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-00820354v1

HAL Id: hal-00820354
https://hal.archives-ouvertes.fr/hal-00820354v1
Submitted on 3 May 2013 (v1), last revised 11 Jun 2014 (v3)
Abstract—Fibered confocal fluorescence microscopy (FCFM) imaging technique is a novel medical imaging technique which interest is yet to be establish for diagnosis problem. 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. For doing so, we have built a pattern recognition scheme which involves a feature extraction and a classification stages. The second contribution relies on the used features for discrimination. Indeed, we have employed the so-called scattering transform for extracting discriminative features robust to small deformations in the images. We have shown that these features yield to better recognition performances than classical yet powerful features like Local Binary Patterns (LBP) for our FCFM image classification problems and are competitive to LBP on other medical imaging classification problems. Another of our findings is that LBP and scattering-based features provides complementary discriminative informations and in some situations, we have empirically established that enhanced performances can be obtained when jointly using LBP and scattering features.

Keywords : FCFM imaging, bronchoscopy, wavelet, scattering transform, textures.

I. INTRODUCTION

Lung cancer is the one of the the most wide-spread form of cancer and unfortunately, it is also one of the major cause of premature decease. Survival to lung cancer is highly correlated with its early diagnosis. From this observation, it is thus of primary importance to develop methods for helping in its diagnostic. Numerous imaging methods have been developed in order to improve the detection of these early cancerization stages [13]. Medical imaging tools for lung cancer diagnosis are essentially based on inspecions of the lung and the chest [24] using X-ray-based systems including computed tomography [3], [26]. For centrally located lung cancer, bronchoscopy is an essential tool for its diagnosis. Until recently, the early diagnosis of lung cancer relied primarily on gross inspection of the bronchial mucosae during an endoscopic procedure and 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 (CIS) [15]. 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) [30], [29].

The FCFM technique (also refered to as probe-based Confocal Laser Endomicroscopy (pCLE), especially in gastrointesti- nal 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, could replace lung biopsy in the future and might prove to be helpful in a large variety of diseases [30]. FCFM has other challenging applications, such as the imaging of colonic polyps in gastroenterology, for which image retrieval and classification works are ongoing [2].

FCFM imaging is thus able to provide in vivo cellular images of the lung and for this reason, it opens the road to novel methods for early diagnosis of lung diseases including cancer. This is one of the objective of this paper: 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 [28], [7] or in computed tomography (CT) [32], [4]. CT is often coupled with advanced numerical techniques for extracting relevant features from the images and as well as with state of the art classifiers [13], [31]. Usually, the systems are based upon a two-stage approach: feature extraction and classifier training [7], [9]. 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. Frequently used features for classification are related to texture of the images [9], [10], [27], described in particular by means of the local binary patterns (LBP) which are known to be very competitive texture feature extractor [22], [12]. 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 a mix between wavelet transform and convolutional neural networks [16].
Indeed, in such networks, layers are successively cascaded through a convolution filter while the scattering transform is obtained from a cascade of wavelet transform, thus a linear filtering, where the input of each cascade stage is the previous stage output’s modulus. While Gabor wavelet transform and convolutional neural networks are known to provide relevant texture features, scattering transform is more robust to deformations and thus it is expected to be more effective in yielding relevant texture features.

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 bring are three-fold:

- at first, we assess the feasibility of lung cancer diagnosis from FCFM images. FCFM imaging is a rather new imaging technique and its importance and applicabilities are yet to be explored. In this paper, we show that lung cancer can be carefully detected.
- then, we show that scattering representations are highly discriminative features for this FCFM-based lung cancer classification problem. When compared to state of the art features like Local Binary Pattern (LBP) features, they achieve strong competitive results as they yield to the best discrimination performance. Yet, we also empirically prove that coupling LBP and scattering features in this classification problem results in enhanced performances.
- in order to have a better insight on the discriminative power of scattering representation, we have also analyzed its behaviour on two other medical image classification problems.

The paper is organized as follows. Section II briefly presents the FCFM imaging technique as well as the image dataset we are working with. A discussion on the scattering transform is given in Section III. This sections also briefly recall how LBP features are constructed as well as the classifier that is used for learning a decision function. Experimental analyses are reported in Section IV.

II. THE DATA

The technique of in vivo bronchial FCFM imaging consists in the introduction of a miniprobe, made of thousands of fiber cores, into the 2-mm working channel of the bronchoscope. The probe tip is applied onto the bronchial wall. However, the bronchial epithelial layer is significantly altered with cancer. But 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/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.

