

Automatic Pleural Line Extraction and COVID-19 Scoring from Lung Ultrasound Data

Leonardo Carrer, *Member, IEEE*, Elena Donini, *Student Member, IEEE*, Daniele Marinelli, Massimo Zanetti, *Member, IEEE*, Federico Mento, Elena Torri, Andrea Smargiassi, Riccardo Inchingolo, Gino Soldati, Libertario Demi, *Member, IEEE*, Francesca Bovolo, *Senior Member, IEEE*, and Lorenzo Bruzzone, *Fellow, IEEE*

Abstract—Recent works highlighted the significant potential of Lung Ultrasound (LUS) imaging in the management of subjects affected by COVID-19. In general, the development of objective, fast, and accurate automatic methods for LUS data evaluation is still at an early stage. This is particularly true for COVID-19 diagnostic. In this paper, we propose an automatic and unsupervised method for the detection and localization of the pleural line in LUS data based on the Hidden Markov Model and Viterbi Algorithm. The pleural line localization step is followed by a supervised classification procedure based on the Support Vector Machine (SVM). The classifier evaluates the healthiness level of a patient and, if present, the severity of the pathology, i.e., the score value for each image of a given LUS acquisition. The experiments performed on a variety of LUS data acquired in Italian hospitals with both linear and convex probes highlight the effectiveness of the proposed method. The average overall accuracy in detecting the pleura is 84% and 94% for convex and linear probes, respectively. The accuracy of the SVM classification in correctly evaluating the severity of COVID-19 related pleural line alterations is about 88% and 94% for convex and linear probes, respectively. The results as well as the visualization of the detected pleural line and the predicted score chart, provide a significant support to medical staff for further evaluating the patient condition.

Index Terms—COVID-19, Lung Ultrasound Imaging, Diagnostic, Viterbi Algorithm, Support Vector Machine, Signal Processing

I. INTRODUCTION

Given its cost-effectiveness, portability, non-invasiveness, and safety, ultrasonography represents an extremely useful instrument to anatomically investigate the human body. Soft-tissues are indeed characterized by very similar speeds of propagation of sound, hence allowing the exploitation of imaging protocols that assume quasi-homogeneous velocity (equal to 1540 m/s [1]) in the volume of interest. Furthermore, since the acoustic impedance of human tissues are also very similar [2], the different acoustic interfaces normally have a high transmission coefficient, thereby allowing the propagation of ultrasound waves.

L. Bruzzone (e-mail: lorenzo.bruzzone@unitn.it), L. Carrer, L. Demi, D. Marinelli, and F. Mento are with the Department of Information Engineering and Computer Science, Trento, University of Trento, 38123, Trento, Italy

F. Bovolo, E. Donini and M. Zanetti are with the Center for Information and Communication Technology, Fondazione Bruno Kessler, 38123 Trento, Italy

E. Torri is with Bresciamed, Brescia, Italy

R. Inchingolo and A. Smargiassi are with the Pulmonary Medicine Unit, Department of Cardiovascular and Thoracic Sciences-Fondazione Policlinico Universitario A. Gemelli IRCCS, Largo Gemelli, 8, 00168 Rome, Italy

G. Soldati is with the Diagnostic and Interventional Ultrasound Unit, Valle del Serchio General Hospital, Lucca, Italy

Nevertheless, these assertions do not hold for the lung, where the presence of air complicates ultrasound propagation. In particular, the high mismatch between the acoustic impedance of intercostal tissues and the air contained in lungs creates an acoustic interface (pleural line) whose reflection coefficient tends to 1 [3], thus making the lung normally impenetrable to ultrasound. Therefore, alternative diagnostic strategies should be considered to examine this organ.

Nowadays, Lung Ultrasonography (LUS) is based on the interpretation of imaging artifacts that appear in the reconstructed image below the pleural line. When a lung is healthy, it behaves as an almost perfect reflector and generates horizontal artifacts (known as A-lines), which are reverberations that appear at multiples of the distance between the probe and the pleural line [4]. In contrast, the vertical artifacts called B-lines correlate with various pathological conditions of the lung [5], [6], [7], [8], [9] and are likely associated to the formation of acoustic channels along the pleural surface [4]. Nonetheless, since the B-line genesis remains unclear [4], LUS is generally based on qualitative and subjective observations. However, some quantitative approaches have been recently proposed [10], [11], even if their use by clinicians is still limited [10], [11]. In contrast, semi-quantitative approaches, mainly based on counting vertical artifacts in the image [12], [13], [14], are the most largely used. Furthermore, the automatic detection of characteristic LUS patterns, such as B-lines, is of growing interest because it provides clinicians with real-time important visual information [15], [16], [17], [18], [19], [20], [21].

Specific LUS patterns have been recently proposed for grading the condition of patients affected by COVID-19 pneumonia [22], [23], [24], which at present time affects more than 5 million people with about 345 thousand deaths worldwide [25]. Patterns of interest are: (i) thickening of the pleural line with pleural line irregularity; (ii) various patterns, including focal and multifocal vertical artifacts; and white lung (iii) consolidations. Other findings such as pleural effusions are uncommon, while A-lines could be helpful in signaling a recovery phase. In this work, we focus on the analysis of the pleural line characteristics, which are all affected by the significant patterns described above. Thickening of the pleural line with pleural line irregularity clearly affects directly the pleural line characteristics. The presence of vertical artifacts also influences the pleural line characteristic, as these patterns do originate from the pleural line itself. Ultimately, consolidations de facto break the pleural line, and strongly impact its intensity and continuity. Moreover, the presence

of A lines implies a well reflecting pleural line, and thus correlate with a higher intensity of the pleural line. Of all the described alterations, pleural effusions are probably the least impactful on the pleural-line characteristics. In this context, the automatic analysis of the pleural line is thus particularly interesting. In fact, the pleural line appears completely continuous in a healthy lung, and becomes gradually more disrupted when the pathological condition worsens, as reported by the 4-level scoring system defined in [24]. The scores range from 0 to 3, where score 0 means absence of pathological signs, and score 3 represents a severe condition. While score 0 represents a continuous and regular pleural line with the associated presence of A-lines, score 1 is associated to a slightly disrupted pleural line with the presence of vertical artifacts below the disruptions [24]. Score 2 is associated with a broken pleural line that presents small or large consolidated areas (darker areas) below the breaking point, whereas score 3 denotes a completely permeable and discontinuous pleural line with large white areas underneath it (with or without large consolidations) [24].

Taking advantage of the above mentioned COVID-19 medical scoring procedure, this paper addresses the problem of automatically and quickly performing the scoring task on LUS data. Although artificial intelligence diagnostic systems based on deep-learning [16], [18] might represent a straightforward choice for the development of very high accuracy automatic detection algorithms, their main drawback is that they require a very large set of annotated data samples that are currently not available for LUS data of COVID-19 patients.

Thus, it is required to define either unsupervised or supervised techniques based on "shallow" machine learning approaches that can properly operate with a limited number of labeled samples. In this paper, we propose a novel system for the automatic estimation of COVID-19 severity (i.e., scoring) in LUS data by adapting and exploiting previous work developed for detecting and characterizing subsurface geological features in radar tomographic data [26], [27]. In detail, the proposed system is composed of two main parts. The first part is an unsupervised method for the automatic detection and characterization of the pleural line. The method is able to extract the pleural line geometric and intensity characteristics for each pixel of a given LUS image by exploiting a combination of a local scale Hidden Markov Model (HMM) and the Viterbi Algorithm (VA). The second part leveraging on the information provided by the pleural line detection method is a supervised classification technique that provides the score for each image of the LUS video. It is based on ad-hoc quantitative metrics based on COVID-19 pulmonary manifestations affecting the pleural line and the area below it and then on a Support Vector Machine (SVM) classifier that performs the automatic scoring (i.e. COVID-19 severity assessment) of each LUS image.

The overall strategy is tested on in vivo lung ultrasound data recently acquired in several Italian hospitals. The data set is very heterogeneous and highlights different severity of the pathology. Moreover, the data are acquired with different probes, namely linear and convex, thus providing an excellent statistical variation of COVID-19 manifestations in LUS data.

