How to Automatically Identify Regions of Interest in High-resolution Images of Lung Biopsy for Interstitial Fibrosis Diagnosis

1st Oscar Cuadros Linares
Institute of Mathematics and Computer Sciences
University of São Paulo
São Carlos, Brazil
ocuadros@icmc.usp.br

2nd Bruno S. Faiçal
Institute of Mathematics and Computer Sciences
University of São Paulo
São Carlos, Brazil
bsfaical@alumni.usp.br

3rd Paulo Barbosa
Institute of Mathematics and Computer Sciences
University of São Paulo
São Carlos, Brazil
paulorcb@usp.br

5th Bernd Hamann
Department of Computer Science
University of California, Davis
Davis, California, U.S.A.
hamann@cs.ucdavis.edu

4th Alexandre T. Fabro
Ribeirão Preto Medical School
University of São Paulo
Ribeirão Preto, Brazil
fabro@fmrp.usp.br

6th Agma J. M. Traina
Institute of Mathematics and Computer Sciences
University of São Paulo
São Carlos, Brazil
agma@icmc.usp.br

***Abstract*—Airway-Centered Interstitial Fibrosis (ACIF) is a described histological pattern of Interstitial Lung Diseases. Its diagnosis must require a multidisciplinary approach, in which diverse information such as clinical data, Computed Tomography, and lung biopsy, is analyzed. Biopsy samples are digitized at a high-resolution file, in which the key finding is broncho- and bronchiolocentric remodeling with extracellular matrix deposition. To analyze such an image, specialists have to analyze it in a low microscope magnification, select a region of interest and export it as a "smaller" sub-image to be analyzed in a higher magnification. This process is performed several times and it requires hours, becoming a tiresome task. In this work, we propose a method aimed at supporting pathologists to identify specific patterns of ACIF in high-resolution images from lung biopsies. This can be done by a) automatic microscope magnification reduction, b) computing the probability of pixels belonging to high-density regions, c) extracting Local Binary Patterns (LBP) of the high- and low-density regions, and d) visualizing them as colormaps. We evaluated our method on nine high-resolution lung biopsies. The LBP features of high- and low-density regions were tested with the kNN algorithm obtaining a classification accuracy of 94.4%, which is the highest one reached so far in the literature.**

***Index Terms*—Airway-Centered Interstitial Fibrosis, High-resolution Lung Biopsy, Image analysis, Quantitative assessment of microscopic images, Interstitial Fibrosis Diagnosis.**

I. Introduction

Pulmonary fibrosis is a heterogeneous group of interstitial lung diseases in which lung tissue becomes thick, stiff and scarred over time. Specifically, the name of fibrosis refers to the formation of those scars. The most serious consequence is that it progressively hinders the proper functioning of the lungs, leading to shorter of breath and hypoxia [1]. So far, these diseases could not be cured and their severity and outcome vary by etiology and person, often with poor prognosis.

Several disorders can be included in the category of the Interstitial Lung Diseases (ILD). Nevertheless, among the most complex to classify and diagnose are the different types of lung fibrosis, which describe a collection of more than 200 disorders, all of them affecting the Interstitium [2]. This complexity derives from the fact that various disorders could be categorized in different ways: diffuse or localized, predominant fibrosis or inflammatory, and other various criteria [3].

In this context, Airway-centered Interstitial Fibrosis (ACIF) is a recently (2002) described histological pattern of ILD with variable etiology and outcome, which can be idiopathic or with a history of environmental exposure, such as smoke, birds, chalk dust, agrochemicals, and cocaine [4]. It is mainly characterized by broncovascular remodeling and metaplastic bronchiolar. Clinically patients present cough and progressive dyspnea. ACIF is a disease with a variable prognosis of 40% mean survival in 10 years [5].

