



# Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images

Alain Rakotomamonjy, Caroline Petitjean, Mathieu Salaün, Luc Thiberville

## ► To cite this version:

Alain Rakotomamonjy, Caroline Petitjean, Mathieu Salaün, Luc Thiberville. Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images. 2013. hal-00820354v3

**HAL Id: hal-00820354**

**<https://hal.science/hal-00820354v3>**

Preprint submitted on 11 Jun 2014

**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.

# Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images

Alain Rakotomamonjy<sup>a,\*</sup>, Caroline Petitjean<sup>a</sup>, Mathieu Salaün<sup>b</sup>, Luc Thiberville<sup>b</sup>

<sup>a</sup>*Laboration Informatique, Traitement de l'Information et Systèmes EA4108, University of Rouen, Avenue de l'université, 76800 Saint Etienne du Rouvray, France.*

<sup>b</sup>*Clinique Pneumologique, Centre Hospitalier Universitaire de Rouen, 1 rue de Germont, 76031 Rouen Cedex, France*

---

## Abstract

**Objective:** To assess the feasibility of lung cancer diagnosis using fibered confocal fluorescence microscopy (FCFM) imaging technique and scattering features for pattern recognition.

**Methods:** FCFM imaging technique is a new medical imaging technique for which interest has yet to be established for diagnosis. This paper addresses the problem of lung cancer detection using FCFM images and, as a first contribution, assesses the feasibility of computer-aided diagnosis through these images. Towards this aim, we have built a pattern recognition scheme which involves a feature extraction stage and a classification stage. The second contribution relies on the features used for discrimination. Indeed, we have employed the so-called *scattering transform* for extracting discriminative features, which are robust to small deformations in the images. We have also compared and combined these features with classical yet powerful features like *local binary patterns* (LBP) and their variants denoted as *local quinary patterns* (LQP).

**Results:** We show that scattering features yielded to better recognition performances than classical features like LBP and their LQP variants for the FCFM image classification problems. Another finding is that LBP-based and scattering-based features provide complementary discriminative information and, in some situations, we empirically establish that performance can be improved when jointly using LBP, LQP and scattering features.

**Conclusions:** In this work we analyze the joint capability of FCFM images and scattering features for lung cancer diagnosis. The proposed method achieves a good recognition rate for such a diagnosis problem. It also performs well when used in conjunction with other features for other classical medical imaging classification problems.

**Keywords :** scattering transform, wavelet transform, local binary pattern, support vector machine, texture

---

\*Corresponding author

Email address: [alain.rakotomamonjy@insa-rouen.fr](mailto:alain.rakotomamonjy@insa-rouen.fr) (Alain Rakotomamonjy )

## 1. Introduction

Lung cancer is one of the most widespread forms of cancer and a major cause of premature death. Survival of lung cancer is highly associated with early diagnosis [1]. It follows that it is of primary importance to develop methods to help to diagnose cancer. Numerous imaging methods have been developed in order to improve the detection of these early cancerization stages [2]. Medical imaging tools for lung cancer diagnosis are mostly based on inspections of the lung and the chest [3] using X-ray-based systems including computed tomography [4, 5]. For centrally located lung cancer, bronchoscopy is an essential tool for diagnosis. Until recently, the early diagnosis of lung cancer relied primarily on a rough inspection of the bronchial mucosae during an endoscopic procedure and on pathology examination of biopsy samples. Classic fiberoptic bronchoscopy, using white-light illumination, has repeatedly shown a low sensitivity for the detection of the early - presumably curable - lesions such as carcinoma in situ [6]. In particular, some techniques have emerged over the past five years that extend the field of exploration of bronchoscopy to the distal lung and to the cellular level, among which fibered confocal fluorescence microscopy (FCFM) [7, 8].

The FCFM technique (also referred to as probe-based confocal laser endomicroscopy, especially in gastrointestinal imaging) 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, whose applications for lung exploration are currently under investigation, can replace lung biopsy in the future and may prove to be helpful in the diagnosis of a number of diseases [7]. FCFM has other challenging applications, such as the imaging of colonic polyps in gastroenterology, for which image retrieval and classification works are ongoing [9]. FCFM imaging is thus able to provide in vivo cellular images of the lung; for this reason, it opens the road to novel methods for early diagnosis of lung diseases including cancer. This is one of the objectives of this paper: to assess the feasibility of computer-aided lung cancer detection based on FCFM images.

Computer-aided diagnosis systems for lung diseases or cancer based on imaging usually involve image processing techniques and machine learning approaches. Many approaches focus on nodule detection in chest radiographs [10, 11] or in computed tomography (CT) [12–14]. CT is often coupled with advanced numerical techniques for extracting relevant features from the images as well as with state-of-the-art classifiers [15, 16]. Usually, the systems are based upon a two-stage approach: feature extraction and classifier training [11, 17]. The first stage aims at building relevant features that help in discriminating tissues while in the second stage,

the classifier learns from examples to automatically assign a class, typically normal or abnormal, from a set of features extracted from an image. These features are frequently related to texture of the images [17–19]. They thus may be described by means of the local binary patterns (LBP)[20] which are known to be competitive texture feature extractor [21–24]. Other classical features can also be generated from gray level co-occurrence matrices (GLCM) [25], and they often achieve state-of-the-art recognition performance [14, 26, 27]. However, we depart from this mainstream use of LBP for texture features and instead investigate the benefit of using a novel wavelet-based transform, named scattering transform, for building discriminative texture features.

Scattering transform can be understood as successive applications of a wavelet-based linear filtering followed by a modulus operator and a local averaging. This cascade of wavelet transform and non-linearity operators makes the scattering transform locally invariant to translation and stable to small deformations (shear, rotations or any other displacement field) [28–30]. Here, stability to deformations means that the norm of the difference between the scattering transform of an image and its deformed version is bounded by a term which depends on the norm of the image and a constant depending on the deformation. Therefore, this theoretical property guarantees that the scattering representation of an image does not change too much if the image is deformed, which is a positive property for reducing intra-class variability in classification problems. As such, features resulting from this transform are expected to be more effective than a classical wavelet transform. The other objective of this work is to assess the relevance of scattering features for lung cancer detection in images. As the FCFM dataset we use is novel, we have also evaluated the value of these features on other well-known, publicly available medical image datasets (the *2D-Hela* and the *Pap smear* datasets).

As far as we know, this work presents the first application of scattering transform to medical images. The contributions we present are three-fold :

- First, we assess the feasibility of lung cancer diagnosis from FCFM images. FCFM imaging is a rather recent imaging technique and its importance and applicability are yet to be explored. In this paper, we show that this image modality can help detecting lung cancer.
- Second, we show that scattering coefficients are highly discriminative features for this FCFM-based lung cancer classification problem. When compared to one of the state-of-the-art features, the LBP feature, and its recent variant local quinary pattern (LQP) [24], they achieve strongly competitive results, yielding the best discrimination performance. Comparisons with GLCM-based features also show that scattering features are efficient.
- Finally, in order to achieve insight on the discriminative power of scattering representation, we further analyze its behavior on two other medical image classification problems. The findings suggest that i)

the choice of the wavelet has an effective impact on the classification performance, ii) a generative classifier adapted to scattering features does not perform necessarily better than a standard discriminative classifier like a support vector machine (SVM), iii) coupling LBP or LQP and scattering features by concatenation may further enhance classification performances, and iv) fusing classifiers trained with single LBP, LQP and scattering features yield to slightly better performances than a single classifier trained with concatenated features.

The paper is organized as follows. Section 2 briefly presents the FCFM imaging technique as well as the image dataset used in the experiments. A discussion on the scattering transform is given in Section 3. This section also briefly recalls how LBP and LQP features are constructed and shortly introduces the classifier that is used for learning a decision function. Experimental analyses are reported in Section 4.

## 2. Materials

FCFM bronchi images are acquired by the application of a probe tip onto the bronchial wall. Even if cancer significantly alters the bronchial epithelial layer, the absence of epithelial cell visualization does not allow FCFM to differentiate between the different grades of progression of the pre-cancer bronchial lesions such as metaplasia, dysplasia and carcinoma in situ. To be successfully applied to the exploration of pre-cancer/cancer bronchial epithelial layer, the FCFM technique needs to be coupled with the use of a nontoxic fluorophore, such as methylene blue (MB). To give a fluorescent signal, MB needs to be excited around 660 nm, and is therefore accessible to FCFM imaging using this excitation wavelength. After topical application of methylene blue and 660 nm excitation wavelength, the technique allows the direct visualization of cell nuclei, as shown in Figure 1. Future studies using this technique could make it possible to differentiate normal, premalignant and malignant alterations at the microscopic level. If this strategy is successful, FCFM may become a very powerful technique for in vivo diagnosis of early malignant and premalignant conditions of the bronchial tree. Therefore, the aim in this paper is to provide the clinician with a computer aided-diagnosis (CAD) tool, in order to help him to analyze these new images and detect cancer lesions.

Towards this purpose, we have acquired a set of FCFM images from healthy volunteers as well as from patients diagnosed with bronchial squamous cell carcinoma. The confocal endomicroscopy system used to produce the FCFM images (Cellvizio; Mauna Kea Technologies, Paris, France) is commercially available. It uses very thin and flexible miniprobes (300  $\mu\text{m}$  to 2 mm in diameter) that can contain up to 30,000 compacted microfibers. The main advantages of this design are its very small size and the flexibility of the probe that can reach the more distal part of the lungs, as well as the high speed image acquisition that helps to avoid artifacts due to tissue movement. Specific miniprobes for bronchial and alveolar imaging have a diameter of 1 mm that can enter the working channel of any adult bronchoscope. These miniprobes have a depth of

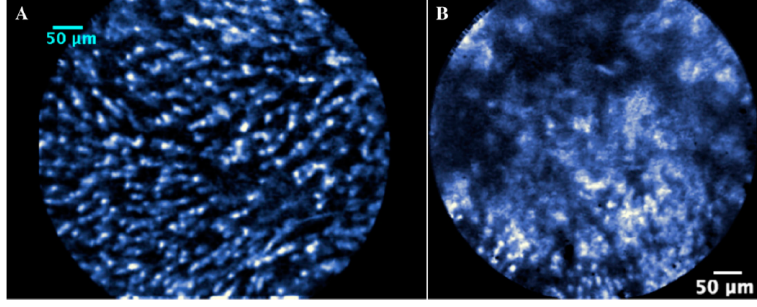


Figure 1: In vivo imaging of the normal epithelial layer after topical application of 0.1% methylene blue, and 660 nm illumination, of a healthy bronchus (left) and a bronchial squamous cell carcinoma (right). Note how cell nuclei are more visible on the healthy image than on the cancer image.