The paper is organized as follows. Section II illustrates the

proposed method. Section III reports the experimental results on both automatic pleural line detection and scoring of LUS data. Finally, Section IV addresses the conclusions of this paper.

II. PROPOSED METHOD

The proposed method analyzes, in an automatic way, LUS videos to detect and characterize the pleural line, on the basis of a scoring system specifically defined for LUS data obtained on COVID-19 patients. This is done by processing the video following an image-by-image approach. Note that in this article we refer to individual frames as images. Let $\mathcal{I} = \{I_1, \dots, I_v, \dots, I_V\}$ be the LUS video where I_v represents the v th image (i.e., one video frame). Figure 1 reports the block scheme of the proposed system for the automatic analysis of a generic image I_v . This system is composed by two main parts: i) Automatic pleural line detection. ii) COVID-19 score classification. The first part aims at detecting on each image $I_v \in \mathcal{I}$ the pleural line by first discriminating it from the background and then reconstructing it by means of a combination of Hidden Markov Model (HMM) and the Viterbi Algorithm (VA). In this way, we obtain a set $\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_v, \dots, \mathbf{p}_V\}$ representing the geometric location of the pleural line at each image. The second part uses the pleural line $\mathbf{p}_v$ to compute relevant features that describe both geometric and radiometric properties of the pleura and the area underneath it. A supervised SVM classifier is used to assign a COVID-19 score to each image $I_v \in \mathcal{I}$. The result is a set of scores $\mathcal{S} = \{s_1, \dots, s_v, \dots, s_V\}$, where s_v is the score of image I_v . Set $\mathcal{S}$ is used to make a binary decision on the positivity or negativity of the patient to COVID-19 LUS patterns. Moreover, it can be represented as a score histogram of the results for all the images in $\mathcal{I}$ to provide a comprehensive patient overview of the score severity. In the following sections the details of the method are presented.

A. Automatic Pleural Line Detection

To efficiently detect and characterize the pleural line in LUS data, an automatic detection method should possess the following requirements: i) it should discriminate between the pleural line and other LUS data features such as the ribs; ii) it should efficiently detect the pleural line under noisy conditions and iii) it should be able to perform inference and identify the pleural line position also when gaps occur in the LUS image (e.g., rupture of the pleura).

To this extent, we exploit and re-adapt a method previously developed for the characterization of radar tomographic data of the subsurface stratigraphy of Mars [26]. The method identifies all the linear or quasi-linear structures providing their intensity and geometrical characteristics by combining a local scale Hidden Markov Model (HMM) and the Viterbi Algorithm (VA).

In this framework, the LUS image of intensity $I_v(j, k)$ with size $J \times K$ can be seen as a trellis (see the example of Fig. 2 where an image portion is shown), where $j = 1, 2, \dots, J$ and $k = 1, 2, \dots, K$ are the image indexes of rows and columns, respectively. Note that to simplify the notation in the following

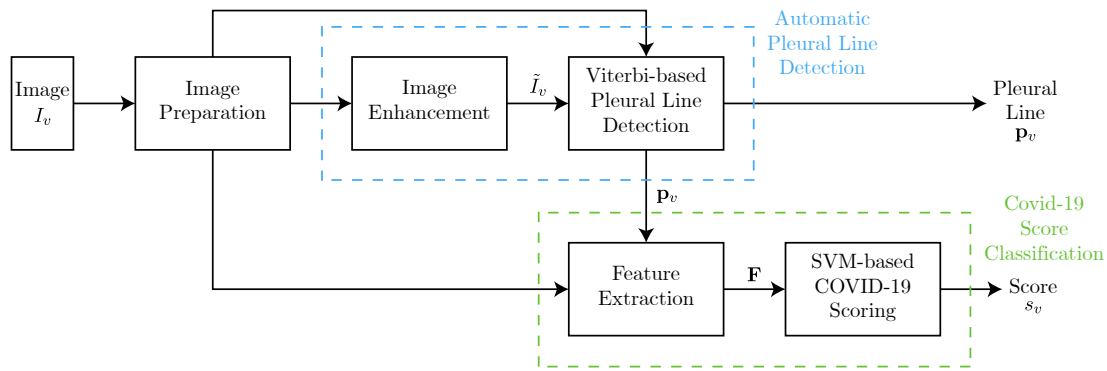


Fig. 1: Block scheme of the proposed method applied to one image I_v of a LUS video.

the image index v is omitted (i.e., $I_v(j, k) = I(j, k)$). The trellis properties are modelled by the HMM on which the VA performs the inference step. Starting from seeding points and based on the observed pixel intensities, the VA finds the most probable path, across the image that corresponds to any given linear structure present in the image (e.g., pleural line, ribs).

The pleural line detection is preceded by a two preliminary processing steps for image preparation and enhancement. The image preparation step is also required for the SVM classifier described in Section II-B, whereas the image enhancement is only relevant for the pleural line detection. The main purpose of the image enhancement processing is to reinforce the relevant features present in a LUS image and reduce the background noise thus easing unsupervised the pleural line detection procedure.

1) *Image Preparation*: LUS data acquired by a convex probe are mapped into a linear grid moving from a polar to a Cartesian coordinate system. This is required for satisfying the assumption that the structures to be detected are linear or quasi-linear. LUS data acquired by linear probes do not require this step.

2) *Image Enhancement*: Firstly, background noise in the image $I(j, k)$ is smoothed out using a circular averaging filter of radius W_1 . The radius value is chosen considering the size of the relevant layering structures and the intensity of the background noise. This is to reduce the noise while preserving the shape and intensity values of the structures.

Secondly, relevant layering structures are separated from the background of the image by means of a statistical method. To this aim, we employ a Rician-based statistical model that is often applied to multiplicative-noise distributed data, especially in medical applications [28], [29], [30]. We model the statistical distribution of a LUS image as a mixture of a background and foreground components, where the background is Rayleigh distributed (magnitude of 0-mean, diagonally co-variated Gaussian) and the foreground (layered structures) is Rician distributed (magnitude of non-null mean diagonally co-variated Gaussian). Unsupervised parameter estimation of this model from data can be done as described in [31], where an automatic iterative implementation of the Expectation Maximization (EM) algorithm is devised specifically for the Rayleigh-Rice mixture. After parameter estimation, the optimal threshold T is obtained by means of Maximum-a-

posteriori approach, as described in [31]. This is done following the Bayes decision rule for minimum error that assigns each pixel to the class (i.e., background or layering structure) that maximizes the posterior conditional probability. Threshold T represents the value of intensity that separates the relevant layering structure from the background noise. A pseudo-image $\hat{I}(j, k)$ is composed by setting to zero background values (i.e., values below T) and retaining foreground values (i.e., values above T), as shown in the example of Figure 5b.

Finally, a skeleton representation $\tilde{I}(j, k)$ of $\hat{I}(j, k)$ is obtained by convolution with a rectangular function as follows [26]:

$$\tilde{I}(j, k) = \hat{I}(j, k) * \text{rect}_{2\delta}(j), \quad (1)$$

where $(*)$ is the convolution operator and δ the half-width of the rectangular function $\text{rect}_{2\delta}(j)$. The convolution described in (1) is performed to avoid nonphysical irregularities in the interface continuity along the k direction. These irregularities could arise from potential imprecisions due to noise in the peak detected values of $\hat{I}(j, k)$. An example of the skeleton image $\tilde{I}(j, k)$ in Figure 5c.

3) *Pleural Line Detection with the Viterbi Algorithm*: In the LUS data context, we interpret the HMM parameters [32] [26] in the following way. The time step corresponds to the column variable k . The hidden states correspond to the candidate linear structure in a particular row j of the LUS image for each time step (Fig. 2). The state transition matrix models the probability that the structure is located in a certain row at a given time step, given it was located at another particular row at the previous time step. For each time step, the observation vector is equal to the LUS image pixel intensities. Given these assumptions, the method described in [26] is applied to the LUS data in a image-by-image fashion of the standard VA approach to reduce the computational complexity by performing the inference on small image sub-blocks which are then chained together to produce the final result for any given tracked feature. Accordingly, multiple features identification can be obtained in near real-time.