To diagnose ACIF, a multidisciplinary approach is required, in which are analyzed in three progressive steps: a) clinical data such as patients age, symptoms, and environmental exposure; b) radiological imaging, X-ray and Computed Tomography (CT). Typical signs of ACIF are often present on High-resolution CT, which can include bronchiectasis, airway-centered reticular infiltrates, ground-glass opacity, and parabronchial interstitial thickening [4]; c) histopathology, usually a surgical lung biopsy is indicated when CT patterns are indeterminate or inconsistent to diagnose of Idiopathic Pulmonary Fibrosis (IPF) or when clinical data suggest an

The authors would like to thank FAPESP for the financial support to the project "Mining, indexing and visualizing Big Data in clinical decision support systems (MIVisBD)" (Grants 16/17078-0, 18/06074-0, and 18/06228-7). The authors also thank CNPq for the financial support grant 144908/2018-2.

alternative diagnostic [6]. Biopsy samples are analyzed at low and high microscope magnification. In ACIF the key histopathological finding is the airway interstitial fibrosis centered on membranous and respiratory bronchioles [5]. At higher magnification, patterns, such as chronic inflammatory infiltrates, and muscle hyperplasia in the wall of bronchioles, can be found.

Advances in medical technology, e.g., X-ray or CT, have allowed specialists to improve their diagnoses and, consequently, patients receive early treatments. Despite this, identify and categorize ILDs is still a very complex, time demanding and tiresome task, since a pathologist could take many hours analyzing a lung biopsy sample. When a biopsy sample is digitized a very high-resolution file is generated, e.g., a digitized biopsy with $40x$ magnification may produce an image of size about $(110,000 \times 110,000)$ or $12,100,000,000$ pixel resolution, which generates a file of approximately 2.0 GB. To analyze such an image, the specialist has to find ACIF patterns, or ROIs, seen in a very low microscope magnification (digitized reduction) and export them to "smaller" images (usually in Tagged Image File Format, (TIFF)) to be analyzed later. This process is performed several times and it requires many hours, not only because the pathologist takes time finding ROIs, but also exporting them requires long computational processing. This process is illustrated in Figure 1. In addition, there does not exist a software, tool or method capable of automatically neither finding ROIs in low microscope magnification nor discovering ACIF patterns in higher magnification.

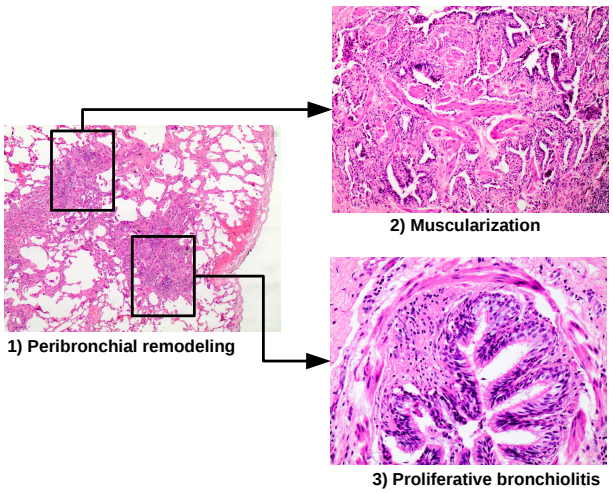


Fig. 1. Pathologists find ACIF patterns in low microscope magnification, see black frames in (1), and export them to isolated images to be analyzed in higher magnification, see (2) and (3).

According to this context, in this manuscript, we present a method to automatically segment regions of interest (ROI) and extract features in order to identify ACIF patterns found in a low microscope magnification. The main goal is to *accurately* and *quickly* highlight the most remodeled regions (mainly fibrosis centered and extending around the bronchioles), the ones candidates to show the diseases. Thus, the specialists can analyze them in high microscope magnification looking for more specific ACIF patterns. By doing so the specialist time and effort is best employed.

II. Related Work

In general, a thorough diagnosis for lung diseases is crucial for giving the best treatment to the patients. However, it is a very complex task and a multidisciplinary approach involving clinicians, radiologists, and often pathologists is demanded. In an effort to contribute to this problem, some studies have been proposed in the field of computerized medical image processing applied to respiratory deceases, more specifically Idiopathic Pulmonary Fibrosis (IPF) and Usual Interstitial Pneumonia (UIP). Most of them extract and process features from CT data, but in many cases, a surgical lung biopsy is indicated when CT patterns are indeterminate or inconsistent [6]. Following we describe representative state of the art computer-based studies in IPF and UIP diagnosis.

A Radial-basis-function neural-network-based method developed to recognize IPF in microscope images was introduced by Ilias et al. [7]. This study was evaluated on 20 samples extracted from mice after the induction of pulmonary fibrosis. The authors evaluated their method by comparing it against the SVM algorithm [8]. This work was one of the first efforts to introduce computer-assisted quantification in lung microscopic images since other previous work are based on CT and MRI.