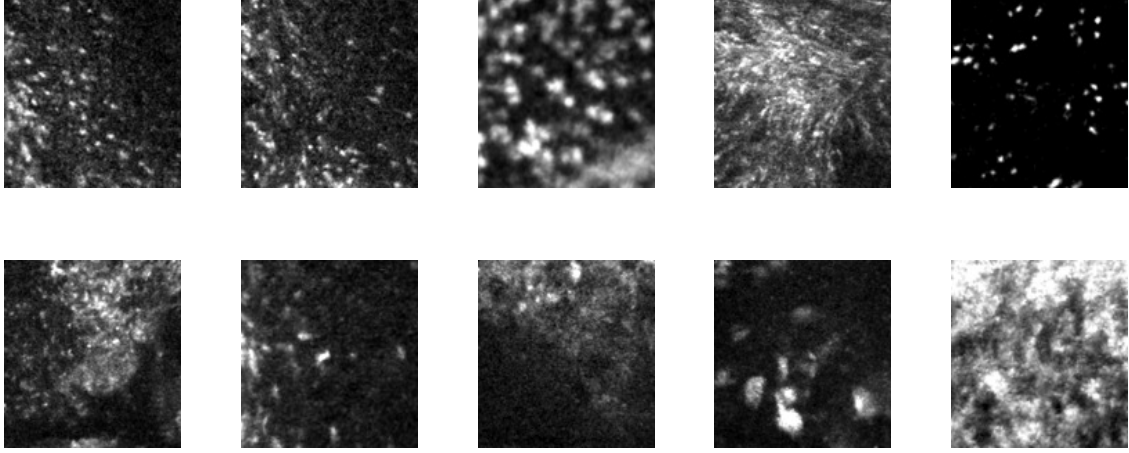


Figure 2: Examples of images from the fibered confocal fluorescence microscopy dataset showing the class variability: non-cancer images (top row) and cancer images (bottom row)

focus of 0 to 50  $\mu m$  and a lateral resolution of 3  $\mu m$ , for a field of view of 600  $\mu m$ . This system produces endomicroscopic imaging in real time at 9 to 12 frames/second [7].

FCFM images for the dataset are obtained by extracting the largest inscribed square from the originally circular images delivered by the system. The dataset is composed of 103 images taken from 8 healthy volunteers, and 70 images from 7 patients with diagnosed bronchial squamous cell carcinoma. Note that these 70 images have been acknowledged by experts as presenting characteristics of squamous cell carcinoma. Image sizes originally vary from  $190 \times 190$  pixels to  $288 \times 288$  pixels, but have been resized to  $128 \times 128$  for a more efficient application of the scattering transform. Figure 2 depicts several examples of images of the two classes, showing that intra-class variability can be important.

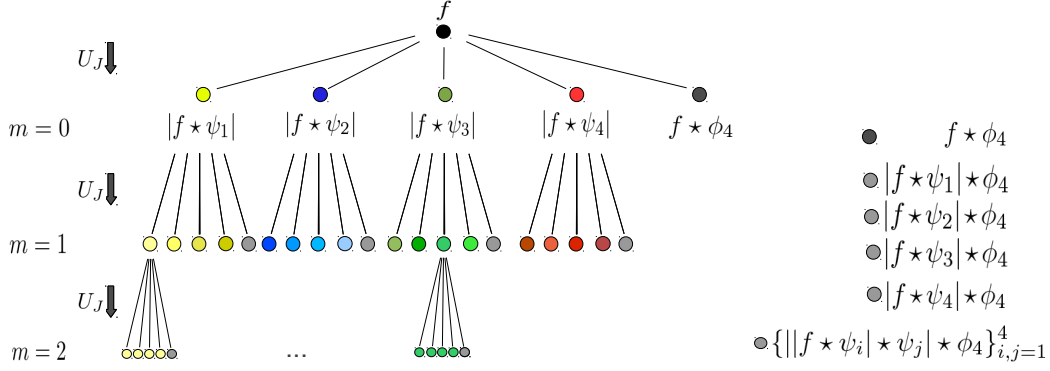


Figure 3: Diagram of a scattering transform of an image  $f$ . For this example, we have supposed that  $J = 2$  and the 2-dimensional wavelet transform is obtained from 4 wavelets  $\{\psi_i\}_{i=1}^4$  with two angles of rotation and two distinct frequencies. All points represent an image obtained from successive applications of  $U_J$ . The grey points, which are averaged versions of a given  $U[p]f$ , are the elements of the transform that are gathered together, after a possible subsampling, to form the scattering transform at depths  $m = 0$ ,  $m = 1$  and  $m = 2$ . Note that because only paths with frequency-decreasing wavelet are of interest, only 4 paths out of the 16 will be finally retained for  $m = 2$ . (best viewed in color).

### 3. Methods

#### 3.1. Scattering representation of images

Scattering transform has recently been introduced by Bruna et al. [28] in order to build representations of images and signals that are stable to deformations. Informally, a scattering transform recursively applies a cascade of wavelet decomposition and modulus operator and in some sense, it mimics convolutional neural networks as it cascades several layers of filtering [31]. We will now provide a more formal definition of the scattering representation of an image. More details can be found in [28, 32].

Let  $\psi_{j,\gamma}(\mathbf{x})$  be a directional wavelet at a scale  $j$ , defined as:

$$\psi_{j,\gamma}(\mathbf{x}) = 2^{2j} \psi(2^j \mathbf{R}_\gamma \mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^2$$

where  $\mathbf{R}_\gamma$  is a rotation matrix of angle  $\gamma$  and  $\psi$  a two-dimensional wavelet. Note that negative and positive values of  $j$  respectively correspond to dilation and contraction of the mother wavelet  $\psi$ . For the sake of simplicity, wavelet parameters  $j$  and  $\gamma$  are gathered into a single parameter  $\lambda$ . Let us also define  $\phi_{2^J}(\mathbf{x}) = 2^{-2J} \phi(2^{-J} \mathbf{x})$ ,  $J \geq 0$  as an averaging spatial window obtained by dilating a function  $\phi$ . The scattering transform of an image  $f$  of size  $2^J \times 2^J$  is obtained from the computation of high-order wavelet coefficients by recursively applying an averaging, a wavelet transform and a modulus operator to the wavelet modulus coefficients obtained at a lower level.

In order to properly define such a transform, we first introduce the operator which computes the modulus

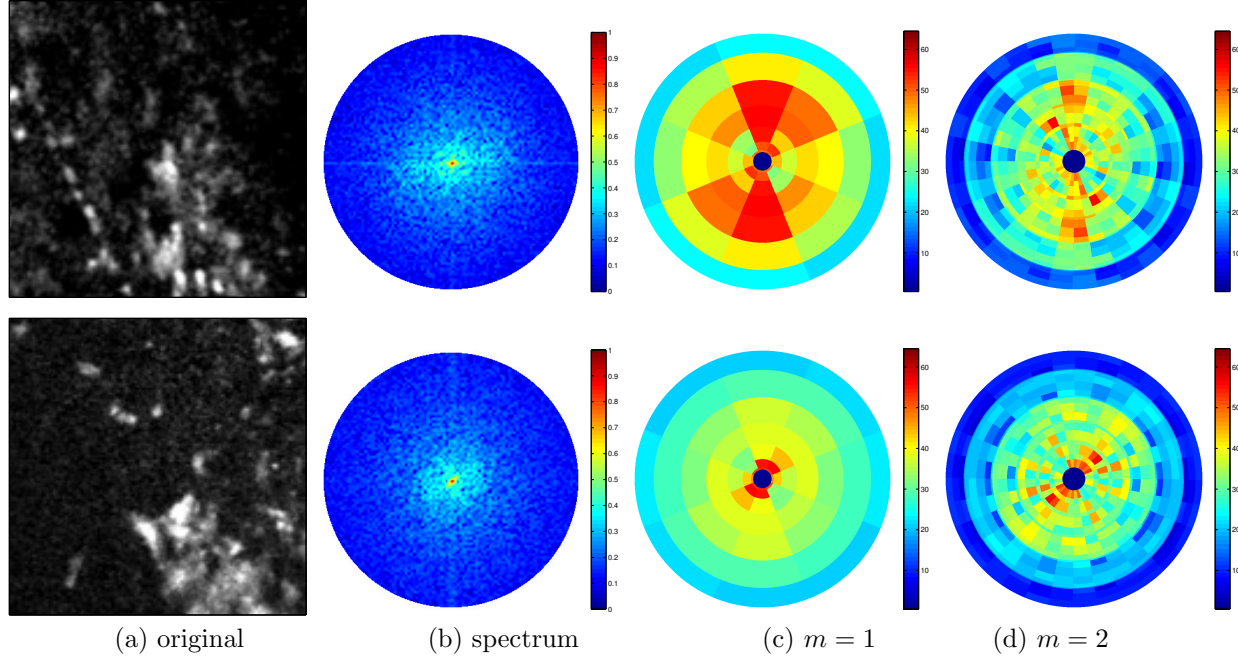


Figure 4: Examples of non-cancer (top) and cancer images (bottom) with their representations. From left to right, (a) the original image. (b) its spectrum. (c)  $m = 1$  scattering coefficients (d)  $m = 2$  scattering coefficients. The original image is of size  $128 \times 128$  and we have chosen  $L = 4$  and  $J = 7$ . The scattering coefficient representation shows, for  $m = 1$ , the amplitude of  $S_J[p]f$  where each quadrant of the diagram is related to a wavelet with a given direction and scale. Wavelet direction and scale divides the diagram respectively according to its radial axis and radius. For  $m = 2$ , each quadrant is again subdivided according to directions and scales of wavelet at the level  $m = 2$ .

wavelet transform and the averaging as:

$$U_J f(\mathbf{x}) = \left\{ f \star \phi_{2^J}(\mathbf{x}); \{|f \star \psi_\lambda(\mathbf{x})|\}_{\lambda \in \Lambda_J} \right\}$$

where  $\star$  is the convolution operator and  $\Lambda_J$  the set of  $j$  and  $\gamma$  defining a 2D wavelet transform at scale  $J$ . Note that  $U_J f(\mathbf{x})$  is composed of the pointwise value at a pixel coordinate  $\mathbf{x}$  of an averaged version of  $f$ , the  $f \star \phi_{2^J}(\mathbf{x})$  part, and the modulus pointwise values of the image convolved with a set of wavelet  $\{\psi_\lambda\}$ .  $U_J f(\mathbf{x})$  is thus a vector of  $\mathbb{R}^{|\Lambda_J|+1}$ .