Fig. 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.

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. Hence, our aim in this paper is to provide the clinician with a computer aided-diagnosis (CAD) tool, so as to help him to analyze these new images and to help in detecting cancer lesions.

The FCFM image dataset is composed of 103 images acquired on 8 healthy volunteers, and 70 images acquired on 7 patients with diagnosed bronchial squamous cell carcinoma. Image sizes originally vary from 190 × 190 pixels to 288 × 288 pixels, but have been resized to 128 × 128 for a more efficient application of the scattering transform.

III. METHODS

A. Scattering representation of images

Scattering transform has been recently introduced by Bruna et al. [5] 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 a convolutional neural networks as it cascades several layers of filtering [17]. Let us now provide a more formal definition of the scattering representation of an image. More details can be found in [5], [6].

We denote as $\psi_{j,\gamma}(x)$ a directional wavelet at a scale $j$ defined as

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

where $R_{\gamma}$ is a rotation matrix of angle $\gamma$ and $\psi$ a one-dimensional wavelet. Note that negative and positive values of $j$ respectively correspond to dilation and contraction of the mother wavelet $\psi$. For a sake of simplicity, we will gather the wavelet parameter $j$ and $\gamma$ in a single parameter $\lambda$. Let also define $\phi_{2^j}(x) = 2^{-2j}\phi(2^{-j}x)$, $J \geq 0$ 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 by computing high-order wavelet coefficients by...
Fig. 2. Diagram of a scattering transform of an image $f$. For this example, we have supposed that $J = 2$ and the 2D 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, eventually after a subsampling, to form the scattering transform at a depth $m = 0$, $m = 1$ and $m = 2$. Note that because only paths with frequency-decreasing wavelet are of interest, at $m = 2$, only 4 paths out of the 16 will be finally kept.

Fig. 3. 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 and (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 depicts, 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 $m = 2$.

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

$$U_J f(x) = \{ f \ast \phi_{2^j}(x); \{ | f \ast \psi_\lambda(x) | \}_{\lambda \in \Lambda_J} \}_{j \in \mathbb{N}}$$

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

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

$$\{ | f \ast \psi_{\lambda_1} | \ast \phi_{2^j}(x); \{ || f \ast \psi_\lambda | \ast \psi_\lambda(x) || \}_{\lambda \in \Lambda_J} \}_{j \in \mathbb{N}}$$

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 this operator $U_J$.

Formally, suppose we are looking for the scattering co-
coefficients of order \( m \) at a scale \( 2^J \) where \( m \) denotes the order of the cascade and we define as \( p \) a sequence \( p = \{\lambda_1, \lambda_2, \cdots, \lambda_m\} \) 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. This path \( p \) defines an unique wavelet modulus cascade output

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

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

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

with and \( S_J[\emptyset]f = f \ast \phi_{2^J} \). 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 over all the image of \( U[p]f \).

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

\[
S_J[\Lambda^m_j]f = \{S_J[p]f\}_{p \in \Lambda^m_j}
\]

which also encompasses the scattering coefficients of \( f \) at lower order.

Another insight on scattering transform can be gained by exhibiting explicitly the cascading scheme involving \( U_J \). Let us first remark that \( U_Jf \) can be rewritten as

\[
U_Jf = \{S[\emptyset]f, U[\Lambda^1_j]f\}
\]

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

\[
U_JU[\Lambda^1_j]f = \{U_JU[p]f\}_{p \in \Lambda^1_j} = \{S_J[\Lambda^1_j]f, U[\Lambda^2_j]f\}
\]

Hence, by iteratively applying this scheme to sequences \( p \) of length \( m^J \leq m \), all the scattering coefficients at order \( m \) can be retrieved. Figure 2 gives a schematic representation of this scheme, for which an efficient algorithm which can be found in [6].

We can also remark from this diagram that the number of resulting coefficients can rapidly grow with respects to \( m \). However, because some paths \( p \) lead to negligible coefficients (those of increasing-frequency), they can be omitted while computing the transform [6]. 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 \( N^{2^J} \sum_{i=0}^{m} L_i \binom{J}{i} \), \( L \) being the number of rotations applied to the wavelet, where the sum counts the number of paths with decreasing frequency and the term \( N^{2^J} \) 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. For illustrating the number of paths, in the example of Figure 2 we have 4 wavelets with 2 different frequencies \( (f_1, f_2) \) and 2 different rotations. At \( m = 1 \), the number of paths with decreasing frequencies is 4 (actually, all the possible paths). At \( m = 2 \), we again have 4 interesting paths, only the two of the form \( \{f_2, f_1\} \) times the two angle of rotations.

