DETECTION OF PATHOLOGICAL CONDITION IN DISTAL LUNG IMAGES

David Hébert[⋆] *Chesner Désir*[⋆] *Caroline Petitjean*[⋆] *Laurent Heutte*[⋆] *Luc Thiberville*[⋆†]

[⋆] Université de Rouen, LITIS - EA 4108, BP 12, 76801 Saint-Etienne-du-Rouvray, France
[†] Clinique Pneumologique, Rouen University Hospital, 76031 Rouen, France

ABSTRACT

Recently, the in vivo imaging of pulmonary alveoli was made possible thanks to confocal microscopy. For these new images, we wish to aid the clinician by developing a computer-aided diagnosis system, able to detect a pathological state in these images. An original approach that combines a texture-based characterization of the images and uses a boosted cascade of classifiers to detect a pathological condition is presented in this paper. We propose and compare two state-of-the-art texture descriptors: cooccurence matrices and local binary patterns (LBP). Recognition rates with LBP reach up to 86.3% and 95.1% for the non-smoking and smoking groups, respectively. Even though tests on extended databases are needed, these preliminary results are encouraging for this challenging task of image classification.

***Index Terms*—** Image classification, pathology detection, boosted cascade of classifiers, endomicroscopic images, lung

1. INTRODUCTION

The lungs are divided into two anatomic and functional regions: the air conduction system (that can be explored during bronchoscopy) and the alveolar region, which is currently investigated only in vitro, using 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, introduced into the 2 mm working channel of a bronchoscope. Alveolar images are continuously recorded at a frame rate of 12 frames per second (fps) and stored for further analysis. This 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].

As shown in Fig. 1, FCFM images represent the alveolar structure, made of elastin fiber, with an approximate resolution of 1 μm per pixel. Images acquired on smoking patients differ from the ones acquired on non-smoking ones, notably because of the presence of macrophages, which appear as white spots (Fig. 2). Macrophages are cells which digest cellular debris; they are normally present in the alveoli. In smoker images, they are made visible because of tobacco-tar induced fluorescence. The elastin fiber framework can be altered by distal lung pathologies and images acquired on pathological patients differ from the ones acquired on healthy volunteers: the alveolar structure in healthy subjects can be described as contrasted continuous lines and curves, whereas the disorganization of the meshing is illustrated by the numerous irregularities and the tangle of the fibered structures in pathological images (Figures 1-2).

One particular task of interest for clinical use is to detect pathological images, while browsing through a pathological lung segment. We describe in this paper an original system that enables to detect such pathological images, based on a boosted cascade of classifiers (BCC), a state-of-the-art classification technique for object detection. Initially proposed by Viola and Jones for face detection [3], the BCC has the practical advantage of being very fast in decision, thus allowing real-time image processing. The design of such a system has also raised the issue of choosing a suitable feature space in which healthy and pathological subjects can be discriminated. As FCFM images images are highly textured, they are described with two well-known texture descriptors (cooccurrence matrices and local binary patterns), which are compared in terms of performance in the experimental section.

2. FCFM IMAGE CLASSIFICATION

2.1. Feature extraction

Literature of image characterization has given rise to many texture descriptors. For FCFM images, we have retained the most standard ones, namely the Haralick parameters computed on Grey Level Cooccurrence Matrices (GLCM) and the Local Binary Patterns (LBP), a powerful texture descriptor that has proven to outperform earlier methods in many applications [4].

GLCM-based feature vector. Cooccurrence matrices provide the joint distribution of gray-level intensities between two image points. These two points are located according to several configurations, that represent different distances and

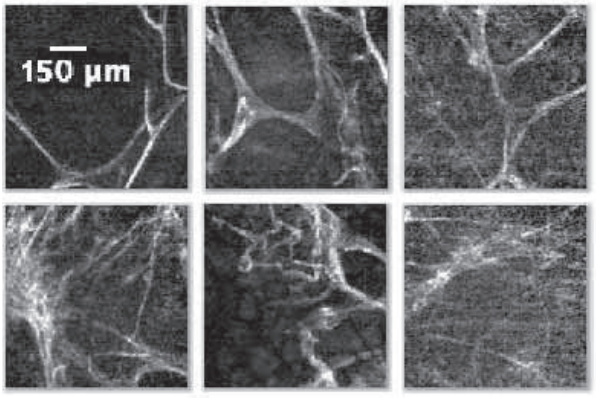


Fig. 1. FCFM images of three non-smoking healthy (top) and pathological cases (bottom)