Su Yeon Kim et al. [9] proposed a machine learning approach to distinguish UIP from other Interstitial Lung Diseases (ILD) in a surgical lung biopsy. Experiments were performed with 86 samples collected from 11 hospitals in North America, 58 samples were identified as UIP, 23 as non-specific UIP, and 16 as hypersensitivity pneumonitis. The main goal of this work was to find a genomic signature of bronchoscopy samples to predict IPF and develop a molecular test in order to avoid surgery in the diagnosis of IPF.

A method to three-dimensional characterization of fibroblast foci in IPF using integrated high-resolution CT and histological data was proposed by Mark G. Jones et al. [10]. The authors studied the morphology of fibroblast foci in 3D data. First, they assessed visualization of diagnostic paraffin-embedded lung tissue 3D microarchitecture by micro-CT. Clear morphological differences between normal and IPF lung tissue are visualized after 3D rendering, normal lung structures such as blood vessels and airways are also visualized. UIP patterns such as heterogeneity of the distribution of interstitial fibrosis, and honeycomb change are also observed.

Pankratz et al. [11] proposed a machine-learning-based method to detect UIP in Transbronchial Biopsies (TBB) [11]. The main goal of this proposal was to distinguish UIP from non-UIP in an effort of reducing the need for surgical lung biopsy. Experiments were performed on 283 TBB from 84 subjects, the algorithm was trained with 53 samples. However, the authors do not provide details about the machine learning algorithm used.

A convolutional neural network (CNN) strategy to segment and identify fibrosis in stained histological images was proposed by Xiaohang et al. [12]. A CNN with 11 convolutional

layers was proposed, the first 9 involve 3×3 filters, and 2 layers of 1×1 convolutions. A dataset of 72 RGB images, 36 from control and DM groups, were used in the experiments. The authors reported a Dice similarity coefficient of 0.947 segmenting the images into three segments: fibrosis, myocytes, and background.

On the other side, a software tool for IPF "virtual" diagnosis has been recently developed by Bennett et al [13]. This tool, called "IPFdatabase", is based on the criteria for IPF diagnosis proposed by Raghu et al. [14]. The main goal of IPFdatabase is to make a "virtual diagnosis" using information provided by specialists, i.e., physicians, radiologists, and pathologists. This information should include medical and patient history such as smoking status, occupational, environmental exposures, and IPF patterns found in CT and, if available, surgical biopsy. Once the data have been provided, the diagnosis is made. Notice this tool does not make any interpretation of medical images, thus radiologists and pathologists have to provide this information. Despite that, this tool is an interesting radiomic approach to simplify, improve, and accelerate the multidisciplinary diagnosis of IPF.

As we can see, more research in the computer-based analysis of lung respiratory deceases is required. The state-of-the-art work focuses mainly on IPF and UIP in CT images. Nevertheless, to the best of our knowledge, research on histopathological analysis, specifically on ACIF is non-existent. Thus, the present work represents the first effort towards a radiomic analysis of ACIF.

III. Airway-centered Interstitial Fibrosis Method

During the process of airway-centered interstitial fibrosis (ACIF) bronchi may exhibit loss of structural integrity, such as muscularized bronchioles, bronchiolar metaplasia, proliferative bronchiolitis, and interstitial thickening. These patterns can be seen, using low microscope magnification, as high-density pixel regions. Examples of these regions are shown in Figure 2.

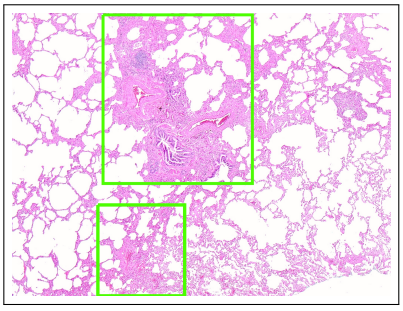


Fig. 2. Example of ACIF patterns in a lung biopsy, see the green frames. Using low microscope magnification ($< 5x$) ACIF patterns can be seen as high-density pixel regions. In this sample, one can see peribronchial remodeling (large frame), and fibroblastic extension (small frame). Both patterns present high concentration of pixels. Notice we consider only "colorized" pixels to define high-density regions.

