ABSTRACT

Lung pleura is a reference structure for the identification of histological findings for the recognition of a pathological interstitial lung disease (ILD) pattern. When a pattern is found, it is important to know whether it is close to the pleura to determine its specific type and severity. This manual process is a tedious and laborious one for the pathologist. Automating this task is important for a complete computer-assisted ILD diagnostic process. We introduce “RadPleura”, a framework for pleura classification of histopathological images using a radiomics-based approach. Our framework performs image pre-processing, region-of-interest segmentation, and extraction of a radiomic-signature suited for ILD classification. To evaluate the radiomic-signature, we classified it into pleura and non-pleura, using two classifiers: Support Vector Machine and Gradient-boosted Decision Trees. Our experiments yield promising results, with F-scores of 92% for SVM, and 91% for GBD. We also created a dataset of lung histopathology images with respective ground truth for pleura classification. To the best of our knowledge, this study is the first disclosed attempt to explore and develop a radiomic-signature for pleura classification. The methods and techniques are integrated into the RadPleura framework developed.

Index Terms— Interstitial lung disease, Lung pleura, Histopathology, Radiomics, Machine learning

1. INTRODUCTION

Interstitial Lung Disease (ILD) belongs to a large group of different diseases related to architecture distortion and further pulmonary fibrosis [1]. The determination of histopathological patterns through Surgical Lung Biopsy (SLB) is crucial to reveal diagnostic clues. When specific patterns are found, it is important to measure their distance to the pleura. Thus, one routine task performed by pathologists is to identify the pleura and determine its distance to the findings, as these measurements could indicate the disease stage and also help to classify it in a specific ILD etiology [2, 3]. Notice that, differently from other medical image exam types e.g., CT or MRI, there are no standard views (e.g., axial or coronal) for an SLB, which could guide the image analysis. Currently, there does not exist a single software tool or a method that aids specialists in analyzing an SLB. Specialists must use many tools, for different parts of the analysis, such as ImageJ, QuPath, Cell profiler, among others, that provide many medical image processing algorithms, mainly for image pre-processing [4, 5, 6]. However, none of them aids in finding the lung pleura. Identifying the pleura serves as a reference to understand what is altered in the interstitium. The pleura is the unique region of the sampled tissue that does not move, i.e., it is always a boundary, so the pleura region is the only guide for specialists. Fig. 1 shows an example of an SLB segment in which we show pleural and non-pleural boundaries.

Fig. 1: Lung pleura in the green square and non-pleura in the blue square. Pleura boundaries are smoother and more regular than non-pleural boundaries.

To classify boundary regions as pleura or non-pleura, we introduce RadPleura, a radiomics-based approach using texture extractors capable of measuring the roughness heterogeneity of the pleura. The Radiomics approach was designed to maximize the extraction of high-dimensional features that include first-, second-, and higher-order statistics, from medical images. Radiomics was a method applied to
find signatures in radiological data to understand tumor phenotypes [7]. Our goal is to establish a “radiomic signature” that captures the pleura heterogeneity in histopathology images. We included many Gray-Level Matrix (GLM) texture descriptors as well as the ones from Local Binary Patterns (LBP). These measurements are well-known for accurately capturing roughness in surfaces [8] and texture in histology images [9]. We use these features to train two classifiers, a Support Vector Machine (SVM) and a Gradient Busted Decision tree (GBD) in order to evaluate the classification properties of the radiomic signature. In addition, we also developed a method for Region of Interest (ROI) segmentation, that includes a pre-processing step for noise elimination and background segmentation. Then, we segment the ROI, defined as the external thick boundaries of each component in the sample, i.e., boundaries that separate the image background and the tissue. These are the main contributions of this work:

- A fast and efficient ROI extraction for pleura detection.
- A radiomic signature for pleura classification.
- A dataset, with respective ground truth, for lung pleura classification.
- A new web-based software tool to aid in ILD diagnosis in which we included our framework for lung pleura classification.

2. FRAMEWORK

A summary of our framework for lung pleura classification is shown in Fig. 2.