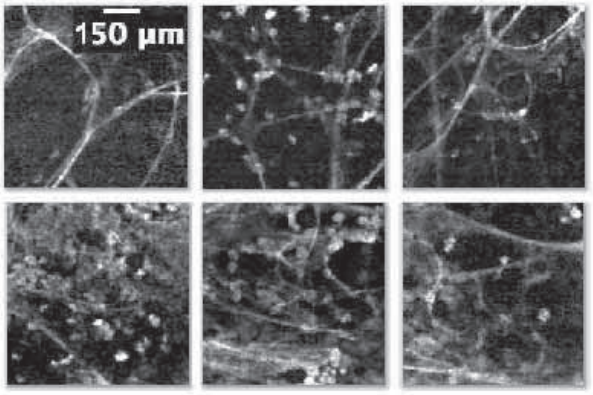


Fig. 2. FCFM images of three smoking healthy (top) and pathological cases (bottom)

rotation angles. We choose the following classical 10 translation vectors: [0 1], [-1 1], [-1 0], [-1 -1], [0 2], [-1 2], [-1 -2], [-2 1], [-2 0], [-2 -1]. Parameters are then computed on each of the 10 co-occurrence matrices. From the features originally proposed by Haralick et al. [5], we retain the following ones: 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 is added dissimilarity, a measure of homogeneity [6]. This GLCM-based characterization of FCFM images leads to a 140-feature vector.

LBP-based feature vector. In the LBP framework, one considers the circular neighborhood of each pixel, characterized by a radius R and a number of neighbors P. A first step consists in computing the distribution of binary patterns in the

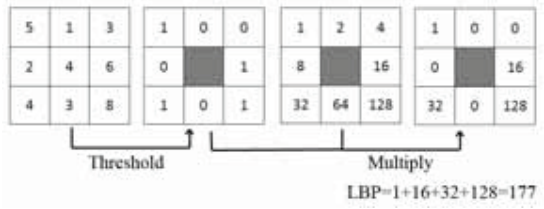


Fig. 3. The principle of LBP operator (from [7])

neighborhood of each pixel [7]. These patterns are obtained by thresholding neighboring pixels compared to the central one and coding them with powers of 2 (Figure 3). The LBP operator provides the distribution of these patterns over the image, in $P + 2$ values. In a second step, the distribution of the gray level variance computed in the pixel neighborhood is added to the LBP operator. Since the variance has a continuous-valued output, it requires to be quantified. Quantization intervals are defined by assigning, between each cut value, the same amount of values from the total distribution of variance. As suggested in [7], we have empirically determined the number of intervals on our training set of images. A number of 15 intervals for quantization of the variance values has finally been chosen, as it provided the best detection rates on our training set of images. Note that theses two descriptors (LBP operator and variance) are by construction rotation-invariant.

In order to have a multiscale description of the image, the feature vector can be made up of a concatenation of both operators for varying values of P and R. In some previous experiments, such a combination with values ((8,1) + (16,2) + (24,3)) for (P, R), provided the best performance on our images [8]. FCFM images are thus characterized by $10 + 18 + 26 + 15 * 3 = 99$ features.

2.2. Boosted cascade of classifiers

Currently a state-of-the-art classifier in face detection, Boosted Cascade of Classifiers (BCC) have also been used for other object detection problems (for example vehicle detection [9]) and, recently, in a medical context for human heart localization [10]. The methodology behind these appplications consists in extracting image subwindows, characterizing them with Haar wavelet based features, and processing them with the BCC, which has been trained to either accept subwindows that contain the object and reject ones that do not. BCC can be adapted to the classification of FCFM images by computing feature vectors on the whole images (either GLCM-based or LBP-based) rather than on subwindows, and feeding the BCC with these feature vectors. Since we want to detect pathological condition, the BCC is trained to accept pathological images and reject healthy images.