Applying  $U_J$  to the output  $|f \star \psi_{\lambda_1}|$  (which is an image), with  $\lambda_1 \in \Lambda_J$ , gives the following set of coefficients

$$\left\{ |f \star \psi_{\lambda_1}| \star \phi_{2^J}(\mathbf{x}); \{| |f \star \psi_{\lambda_1}| \star \psi_\lambda(\mathbf{x}) | \}_{\lambda \in \Lambda_J} \right\}$$

This latter set of coefficients is denoted as a two-level cascade of modulus wavelet transform and forms the basis of the so-called scattering transform. Indeed, the coefficients of a scattering transform are obtained by cascading several times the  $U_J$  operator.

Formally, let us assume we are looking for the scattering coefficients of order  $m$  at a scale  $2^J$ . We denote by  $m$  the order of the cascade and we define  $p$  as a sequence  $p = \{\lambda_1, \lambda_2, \dots, \lambda_m\}$  of  $m$  wavelet parameters, where each  $\lambda_i$  belongs to  $\Lambda_J$ .  $p$  can thus be understood as a specific path of length  $m$  among the  $|\Lambda_J|^m$



possible ones available in the set denoted as  $\Lambda_J^m$ . A path  $p$  defines an unique wavelet modulus cascade output

$$U[p]f = ||f \star \psi_{\lambda_1} \star \psi_{\lambda_2} \cdots \star \psi_{\lambda_m}|$$

Note that for a given  $p$ ,  $U[p]f$  is an image obtained by successively convolving  $f$  with a wavelet  $\psi_\lambda$  and by taking the modulus of the resulting coefficient. By averaging these coefficients at the scale  $2^J$ , we can define

$$S_J[p]f = U[p]f \star \phi_{2^J}$$

with  $S_J[\emptyset]f := f \star \phi_{2^J}$ . We can thus interpret  $S_J[p]f$  as an averaged version of the wavelet modulus cascade coefficients  $U[p]f$ . After convolution with  $\phi_{2^J}$ , the output is usually subsampled at intervals  $2^J$ . This means that if the image size is  $2^J$  the averaging produces a single coefficient and thus  $S_J[p]f$  is a single coefficient obtained as the average of  $U[p]f$  over the whole image.

From all these equations, we can now define the scattering transform  $S_J[\Lambda_J^m]$  of  $f$  at order  $m$  and scale  $2^J$  as the set of coefficients

$$S_J[\Lambda_J^m]f = \{S_J[p]f\}_{p \in \Lambda_J^m}$$

which also includes the scattering coefficients of  $f$  at lower order.

Another insight on scattering transform can be gained by explicitly exhibiting the cascading scheme involving  $U_J$  [32]. Let us first remark that  $U_J f$  can be rewritten as

$$U_J f = \{S[\emptyset]f, U[\Lambda_J^1]f\}$$

where  $U[\Lambda_J^1]f = \{U[p]f\}_{p \in \Lambda_J^1}$ .  $U[\Lambda_J^1]f$  is thus the set of all wavelet modulus coefficients, since  $p$  is of length 1 and consequently  $\Lambda_J^1 = \Lambda_J$ .  $S_J[\Lambda_J^1]$  can now be obtained by averaging these images through  $\phi_{2^J}$ . This can be done by applying  $U_J$  to  $U[\Lambda_J^1]f$ , yielding:

$$\begin{aligned} U_J U[\Lambda_J^1]f &= \{U_J U[p]f\}_{p \in \Lambda_J^1} \\ &= \{S_J[\Lambda_J^1]f, U[\Lambda_J^2]f\} \end{aligned}$$

Hence, by iteratively applying this scheme to sequences  $p$  of length  $m' \leq m$ , all the scattering coefficients at order  $m$  can be retrieved. Figure 3 gives a schematic representation of this scheme, for which an efficient algorithm can be found in [32].

We can also see from this diagram that the number of resulting coefficients can rapidly grow with respect to  $m$ . However, because some paths  $p$  lead to negligible coefficients (those of increasing-frequency), they can be omitted while computing the transform [32]. Hence, according to Bruna et al., for an image with  $N \times N$  pixels, scattering transform leads to a novel representation at depth  $m$  of size  $N^2 2^{-2J} \sum_{i=0}^m L^i \binom{J}{i}$ ,  $L$  being the

number of rotations applied to the wavelet, where the sum counts the number of paths with decreasing frequency and the term  $N^2 2^{-2J}$  counts the number of coefficients after averaging and subsampling. Hence, if  $J$  is chosen so that  $N = 2^J$ , which means that the averaging with  $\phi_{2^J}$  applies on the whole image, only a single coefficient results from the averaging and the subsampling. This considerably reduces the number of coefficients in the scattering transform at the expense of losing some spatial resolution. The number of paths is illustrated in Figure 3, where 4 wavelets with 2 different frequencies ( $f_1$  and  $f_2$  with  $f_1 > f_2$ ) and 2 different rotations are used. For  $m = 1$ , the number of paths with decreasing frequencies is 4 (i.e. all possible paths). For  $m = 2$ , we still have 4 interesting paths, only the two of the form  $\{f_2, f_1\}$  times the two angles of rotation.

The strengths of scattering representation for classification problems come from several properties it is endowed with. Among those of paramount importance, we can mention the following ones:

- Its stability to small deformations. Let us assume a small, spatial or iconic (i.e., image intensity related) deformation is applied to an image. The scattering representation of the deformed image is expected not to differ much from the one of the original image [30, 32]. This is essential in classification tasks since images from same classes are supposed to live in a smooth manifold and thus, we expect a representation to also vary smoothly. Note that in this work, we will not carry out experiments that support this claim on robustness to deformation.
- The modulus operator and the averaging step provide local translation invariance to the representation. This local invariance acts at the level of the scaling  $2^J$ . Hence, if  $\phi_{2^J}$  covers the whole image, we have a scattering representation that is fully translation invariant at the cost of poor spatial resolution.
- Under mild conditions on the wavelet, it can be shown [32] that the norm of the scattering coefficients is equal to the one of the original signal and that most signal energy can be retrieved by considering scattering transform of depth  $m = 3$ .
- The scattering transform also bears an interesting property related to textures that can be considered as a stationary process. Indeed, Bruna et al. have proved that scattering coefficients provide information about high-order statistical moments of textures. Hence, scattering representation has the ability to discriminate textures that have similar moments, up to a certain order.

For the sake of representation, Bruna and Mallat introduced a polar representation of the scattering transform that can handle depth up to  $m = 2$ . As shown in Figure 4, the plot uses a gray-scale for representing coefficient amplitudes in regions depending on the scale and rotation angle of the wavelets. We remark in this figure that the scattering representations discriminate quite well between the normal and the cancer FCFM lung image examples, and that a good discriminative power of the resulting features may be expected.

### 3.2. Local binary pattern and variants

Since the objective is to assess whether scattering coefficients can be useful for medical imaging texture classification problems, we have compared their discriminative power with those of classical texture descriptors, namely the Local Binary Patterns, as well as their recent variants. These LPB have been popularized by Ojala et al. [20] and have been extensively used for image classification problems [21] [33]. In this work, we have used them as baseline features for comparison.

The LBP is a texture operator that extracts some local information from the neighborhood of a given pixel. In the original LBP formulation, this neighborhood is defined as a circle of radius  $R$  around the pixel. The LBP operator assigns to that center pixel a label obtained by comparing the gray value of each neighbor with the one of the center. The binary number resulting from the concatenation of all 0 or 1 output by the comparison forms the label of the center pixel. Formally, the LBP pattern for a given pixel is defined as:

$$LBP(c, P) = \sum_{p=0}^{P-1} \sigma(I_p - I_c) 2^p$$

where  $\sigma(u) = 1$  if  $u \geq 0$  and 0 otherwise,  $I_p$  and  $I_c$  being respectively the gray value of one of the  $P$  neighbor pixels and the center pixel. The LBP texture feature is defined as the histogram of the labels assigned to all image pixels. More details on LBP features may be obtained in these references [20, 21].

The variants of LBP consider different codings of the gray level difference. Instead of a binary pattern, a quinary one involving 2 thresholds is used in [24]. For this pattern, the coding of the difference is as follows:

$$\sigma(u) = \begin{cases} 2 & u \geq \tau_2 \\ 1 & \tau_1 \leq u \leq \tau_2 \\ 0 & -\tau_1 \leq u \leq \tau_1 \\ -1 & -\tau_2 \leq u \leq -\tau_1 \\ -2 & u \leq -\tau_2 \end{cases}$$

The quinary pattern is then transformed into 4 binary patterns and the resulting histograms are concatenated together to form a feature vector. One of the claimed advantages of considering these codings compared to the binary ones is their better robustness to noise [24].

### 3.3. Classification scheme

For evaluating the discriminative power of the scattering coefficients and the LBP features, we have employed the following simple classification scheme.