Given the enhanced LUS image $\tilde{I}(j, k)$, the linear structures are sequentially tracked by the VA starting from intensity values denoted as seeding points taken from $\tilde{I}(j, k)$ (i.e., values above threshold). For each time step, a number of states equal to β around the seeding point is evaluated. Accordingly, β also defines the maximum allowable variation, in terms of

index j , of the identified linear structure from one time step to another. The method identifies a set of N linear structures $\mathcal{L} = \{l_1, \dots, l_n, \dots, l_N\}$ (see the example of Fig. 2 based on a portion of $I(j, k)$). These objects in $\mathcal{L}$ could correspond to any given actual linear structure present in the image. Each structure l_n is characterized by its i) geometric position expressed as a set of couples (j, k) (i.e., pixel wise detection) within the LUS image and ii) corresponding intensity extracted from $I(j, k)$.

With respect to [26], where all the identified structures are equally important, here, we aim to identify only the pleural line (see Fig. 2). To this purpose, ad-hoc automatic detection conditions are specified as following. Let us define as: i) η_n the average deepness (in terms of row coordinates j) of the n th structure; ii) λ_n and Λ the average intensity of the n th structure and the average intensity of all the structures in $\mathcal{L}$ (i.e., $\Lambda = \sum_{n=1}^N \lambda_n / N$), respectively, and iii) ζ_n and Z the length of the n th structure and the average length of all the structures, (i.e., $Z = \sum_{n=1}^N \zeta_n / N$), respectively. First we identify a subset of linear structures $\mathcal{L}'$ such that:

$$\mathcal{L}' = \{l_n \mid \lambda_n > \Lambda \ \& \ \zeta_n > Z\}, \quad (2)$$

to select only those having length and intensity above the average. Then, the pleural line $\mathbf{p}$ is automatically identified within set $\mathcal{L}'$ as the deepest one, i.e.,

$$\mathbf{p} = l_n : \eta_n > \eta_x, \forall l_x \in \mathcal{L}', l_n \neq l_x. \quad (3)$$

In order to refine the estimate of geometric and intensity characteristics of $\mathbf{p}$, a second and finer iteration of the VA is applied only to the surroundings of $\mathbf{p}$. This is done directly on the image $I(j, k)$ pleural portion, without enhancement, and by applying a circular filter with radius W_2 (instead of W_1).

The detected pleural line $\mathbf{p}$ (see an example of detected pleural line in Figure 5d) is described with the graph of a function ρ defined over the domain of the columns of each image I . The function $\rho(\cdot)$ is defined as $\rho : \{k_s, k_e\} \rightarrow \{1, J\}$, where $1 \leq k_s < k_e \leq K$ and the general notation $\{a, b\}$ indicates the set of all integer numbers from a to b , inclusive. The function $\rho(\cdot)$ associates a given column index k to a certain row $\rho(k)$. Therefore, the pair $(\rho(k), k)$ therefore identifies the pleural line in image coordinates. The condition on the column index bounds k_s and k_e is as such because the pleural line may not span the entire horizontal dimension of the LUS image $I(j, k)$ for each given image.

By applying the proposed approach to all the images in $\mathcal{I}$ the set of pleural lines $\mathcal{P}$ is obtained. The automatic determination of the position and intensity characteristics of the pleural line described by $\mathbf{p}$ allows us to quantitatively analyze its properties and to define a scoring procedure based on the classification approach that will be described in the next section.

B. Classification-based COVID-19 Scoring Procedure

The geometric and intensity properties of the detected pleural line $\mathbf{p}$ and the underneath region (i.e., the area below the pleural line) are further modeled by a set of features

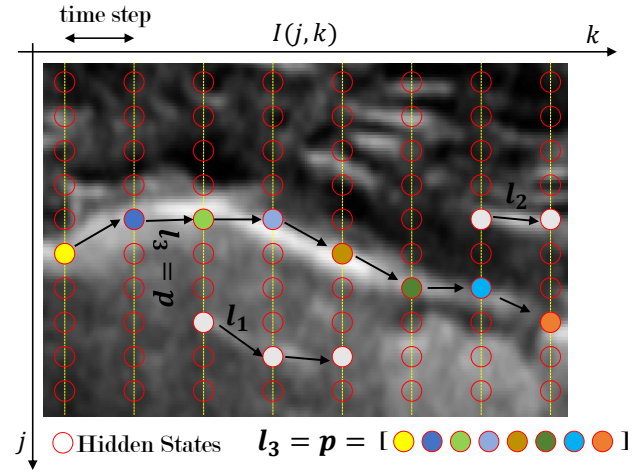


Fig. 2: Visualization of a portion of the LUS image $I(j, k)$ as a trellis. The geometrical position and intensity of the pleural line $\mathbf{p}$ is retrieved by a combination of a Hidden Markov Model and the Viterbi Algorithm.

(i.e., metrics) to be analyzed by the means of a supervised classification approach. The classification goal is to provide a prediction on the score (i.e., the diagnosis) for each image. Indeed, the score value is directly connected to the presence of structures linked to COVID-19 and its severity. In this method we consider the scoring system recalled in Section I and described in [24], which consists of four possible scores s ranging from 0 to 3. Therefore, each score value can be seen as a class.

For each image, the feature vector $\mathbf{F}$ is defined as $\mathbf{F} = [f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8]$. Features f_1, f_2, f_3 are related to the pleural line intensity, while features f_4, f_5, f_6, f_7, f_8 extract relevant statistical information on the intensity of $I(j, k)$ below the pleural line. The features f_1, f_2 are novel and ad-hoc defined for the COVID-19 scoring, while others are adapted from previous works in the radar domain. Given the above, we define each feature in $\mathbf{F}$ as follows.

f_1 - *discontinuities in the pleural line*: Let us define a metric M_1 representing the intensity of the pleural line for each column normalized by the pleural line average intensity (across image columns).

$$M_1(k) = \frac{I(\rho(k), k)}{\sum_{i=k_s}^{k_e} I(\rho(i), i) / (k_e - k_s + 1)}. \quad (4)$$

Large variations of $M_1(k)$ are potential indicators of pleural line discontinuities. A steady value of the metric across k is representative of healthy patients. Accordingly, the feature f_1 is defined as the average value of M_1 , i.e.,

$$f_1 = \sum_{k=k_s}^{k_e} \frac{M_1(k)}{(k_e - k_s + 1)} \quad (5)$$

f_2 - *presence of consolidations*: The second feature quantifies the average intensity of the portion of the image $I(j, k)$ below the pleural line with respect to the intensity value of

the pleural line $I(\rho(k), k)$ for each k . Let $M_2(k)$ be defined as follows

$$M_2(k) = \frac{1}{I(\rho(k), k)} \sum_{j=\rho(k)+1}^{\rho(k)+1+K_0} \frac{I(j, k)}{K_0}, \quad (6)$$

where the number of pixels under the pleural line considered for integration is given by the constant K_0 . f_2 is defined as the average of the metric M_2 , i.e.,

$$f_2 = \sum_{k=k_s}^{k_e} \frac{M_2(k)}{(k_e - k_s + 1)} \quad (7)$$

The feature f_2 quantifies phenomena such as white lungs and consolidations. Hence, steep peaks of this metric indicate consolidations in the lungs, while stable values are associated with lower scores.

f_3 - *total variation*: This feature models strong variations in the pleural line intensity $I(\rho(k), k)$. It is defined as follows:

$$f_3 = \sum_{k=k_s}^{k_e-1} |I(\rho(k+1), k+1) - I(\rho(k), k)| \quad (8)$$

High values of f_3 correspond to strong variations in the pleural line possibly indicating large pleural disruptions.