In this work, we propose a method aimed at supporting pathologists to identify specific pattern lesions of ACIF. The traditional specialist's approach searches for important patterns in large microscopic images, which is an extremely time-consuming and tiring task. Thus, our main goal is to highlight regions of interest with relevant information to direct the specialist's attention to them.

Relevant information is defined by high-density pixel regions, which are automatically searched, thereby leading to a significantly reduction of the time invested by specialists, and enabling a more accurate diagnosis. Our method is divided into four main steps:

1) Automatic high to low microscope magnification reduction, which produces a lower pixel resolution image.
2) Computation of probability of a pixel belonging to a high-density region, defined as ROI.
3) Extraction of features of high- and low-density regions.
4) Visualization of ROIs using a colormap.

It is very important to note that the original image is not affected. Our method creates a new image with reduced resolution. Thus, it is possible to extract the ROIs, found in low magnification, from the original image with their original resolution/magnification.

A. Magnification reduction

In this step, we reduce the microscope magnification to $< 5x$, which produces a new smaller (pixel resolution) image. The goal of this step is to reduce the huge number of pixels in the input image, thus next steps can be processed in an acceptable computational time.

Let Ω be an image consisting of n pixels, in which the value of the i^{th} pixel is defined as Ω_i. The spatial location of pixel i is $\Omega_i(\mathbf{x}, \mathbf{y})$, the physical isotropic pixel sizes, in micrometers, are defined as $\{PhysicalX, PhysicalY\}\mu m$, in the $X-$ and $Y-$axis, respectively. The sizes, in pixels units, of Ω are defined as $\{SizeX, SizeY\}$.

We guide the adequacy of the original image magnifications using the physical isotropic pixel size. The magnification reduction is performed based on the following computations:

$$\begin{aligned} \Delta M &= ((M_{in} \times ep) - (M_{out} \times ep)) \\ PhysicalX_{out} &= (PhysicalX_{in} \times \Delta M) + PhysicalX_{in} \\ PhysicalY_{out} &= (PhysicalY_{in} \times \Delta M) + PhysicalY_{in} \\ SizeX_{out} &= \frac{SizeX_{in}}{PhysicalX_{out}} \\ SizeY_{out} &= \frac{SizeY_{in}}{PhysicalY_{out}} \end{aligned} , \tag{1}$$

where ΔM is the microscope magnitude reduction, M_{in} is the original magnitude, M_{out} is the target magnitude, $ep = 10x$ is the microscope eyepiece, $\{PhysicalX_{in}, PhysicalY_{in}\}\mu m$ are the original physical pixel sizes in X and Y directions, $\{PhysicalX_{out}, PhysicalY_{out}\}\mu m$ are the resulting physical pixel sizes after reduction, $\{SizeX_{in}, SizeY_{in}\}$ are the original sizes in pixel units, and $\{SizeX_{out}, SizeY_{out}\}$ are the resulting sizes.

Due to the very high resolution of the input image, we divide it into tiles for both reading and processing. We then reduce the size of every tile using an interpolation method, called "INTER_AREA" provided in OpenCV [15]. The INTER_AREA method receives the $\{SizeX_{out}, SizeY_{out}\}$ (Equation 1) values as input parameters and uses the pixel area relation to re-sample the image. Finally, we build a reduced image Ω' as a mosaic of reduced tiles. The pixel values of the reduced image Ω' are given in gray-scale. Notice, the pixel values of the input image Ω represent color (blue, violet and red) of Hematoxylin and Eosin stain (H&E), used to color the biopsy. In order to categorize texture and morphological patterns, we use the gray-scale value instead of the H&E value (usually an RGB triple).

B. High-density region identification

In order to identify high-density regions, we first segment the white background, see Figure 3, by applying a threshold-based approach to Ω' using the Otsu algorithm [16]. This step produces a binary representation Ω^b of Ω'. We define a circular mask $\omega \in \Omega^b$ of radius r and, for each pixel $i \in \Omega^b$, we compute the probability P_i of it belonging to a high-density region, as follows:

$$P_{i \in \Omega^b} = \frac{(\sum_{j \in \omega} j) \times 100}{\pi r^2}, \qquad (2)$$

where, $j \in \omega$ is a neighbor pixel of i. We use a threshold t to divide the high (H) and the low (L) density regions. We consider $P_i > t$ to separate them. Figure 3 illustrates the high-density region identification step of our method.