Suppose we have  $\{\tilde{\mathbf{x}}_i, y_i\}_{i=1}^n$  images and their associated labels for learning a decision function  $f(\cdot)$  that predicts the label of an unseen image  $\tilde{\mathbf{x}}$ . For this purpose, a feature vector  $\mathbf{x}_i$  is computed from each image

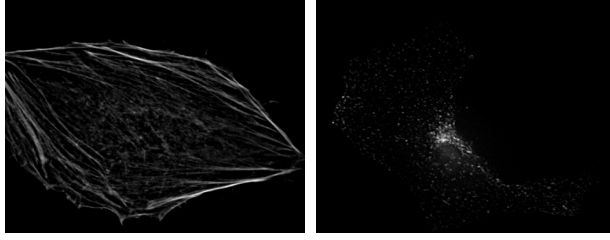


Figure 5: Examples of images from the *hela* dataset, “actin” class (left) and “endosome” class (right)

$\tilde{\mathbf{x}}_i$ , the features being obtained either through a scattering transform or the computation of LBP. For binary classification problems ( $y_i = \pm 1$ ), the decision function is chosen to be an SVM-based decision function [34].

It is defined as

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i y_i K(\mathbf{x}, \mathbf{x}_i) + b$$

where  $K(\mathbf{x}, \mathbf{x}')$  is a positive-definite kernel function such as the Gaussian kernel  $e^{-\frac{\|\mathbf{x}-\mathbf{x}'\|_2^2}{2\sigma^2}}$  and  $\{\alpha_i\}_{i=1}^n$  being the solution of the quadratic programming problem

$$\begin{aligned} \max_{\alpha} \quad & \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{st} \quad & \sum_i \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, \quad \forall i \end{aligned}$$

$C$  is a parameter that penalizes mis-classified examples during the learning procedure.

When a multi-class problem is in play, a *one-against-one* SVM has been deployed using the same training procedure as above for each binary classification problem.

#### 4. Experimental results and discussions

The aims of the experimental results are two-fold. Primarily, we want to empirically evaluate the efficiency of scattering-operator-based features on classical and well-known medical imaging classification problems. Afterwards, we focus on the FCFM dataset and provide an in-depth empirical analysis of scattering features for this lung imaging cancer detection problem. The Matlab code used for producing these results is available on the author’s website<sup>1</sup>.

##### 4.1. Description of image datasets

In order to better understand the generic discrimination power of the scattering transform features, we have also benchmarked their performances on the *2D-Hela* dataset and *Pap smear* dataset (both available online<sup>2</sup>)

<sup>1</sup><https://sites.google.com/site/alainrakotomamonjy/home/publication/CodeScattLung.zip>, last accessed April 1st 2014

<sup>2</sup><http://labs.fme.aegean.gr/decision/downloads> last accessed January 27 2014.

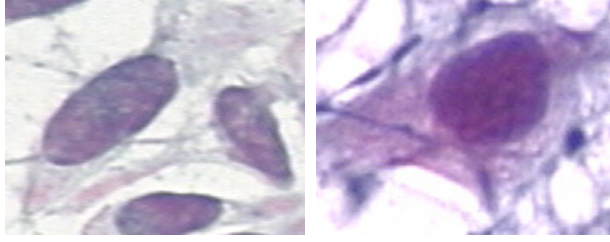


Figure 6: Examples of images from the *pap smear* dataset normal (left) abnormal (right)

in addition to the FCFM images.

The *2D-Hela* dataset is a cell phenotype multi-class, image classification problem [35]. It consists of a 10-class problem, each class being composed of a number of images ranging from 73 to 98 for a total of 862 images. For this problem, images have been resized to  $382 \times 382$ . Some examples of images are presented in Figure 5. We note that in these images the texture seems to be located at the fine scales.

The *Pap smear* dataset aims at detecting pre-malignant cells extracted from the uterine cervix [36, 37]. The original dataset includes 7 different classes which can be grouped into normal and abnormal classes, composed of respectively 242 and 675 images of size  $128 \times 128$ . We have thus turned the problem into a binary classification problem. Some examples of cell images are presented in Figure 6. Unlike the *2D-Hela* dataset, the texture energy of *Pap smear* dataset images is located at coarse scales.

We have focused interest on these datasets as they serve as reference datasets on texture-based image classification problems [21]. Note that images have not been pre-processed before feature extraction.

#### 4.2. Experimental set-up

Scattering representations of images have been obtained by means of the Matlab toolbox<sup>3</sup> provided by Bruna et al. [28]. Parameters of the scattering transforms have been fixed as follows. The number of orientations  $|\gamma|$  as well as the representation depth  $m$  have been set by default to 4 and 2. The scale  $J$  is adapted according to image size and will be reported for each experimental analysis.

Different features will be compared throughout the experiments. We do not make any specific choice on the wavelet for the scattering representation but instead, we will investigate their specific discriminative power. The extraction of LBP features also relies on a Matlab code available online<sup>4</sup>. In order to enrich the LBP representation, we have concatenated the normalized histogram obtained from neighbourhoods of radius 1, 2 and 4 with a respective sampling of 8, 16 and 24 pixels on the circle. We have also built rotation invariant

<sup>3</sup><http://www.di.ens.fr/data/software> last accessed January 27 2014.

<sup>4</sup><http://www.cse.oulu.fi/CMV/Downloads/LBP Matlab> last accessed January 27 2014.

uniform LBP features and uniform LBP of respective sizes 54 and 857. For the LQP features, we have proceeded similarly as above. The Matlab code we used is the companion code<sup>5</sup> of the works of Paci et al. [24]. Here, we have concatenated the features obtained from neighborhood of radius 1 and 2 with 8 and 16 pixels on the circle. For each neighbourhood, the quinary patterns yielded from several values of threshold  $\tau_1$  and  $\tau_2$  have been computed. In these experiments  $\tau_1$  takes value from  $1, 3, \dots, 9$  while  $\tau_2$  ranges from  $\tau_1 + 2, \tau_1 + 4, \dots, 11$ .

Once the features have been obtained, they are fed to a classifier which is a Gaussian kernel Support Vector Machine [34]. In order to evaluate the couple feature-classifier, we ran 30 trials where each trial consists of a random split of the available examples in training examples and test examples. Sizes of the split will vary from 30% to 80% of the number of examples. The SVM hyperparameters have been tuned by means of a validation step which randomly splits the training examples 50% – 50% into a learning and validation sets. Parameters are the kernel bandwidth  $\sigma$  and the SVM slack parameter  $C$ . They are tuned respectively in the range  $[0.01, 0.1, \dots, 1000]$  and  $[1, 2.5, 5, 7.5, 10, 12.5, 15, 25, 30, 35]$ . This random split is performed 5 times and the best averaged result defines the best SVM hyperparameters. Finally, a SVM with these optimal hyperparameters is learned based on the full training examples and evaluated on the test set. Performances are then averaged over the 30 trials. Statistically significant differences in performances have been evaluated according to a Wilcoxon signrank test at a level of 0.05.

In the experiment with the *FCFM* dataset, performances are evaluated according to the area under roc curve (AUC) which is a more significant performance criterion than accuracy for medical classification problems. For *Pap-smear* and *2D Hela*, we have respectively used AUC and the classification accuracy in order to be consistent with some recently published works [21, 24].

While the *2D Hela* and *Pap smear* datasets do not contain any information about patients, the *FCFM* dataset has these information. Consequently, for the in-depth analysis of this dataset, we will consider a *leave-one-patient-out* (LOPO) evaluation procedure.

#### 4.3. On the choice of the wavelet for scattering features

The first experiment consists in evaluating the discriminative power of scattering features obtained from different wavelets in the scattering transform.

Parameters for the scattering representation have been set as  $m = 2$  and  $L = 4$ . These default values are expected to already provide a good classification accuracy since with paths of that depth, the scattering transform captures most of the image energy. The scaling factor  $J$  is chosen to fit the entire image, which

---

<sup>5</sup><http://bias.csr.unibo.it/nanni/TernaryCoding.rar> last accessed January 27 2014.

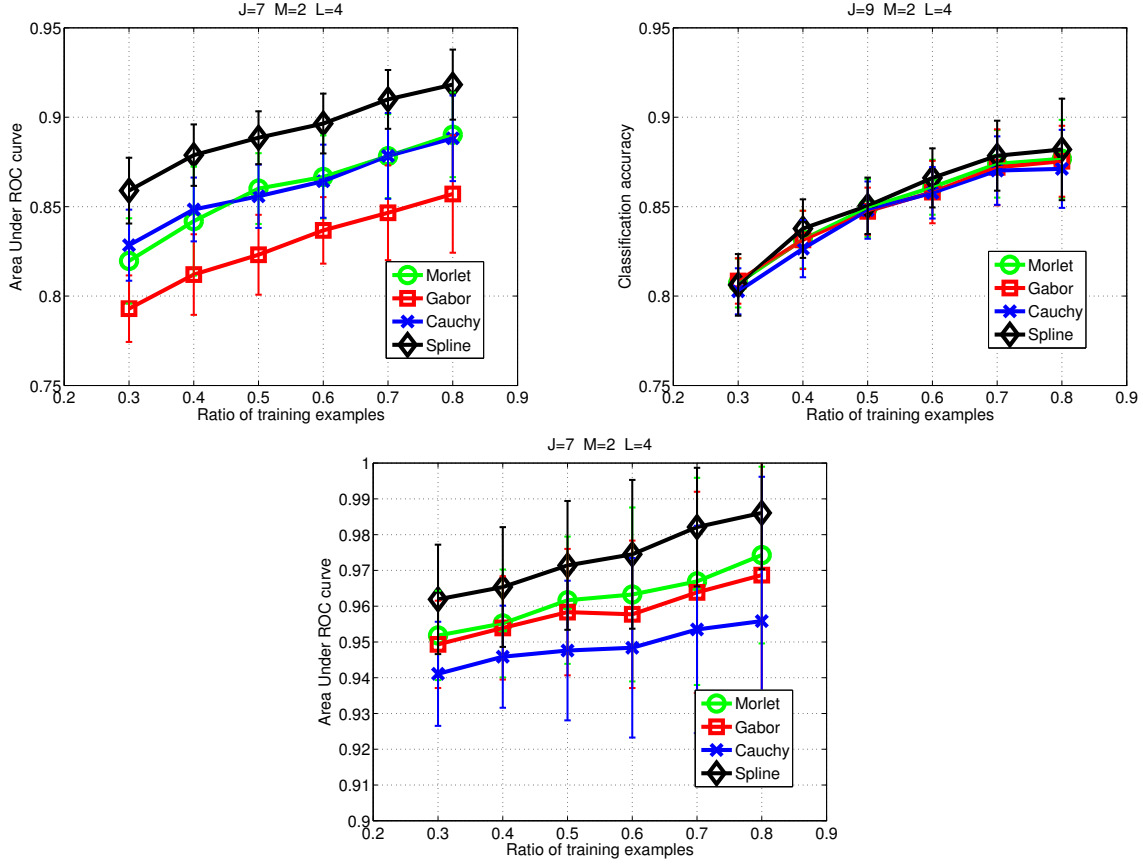


Figure 7: Influence of the wavelet on the classification performances for increasing ratio of training examples. Pap-smear dataset (top-left), 2-dimensional Hela dataset (top-right), FCFM images (bottom).

means that from each sequence  $p$ , only a single coefficient is obtained, resulting in a full translation invariance at the expense of a complete loss of spatial resolution.