f_4, f_5, f_6, f_7, f_8 - *statistical characterization of the image area below the pleural line*: Let us denote as $\mathbf{I}_b = I(j_b, k)$, $j_b \in [\rho(k), \rho(k) + 1 + K_0]$, $k \in [k_s, k_e]$ the matrix containing the intensity values of $I(j, k)$ below the pleural line. The area below the pleural line is of particular importance as it contains several indicators of the pathological conditions (e.g., A and B lines). Feature f_4 describes the distribution parameters that best model the intensity values in $\mathbf{I}_b$. As a preliminary step, $\mathbf{I}_b$ is fitted to several distributions [27] to identify the best-fitting distribution according to the value of the root mean square error (RMSE) and the Kullback-Leibler (KL) divergence between the original and the fitted data. The fitting is performed by applying a sliding window ψ to $\mathbf{I}_b$ in both the horizontal and vertical directions. For each position of ψ , the fitting is done by computing the statistical parameters of the previously-estimated best fitting distribution of the pixels in the window. Accordingly, feature f_4 includes the average best-fit distribution parameters computed for each position of ψ . The number of parameters depends on the best-fitting distribution: the Rayleigh distribution has one parameter, while the Gamma distribution has two parameters that defines the shape and the rate (i.e., the inverse of the scale) of the function. Features f_5, f_6, f_7, f_8 are defined as the first four statistical moments (i.e., average, standard deviation, skewness, and kurtosis) evaluated on the intensity values in $\mathbf{I}_b$ for each image. High values of f_6, f_7 , and f_8 are related to pleural ruptures and consolidations as they provide a quantification of the amount of scattered ultrasound energy that originated from below the pleura as per definition of $\mathbf{I}_b$. For each available LUS image $I(j, k)$, the feature vector $\mathbf{F}$ is computed and then normalized.

Then, it is given as input to a Support Vector Machine (SVM) classifier. The choice of an SVM over a deep learning approach is mainly related to the size of the training set

required by the two types of classifiers. Indeed, a deep neural network requires a very large number of annotated samples in the training dataset to estimate all the network parameters. This is not the case for LUS data of COVID-19 patients which, due to the recent nature of the virus, are to this date very scarce. A deep network trained on such few data is likely to show poor performance in terms of generalization capabilities, and thus, accuracy. In contrast, SVM classifiers show good generalization properties even when trained with a limited number of samples [33]. Furthermore, SVMs: i) can solve strongly non-linear problems in the feature space, ii) provide a sparse and unique solution to the learning problem and iii) are fast in the training and testing phases and require low memory usage. To improve the discrimination performance and transform a non-linear problem in a linear one, non-linear kernel functions are applied to the input data. Here, we apply the Gaussian radial basis function (RBF) kernel as it is the most widely used one in general problems and has a better convergence time than the polynomial one [34].

By applying the scoring procedure to all the images of a video, a set of score $\mathcal{S} = \{s_1, \dots, s_v, \dots, s_V\}$ is obtained.

III. EXPERIMENTAL RESULTS

A. Dataset Description

The proposed method has been tested on a subset of the Italian COVID-19 Lung Ultrasound Data Base (ICLUS-DB) [35]. The data have been acquired in multiple clinical structures (BresciaMed, Brescia, Italy, Valle del Serchio General Hospital, Lucca, Italy, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy, Fondazione Policlinico Universitario San Matteo IRCCS, Pavia, Italy, Tione General Hospital, Tione (TN), Italy). Data have been acquired with different types of ultrasound scanners (MindrayTM DC-70 Exp[®], EsaoteTM MyLabAlpha[®], ToshibaTM Aplio XV[®], WiFi Ultrasound Probes - ATL) using both linear and convex probes depending on the needs. Every image in a LUS video has been assigned a score ranging from 0 to 3 according to the severity of the illness, where 0 is a negative patient and 3 identifies a severe form of lung surface alteration connected to COVID-19 [24]. To guarantee objective annotation, the labelling process was stratified into 4 independent levels: i) score assigned image-by-image by 4 master students with ultrasound background knowledge, ii) validation of the assigned scores performed by a PhD student with expertise in LUS, iii) second level of validation performed by a biomedical engineer with more than 10 year of experience in LUS, iv) third level of validation and agreement between clinicians with more than 10 years of experience in LUS. The exploited dataset is composed by 29 cases (10 negative, 15 confirmed positive to COVID-19 by swab technique, and 4 suspected positive to COVID-19), for a total of 58 videos. 20 videos have been acquired by a linear probe and 38 by a convex one. The data, for each acquisition, is provided as a sequence of images with intensity coded in the range [0-255]. The average number of processed images per video is about 60.

Table I shows the experimental setup of the parameters for the proposed method. Data acquired by either linear and

TABLE I: Parameters of the proposed method for linear and convex probes.

Parameter	Linear	Convex
Filter radius W_1	3	8
Filter radius W_2	3	4
Rectangle Half Width δ	16	8
Viterbi State Bound β	10	10
Pixels under the pleural line K_0	100	100
Sliding windows ϕ size	20×10	20×10

TABLE II: Parameters of the classification: total number of samples, numbers of the training and test samples, and best parameters τ and γ for the RBF kernel of the SVM models.

Parameter	Model				
	B_L	B_C	R_L	R_C	L_C
$ \Omega $	482	1246	102	102	278
$ \Omega_{Le} $	241	623	51	51	139
τ	32	32	256	16	16
γ	2	2	4	4	4

convex probes have different characteristics. Here, we defined five groups as the dataset that contains videos from linear and convex probes from the hospitals in Brescia and Rome and the data from the convex probe from the hospital in Lucca. We analyzed these five groups of video and generated five SVM models, called model B_L , B_C , R_L , R_C , L_C , where the letters B , R , and L indicate the different hospitals (Brescia, Rome, and Lucca) and the subscript L and C designates the acquisition with the linear and convex probes, respectively. For each model, the number of samples (i.e., feature vectors $\mathbf{F}$) denoted as Ω_{Le} is uniformly extracted from the overall available reference samples Ω and exploited for the learning stage. The remaining samples $\Omega_{test} = \Omega - \Omega_{Le}$ are used for predicting the score of unseen patients and testing the generalization capability of the model. Here, we train the classifier with 50% of the reference samples (Table II). To determine the parameters of the RBF kernel, we apply a 10-fold cross-validation considering the range of the parameters $\tau \in [10^{-2}, 10^{10}]$ and $\gamma \in [10^{-2}, 10^{10}]$. τ is the regularization parameter that controls the trade off between having a low training accuracy and a low testing accuracy, hence, τ is related to the generalization capability of the model. γ is the inverse of the standard deviation of the Gaussian function, hence it indicates how much influence a training sample has in the classification phase. Table II shows the optimal kernel parameters τ and γ for the SVM models.

B. Automatic Pleural line Detection

In this Section, we discuss the results of the method described in Section II-A for pleural line detection.

First, the preprocessing step is applied to the LUS image for feature highlighting and background noise suppression. Fig. 3 shows some examples of the results obtained by applying the circular filter with different values of the radius W_1 . Fig. 4 reports an example of probability density estimation by the EM algorithm. The algorithm models the LUS image intensity

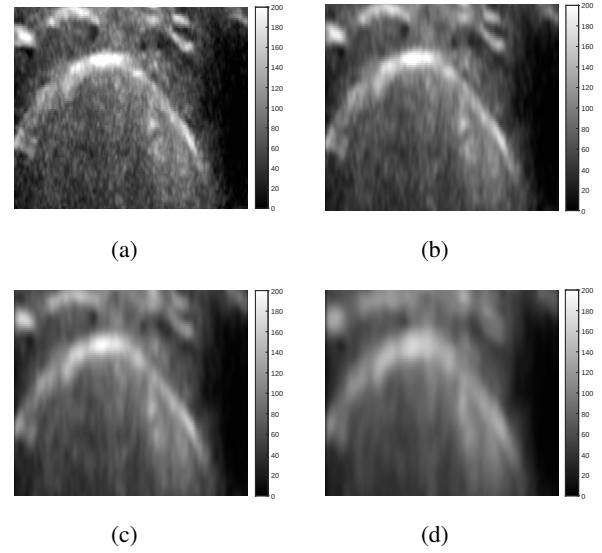


Fig. 3: Examples of results obtained by applying the circular filter with radius W_1 to a convex LUS image. (a) Unfiltered image (b) $W_1 = 3$, (c) $W_1 = 5$ and (d) $W_1 = 8$.