The pixel probabilities P_i can be visualized as a colormap, in which all high-density regions are represented with a red color. In this work we use the "Jet" colormap, which is represented by colors from blue to red. Examples of the generated colormaps can be seen in Figure 6.

On the other hand, notice that by using Otsu to remove the background in Ω' (first step in Figure 3), some regions may wrongly be labeled as background, as shown in Figure 4. This problem can also arise when using other threshold-based methods (simple threshold methods, triangle or statistical moment-based methods).

We solve this problem by computing the convex hulls of all components in H, filling them with the respective pixel values in Ω', which includes background pixels (step 3 in Figure 3). Finally, image L can be viewed as the input Ω' with all convex hulls removed. The resulting images are shown in Figure 5.

IV. Feature Extraction

Several features can be extracted from histopathological images, e.g., gray-scale histograms, statistical moments, co-occurrence matrices, Local Binary Patterns (LBP), among others. We consider LPB features as suggested by Kumar et al. [17], where a comparison of LBPs, Convolutional Neural Networks (CNNs), and a Bag of Visual Words (BoVWs) showed that LBP features led to high-accuracy results for classifying histopathological images. The main goal of this

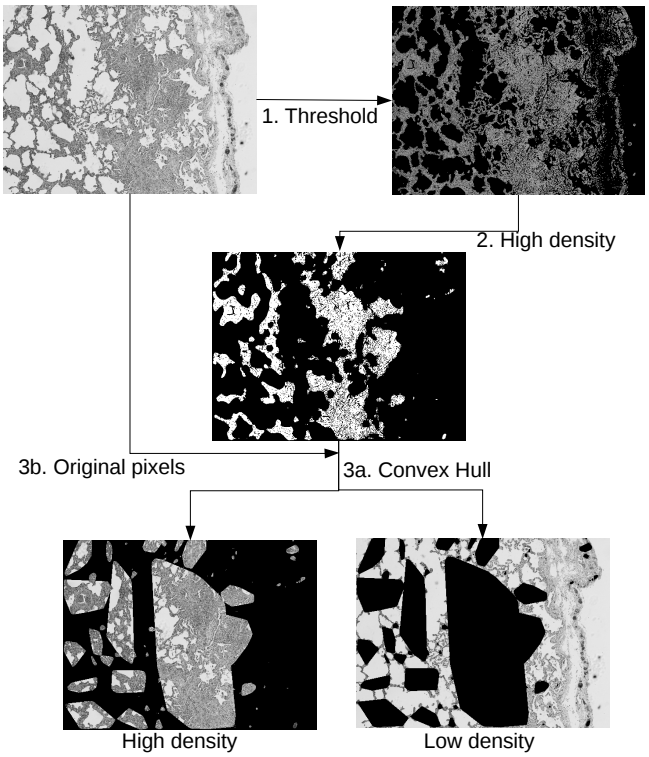
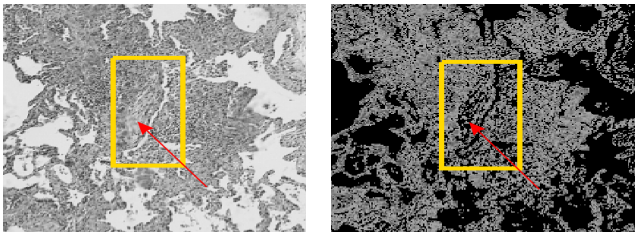
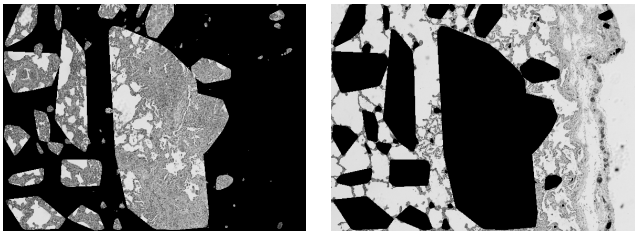


Fig. 3. First, high-density regions are identified. Two images are generated with high- and low-density regions. The high-density image is divided into connected components in order to extract features and high-light severely damaged regions.



(a) Input Image (b) Otsu

Fig. 4. Regions can be wrongly labeled as background (black color) when using Otsu, see yellow frames in (a) and (b). Proliferative bronchiolitis pointed out with red arrows, notice that Otsu misses the remodeled tissue.