2.1. Pre-processing

Histopathological images present problems, such as color irregularities due to staining procedures, irregular illumination, and many artifacts such as hand-made annotations, ink stains, and collage-like (straight lines) due to the tile staking process used during the digitizing process. Fig. 2.1, shows examples of such problems.

To eliminate these problems, we apply a CIELab-based method to segment the tissue pixels (foreground) from all other elements including noise and artifacts present in the images and, at the same time to correct intensity-related issues. An important color-based feature of tissue in H&E staining histology images is the pink-scale intensity [10]. Pathologists examine the pink intensity of cells, for example, to analyze the cell properties. We can filter the tissue pixels by selecting the pixels within the pink region in the CIELab color space, i.e., the pixels should be greater than zero for channel $a$ and less than zero for channel $b$. Some noise may remain, nevertheless, most of them are filtered out which is sufficient for our application. All non-tissue pixels are replaced by a “white” pixel. After filtering out the non-desired pixels, we convert the CIELab image into a single RGB gray-scale image. Fig. 2.2, shows an output example of this process.

2.2. Region of Interest Extraction

The Lung pleura is always located on the tissue boundaries that we define as the ROI, and it is a thick boundary, “a belt,” of the tissue. Therefore, we detect only “external” boundaries by first creating a binary version of the filtered image. This becomes to be a simple task since we previously set all background pixels with the same white color. We consider the im-
We classify the feature vectors into two classes: pleura and non-pleura using two classifiers, an SVM and a GBD. The goal of evaluating these classifiers is not only to evaluate their performance but also to better understand the radiomic signature of the lung pleura to further improvements. This also becomes an important experiment since there is no related tool/system to build on or to compare with.

On the other hand, features can be redundant, irrelevant, and/or linearly dependent. To select relevant features we apply a Sequential Floating Forward Search (SFFS) method [11]. SFFS is a heuristic greedy algorithm commonly applied in histology analysis [12] that consists in sequentially adding the features that increment the classification score.

3. EXPERIMENTS AND RESULTS

We have evaluated our method with histopathology images of human lungs SLB of subjects with a confirmed ILD (by the Department of Pathology and Forensic Medicine, University of Sao Paulo, Ribeirao Preto, Brazil). The SLBs were digitized using a microscope power standardization of 20x. We used an Olympus BX61 VS. microscope with pixel resolutions of (110, 019 × 100, 196) and (128, 018 × 56, 878), and physical isotropic pixel size of (0.172μm × 0.172μm). Each scanned slice was verified by three specialists.

To reduce the samples size we re-scaled the dataset, from 20x to 1x. We generated images of about 3k×3k pixel resolution, which is still large. However, at this lower resolution we can apply our method in a practically acceptable time, requiring about 20 seconds on average for processing one image.

We performed our experiments on a Linux workstation (Intel Core i7-2600 CPU 3.40GHz x 4 with 16GB memory).

The Ground Truth (GT) data were manually created by physician specialists in pulmonary pathology. The process of labeling consisted of painting the lung pleura using a brush tool. This is a highly delicate task that must be done carefully, since many non-tissue (background) pixels can be wrongly labeled even by an experienced specialist. Therefore, to improve the labeling, we only considered as lung pleura the painted tissue pixels and deleted painted background pixels.

To perform this task, we unpainted the pixels outside the ROI using the method described in Section 2.2, which produced the final GT image data. Our dataset along with the GT is freely available and can be download, to encourage further research for lung pleura classification and computer-based histopathology analysis supporting ILD diagnosis. Please contact the corresponding author.

3.1. Histopathology Dataset

We created a dataset of 318 SLB of subjects with a confirmed ILD (by the Department of Pathology and Forensic Medicine, University of Sao Paulo, Ribeirao Preto, Brazil). The SLBs were digitized using a microscope power standardization of 20x. We used an Olympus BX61 VS. microscope with pixel resolutions of (110, 019 × 100, 196) and (128, 018 × 56, 878), and physical isotropic pixel size of (0.172μm × 0.172μm).

3.2. Local Binary Patterns

We tested the SVM and GBD using LBP for different radius, ranging from 6 to 20, and a tile side size of 400 pixels. Radius values above 20 pixels are unsuitable for a practical application due to the high processing time required to extract the LBP patterns, about 125 seconds. The best F-score obtained using LBP was 0.87, for a radius value of 10.