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

- its stability to small deformations. This means that if a small deformation, either spatial or in the intensity of a pixel, is applied to an image then it is expected that the scattering representation of that deformed image does not differ much from the one of the original image. 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.
- the modulus operator and the averaging step brings local translation invariance to the representation. This local invariance acts at the level of the scaling \( 2^J \). Hence, if \( \phi_{2^J} \) spans 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 [6] that the norm of the scattering coefficients is equal to the one of the original signal and that most signal energies can be retrieved by considering scattering transform of depth 3.
- scattering transform also bears an interesting property related to textures that can be considered as stationary process. Indeed, Bruna et al. have proved that scattering coefficients brings 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 a sake of representation, Bruna and Mallat introduced a polar representation of scattering representations that can handle depth up to \( m = 2 \). As shown in Figure 3 the plot uses a gray-scale for representing coefficient amplitude in regions related to scale and rotation of the wavelets. We remark in this figure that the scattering representations discriminate pretty well the two FCFM lung image examples so that we can expect a good discriminative power of the resulting features.

**B. Local Binary Patterns**

Since our objective is to assess whether scattering coefficients can be useful for medical imaging texture classification problems, we have compared their discriminative power with the ones of a classical feature for texture, namely the Local Binary Patterns (LPB). These LPB have been popularized by Ojala et a. [23] and since then they have been extensively used
for building features for classification problems [22]. In this work, we have used them as a baseline feature comparison.

LBP is a texture operator that extracts some local information around the neighborhood of a given pixel. In the original LBP formulation, the one that we are considering here, this neighborhood is defined as circle of radius \( R \) around that pixel. The LBP operator assigns to that center pixel a label obtained by comparing the gray value of each neighbour with the one of the center. The binary number resulting from the concatenation of all 0 or 1 output by the comparison forms the the label of the center pixel. The LBP texture feature is obtained as the histogram of all these labels assigned to the image pixels. We will not go into deeper details about LBP features as they can be obtained in these references [23], [22].

C. 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 at our disposal \( \{\tilde{x}_i, y_i\}_{i=1}^n \) images and their associated labels for learning a decision function \( f(\cdot) \) that predicts the label of an novel image \( \tilde{x} \). For this purpose, a feature vector \( x_i \) is computed from each image \( \tilde{x}_i \), 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 [25]. It is defined as

\[
f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + b
\]

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

\[
\begin{align*}
\max & \quad \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\
\text{st} & \quad \sum_{i} \alpha_i y_i = 0 \\
& \quad 0 \leq \alpha_i \leq C, \quad \forall i
\end{align*}
\]

\( C \) 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.

IV. EXPERIMENTAL RESULTS

The aims of the experimental results we present in this section are two-fold. At first, we want to empirically evaluate the efficiency of scattering operator based features on some 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.

A. 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 [23] in addition to our FCFM images.

The 2D-Hela dataset is a cell phenotype multi-class classification problem from images [21]. It consists in a 10-class problems, each class being composed of a number of images ranging from 73 to 98 with a total of 862 images. For this problem, images have been resized to 382 \( \times \) 382. Some examples of images are presented in Figure 4.

The Pap smear dataset aims at detecting pre-malignant cells extracted from the uterine cervix [20], [1]. The original dataset provides 7 different classes which can be grouped into a normal and abnormal classes. In our work, we have turned the problem into a binary classification problem and the normal and abnormal classes are composed of respectively 242 and 675 images of size 128 \( \times \) 128. Some examples of cell images are presented in Figure 5.

We have focused our interest on these datasets as they serve as reference datasets on texture-based image classification problems [22].

B. Experimental set-up

Scattering representations of images have been obtained by means of the Matlab toolbox provided by Bruna et al.. The chosen wavelet is a Morlet wavelet [19]. 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 as 4 and 2. The scale \( J \) is adapted according to image size and will be reported for each experimental analysis.

http://www.cmap.polytechnique.fr/scattering/

http://labs.fme.aegean.gr/decision/downloads
For extracting LBP features, we have also employed a Matlab code available online. In order to enrich the LBP representation, we have concatenated the normalized histogram obtained from neighborhoods of radius respectively 1, 2 and 4 with a sampling of 8, 16 and 24 pixels on the circle. Resulting LBP features is thus of size 54.