Performances of the different scattering features obtained from different wavelets for the 3 datasets of interest are depicted in Figure 7. Results can be interpreted as follows: first, we can note that for all datasets, the best performing wavelet is the Spline wavelet. For the *Pap smear* and the *FCFM* datasets, these differences in performance are statistically significant ( $p\text{-value} < 10^{-3}$ ). Except for this trend, it is difficult to extract a general behaviour of the wavelet on the scattering feature performances as wavelet performances vary with the dataset. This suggests that while Spline wavelet can be a good default choice, it is preferable to consider the type of wavelet as a parameter to be optimized, for a problem at hand.

#### 4.4. Affine classifier vs SVM

As suggested by Bruna et al. [32], an affine classifier may perform better than a discriminative classifier like SVM for scattering features. The goal in the next experiment is to compare the performance of an SVM

classifier to one of such generative affine classifiers. For this affine classifier, the predicted label depends on the subspace (we have one subspace per class) which best approximates the scattering feature vector representing a novel image to classify. Formally, the assigned label of a novel vector  $\mathbf{x}$  is defined as:

$$\operatorname{argmax}_{k=1,\dots,N_C} \|\mathbf{x} - P_{A_k}(\mathbf{x})\|_2$$

where  $N_C$  is the number of possible classes and  $P_{A_k}(\mathbf{x})$  is the projection of  $\mathbf{x}$  on the subspace spanned by the  $d'$  largest eigenvectors of the empirical covariance matrix of the training examples of class  $k$ . For the experiment we have kept the setting as above, and we have found the best dimension  $d'$  by a cross-validation procedure over the values  $[10, 20, \dots, 100]$ . The wavelet used in the scattering transform is now the Spline wavelet, since it is the best performing one for all considered datasets.

Because the affine classifier directly assigns a class to a novel example, reported results are based on classification accuracy. As shown in Figure 8, the SVM classifier performs better than the affine one. Indeed, performances of the SVM are always higher than those of the affine classifier, with statistical significance (p-value  $< 10^{-3}$ ), except when both are trained on a small number of examples.

#### 4.5. Comparing scattering features, local binary patterns and local quinary patterns

This experiment aims at comparing the discriminative power of LBP, LQP and scattering representations obtained from a Spline wavelet. We have also concatenated LBP and scattering as well as LQP and scattering, so as to evaluate their joint discriminative powers. Experimental set-up is similar as above and the averaged performances over 30 random split of the examples are reported in Figure 9.

Performances on the *Pap smear* dataset are reported in the top-left panel of Figure 9. We note that scattering features lead to better performance than LBP features, and these differences in performance are statistically significant when the ratio of training examples is larger than 0.6. On the contrary, the LQP features perform better than the scattering features, with statistically significant differences when the ratio of training examples is smaller than 0.6. Combining LBP and LQP features with scattering yields in this case to enhanced performances ranging from 2% to 4% of AUC (p-value  $\approx 10^{-5}$ ) compared to the use of scattering features.

For the *2D Hela* dataset (top-right panel in Figure 9), rotation invariant LBP and LQP perform better than scattering features with statistically significant difference (p-value  $< 10^{-3}$ ). For this problem too, feature combination increases performance (p-value  $\approx 10^{-2}$  when large ratio of training examples are considered) compared to the single use of LBP features. As far as we know, the best reported performance on this dataset reaches 95.3% of recognition rate [38]. While this latter work uses a slightly different protocol than the one we use, making results incomparable, they achieve this state-of-the-art result by considering a classical texture feature and a multiresolution approach. We thus believe that integrating the proposed features into such a multiresolution framework may further improve classification accuracy.



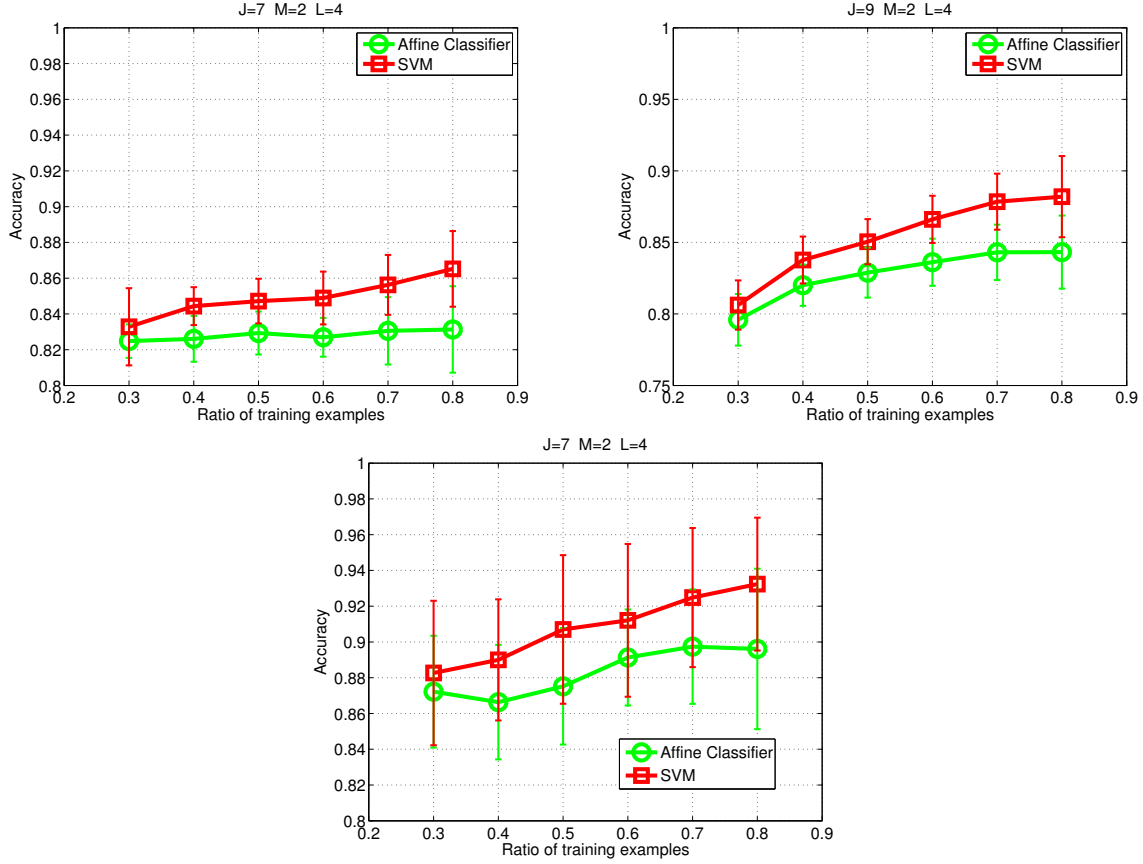


Figure 8: On the impact of the classifier when using Spline-wavelet-based scattering features for increasing number of training examples. Pap-smear dataset (top-left), 2-dimensional Hela dataset (top-right), FCFM images (bottom).

For the FCFM dataset of interest, we first note that the recognition problem is easier than the two previous ones, since both LBP, LQP and scattering coefficients achieved AUC above 0.94 for a sufficiently large ratio of training/test examples. When 80% of the dataset is used as training examples, all LBP and LQP features yield to an AUC of 0.97 whereas scattering features achieve AUC of 0.98. Interestingly, the scattering features perform better than LBP and LQP with statistical differences ( $p\text{-value} < 10^{-2}$ ) for all ratio of training/test examples. Although performances are already excellent, slight performance gains can still be achieved by combining LBP or LQP to scattering features, though the differences are not statistically significant.

