Semi-Automatic A-Line Detection and Confidence Scoring in Lung Ultrasound

Oriane Thiery, Garance Martin, Isabelle Bloch, Martin Dres, Umar Saleem, Andrea Pinna

Semi-Automatic A-Line Detection and Confidence Scoring in Lung Ultrasound

Oriane Thiery
Sorbonne Université, CNRS, LIP6
Paris, France
oriane.thiery@lip6.fr

Garance Martin
Sorbonne Université, CNRS, LIP6
Paris, France
garance.martin@lip6.fr

Isabelle Bloch
Sorbonne Université, CNRS, LIP6
Paris, France
isabelle.bloch@lip6.fr

Martin Dres
Médecine intensive - Réanimation (Département R3S), APHP, Sorbonne Université, Hôpital Pitié-Salpêtrière
Neurophysiologie respiratoire expérimentale et clinique, INSERM UMR_S 1158, Sorbonne Université
Paris, France - martin.dres@aphp.fr

Umar Saleem
Bioserenity
Paris, France
umar.saleem@bioserenity.com

Andrea Pinna
Sorbonne Université, CNRS, LIP6
Paris, France
andrea.pinna@lip6.fr

***Abstract*—Weaning a patient from mechanical ventilation is a critical task in Intensive Care Units, but it can be made safer by using Lung Ultrasound scoring. This scoring is currently done visually by specialists based on Lung Ultrasound artifacts among which are A-lines. Automating this scoring may help standardizing results and saving time for clinicians. In this paper, we propose a method to automatically detect A-lines by using both the intensity profile of the LUS image and morphological operations. A score is then assigned to significant lines and represents the possibility of them being A-lines. The proposed method shows promising results in differentiating Alines from other elements with an Area Under the Curve of 0.95; furthermore, using a threshold at 0.5 to detect A-lines leads to very good performances with an accuracy of 95% and a F0.5- score of 0.84.**

***Index Terms*—Lung Ultrasound, A-lines, automatic detection, confidence estimation**

I. Introduction

Knowing when to wean a patient from mechanical ventilation is a critical task in Intensive Care Units (ICU): waiting too long makes the patient less likely to manage natural breathing, and weaning him too quickly is associated with an increased risk of postextubation distress. One way to predict whether the patient is ready for the weaning is to use Lung Ultrasound (LUS) to estimate aeration changes during the spontaneous breathing trial [1]. The estimation of aeration changes in the lung is based on the scoring of LUS images according to the following criteria: 1) normal aeration: presence of lung sliding with A-lines (thin, horizontal or slightly curved artifacts), or less than two isolated B-lines (vertical artifacts originating from the pleural line); 2) moderate loss of lung aeration: multiple, well-defined B-lines; 3) severe loss of lung aeration: multiple coalescent B-lines; and 4) lung consolidation: presence of a tissue pattern characterized by dynamic air bronchograms [1]. However, this scoring, being done visually, requires both formation and experience for the clinician, and may vary depending on the operator. Therefore, image processing methods that identify the significant lines on LUS images could be a great help for both novice practitioners and for specialists to save time.

These last years, there has been a promising increase in the research for automating the evaluation of LUS images because (among other reasons) of the outbreak of COVID-19, which can be detected by LUS [2], [3]. However, while some research proposes automatic scoring of the LUS images [4]–[8], none of these methods is based on the scoring system used to predict the risk of postextubation distress. Some other approaches are directed towards the identification of special patterns on LUS images, but most of this research is focused on the detection and evaluation of B-lines, these being one of the most significant elements to estimate the state of a lung in LUS images [8]–[18]. Only few approaches try to segment the A-lines, mostly by using the Radon transform of the LUS image [11]–[13], or a morphological top-hat operation followed by an adaptive filter [19]. These methods have not been evaluated quantitatively on their performances in A-lines detection, either because it was not the main objective of the study, or because it was a more qualitative evaluation.

In this context, we propose a method to identify the A-lines in LUS images in a flexible way: instead of a binary estimation of the presence of an A-line, the method presented in this paper gives a level of confidence related to the possibility of a significant line detected in the LUS image being an A-line. This allows clinicians to know how much confidence they can have in the given result and whether or not they should double-check it.