Once the features have been obtained, they are fed to a classifier which in our case is a Gaussian kernel Support Vector Machine. For evaluating the couple feature-classifier, we have run 30 trials where each trial consists in 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. Since the SVM classifier has some hyperparameters, they have been tuned by means of a validation step which again randomly splits the training examples 50%−50% in a learning and validation sets. The parameters are the kernel bandwidth \( \sigma \) and the SVM slack parameter \( C \). They are tuned respectively in the range \([0.01, 0.1, \ldots, 1000]\) and \([1, 2, 5, 7.5, 10, 12.5, 15, 25, 30, 35]\). This random split is performed 5 times and the best averaged result defines the best SVM hyperparameters. Finally, a SVM with those 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.

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

C. Feature performance comparisons

This first experiment aims at comparing the discriminative power of LBP and scattering representations. Furthermore, we have also combined these two types of features and evaluated their joint discriminative power.

Parameters for the scattering representations have been set as \( m = 2 \) and \( L = 4 \). These values have been chosen by default but they are expected to already provide good classification accuracy since with paths of that depth, the scattering transform captures most of the image energy. The scaling factor \( J \) is chosen so as to fit the entire image, which means that from each sequence \( p \), only a single coefficient is obtained, resulting in a full translation invariance at the expense of a loss of spatial resolution.

We have compared LBP and scattering features for increasing ratio of training points. Averaged performances over 30 random split of the examples are reported in Figure 6, 7 and 8.

Performances on the Pap smear dataset are reported in Figure 6. We can note that scattering features lead to slightly better performance than LBP features, but these differences in performance are not statistically significant. Combining both features yields in this case to an enhanced performance of about 4% of AUC (p-value \( \approx 10^{-5} \)) compared to the use of scattering features.

For the 2D Hela dataset (see Figure 7), LBP performs slightly better than scattering features with statistically significant difference up to a ratio of training examples of 0.6. For this problem too, feature combination drastically increases performance (p-value \( \approx 10^{-3} \)). As far as we know, the best reported performance on this dataset reaches 95.3% of recognition rate. While this latter work uses a slightly different protocol than ours, making results hardly comparable, they achieve this state of the art result by considering a classical texture feature and a multiresolution approach. We thus believe that integrating our features into that multiresolution framework may also improved classification accuracy.

For our dataset of interest, we can first note that the recognition problem is easier than the two previous ones since both LBP and scattering coefficients achieved AUC above 0.93. When 80% of the dataset is used as a training examples, LBP yields to an AUC of 0.97. Interestingly, the scattering coefficients still perform slightly better than LBP although with no statistical differences. Although performances are already high, gains can be achieved again by combining LBP and scattering features and in this case these good performances are significantly better (p-value=0.02) than the single use of scattering features.

These are encouraging results since the scattering representation parameters have not been tuned to each dataset as they have been selected by default. In addition, the LBP features we used are somehow enriched as several features resulting from different parametrizations of the feature extractor have been concatenated. These results, while far from being exhaustive, make us believe that the scattering representations achieve discriminative power at least similar to state-of-the-art methods of LBP.

Interestingly, the fusion of the two types of features always leads to enhanced performances for the three problems we consider. This suggests that these features bring complementary information that are essential for classification. Understanding how these two features interact is of primary interest that we postpone for future works.

D. In-depth analysis of performances on FCFM dataset

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

1) LOPO experimental-setup: Since we consider a LOPO setting, performances reported in this subsection do not include standard deviation as each image has been used once in the test set. For selecting the model parameters, a second-stage of

\[ \text{http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab} \]
LOPO has been run and hyperparameters have been selected as the best performing one in that validation LOPO stage. A SVM with a Gaussian kernel has been used as a classifier.

2) Influence of the scattering transform parameters: Performances for different choices of the scattering transform parameters are reported in Table I. We observe that most of them lead to performances ranging in between 59% and 81%. We can note that a slight decrease of performance appears when \( m = 3 \) or \( L = 2 \), \( L \) being the number of angular rotation in the wavelet transform. For \( m = 3 \), we conjecture that this poor performance although still valuable due to the curse of dimensionality since more than 4000 features have been generated by the scattering transform. For \( L = 2 \), we imagine that this decrease is due to a too low number of rotations, which does not allow for a powerful discrimination. For \( m = 4 \), the number of features (up to 21000) make the problem harder and our classifier clearly overfits.