These are encouraging results since the scattering representation parameters have been set by default and not tuned to each dataset. In addition, the LBP and LQP features we used are enriched as several features resulting from different parametrizations of the feature extractor have been concatenated. These results, while far from being exhaustive, lead us to believe that the scattering representations achieve discriminative power at least similar to state-of-the-art methods like LBP and its variants. Interestingly, the fusion of the two types of features leads to enhanced performances for the three problems we consider. This suggests

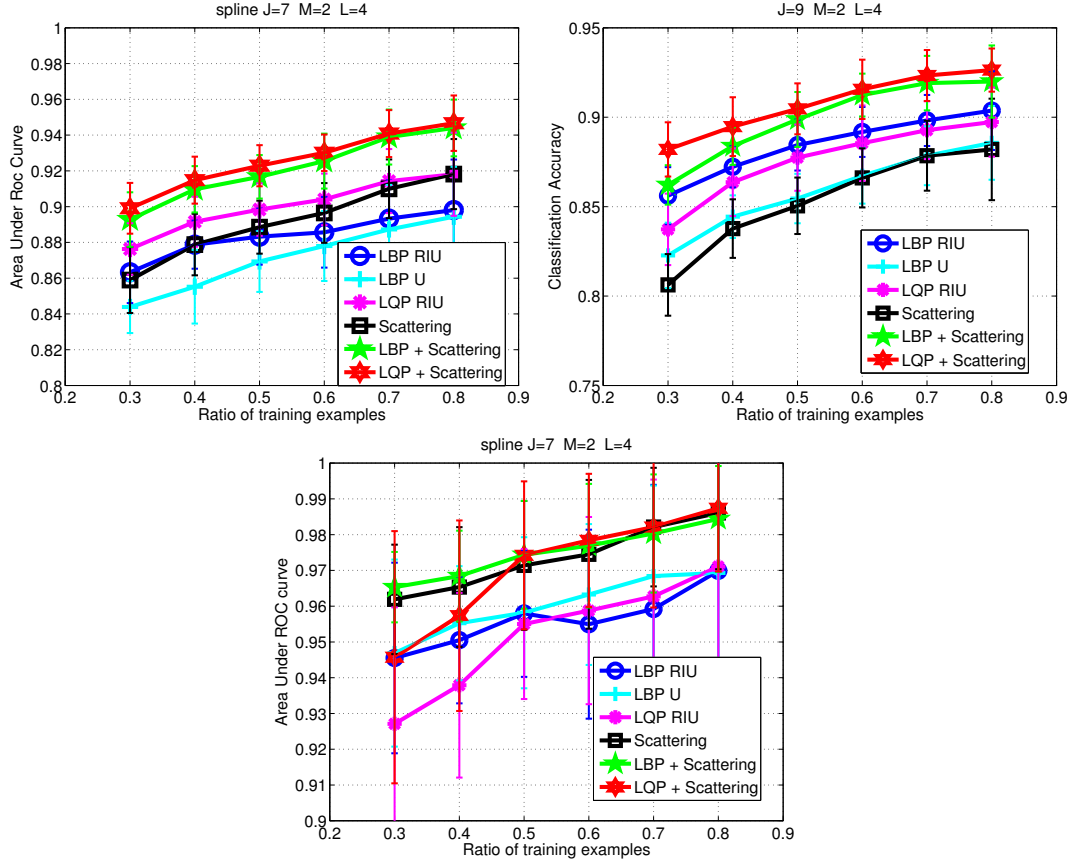


Figure 9: Comparing performances of different local binary pattern, local quinary pattern features and scattering features for the three datasets. Pap-smear dataset (top-left), 2-dimensional Hela dataset (top-right), FCFM images (bottom). LBP RIU and LQP RIU are related to rotation invariant uniform LBP and LQP whereas LBP U depicts uniform LBP.

that these features bring complementary information that are essential for classification. In what follows, we provide intuitions on why and when these two types of features are expected to be complementary and have to be combined.

Note that we have also tested gray-level co-occurrence matrix features but since their performances are poor (always lower than 84%, 61% and 96.5% respectively for *pap-smear*, *2-dimensional hela* and the *fibred confocal fluorescence microscopy* datasets), we have not included them in the figures for a sake of readability.

#### 4.6. Fusing classifiers trained on different wavelet scattering features

Following the idea of combining features, we have investigated the performances of fused classifiers where each classifier is trained on each of the four different wavelets used for building scattering features. We have analyzed the performances of three different fusion rules, namely the *sum*, *max* and *median* rules which respectively consider the sum, the maximum and the median of the SVM scores output by the four classifiers of each one-vs-one scheme.

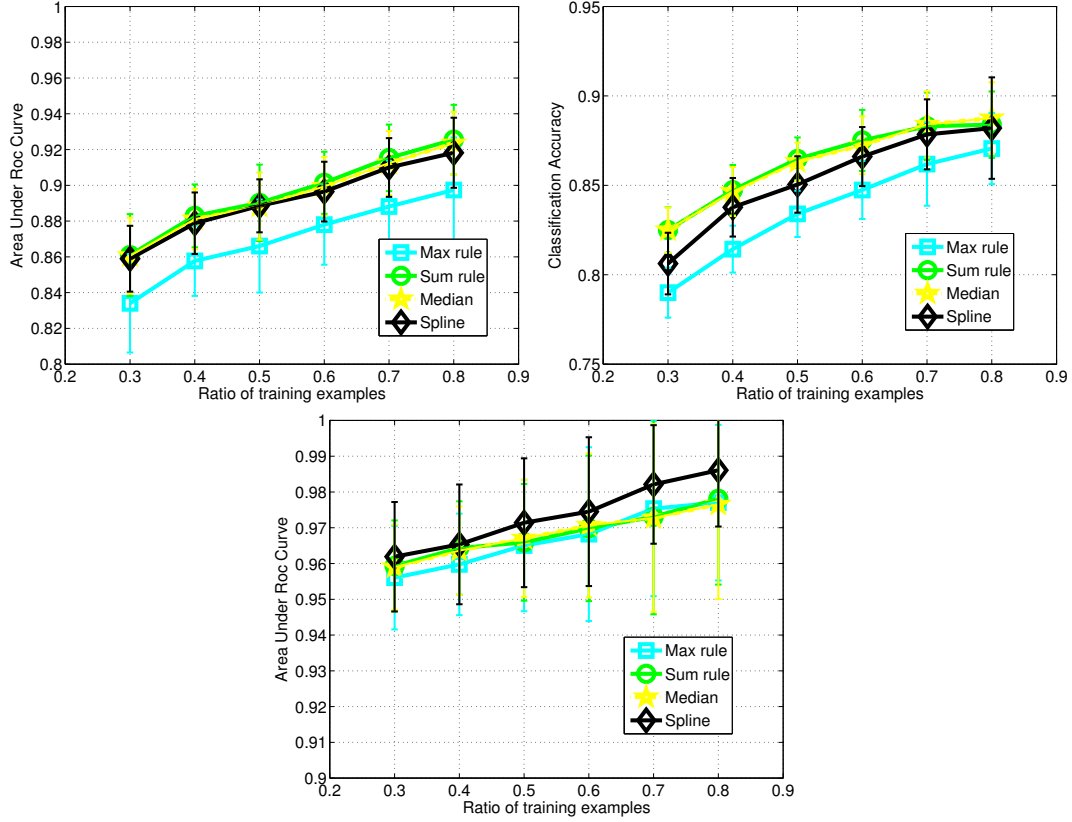


Figure 10: Comparing performances of different fusion rules (maximum, median and sum) of classifiers trained on different wavelet scattering features. Pap-smear dataset (top-left). 2D Hela dataset (top-right). FCFM images (bottom). Results of the best performing classifier trained on a single wavelet (Spline) scattering feature is also reported.

Figure 10 reports the results we obtain and compares the performances of these fused classifiers to those obtained with the Spline scattering features. The results are somewhat disappointing since a fused classifier yields to a statistically significantly better performance than the best single feature only for the 2D Hela dataset. For the FCFM dataset of interest, the single Spline scattering features achieves higher performances.

#### 4.7. Fusing classifiers trained on local binary pattern, local quinary pattern and scattering

In this experiment, we conducted the same fusing classifier protocols as mentioned above, but instead the classifiers are trained respectively on rotation invariant LBP, rotation invariant LQP and Spline scattering features. Figure 11 shows more encouraging results. Indeed, for the Pap-smear dataset, the sum rule fused classifier yields to slightly better performances than the combination of LQP and scattering (with p-value ranging from  $10^{-3}$  to 0.13).

The most impressive gain in performance is obtained for the 2D Hela dataset. The sum rule classifier performs better than LQP and scattering with about 3% gain in classification accuracy. For this problem, the median

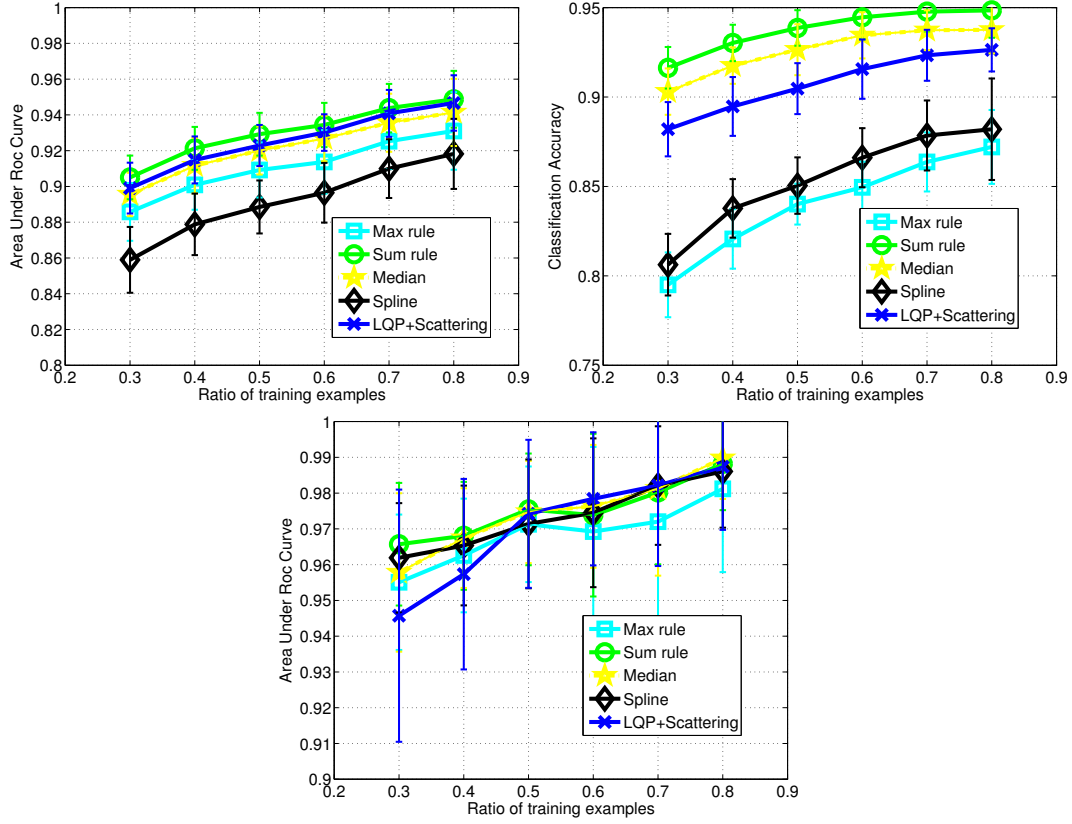


Figure 11: Comparing performances of different fusion rules (maximum, median and sum) of classifiers trained on Spline wavelet scattering feature, LBP and LQP (top-left) Pap-smear dataset. (top-right) 2D Hela dataset. (bottom) FCFM images. Results of the best performing classifier trained on a single wavelet (Spline) scattering feature and trained on LQP+Scattering are also reported.

rule classifier also improves performances. For the *FCFM* dataset, results are less clear. We can conclude that the *sum* and *median* rule classifiers tend to achieve better performances than the Spline scattering and the LQP and scattering features. However, these differences in performance are not statistically significant.