A BCC is made of K stages, each stage including a varying number of weak classifiers, which are typically discrimi-

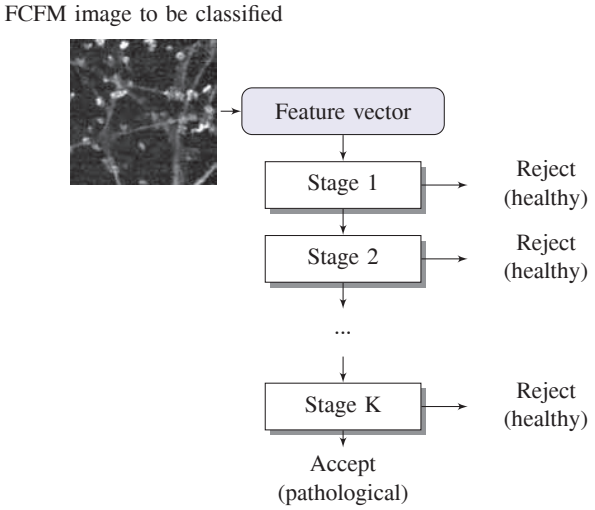


Fig. 4. Detection process using a BCC

native thresholds on one single feature. The idea of the BCC is to train the first stages to rapidly reject "easy to classify as healthy" images, so as to focus in later stages on images which are ambiguous and more difficult to discriminate. Inside each stage, the decisions of all weak classifiers are combined into one single decision, that leads to either reject the image or proceed to the next stage. As a consequence, an input image that has gone through all stages without being rejected will be accepted as pathological (Fig. 4).

The training of the BCC is made stage-by-stage. The first stage is trained on all training set images. During the training of this stage, a first weak classifier is trained to select the best discriminative feature. Weak classifiers are added to the stage until almost all pathological images are detected and at least half of healthy images are rejected, these figures being computed on a validation set. Then, a new stage is created, and is trained on the remaining healthy images that were incorrectly classified by the preceding stage. This multistage process allows to focus on images which are difficult to classify. For more information on the BCC, the reader is referred to [3].

3. EXPERIMENTAL RESULTS

Our experiments have been conducted on a database obtained from an ongoing clinical trial. This trial collects FCFM images in several pathological conditions of the distal lung and on healthy volunteers, providing a total of 133 images for the non-smoking group and 93 images for the smoking group. Images 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 the ground truth is thus not established from FCFM images but relies on external elements.

Table 1. Recognition rates of the 1-NN on the whole image sets (133 non-smoking images and 93 smoking images)

	LBP	GLCM
Smoker	94.6%	91.2%
Non-Smoker	91.7%	84.7%

Table 2. Number of images in training (T), validation (V) and evaluation (E) sets, for the smoker (S) and non-smoker (NS) group

		T	V	E	Per class	Total
S	Healthy	30	18	12	60	93
	Pathological	18	9	6	33	
NS	Healthy	16	9	6	31	133
	Pathological	52	30	20	102	

3.1. Feature vector comparison

In order to compare the LBP-based and the GLCM-based feature vectors, a 1-Nearest Neighbour (1-NN) classifier has been implemented based on a leave-one-out protocol using the complete FCFM image database (133 non-smoking images and 93 smoking images). The 1-NN classifier allows us indeed to gain some insight on the discriminative power of each of the feature vectors. It is shown in Table 1 that the LBP achieves better performance than the GLCM and is thus retained for further experiments with the BCC. This might be explained by the ability of the LBP features to capture discriminative information thanks to its multiscale approach.

3.2. BCC performance analysis

The numbers of images for both non-smoker and smoker groups and their repartition in the different sets are given in Table 2. Approximately 50% of the available samples are included in the training set, 30% in the validation set and the other 20% in the test set. These three databases are built by random selection in the original set of samples. To make accuracy more reliable and reduce the bias, the random drawing that provides a 50%-30%-20% repartition is repeated numerous times. The number of repetitions has been assessed experimentally by comparing results obtained with M trials and $M + 500$ trials: since results do not significantly change between 1500 and 2000 repetitions, a number of 1500 random selections has been retained.