(a) High density (b) Low density

Fig. 5. Image Ω' divided into high- and low-density regions. See lighter colors. B

step is to compare the LBP features between the high- and low-density regions identified before, in order to quantify the accuracy of our high-density identification method.

V. Results

We evaluated our method for nine histopathological images in Virtual Slide Image (VSI) format of the lung, of resolutions $(110,019 \times 100,196)$ and $(128,018 \times 56,878)$, physical isotropic pixel sizes of $(0.172\mu m \times 0.172\mu m)$, and microscope magnifications of $40x$ and $20x$. We reduced the original image magnifications to $5x$, leading to a resolution of (1611×1467) and physical isotropic pixel sizes of $(68.309\mu m \times 68.309\mu m)$. All images were obtained with the Microscope Olympus BX 61 VS with an eyepiece of magnification $10x$, using biopsies from patients with confirmed diagnosis of ACIF.

We have performed our experiments on a Linux workstation, with Intel Core i7-2600 CPU 3.40GHz x 4 and 16GB memory. The main average processing time was 15.13 minutes. This short processing time is due to the magnification reduction step and because we have implemented our method using multiple threads. Magnification and high-density region identification are steps that can be parallelized in a straightforward manner once the image is divided into and processed as tiles.

A. Colormap

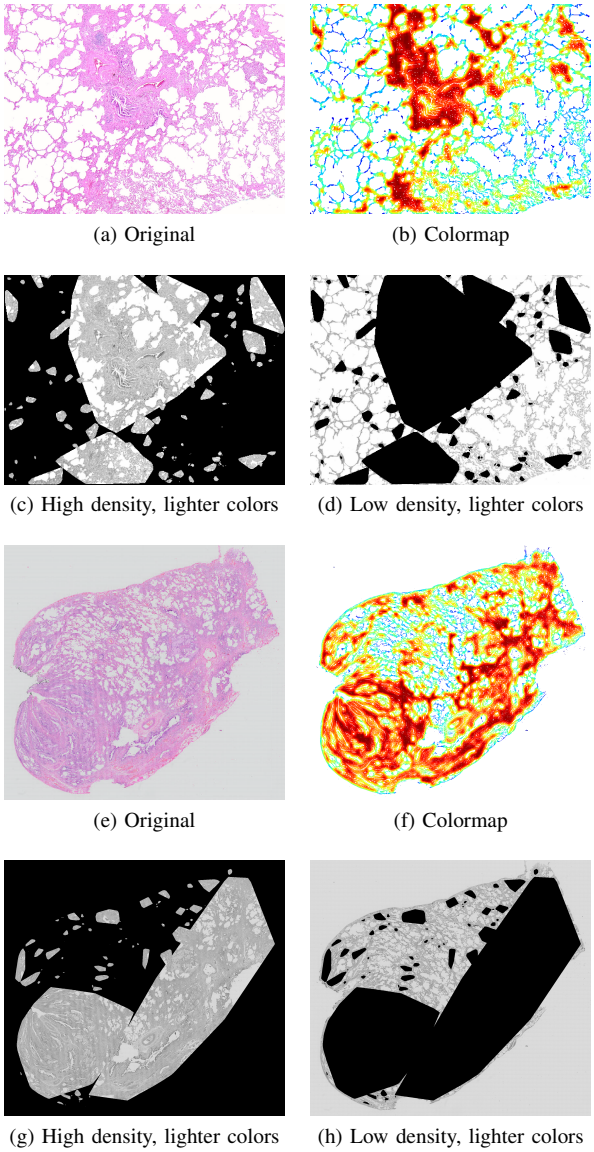
After identification of high-density regions in all samples, results are shown in high- and low-density regions, using a colormap. Pathologists, trained in ACIF identification, have analyzed these results and were able to identify ACIF patterns in the high density regions, demonstrating that our segmentation correctly identifies and captures all ACIF patterns in high-density regions. According to the specialists, the colormap helps to rapidly focus on important areas and increase microscope magnification for them, to search for more specific patterns. Figure 6 shows two results along with the patterns identified by the pathologist.

B. Local Binary Pattern

In order to verify and quantify the precision of our method, we have computed the LBP pattern of both high- and low-density regions. Figure 7 shows the resulting curves of the LBP histograms. Observe that all LBP curves for high-density and low-density are very close to the other of the same category. This result is independent of the size and shape of the ROIs.