From this experiment, we may conclude that fusing the three classifiers trained on LBP, LQP and scattering features using the *sum* rule seems to yield to consistently slightly better performances than concatenated LQP and scattering features.

#### 4.8. Discussions on local binary patterns (and their variants) and scattering features

The results we have reported above show that LBP and scattering features are both capable of encoding discriminative information for texture classification and it seems that the information they deliver are complementary. Remember that LBP and LQP capture very local information and they operate at the finest scale of the images. Typically, a binary pattern in the LBP encodes the variation of grey level around a given pixel and in these experiments, we have considered a pixel neighbourhood which has at most a 4 pixel

radius. Locality of the grey level variation information is then lost by computing the histogram of the binary patterns. While spatial information is lost, there is the advantage of providing translation invariant features and of compacting the discriminative information, reducing thus the risk of overfitting. Unlike LBP, scattering features are extracted at all scales (from the finest to the coarsest) and thus capture more discriminative information than LBP or LQP, especially if this information is located at coarse scales. Another advantage of scattering features is their robustness to deformation, at the contrary of LBP. However, for scattering representations, these discriminative features have to be estimated from few examples and thus they have to be estimated consistently. This can be done by considering large values of  $J$  at the expense of loss of spatial information [32].

From these discussions, the conclusions we may gather from the differences in LBP and scattering representation are that if the discriminative information in the textures is located in the coarsest scales, then scattering representations should be more efficient than LBP. However, if they are located at the finest scales, then LBP may be an appropriate tool for capturing these discriminative information. These conclusions are in agreement with the results shown in Figure 9. Indeed, according to the image examples provided in Figure 2, 5 and 6, we can consider that for both the *Pap smear* and *FCFM* datasets, discriminative features are at coarse scales whereas for *2D Hela*, they are located at finer scales. In accordance, LBP and LQP features are more discriminative for the latter dataset than for the two former ones. Also from this perspective, the gain in performances obtained when using both features can also be considered likely as they provide complementary information. Indeed, we may conclude that LBP has a better capability of discriminating fine-scale (at pixel level) textures than dyadic wavelets. However, we suggest other practitioners to validate the benefit of combining these features by means of a validation scheme.

#### 4.9. In-depth analysis of performances on fibered confocal fluorescence microscopy dataset

In order to get a better insight on how scattering representations performs on the FCFM dataset of interest, we have run two experiments. The first one provides some detailed performances of the method in a LOPO setting while the second one explores the influence of the scattering representation parameters.

##### 4.9.1. Leave-one-patient-out experimental-setup

Since we use a LOPO setting, performances reported in this subsection do not include standard deviation as each image has been used once in the test set. In order to select the model parameters, a second validation stage of LOPO has been run and hyperparameters have been selected as the best performing one in this stage. A SVM with a Gaussian kernel has been used as a classifier.

Table 1: Influence of the scattering representation parameters on the image correct classification rate. The first row of each cell gives the parameter value. The second one is the *leave-one-patient-out* image recognition rate (RR) in (%) and the last one is the dimensionality *dim* of the problem.

Default parameters  $m = 2$ ,  $J = 7$ ,  $L = 4$

$J :$	4	5	6	7
RR :	68.79	77.46	77.46	<b>80.35</b>
dim :	7232	2896	1060	365
$L :$	2	4	6	8
RR :	74.57	<b>80.35</b>	<b>80.35</b>	<b>80.35</b>
dim :	99	365	799	1401
$m :$	1	2	3	4
RR :	78.03	80.35	78.61	75.14
dim :	29	365	2605	11565

#### 4.9.2. Influence of the scattering transform parameters

Performances for different choices of the scattering transform parameters are reported in Table 1. We observe that most of them lead to performances ranging from 68% to 81%. We can note that a slight decrease of performance appears when  $m = 3$  or  $L = 2$ ,  $L$  being the number of angular rotations in the wavelet transform. For  $m = 3$ , we conjecture that this poor performance, although still valuable, is caused by the curse of dimensionality since more than 2600 features have been generated by the scattering transform. For  $L = 2$ , we imagine that this decrease is due to a low number of rotations, which does not allow for powerful discrimination. For  $m = 4$ , the number of features (up to 11000) makes the the problem harder and the classifier slightly overfits.

The scale of the averaging  $J$  also seems to have an impact on the classifier performances. Discriminative patterns are thus non-spatially localized since a global averaging over the full scale of the images results in equivalent performance than a more localized averaging. It also appears to us that the number of wavelet directions  $L$  has a slight influence on the performances with variations ranging from 74% to 81%.

Interestingly, for this dataset, the best performance is achieved for  $m = 2$ ,  $L = \{4, 6, 8\}$  and  $J = 7$ . Hence, it seems that paths of depth  $m = 2$  are sufficient to extract all relevant discriminative information from the images and that the averaging scale is the most critical parameter.

Note that for this problem, the computational burden needed to classify an image is essentially due to the scattering transform. This latter takes less than 1 second and about ten seconds for respectively  $m = 2$  and  $m = 4$  on a Linux machine with an Intel i5 processor clocked at 2.5 GHz.

Table 2: Performances in (%) on the FCFM dataset in a *leave-one-patient-out* context of the scattering and local binary pattern features. Recognition rate (RR) has been computed with respect to all images (173) while other rates (TPR, TNR, FPR, FNR) have been computed with respect to the number of patients. Scattering parameters are  $J = 7$ ,  $L = 4$ , and  $m = 2$ .

Features	Image RR	At patient level			
	Reco. rate	TPR	TNR	FPR	FNR
LBP	77.46	100.00	62.50	37.50	0.00
Scattering	<b>79.77</b>	100.00	75.00	25.00	0.00
LBP+Scatt.	79.19	100.00	75.00	25.00	0.00

#### 4.9.3. Comparison with local binary pattern

Table 2 provides a performance comparison with LBP in a LOPO setting (we have omitted LQP since they do not perform better than LBP for this dataset). Two kinds of performance measures have been reported. The first one evaluates the number of images correctly classified as presenting normal or abnormal tissues, denoted as *Image RR*. The second performance criterion gives the true positive, true negative, false positive and false negative rate of patients. For this, we have assigned to a patient the label of the majority class obtained from his images. A positive patient refers to a patient considered as presenting cancer lung tissues. We call this classification at “a patient level” to distinguish from the classification at “an image level” which consists of assigning a label to an image of a patient independently to the others. Note that “at a patient level”, we have only 15 examples (patients) to classify.

Results show that in this LOPO context, LBP and scattering coefficients provide similar results with a slight advantage to the scattering transform at the image classification level and at the patient classification level. For the scattering transform ( with  $J = 7$ ,  $L = 4$  and  $m = 2$ ), we can achieve a recognition performance of about 80% at the image level and only two patients are incorrectly diagnosed as having abnormal tissues out of the 15 patients. These are very encouraging results as no false negative has been detected. Figure 12 shows examples of misclassified images of one false-positive patient. Visually, they are difficult to differentiate from the abnormal tissues presented in Figure 2.

Note that for this *LOPO* analysis, combining LBP and scattering features does not lead to enhanced performance. After careful analysis of the results, we found out that this loss in performance is essentially due to one patient for which the combination induces overfitting and poorer performances than using LBP alone. We suspect that such overfitting occurs mainly because of the *LOPO* context in which data from one patient may follow a different distribution than those of the other patients. Such phenomenon can be avoided by either employing machine learning techniques able to handle covariate shift and to perform domain adaptation, or by building a dataset with a larger number of patients.

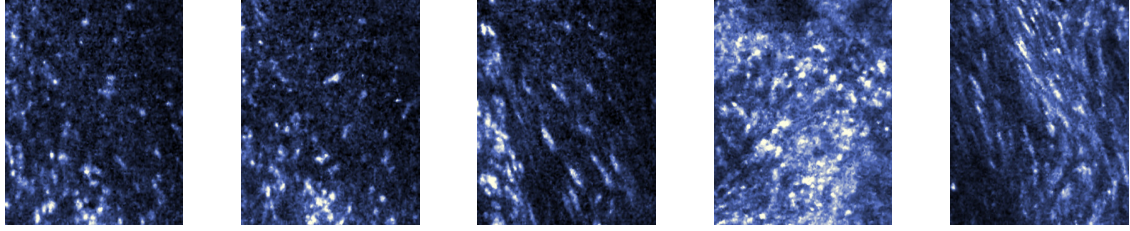


Figure 12: Examples of mis-classified images from one of the false positive patient.

## 5. Conclusions

The objective was to assess the feasibility of lung cancer diagnosis with FCFM imaging techniques. We have shown that when features obtained from scattering transform are extracted from these images, it is possible to learn a classifier able to achieve image recognition rate score as high as 80% in a *Leave-One-Patient-Out* setting. Similar performances, although slightly less successful, can also be obtained using classical feature extraction techniques such as Local Binary Patterns. Interestingly, scattering transform based feature leads to classifiers which achieve no false negative error. These findings urge us to build a larger dataset involving more patients in the protocol so as to confirm this promising result.

Analysis on other medical imaging datasets also corroborate the fact that scattering features are competitive discriminative features for texture classification problems and that they should be integrated in the set of tools to be considered when addressing such problems. We also conclude that LBP, LQP and scattering features seem, in many situations, to provide complementary discriminative information to the classifiers, resulting in enhanced performances when used together either by concatenation or by classifier fusion.

Future works essentially focus on possible refinements of FCFM images for detecting lung cancer. We plan to investigate whether it is possible to correctly recognize the different grades of progression of the pre-cancer bronchial lesions such as metaplasia, dysplasia and carcinoma in situ. Solving this challenging problem naturally raises the problem of the discriminative features to be considered.