Results presented in Table 3 show mean and standard deviation on the BCC recognition rates with the LBP feature vector, computed over the 1500 trials. This table also includes the results of the 1-NN, and results of a manual classification experiment, organized in order to quantify the difficulty of an-

Table 3. Recognition rates of the BCC and the 1-NN on the LBP feature vector, along with manual recognition rates

	1-NN	BCC	Manual
Smoker	94.6%	95.1%±5.0%	63.3%
Non-Smoker	91.7%	86.3%±5.8%	77.2%

alyzing FCFM images [11]. Indeed one can wonder if a medical expert, who is well aware of alveolar structure, would perform better (after specific training) than an automatic classification system. A manual classification session of healthy vs. pathological FCFM images was thus carried out by 7 pneumologists, on the same image collection as the one used in this paper, after a minimal learning step. The pneumologists had no prior experience of examining FCFM images. The relatively low recognition rates obtained by humans show the difficulty of the task.

Regarding the BCC performance, one can note that it outperforms manual classification. When compared to the reference performance of the 1-NN, the BCC results are satisfying for the smoker group but leave room for improvement for the non-smoker group, even though full comparison cannot be made because of the different protocols (leave-one-out vs. 1500 random splits). Note that whereas humans are better in classifying non-smoker images than smoking ones, automatic classification exhibits opposite results. This could be related to the presence of several additional corners in smoker images, because of the macrophages: they are clearly visible on healthy images and somehow hidden when the alveolar structure is disorganized because of a pathology. This information seems to be captured by the LBP very efficiently, maybe because it corresponds to particular binary patterns.

As stated in the Introduction, one of the interests of the BCC is its low computational cost. The BCC algorithm is very fast, as it requires about 3 ms on a regular PC hardware to decode 12 positive (pathological) images and less than 1 ms for 6 negative (healthy) images. These times are compatible for real-time image processing with the frame rate of the image sequence, equal to 12 frames per second.

4. CONCLUSION

The present work deals with the detection of pathological condition in a new category of images from the distal lung. Two texture-based feature vectors, namely the LBP operator and Haralick parameters on GLCM, are used to describe these images. An image classification system based on a BCC is used to detect pathological images. Experiments show that FCFM images are more adequately described with the LBP approach than with the GLCM. The use of the BCC allows to obtain recognition rates of 86.3% and 95.1% for non-smoking and smoking groups, respectively. This performance makes the automatic system useful to assist the clinician, especially since the detection is compatible with real time processing. Nonetheless our approach required to be validated on larger and more balanced databases.

5. REFERENCES

[1] L. Thiberville, S. Moreno-Swirc, T. Vercauteren, E. Peltier, C. Cave, and G. Bourg-Heckly, "In vivo imaging of the bronchial wall microstructure using fibered confocal fluorescence microscopy," *Am. J. Respir. Crit. Care*, vol. 175, pp. 178–187, 2007.

[2] L. Thiberville, M. Salaün, S. Lachkar, S. Dominique, S. Moreno-Swirc, C. Vever-Bizet, and G. Bourg-Heckly, "Human in vivo fluorescence microimaging of the alveolar ducts and sacs during bronchoscopy," *Eur Respir J.*, vol. 33, no. 5, pp. 974–85, 2009.

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *In IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 511–518, 2001.

[4] M. Pietikäinen, "Image analysis with local binary patterns," in *Scandinavian Conference on Image Analysis*. 2005, vol. 3540 of *LNCS*, pp. 115–118, Springer.

[5] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics*, vol. 3(6), pp. 610–621, 1973.

[6] W.K. Pratt, *Digital Image Processing, 2nd Edition*, John Wiley & Sons, 1991.

[7] 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. 971987, 2002.

[8] A. Saint-Réquier, B. Lelandais, C. Petitjean, C. Désir, and L. Heutte, "Characterization of endomicroscopic images of the distal lung for computer-aided diagnosis," in *International Conference on Intelligent Computing (ICIC), Ulsan, South Korea*. 2009, vol. 5754 of *LNCS*, pp. 994–1003, Springer.

[9] B. Alefs, "Embedded vehicle detection by boosting," in *Intelligent Transportation Systems Conference, Vienna, Austria*, 2006, pp. 536–541.

[10] S.-K. Pavani, D. Delgado, and A.F. Frangi, "Haar-like features with optimally weighted rectangles for rapid object detection," *Pattern Recognition*, vol. 43, no. 1, pp. 160–172, 2010.

[11] C. Petitjean, C. Désir, G. Lettier, and L. Heutte, "Manual vs. automatic classification of endomicroscopic images of the distal lung," in *Computer-Assisted Radiology and Surgery (CARS), Berlin, Germany*, 2011.