Kumar's study showed good results using LBP combined with Support Vector Machine (SVM). However, in our case, the number of available images do not allow to train a SVM. Therefore, we used the LBP histograms as input of a k-Nearest Neighbor algorithm (k-NN). We have obtained a classification accuracy of 94.44%, which is the highest one reached so far in the literature.

In Table I, we present all parameters used in our experiments.



(a) Original (b) Colormap

(c) High density, lighter colors (d) Low density, lighter colors

(e) Original (f) Colormap

(g) High density, lighter colors (h) Low density, lighter colors

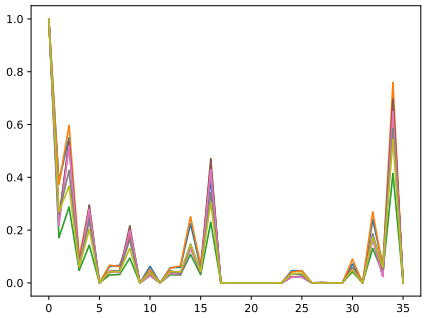
Fig. 6. Two results shown in (a) to (d), and (e) to (h), respectively. High-density regions are shown in dark red, which are extracted as ROIs by computing their respective convex hulls, see (c) and (g).

TABLE I
Parameters used in our experiments.

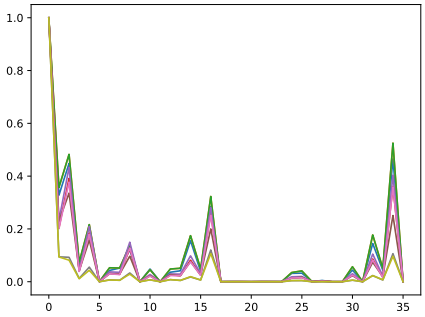
Description	Value	Units
Density Radius	7	pixels
Density Threshold	60	grays-cale
Magnification	5x	microscope power
LBP Radius	1	pixels
Number of LBP points	8	integer
KNN	3	number k of neighbors
Cross validation	9	number of folds

VI. Conclusions and Future Work

We have presented a method to automatically identify high-density regions in very high-resolution lung biopsies. During the process of Airway-centered interstitial fibrosis deceased bronchi may suffer severe remodeling. This pattern can be



(a) High density



(b) Low density

Fig. 7. LBP histograms of high- and low-density regions of nine biopsies. Note that all 9 curves for high- and low-density histograms closely ties.

seen, in a low microscope magnification, as regions of high concentration of pixels. Our method has been shown to be well suited to the identification of those regions in biopsies of patients with confirmed ACIF.

Besides, we compute local binary pattern histograms of high- and low-density regions in order to quantify the accuracy of the identification process. We obtained accuracy results up to 94.44%, showing its effectiveness to highlight the most remodeled regions so that the specialist can analyze them in the following step, for looking more patterns to diagnose ACIF in higher microscope magnification.

Since our method is a fully automatic approach, the interaction with the specialist results in a very straightforward task. Our implementation not only shows to the pathologist the high-density regions as colormaps, but also extracts them (showing their convex hulls) from the original input image Ω as separate images in Tiff format. Having the images separated, as is usually done do by specialists, allows them to analyze the images in higher magnification as the process demands.

This work is, to the best of our knowledge, the first effort to deal with images of ACIF, automatically highlighting regions of interest for the pathologist analyze and build the disease report. It is important to stress that this approach enables improvements not only in the accuracy, but also in the analysis time. That is a procedure that takes hours can be done in minutes.

Our future investigations will include analyzing various strategies for automatic identification of ACIF patterns on histopathological images in high microscope magnification. For that, we intend to use the ROIs automatically identified (method presented in this manuscript) as input to look for specific ACIF patterns in higher microscope magnification. Currently, this task is performed by the pathologist by analyzing the ROIs found in a low microscope magnification.

References

[1] K. Ley and A. Zarbock, "From lung injury to fibrosis," *Nature Medicine*, vol. 14, no. 1, p. 20, 2008.

[2] D. Schraufnagel, *Breathing in America: diseases, progress, and hope*. American Thoracic Society, 2014.