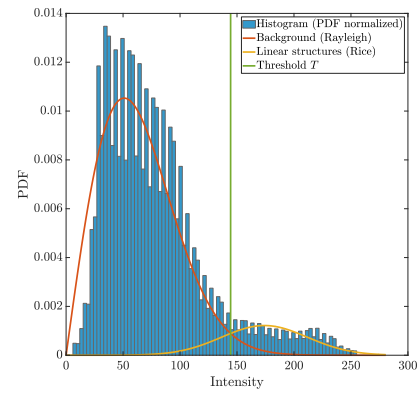


Fig. 4: Example of the EM algorithm applied to a LUS image.

histogram as a mixture of two distributions, namely Rayleigh for the background and Rice for the putative features. An example of the results of the application of the threshold T automatically derived by using the Bayes rule for minimum error is shown in Fig. 5b, where a comparison of the obtained $\hat{I}_v$ with the original image I_v is provided (see Fig. 5a). Fig. 5c reports the enhanced LUS image $\tilde{I}_v$ obtained by processing $\hat{I}_v$ in Fig. 5b. The thresholding procedure clearly highlights the image features. This reduces the background noise and is beneficial for the automatic pleural line detection performed with the Viterbi algorithm.

The parameter settings for the actual pleural line detection method are summarized in Tab. I. The state transition matrix (see Section II-A) is assumed to be a triangular function with size equal to β for each given hidden state [26]. Fig. 6 reports some examples of results of pleural line detection for both convex (after projection) and linear probes and for different scores. The figures display the detected pleural line, superimposed to the LUS image, and its associated intensity. The results highlight the method capability in accurately

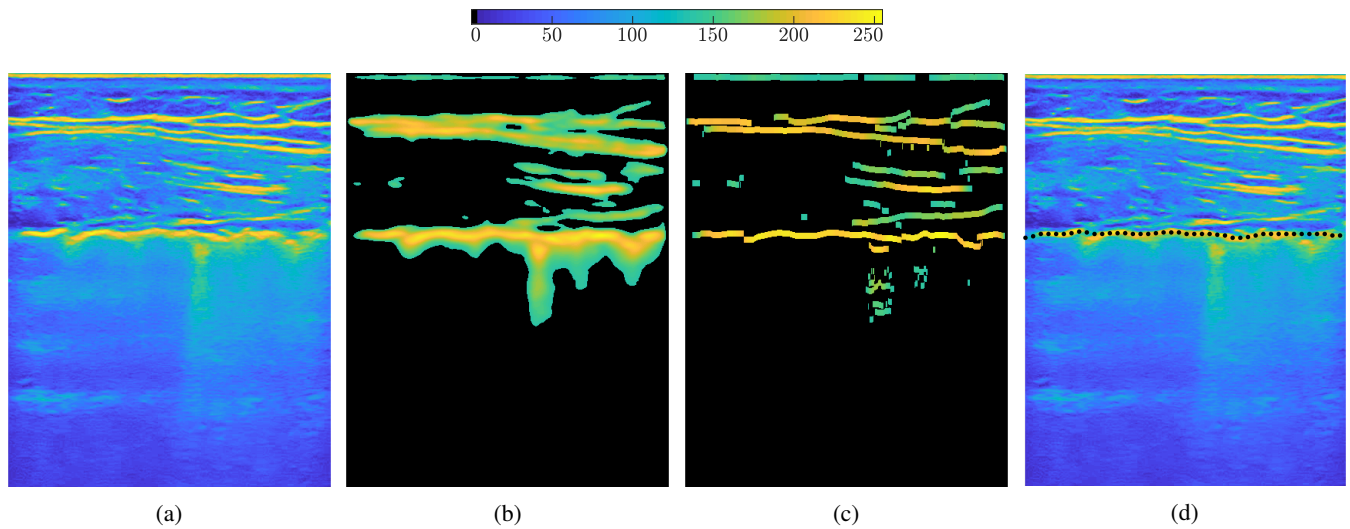


Fig. 5: Examples of results of different steps of the automatic pleural detection: (a) original image I_v , (b) masked image by EM-based thresholding $\hat{I}_v$, (c) enhanced LUS image $\tilde{I}_v$, (d) detected pleural line p_v (black dotted line).

detecting the pleural line geometric position also in the presence of pleural fragmentation, which is typical of score 2 and 3. Indeed, the quantitative characterization of the pleural line allows the determination of its subtle variations in both intensity and size of the pleural gaps. Score 0 patients (i.e., healthy patients) always show a stable value of the pleural line intensity in both the linear and convex case (see Figs. 6a and 6e).

Tab. III reports the method accuracy for different types of probes and scores. The accuracy is computed as the number of images where the pleura is correctly detected over the total number of images for each analyzed acquisition (i.e., one video). The number of test cases is reported in Section III-A. The global accuracy (i.e., computed independently from the score value) is 0.92 for the linear case and 0.84 for the convex one. The method has a high accuracy for LUS images up to Score 2. The accuracy decreases for Score 3. This is expected as for some patients the severity of the pleural line anomaly (e.g., multiple large ruptures) is such that a meaningful detection becomes problematic. It is worth mentioning that in a very small number of experiments the proposed method was not able to correctly detect the pleural line for any given image of the video. The extremely poor image contrast resulted in a wrong estimation of the threshold T by the EM as a result of the impossibility of modelling the background and the features intensity distribution as the sum of two separable distributions. As a lesson learned, a proper setting of the contrast before acquisition would be desirable for improving the algorithm detection performance.

C. Classification-based Scoring Procedure

This section describes the results of the automatic scoring procedure based on the supervised SVM classification. First, the results of the statistical analysis on the LUS images for the determination of the relevant distribution model of the intensity of the area I_b below the pleural line (see Section

TABLE III: Accuracy in the unsupervised detection of the pleural line (overall and separated by score) evaluated at a single image level.

Type	Score				Overall Accuracy
	0	1	2	3	
Linear	0.93	0.92	0.98	0.77	0.92
Convex	0.83	0.93	0.81	0.78	0.84

TABLE IV: Fitting performance of the Rayleigh, Gamma, Nakagami, and Weibull distributions to the LUS data.

Score	Metric	Rayleigh	Gamma	Nakagami	Weibull
0	KL distance	0.1653	0.1067	0.1683	0.1851
	RMSE	0.0075	0.005	0.0069	0.0068
1	KL distance	0.0441	0.0439	0.0457	0.0884
	RMSE	0.1535	0.0039	0.0044	0.0059
2	KL distance	0.0621	0.0197	0.0672	0.0834
	RMSE	0.0725	0.0081	0.0628	0.0262
3	KL distance	0.0086	0.0173	0.0536	0.0575
	RMSE	0.0125	0.0017	0.0024	0.0068

II-B) as a function of the score are illustrated. Then, we present the classification results based on the SVM.

The statistical analysis is performed by fitting the Gamma, Nakagami, Rayleigh, and Weibull distributions to the distribution of the intensities in I_b (i.e., the portion of $I(j, k)$ below the pleural line). The Kullback–Leibler (KL) distance and the root square mean error (RMSE) are exploited as performance metrics for evaluating the quality of the fit. Table IV shows the results of this analysis. The Gamma distribution shows better performance than the other distributions for all the scores in terms of both the KL distance and the RMSE. Hence, for the computation of the feature f_4 , we assume that I_b is Gamma-distributed.