Another important point to address is raised by the protocol used for acquiring FCFM images. Because of the lack of control of the miniprobe, its positioning also affects the orientation and the positioning of the image, and thus introduces rotation, scale and shear variabilities among the images. Hence, we believe that it is of primary interest to go beyond features that are translation-invariant but that are also invariant to several other transformations. One possible direction to investigate is a specific and adapted scattering transform like those theoretically introduced in [30].

In conclusion, we want to state that FCFM imaging is still a recent technology compared to the more established CT imaging. Hence, its full potential is yet to be explored and many advances will still need to



be made before it can be considered a gold standard for lung examination. Regarding its clinical use, over more than 200 FCFM examinations have been performed in the Rouen University Hospital, and none of them has induced serious complications due to the in-vivo examination. Regarding its practicability, the FCFM procedure adds ten minutes to a standard endoscopy and the probe cost, which is usable for 20 examinations, is about 6000 euros. Taking into account all these parameters as well as the diagnostic capabilities of FCFM imaging help us to understand the progress that needs to be made before the FCFM procedure can become a gold standard. This work contributes to such efforts as it validates the automated diagnosis feasibility of lung cancer from FCFM images.

## Acknowledgments

This work was partly supported by grants from ASAP ANR-09-EMER-001.

## References

- [1] F. Hirsch, W. Franklin, A. Gazdar, P. Bunn, Early detection of lung cancer : Clinical perspectives of recent advances in biology and radiology, *Clinical Cancer Research* 7 (2001) 5–22.
- [2] L. Fass, Imaging and cancer : a review, *Molecular oncology* 2 (2008) 115–152.
- [3] E. Patz, Imaging lung cancer, *Semin. Oncology* 26 (1999) 21–26.
- [4] D. D. Boo, M. Prokop, M. Uffmann, B. van Ginneken, C. Schaefer-Prokop, Computer-aided detection (cad) of lung nodules and small tumours on chest radiographs, *European Journal of Radiology* 72 (2) (2009) 218–225.
- [5] I. Sluimer, A. Schilham, M. Prokop, B. van Ginneken, Computer analysis of computed tomography scans of the lung: a survey, *IEEE Trans. Medical Imaging* 25 (4) (2006) 385–405.
- [6] N. Ikeda, A. Hayashi, K. Iwasaki, H. Honda, M. Tsuboi, J. Usuda et al. Comprehensive diagnostic bronchoscopy of central type early stage lung cancer, *Lung Cancer* 56 (3) (2007) 295 – 302.
- [7] L. Thiberville, M. Salaün, S. Lachkar, S. Dominique, S. Moreno-Swirc, C. Vever-Bizet et al., Confocal fluorescence endomicroscopy of the human airways, *Proceeding American Thoracic Society* 6 (5) (2009) 444–449.
- [8] L. Thiberville, M. Salaün, Bronchoscopic advances: on the way to the cells., *Respiration* 79 (6) (2010) 441–449.

- [9] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, N. Ayache, A smart atlas for endomicroscopy using automated video retrieval, *Medical Image Analysis* 15 (4) (2011) 460–476.
- [10] K. Suzuki, H. Abe, H. MacMahon, K. Doi, Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network, *IEEE Trans. Medical Imaging* 25 (4) (2006) 406–416.
- [11] P. Campadelli, E. Casiraghi, D. Artioli, A fully automated method for lung nodule detection from postero-anterior chest radiographs, *IEEE Trans. Medical Imaging* 25 (12) (2006) 1588–1603.
- [12] X. Ye, X. Lin, J. Dehmshki, G. Slabaugh, G. Beddoe, Shape-based computer-aided detection of lung nodules in thoracic ct images, *IEEE Trans. Biomedical Engineering* 56 (7) (2009) 1810–1820.
- [13] M. Brown, M. McNitt-Gray, J. Goldin, R. Suh, J. Sayre, D. Aberle, Patient-specific models for lung nodule detection and surveillance in ct images, *IEEE Trans. on Medical Imaging* 20 (12) (2001) 1242–1250.
- [14] U. Bagci, J. Yao, A. Wu, J. Caban, T. Palmore, A. Suffredini et al., Automatic detection and quantification of tree-in-bud (tib) opacities from CT scans, *IEEE Transactions on Biomedical Engineering* 59 (6) (2012) 1620–1632.
- [15] M. Lee, L. Boroczky, K. Sungur-Stasik, A. Cann, A. Borczuk, S. Kawut et al., Computer-aided diagnosis of pulmonary nodules using a two-step approach for feature selection and classifier ensemble construction, *Artificial Intelligence in Medicine* 50 (1) (2010) 43–53.
- [16] T. Way, B. Sahiner, H. Chan, L. Hadjiisjki, P. Cascade, A. Chughtai et al., Computer-aided diagnosis of pulmonary nodules on ct scans: improvement of classification performance with nodule surface features., *Medical Physics* 36 (7) (2009) 3086–3098.
- [17] A. Depeursinge, D. V. de Ville, A. Platon, A. Geissbuhler, P. Poletti, H. Muller, Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames., *IEEE Trans. Inf. Technology Biomedical* 16 (4) (2012) 665–675.
- [18] A. Depeursinge, A. Foncubierta-Rodriguez, D. V. de Ville, H. Muller, Lung texture classification using locally-oriented riesz components, in: T. Peters, G. Fichtinger, A. Martel (Eds.), *Proceedings of MICCAI 2011*, Springer (New-York), 2011, pp. 231–238.
- [19] L. Sorensen, S. Shaker, M. de Bruijne, Texture classification in lung CT using local binary pattern, in: D. Metaxas, L. Axel, G. F. ang G. Szekely (Eds.), *Proceedings of MICCAI 2008, Part I*, LNCS 5241, Springer (New-York), 2008, pp. 934–941.

- [20] T. Ojala, M. Pietikäinen, T. Mäenpää, Gray scale and rotation invariant texture classification with local binary patterns, in: D. Vernon (Ed.), *Proceedings of the European conference on Computer Vision*, Springer-Verlag, (Berlin), 2000, pp. 404–420.
- [21] L. Nanni, A. Lumini, S. Brahmam, Local binary patterns variants as texture descriptors for medical image analysis, *Artificial Intelligence in Medicine* 49 (2010) 117–125.
- [22] C. Désir, C. Petitjean, L. Heutte, L. Thiberville, M. Salaün, An svm-based distal lung image classification using texture descriptors, *Computer Medical Imaging Graphics* 36 (4) (2012) 264–270.
- [23] C. Désir, C. Petitjean, L. Heutte, M. Salaün, L. Thiberville, Classification of endomicroscopic images of the lung based on random subwindows and extra-trees, *IEEE Transactions on Biomedical Engineering* 59 (9) (2012) 2677–2683.
- [24] M. Paci, L. Nanni, A. Lahti, K. Aalto-Setälä, J. Hyttinen, S. Severi, Non-binary coding for texture descriptors in sub-cellular and stem cell image classification, *Current Bioinformatics* 8 (2) (2013) 208–219.
- [25] R. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, *IEEE Transactions on Systems, Man and Cybernetics* 6 (1973) 610–621.
- [26] H. Wu, T. Sun, J. Wang, X. Li, W. Wang, D. Huo et al., Combination of radiological and gray level co-occurrence matrix textural features used to distinguish solitary pulmonary nodules by computed tomography, *Journal of Digital Imaging* 26 (4) (2013) 797–802.
- [27] A. Pires, H. Rusinek, J. Suh, D. Naidich, H. Pass, J. Ko, Clustering of lung adenocarcinomas classes using automated texture analysis on ct images, in: S. Ourselin, D. Haynor (Eds.), *Proceedings of the SPIE Medical Imaging: Image Processing*, SPIE (Bellingham, Washington), 2013, pp. 25–32.
- [28] J. Bruna, S. Mallat, Classification with scattering operators, in: P. Felzenszwalb, D. Forsyth, P. Fua (Eds.), *Proceedings of the IEEE CVPR 2011*, IEEE (New-York), 2011, pp. 1561–1566.
- [29] L. Sifre, S. Mallat, Rotation, scaling and deformation invariant scattering for texture discrimination, in: M. Hebert, B. Freeman, R. Szeliski, G. Hager (Eds.), *Proceedings of the IEEE Computer Vision and Pattern Recognition Conference*, IEEE (New-York), 2013, pp. 1233–1240.
- [30] S. Mallat, Group invariant scattering, *Communications in Pure and Applied Mathematics* 65 (10) (2012) 1331–1398.
- [31] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient based learning applied to document recognition, *Proceedings of IEEE* 86 (11) (1998) 2278–2324.

- [32] J. Bruna, S. Mallat, Invariant scattering convolution networks, *IEEE Trans. in Pattern Analysis and Machine Intelligence* 35 (2013) 1872–1886.
- [33] M. Heikkilä, M. Pietikäinen, C. Schmid, Description of interest regions with local binary patterns, *Pattern Recognition* 42 (3) (2009) 425–436.
- [34] B. Schölkopf, A. Smola, *Learning with Kernels*, MIT Press (Cambridge, MA), 2001.
- [35] R. Murphy, M. Velliste, G. Porreca, Robust numerical features for description and classification of subcellular location patterns in fluorescence microscope images, *Journal of VLSI Signal Processing* 35 (2003) 311–321.
- [36] Y. Marinakis, G. Dounias, J. Jantzen, Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification, *Computers in Biology and Medicine* 39 (2009) 69–78.
- [37] N. Ampazis, G. Dounias, J. Jantzen, Efficient second-order neural network training algorithms for the construction of a pap-smear classifier, in: G. Vouros, V. Panayiotopoulos (Eds.), *Proceedings of the third Hellenic conference on AI (LNAI 3025)*, Vol. 3025, 2004, pp. 230–245.
- [38] A. Chebira, Y. Barbotin, C. Jackson, T. Merryman, G. Srinivasa, R. F. Murphy et al. A multiresolution approach to automated classification of protein subcellular location images, *BMC bioinformatics* 8 (1) (2007) 210.