[3] R. K. Virk and A. E. Fraire, "Interstitial lung diseases that are difficult to classify: a review of bronchiolocentric interstitial lung disease," *Archives of Pathology & Laboratory Medicine*, vol. 139, no. 8, pp. 984–988, 2015.

[4] E. Silbernagel, A. Morresi-Hauf, S. Reu, B. King, W. Gesierich, M. Lindner, J. Behr, and F. Reichenberger, "Airway-centered interstitial fibrosis–an under-recognized subtype of diffuse parenchymal lung diseases," *Sarcoidosis Vasculitis and Diffuse Lung Disease*, vol. 35, no. 3, pp. 218–229, 2018.

[5] A. Churg, J. Myers, T. Suarez, M. Gaxiola, A. Estrada, M. Mejia, and M. Selman, "Airway-centered interstitial fibrosis: a distinct form of aggressive diffuse lung disease," *The American Journal of Surgical Pathology*, vol. 28, no. 1, pp. 62–68, 2004.

[6] D. A. Lynch, N. Sverzellati, W. D. Travis, K. K. Brown, T. V. Colby, J. R. Galvin, J. G. Goldin, D. M. Hansell, Y. Inoue, T. Johkoh *et al.*, "Diagnostic criteria for idiopathic pulmonary fibrosis: a fleischner society white paper," *The Lancet Respiratory Medicine*, 2017.

[7] I. Maglogiannis, H. Sarimveis, C. T. Kiranoudis, A. A. Chatziioannou, N. Oikonomou, and V. Aidinis, "Radial basis function neural networks classification for the recognition of idiopathic pulmonary fibrosis in microscopic images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 1, pp. 42–54, 2008.

[8] B. Scholkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.

[9] S. Y. Kim, J. Diggans, D. Pankratz, J. Huang, M. Pagan, N. Sindy, E. Tom, J. Anderson, Y. Choi, D. A. Lynch *et al.*, "Classification of usual interstitial pneumonia in patients with interstitial lung disease: assessment of a machine learning approach using high-dimensional transcriptional data," *The Lancet Respiratory Medicine*, vol. 3, no. 6, pp. 473–482, 2015.

[10] M. G. Jones, A. Fabre, P. Schneider, F. Cinetto, G. Sgalla, M. Mavrogordato, S. Jogai, A. Alzetani, B. G. Marshall, K. M. OReilly *et al.*, "Three-dimensional characterization of fibroblast foci in idiopathic pulmonary fibrosis," *JCI insight*, vol. 1, no. 5, 2016.

[11] D. G. Pankratz, Y. Choi, U. Imtiaz, G. M. Fedorowicz, J. D. Anderson, T. V. Colby, J. L. Myers, D. A. Lynch, K. K. Brown, K. R. Flaherty *et al.*, "Usual interstitial pneumonia can be detected in transbronchial biopsies using machine learning," *Annals of the American Thoracic Society*, vol. 14, no. 11, pp. 1646–1654, 2017.

[12] X. Fu, T. Liu, Z. Xiong, B. H. Smaill, M. K. Stiles, and J. Zhao, "Segmentation of histological images and fibrosis identification with a convolutional neural network," *Computers in Biology and Medicine*, vol. 98, pp. 147–158, 2018.

[13] D. Bennett, M. A. Mazzei, B. Collins, E. Bargagli, S. Pipavath, D. Spina, M. L. Valentini, C. Rinaldi, G. Bettini, A. Ginori *et al.*, "Diagnosis of idiopathic pulmonary fibrosis by virtual means using ipfdatabase-a new software," *Respiratory Medicine*, vol. 147, pp. 31–36, 2019.

[14] G. Raghu, H. R. Collard, J. J. Egan, F. J. Martinez, J. Behr, K. K. Brown, T. V. Colby, J.-F. Cordier, K. R. Flaherty, J. A. Lasky *et al.*, "An official ats/ers/jrs/alat statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management," *American Journal of Respiratory and Critical Care Medicine*, vol. 183, no. 6, pp. 788–824, 2011.

[15] G. Bradski and A. Kaehler, "Opencv," *Dr. Dobbs Journal of Software Tools*, vol. 3, 2000.

[16] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[17] M. D. Kumar, M. Babaie, S. Zhu, S. Kalra, and H. R. Tizhoosh, "A comparative study of cnn, bovw and lbp for classification of histopathological images," in *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017, pp. 1–7.