For each image of each video, we extracted the feature vector $\mathbf{F}$ as per definition in Section II-B. As an example,

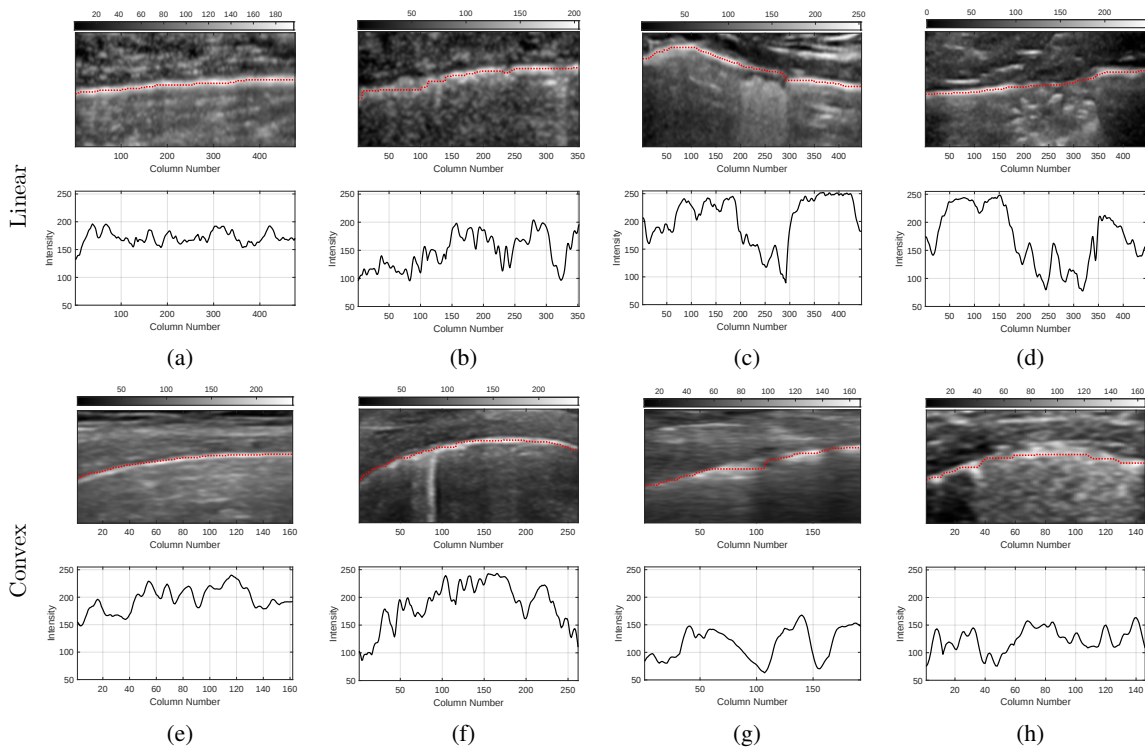


Fig. 6: Examples of experimental results on pleural line automatic detection (the pleural line detected by the method is shown in red) and intensity retrieval for linear and convex data. (a) Linear, score 0 (i.e., healthy patient). (b) Linear, score 1. (c) Linear, score 2. (d) Linear, score 3. (e) Convex, score 0. (f) Convex, score 1. (g) Convex, score 2. (h) Convex, score 3.

Figure 7 shows the values of the metrics M_1 and M_2 (exploited for defining f_1 and f_2) for the linear and convex probe data in Fig. 6. When the values of M_1 and M_2 are stable, the score associated with the image is 0. On the contrary, drops and peaks in the features indicate images with a higher score. While not explicitly exploited in this paper, it is also interesting to note that a given sequence of peaks or drops in the metrics can be used for evaluating the size of the pleural line anomaly.

Table V reports the SVM classification performance for each model in terms of the sensitivity (i.e., true positive), specificity (i.e., true negative), and overall accuracy for each score. The general effectiveness of the classification method is indicated by the high values of the sensitivity and the specificity of all the SVM models. Furthermore, the results show that the accuracies of correctly assigning the different scores are in the range of 80-95%.

To better understand the potential of the proposed approach, we tested the method with the data of three patients, here called Patient B6, Patient R1, and Patient L2, all screened 3 times. Patient B6 is patient 6 in the Brescia dataset and has been screened two times by a convex probe and one time with the linear probe. Patient R1 is patient number 1 of the Rome dataset and has been screened one time by the convex probe and two times by the linear probe. Finally, Patient L2 is patient number 2 of the Lucca dataset. Patient L2 was imaged with a convex probe. Table VI shows the overall accuracy for each video and the overall accuracy for each patient. Each video is processed and classified with the SVM model defined considering the specific probe and hospital. The overall video

TABLE V: Classification accuracy for each the model – sensitivity, specificity, and overall accuracy (OA). B,R and L stands for Brescia, Rome and Lucca data, respectively, while the pedix L or C indicates either linear or convex data.

Model	Metric	Score				Average
		0	1	2	3	
B_L	Sensitivity	0.8333	0.9394	0.9923	0.9663	0.9328
	Specificity	0.9979	0.9955	0.9641	0.9967	0.9885
	OA	0.8333	0.9394	0.9923	0.9663	0.9751
B_C	Sensitivity	0.8010	0.8155	0.9667	0.8107	0.8507
	Specificity	0.9838	0.9779	0.8669	0.9831	0.9529
	OA	0.8010	0.8155	0.9667	0.8107	0.8484
R_L	Sensitivity	0.9012	0.9079	0.9691	0.9326	0.9277
	Specificity	0.9979	0.9978	0.9103	0.9737	0.9699
	OA	0.9012	0.9079	0.9691	0.9326	0.9177
R_C	Sensitivity	0.8824	0.9333	0.9508	0.8222	0.8971
	Specificity	0.9882	0.9885	0.8800	0.9881	0.9612
	OA	0.8824	0.9333	0.9508	0.8222	0.8971
L_C	Sensitivity	0.9286	0.8696	0.9315	0.8874	0.9042
	Specificity	0.9712	0.9617	0.9189	0.9725	0.9560
	OA	0.9286	0.8696	0.9315	0.8874	0.9042

accuracy is defined as the average of the accuracies of each score in the video. By averaging the accuracies for the videos of the same patient, we can retrieve the overall accuracy for that patient. Fig. 8 shows the examples of the output of the scoring procedure based on the SVM classifier. The score is represented by the color of the square box around the image.

The results for each patient can be further analyzed by

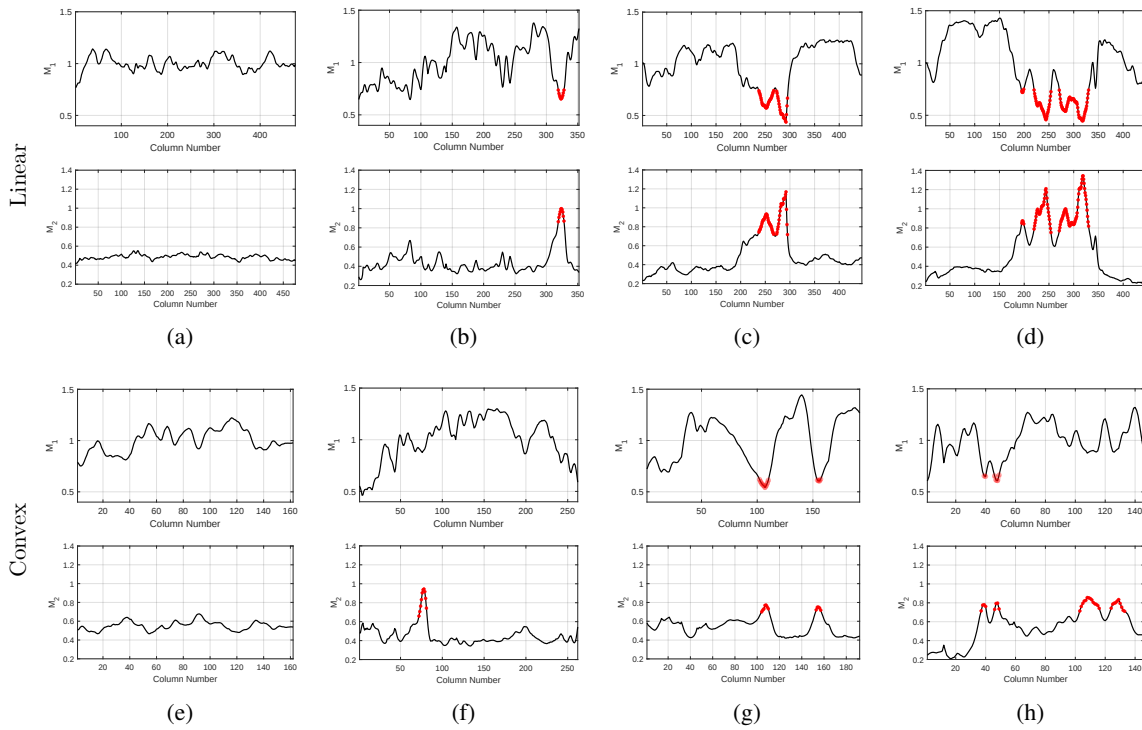


Fig. 7: Values of M_1 and M_2 for the LUS data of Fig. 6. (a) Linear, score 0 (i.e., healthy patient). (b) Linear, score 1. (c) Linear, score 2. (d) Linear, score 3. (e) Convex, score 0. (f) Convex, score 1. (g) Convex, score 2. (h) Convex, score 3. Red dots indicate potential anomalies. M_1 relates to pleural line intensity, while M_2 provides information by comparing the pleural line intensity to the intensity of the image area below it.

