of endomicroscopic images of the distal lung for computer-aided diagnosis, in: *International Conference on Intelligent Computing*, Ulsan, South Korea, Vol. 5754 of LNCS, 2009, pp. 994–1003.
- [14] V. Vapnik, *The nature of statistical learning theory*, Springer-Verlag, 1995.
- [15] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, *Journal of Machine Learning Research* 3 (2003) 1157–1182.
- [16] A. Rakotomamonjy, Variable selection using SVM-based criteria, *Journal of Machine Learning Research* (2003) 3:1357–1370.
- [17] S. Canu, Y. Grandvalet, V. Guigue, A. Rakotomamonjy, SVM and kernel methods Matlab toolbox, *Perception Systemes et Information*, INSA de Rouen, France.