This article is organized as follows: Section II describes the proposed method, Section III presents the used dataset, and finally, Section IV discusses the results of the testing phase.

II. Method

This section describes the proposed method for A-lines detection and confidence estimation. Before anything else, pre-

processing is necessary in order to remove any annotation added by the imaging ultrasound machine on the LUS frame, to rescale the intensity values between 0 and 255, and to crop the frame to keep only the smallest rectangular image containing all the LUS information. The result of this operation is presented in Figure 1a. In our dataset, A-lines appear as arcs of a circle of 3-7 pixels thickness. Our method for their detection is thus based on the detection of bright arcs of a circle, which are considered A-line candidates, and the evaluation of three of their characteristics: their position, the contrast with their neighborhood and the continuity of the brightness in the detected line.

In order to study these characteristics, five steps are considered: pleural line detection, study of the vertical intensity profile of the LUS image. morphological operations to enhance potential lines in the image, scoring of each candidate line, and final decision among the candidates lines using a threshold. These steps are described hereafter.

1) ***Pleural line detection****:* as the following steps rely on the pleural line position to estimate if the detected line position makes it likely to be an A-line, the pleural line is identified by the clinician (cf Figure 1b). Doing it manually prevents errors on the detection of this element which is crucial to identify the A-lines, and in extreme cases, where the pleural line cannot be detected in the image, it leads to the immediate conclusion that the LUS image contains no A-lines (because they are reflections of the pleural line that cannot occur in the absence of pleural line).

From this point on, everything is done automatically.

2) ***LUS vertical intensity profile****:* in order to determine A-line candidates, the intensity profile of the LUS image is studied so as to identify any peak of brightness (relative to its surroundings). To this end, the pleural line being identified, the area under it is scanned to get the profile of the mean intensity of each arc of a circle having for center the peak of the LUS image cone (cf Figure 1c). The resulting profile is smoothed with a Gaussian filter ($\sigma = 5$) to remove the high-frequency variations due to the noisy nature of the LUS images, while preserving the lines (cf Figure 1c). Every maximum detected on the final curve is a candidate line for being an A-line.

3) ***Morphological operations****:* In order to evaluate if the candidate lines have not been detected thanks to a localized peak of brightness (and that they do present a long and thin bright intensity profile), some morphological operations are applied to the LUS frame to highlight long and thin elements. First, the contrast of the LUS frame is greatly increased by setting to 0 every pixel whose value is under the 98th percentile of intensity of the image, and rescaling the rest between 0 and 255 (cf Figure 1d). Then, an Alternate Sequential Filter (ASF) is applied to this image (cf Fig. 1e). This step consists in alternating morphological openings and closings with a structuring element of growing size. Here, the structuring element is chosen as an horizontal line (because most of the A-lines, independently of their orientation, contain at least an horizontal part) of length increasing from 2 to 30 pixels. To identify more precisely the long bright regions characteristic of A-lines, an opening with a structuring element of size 5×40 is then applied on the result of the ASF (cf Figure 1f).