TABLE VI: Overall accuracy of the prediction of the score for the analysis of the videos of patients B6, R1, and L2.

Patient	Overall Accuracy			
	Video 1	Video 2	Video 3	Patient accuracy
B6	0.873	0.876	0.941	0.896
R1	0.938	0.952	0.901	0.930
L2	0.913	0.926	0.907	0.915

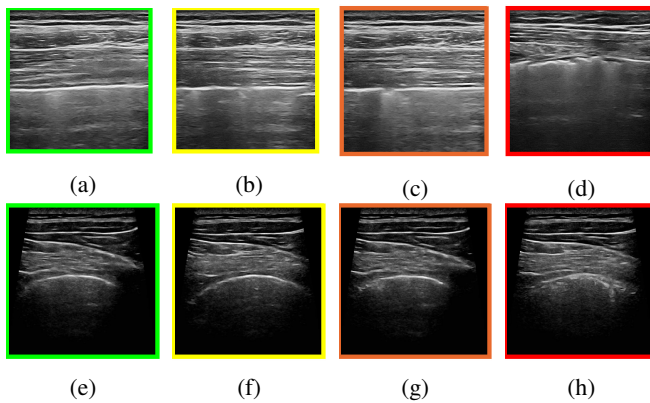


Fig. 8: Examples of the scoring procedure based on the SVM classifier. (a)-(d) data are acquired by the linear probe, (e)-(h) data are acquired by the convex probe. The columns and the color of the box around the image indicates the predicted score value. Colors green, yellow, orange, and red correspond to score 0, 1, 2, and 3, respectively.

evaluating the distribution of the predicted scores for each video (i.e., histogram of the set S). The first three columns of Fig. 9 report the distribution of the scores for the three patients and for each video separately. The predicted score and reference score are reported in blue and in red, respectively. The last column of Figure 9 reports the histograms of the scoring as a result of the joint analysis of the three videos for each patient. The individual distribution of the predicted scores (i.e., one video) is a potential indicator of the average health status of the patient in the anatomical region where the video has been acquired. The sum of the scoring distribution indicates the overall health status with respect to COVID-19 LUS manifestations of the patient.

IV. CONCLUSION

Lung ultrasound imaging is a very promising technology for identifying and monitoring patients affected by COVID-19. The development of detection methods and techniques capable to automatically perform a fast and accurate diagnosis, without requiring the assistance of a trained expert, is still at an early stage. In this paper, we have proposed a method for: i) an unsupervised and automatic detection of the pleural line and the extraction of its geometric and intensity characteristics and ii) an automatic and supervised classification of each LUS image in terms of score (i.e. severity of COVID-19 pulmonary manifestation).

For the task of the pleural line detection, we exploited a method based on a HMM and the Viterbi Algorithm. The pleural line and its underneath region information extracted by

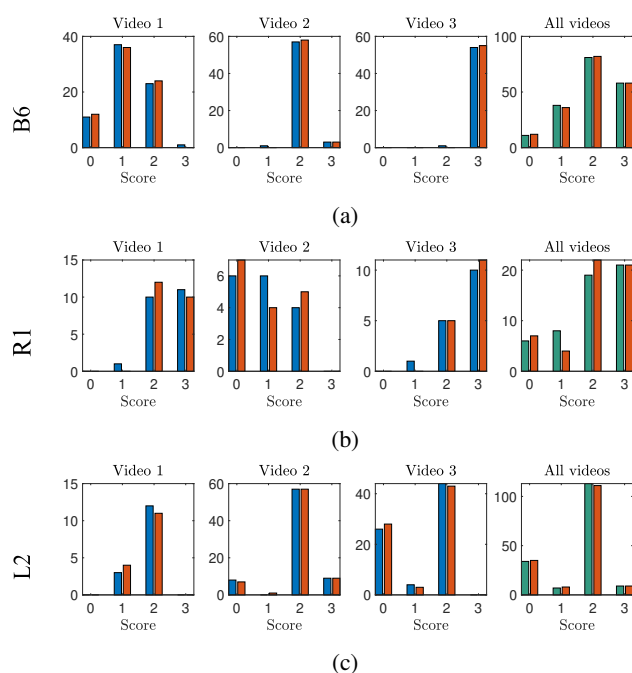


Fig. 9: Scoring histograms of the frame of each video for patients (a) B6, (b) R1, (c), and L2. The first three columns show the distribution of the score values for each video. The last column shows the histogram of the scoring evaluated over three videos. Red bars represent the reference data, while the predicted labels are in blue and green.

the method are then modeled by ad-hoc metrics (i.e., features) that are given as input to the automatic scoring classification procedure.

The proposed scoring is based on a supervised SVM classifier that is applied to each image after the extraction of a proper feature vector.

The proposed method has been tested on the Italian COVID-19 Lung Ultrasound Data Base. The dataset is very heterogeneous being acquired in different hospitals (thus different operators) and with different probes (i.e., linear and convex). The data are accompanied by the scoring information of each image that we exploit as reference dataset. The experimental results show that the pleural line detector can accurately retrieve the pleural line characteristics (i.e., pixel by pixel geometric position and associated intensity) in an image with an overall accuracy of 92% and 84% for linear probes and convex probes, respectively. The SVM classifier provides an image-by-image evaluation of COVID-19 related LUS patterns, accompanied by the predicted score value with an average accuracy of 94% and 88% for linear and convex probes, respectively. The scoring results can both be visually displayed on each image and also provided in a histogram indicating the distribution of the predicted scores for the entire set of acquisitions of a patient in a video for further medical evaluation.

As a final remark, the proposed method has the potential for real-time implementation given the relatively low complexity of the proposed algorithms. The method has been implemented in MATLABTM environment without any particular optimiza-

tion. For each image, the algorithm identifies the pleural line in less than two seconds on a DellTM XPSTM 9550 laptop. The algorithm for the scoring was tested on an ASUSTM F555UJ-XX006TTM laptop. The feature extraction on a frame and the testing phase of a sample take each less than two seconds, while the learning phase time depends on the number of samples. In our case, the learning time was in the range of few hours. However, this is an offline phase that is required only in the set up of the method. This opens the possibility of implementing the method directly on the ultrasound scanners, therefore, providing the user with both tools that support a post-acquisition diagnosis and also a stream of enhanced information during the acquisition process.

As future work, we plan to extend the validation of the proposed method and the application to automatically identify other pulmonary diseases based on their appearance in LUS data.

ACKNOWLEDGMENT

Authors want to acknowledge the financial contribution of the VRT Foundation for this research (COVID-19 Call 2020 Grant #1).