The scale of the averaging \( J \) seems to have only a slight impact on the classifier performances. Discriminative patterns are thus not spatially localized since a global averaging over the full scale of the images results in equivalent performance than a more localized averaging. It appears to us that the number of wavelet direction \( L \) has the largest influence on the performances with performances ranging from 76% to 81%.

Interestingly, for this dataset, the best performance is achieved for \( m = 2, L = 6 \) and \( J = 8 \). Hence, it seems that paths of depth \( m = 2 \) are sufficient for extracting all relevant discriminative informations from the images and the number of directional wavelet can be a critical parameter.

Note that for this problem, the computational burden needed for classifying 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.

3) Comparison with LBP: Table II provides a performance comparison with LBP in a LOPO setting and with scattering features using default values. Two kinds of performance measures have been reported. The first one evaluates the number of images correctly classified as obtained from patients 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 supposed that a patient has been assigned the label of the majority labels obtained from its images. A positive patient refers to a patient considered as presenting cancer lung tissues.

![Fig. 6. Pap-smear dataset. AUC with respects to the ratio of training examples. The scattering representation parameters are \( J = 7, M = 2, L = 4 \).](image1)

![Fig. 7. 2D Hela dataset. Accuracy with respects to the ratio of training examples. The scattering representation parameters are \( J = 9, m = 2, L = 4 \).](image2)

![Fig. 8. FCFM dataset. AUC with respects to the ratio of training examples. The scattering representation parameters are \( J = 7, M = 2, L = 4 \).](image3)

<table>
<thead>
<tr>
<th>( J )</th>
<th>( M )</th>
<th>( L )</th>
<th>RR</th>
<th>( \text{dim} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>( L )</td>
<td>( 2 )</td>
<td>( 4 )</td>
<td>( 6 )</td>
<td>( 8 )</td>
</tr>
<tr>
<td>( \text{RR} )</td>
<td>77.46</td>
<td>79.77</td>
<td>79.19</td>
<td>78.61</td>
</tr>
<tr>
<td>( \text{dim} )</td>
<td>2896</td>
<td>1060</td>
<td>365</td>
<td>481</td>
</tr>
<tr>
<td>( m )</td>
<td>( 1 )</td>
<td>( 2 )</td>
<td>( 3 )</td>
<td>( 4 )</td>
</tr>
<tr>
<td>( \text{RR} )</td>
<td>76.88</td>
<td>78.61</td>
<td>75.14</td>
<td>59.54</td>
</tr>
<tr>
<td>( \text{dim} )</td>
<td>33</td>
<td>481</td>
<td>4065</td>
<td>21985</td>
</tr>
</tbody>
</table>
Results show that in this LOPO context, when default values are used, LBP and scattering representations provide similar results with a slight advantage to the scattering transform at the image classification level. At the patient classification level however, both methods give exactly the same performance measures. When optimal parameters of the scattering transform are used (J = 8, L = 6 and m = 2), we can achieve a recognition performance about 80% at the image level and only two patients are incorrectly diagnosed as having abnormal tissues out the 15 patients. These are very encouraging results as no false negative has been detected. Note that for this LOPO analysis, combining LBP and scattering features do not lead to enhance 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 fixed by considering machine learning techniques able to handle covariate shift and perform domain adaptation or by building a dataset with a larger number of patients.

V. CONCLUSIONS

Our objective was to assess the feasibility of lung cancer diagnosis with FCFM imaging techniques. We showed that when features obtained from scattering transform are extracted from these images, then 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 lower, can also be obtained using classical feature extraction techniques like Local Binary Pattern. Interestingly, using scattering transform based feature leads to classifiers which achieve no false negative error. This findings urge us to the building of a larger dataset involving more patients in the protocol.

Analyses on other medical imaging datasets also corroborate the fact that scattering features are competitive discriminative features for texture classification problems and they should be integrated in the set of tools to be considered when addressing such problems. Another of our conclusions is that LBP and scattering seem, in many situations, to provide complementary discriminative informations to the classifiers, resulting in enhanced performances when both features are mixed together.

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/carcinoma in situ. Solving this challenging problem naturally poses the problem of the discriminative features to be considered.

ACKNOWLEDGMENTS

This work was supported in part by the IST Program of the European Community, under the PASCAL2 Network of Excellence, IST-216886. This publication only reflects the authors views. This work was partly supported by grants from ASAP ANR-09-EMER-001.

REFERENCES