According to pathologists, this score is sufficient and acceptable for a practical application.

3.3. Gray-level Texture Descriptors

We used the default parameters value settings suggested in [13] to evaluate the texture descriptors to be compatible with other published methods. We also evaluated the influence of tile side size, using values of 200, 300, 400, and 500 pixels. The F-scores for the SVM and GBD were 0.88 and 0.86, respectively, for a tile size of 500 pixels.

1GNU Image Manipulation Program (GIMP)
3.4. LBP with GLM Features

We concatenated the LBP histograms and the GLM vectors into one single vector of 175 dimensions (82 LBP and 93 GLM). We also analyzed the influence of tile size on the classification score. Table 1 shows the resulting F-score, precision, and recall values for an LBP radius of 10 and isotropic tiles of side sizes 200, 300, 400, and 500 pixels, used for the SVM and GBD classifiers. One sees that the combination of LBP and GLM features produced improved classification scores of 0.92 (SVM) and 0.91 (GBD).

<table>
<thead>
<tr>
<th>classifier</th>
<th>tile</th>
<th>F-score</th>
<th>precision</th>
<th>recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>300</td>
<td>0.90</td>
<td>0.88</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0.91</td>
<td>0.89</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.92</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>GBD</td>
<td>300</td>
<td>0.89</td>
<td>0.87</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td>0.90</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>0.91</td>
<td>0.90</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 1: F-score, precision, and recall values for SVM and GBD classifiers using LBP and GLM features (10 k-fold).

3.5. Feature Selection

We used the Sequential Floating Forward Search (SFFS) Algorithm [11] to determine the number of selected features. The SVM algorithm was based on parameter values that produced the best results when using all features, i.e., radius value 10 and tile size 500. We sequentially applied the SFFS algorithm, reducing the number of features by five. The minimal number of selected features that ensures an F-score of (at least) 0.9 is 30.

3.6. Segmentation Quality Evaluation

In addition, we evaluated the segmentation quality for each individual image, we computed the DICE coefficient for pleura and non-pleura. The average coefficient values were 0.90 and 0.88, respectively. Fig. 3 shows two example results with the respective DICE coefficients.

3.7. Web Software Tool for ILD Diagnosis

We are developing a web-based software tool that includes RadPleura, in addition to other methods like automatic cell-nuclei segmentation and ILD pattern ROI segmentation, see [14, 15]. It is available here: https://ivarvb.github.io/RadPleura/.

4. CONCLUSIONS

We have introduced a method for lung pleura classification in histopathological images. Our method includes the main steps of a machine learning approach: image pre-processing, background segmentation, ROI extraction, feature extraction, classification, and feature selection. The radiomics approach makes it possible the extraction of many meaningful texture features to define a robust image signature. We used the feature vectors to evaluate two classifiers, SVM and GBD. The F-score values were up to 92%, demonstrating that a radiomics signature has the potential to characterize lung pleura for classification purposes.

Less than 20 seconds are needed for processing 500 × 500 tiles. We have used a feature selection algorithm, allowing us to reduce the size of the original feature vector from 175 to 30 dimensions, without significant score reduction.

An important contribution of our research is the provision of the dataset for pleura classification in histopathological images. There is no cure for ILD; it is not possible to treat or heal damaged tissue. A patient diagnosed with ILD must live with the disease for the rest of his/her life. With our research and provided dataset we hope to encourage more research to develop computer-based methods for ILD analysis. Finally, our method is adequate for practical use as a guide for specialists.

Compliance with Ethical Standards: This work is a study for which no ethical approval was required.
Acknowledgments: The authors thank FAPESP for financial support of the project “Mining, indexing and visualizing Big Data in clinical decision support systems (MIVisBD)” (Grants 16/17078-0, 18/06074-0, 2021/00360-3, 2023/03462-7, and 2020/07200-9). The authors also thank the Brazilian agencies CNPq and CAPES for the promotion of scientific and technological research.

5. REFERENCES