REFERENCES

- [1] T. L. Szabo, "Chapter 1 - introduction," in *Diagnostic Ultrasound Imaging: Inside Out (Second Edition)*, T. L. Szabo, Ed. Boston: Academic Press, 2014, pp. 1 – 37. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B978012396487800001X>
- [2] —, "Appendix b," in *Diagnostic Ultrasound Imaging: Inside Out (Second Edition)*, T. L. Szabo, Ed. Boston: Academic Press, 2014, pp. 785 – 786. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/B97801239648780000306>
- [3] M. Demi, R. Prediletto, G. Soldati, and L. Demi, "Physical mechanisms providing clinical information from ultrasound lung images: hypotheses and early confirmations," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 2019.
- [4] G. Soldati, M. Demi, A. Smargiassi, R. Inchingolo, and L. Demi, "The role of ultrasound lung artifacts in the diagnosis of respiratory diseases," *Expert Review of Respiratory Medicine*, vol. 13, no. 2, pp. 163–172, 2019, pMID: 30616416. [Online]. Available: <https://doi.org/10.1080/17476348.2019.1565997>
- [5] D. Lichtenstein, G. Mézière, P. Biderman, A. Gepner, and O. Barré, "The comet-tail artifact," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, no. 5, pp. 1640–1646, 1997, pMID: 9372688. [Online]. Available: <https://doi.org/10.1164/ajrcm.156.5.96-07096>
- [6] R. Copetti, G. Soldati, and P. Copetti, "Chest sonography: a useful tool to differentiate acute cardiogenic pulmonary edema from acute respiratory distress syndrome," *Cardiovascular ultrasound*, vol. 6, pp. 16–16, Apr 2008, 18442425[pmid]. [Online]. Available: <https://doi.org/10.1186/1476-7120-6-16>
- [7] E. Picano and P. A. Pellikka, "Ultrasound of extravascular lung water: a new standard for pulmonary congestion," *European heart journal*, vol. 37, no. 27, pp. 2097–2104, Jul 2016, 27174289[pmid]. [Online]. Available: <https://doi.org/10.1093/eurheartj/ehw164>
- [8] L. Gargani, M. Doveri, L. D'Errico, F. Frassi, M. L. Bazzichi, A. D. Sedie, M. C. Scali, S. Monti, S. Mondillo, S. Bombardieri, D. Caramella, and E. Picano, "Ultrasound lung comets in systemic sclerosis: a chest sonography hallmark of pulmonary interstitial fibrosis," *Rheumatology*, vol. 48, no. 11, pp. 1382–1387, 08 2009. [Online]. Available: <https://doi.org/10.1093/rheumatology/kep263>
- [9] G. Soldati, M. Demi, R. Inchingolo, A. Smargiassi, and L. Demi, "On the physical basis of pulmonary sonographic interstitial syndrome," *Journal of Ultrasound in Medicine*, vol. 35, no. 10, pp. 2075–2086, 2016. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.7863/ultra.15.08023>
- [10] L. Demi, T. Egan, and M. Muller, "Lung ultrasound imaging, a technical review," *Applied Sciences*, vol. 10, no. 2, p. 462, Jan 2020. [Online]. Available: <http://dx.doi.org/10.3390/app10020462>

- [11] X. Zhang, T. Osborn, B. Zhou, D. Meixner, R. Kinnick, B. Bartholmai, J. Greenleaf, and S. Kalra, "Lung ultrasound surface wave elastography: A pilot clinical study," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, pp. 1298–1304, 09 2017.
- [12] F. Corradi, C. Brusasco, A. Vezzani, G. Santori, T. Manca, L. Ball, F. Nicolini, T. Gherli, and V. Brusasco, "Computer-aided quantitative ultrasonography for detection of pulmonary edema in mechanically ventilated cardiac surgery patients," *CHEST*, vol. 150, no. 3, pp. 640–651, Sep 2016. [Online]. Available: <https://doi.org/10.1016/j.chest.2016.04.013>
- [13] L. Gargani, "Lung ultrasound: a new tool for the cardiologist," *Cardiovascular ultrasound*, vol. 9, pp. 6–6, Feb 2011, 21352576[pmid]. [Online]. Available: <https://doi.org/10.1186/1476-7120-9-6>
- [14] —, "Ultrasound of the lungs: More than a room with a view," *Heart Failure Clinics*, vol. 15, no. 2, pp. 297 – 303, 2019, imaging the Failing Heart. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1551713618301223>
- [15] R. Moshavegh, K. L. Hansen, H. Møller-Sørensen, M. B. Nielsen, and J. A. Jensen, "Automatic detection of b-lines in *in vivo* lung ultrasound," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 66, no. 2, pp. 309–317, 2018.
- [16] R. J. van Sloun and L. Demi, "Localizing b-lines in lung ultrasonography by weakly-supervised deep learning, in-vivo results," *IEEE Journal of Biomedical and Health Informatics*, 2019.
- [17] M. Correa, M. Zimic, F. Barrientos, R. Barrientos, A. Román-Gonzalez, M. J. Pajuelo, C. Anticona, H. Mayta, A. Alva, L. Solis-Vasquez *et al.*, "Automatic classification of pediatric pneumonia based on lung ultrasound pattern recognition," *PloS one*, vol. 13, no. 12, 2018.
- [18] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli *et al.*, "Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound," *IEEE Transactions on Medical Imaging*, 2020.
- [19] N. Anantrasirichai, W. Hayes, M. Allinovi, D. Bull, and A. Achim, "Line detection as an inverse problem: application to lung ultrasound imaging," *IEEE transactions on medical imaging*, vol. 36, no. 10, pp. 2045–2056, 2017.
- [20] S. Kulhare, X. Zheng, C. Mehanian, C. Gregory, M. Zhu, K. Gregory, H. Xie, J. M. Jones, and B. Wilson, "Ultrasound-based detection of lung abnormalities using single shot detection convolutional neural networks," in *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation*. Springer, 2018, pp. 65–73.
- [21] F. Mojoli, B. Bouhemad, S. Mongodi, and D. Lichtenstein, "Lung ultrasound for critically ill patients," *American journal of respiratory and critical care medicine*, vol. 199, no. 6, pp. 701–714, 2019.
- [22] Q.-Y. Peng, X.-T. Wang, L.-N. Zhang, C. C. C. U. S. Group *et al.*, "Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic," *Intensive care medicine*, p. 1, 2020.
- [23] G. Soldati, A. Smargiassi, R. Inchingolo, D. Buonsenso, T. Perrone, D. F. Briganti, S. Perlini, E. Torri, A. Mariani, E. E. Mossolani *et al.*, "Is there a role for lung ultrasound during the covid-19 pandemic?" *Journal of ultrasound in medicine: official journal of the American Institute of Ultrasound in Medicine*, 2020.
- [24] —, "Proposal for international standardization of the use of lung ultrasound for covid-19 patients; a simple, quantitative, reproducible method," *Journal of Ultrasound in Medicine*, 2020.
- [25] W. H. Organization *et al.*, "Coronavirus disease 2019 (covid-19): Situation report–95," 2020. [Online]. Available: <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200424-sitrep-95-covid-19.pdf>
- [26] L. Carrer and L. Bruzzone, "Automatic enhancement and detection of layering in radar sounder data based on a local scale hidden markov model and the viterbi algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 962–977, 2016.
- [27] A.-M. Ilisei and L. Bruzzone, "A system for the automatic classification of ice sheet subsurface targets in radar sounder data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3260–3277, 2015.
- [28] O. T. Karlsen, R. Verhagen, and W. M. Bovée, "Parameter estimation from rician-distributed data sets using a maximum likelihood estimator: application to t1 and perfusion measurements," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 41, no. 3, pp. 614–623, 1999.
- [29] J. Sijbers, A. J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck, "Estimation of the noise in magnitude mr images," *Magnetic resonance imaging*, vol. 16, no. 1, pp. 87–90, 1998.
- [30] R. Maitra and D. Faden, "Noise estimation in magnitude mr datasets," *IEEE Transactions on Medical Imaging*, vol. 28, no. 10, pp. 1615–1622, 2009.
- [31] M. Zanetti, F. Bovolo, and L. Bruzzone, "Rayleigh-rice mixture parameter estimation via em algorithm for change detection in multispectral images," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5004–5016, 2015.
- [32] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [33] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on geoscience and remote sensing*, vol. 42, no. 8, pp. 1778–1790, 2004.
- [34] A. Ben-Hur and J. Weston, "A user's guide to support vector machines," in *Data mining techniques for the life sciences*. Springer, 2010, pp. 223–239.
- [35] (2020, Mar.) Italian covid-19 lung ultrasound data base (ICLUS-DB). [Online]. Available: <https://covid19.disi.unitn.it/iclusdb/login>