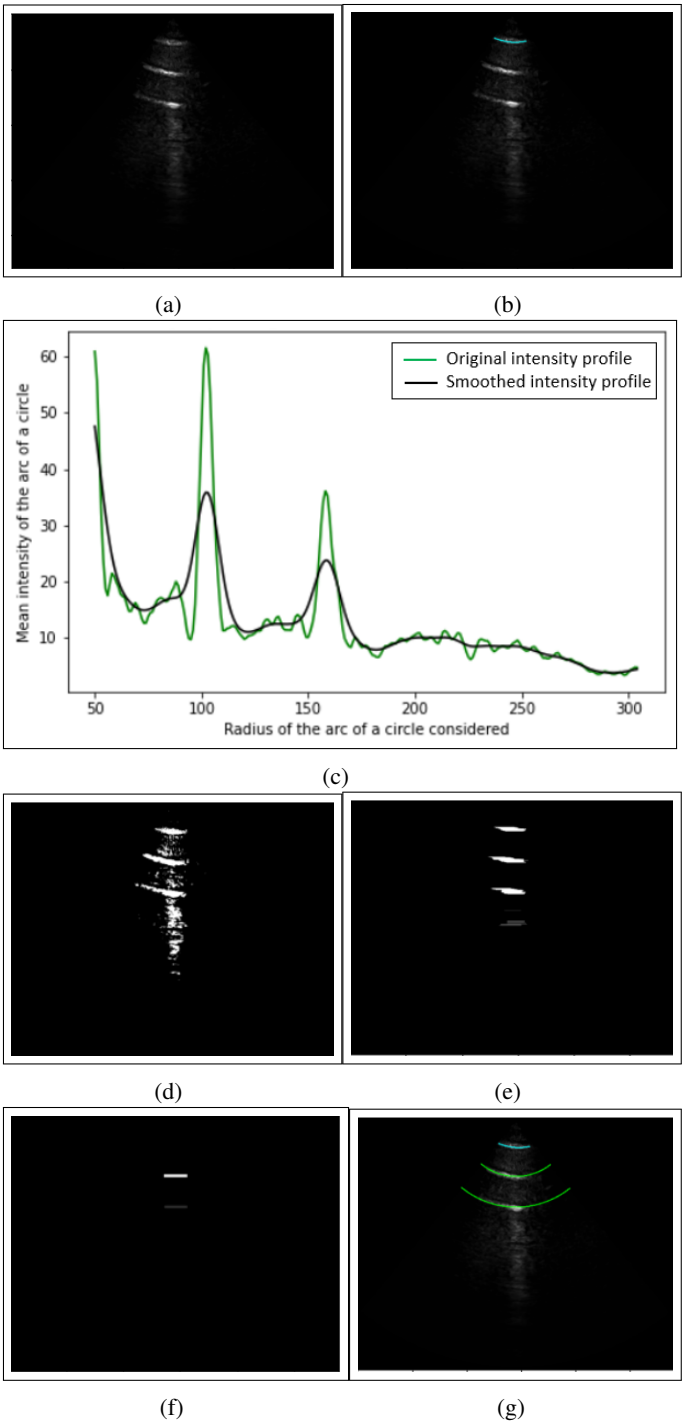

Fig. 1: Steps of the proposed method: (a) Pre-processed image; (b) Pleural line (in blue) identification; (c) Original and smoothed intensity profile of the image; (d) Image after contrast adjustment; (e) Image after the ASF operation; (f) Image after the opening operation; (g) Result of the detection with the detected A-lines displayed in green.

*4) **Scoring**:* To estimate the possibility that the candidate lines are A-lines, both the intensity profile of the frame and the result of the morphological operations are used. A score between 0 and 1 is given to each line for each of the three parameters taken into account, and these scores are then aggregated. The three considered parameters are:

• position of the candidate line: it corresponds to the radius of the arc of a circle considered as a candidate line. As this position p should be a multiple of the distance between skin and the pleural line (denoted by d_{sp} here, and typically between 40 and 130 pixels in our dataset), the score s_1 given to this parameter is determined by Algorithm 1 (with m being an imprecision margin considered on each side of the position of the pleural line, equal to 3 pixels).

Algorithm 1 Scoring of the position of a candidate line

Input: p, m, d_{sp}
Output: s_1
$s_1 \leftarrow 0$
for k in [2,3,4] **do**
  if $p \geq k(d_{sp} - m)$ & $p \leq k(d_{sp} + m)$ **then**
    $s_1 \leftarrow 1$
  else if $p > k(d_{sp} + m)$ & $p < (k + 0.5)d_{sp}$ **then**
    $s_1 \leftarrow \frac{p - d_{sp}(k+0.5)}{-0.5d_{sp} + km}$
  else if $p < k(d_{sp} - m)$ & $p > (k - 0.5)d_{sp}$ **then**
    $s_1 \leftarrow \frac{p - d_{sp}(k-0.5)}{0.5d_{sp} - km}$
  end if
end for

The score is 1 when the position of the A-line is a multiple of the distance between the skin and the pleural line, and it decreases as the line goes away from these positions until it reaches 0;

• contrast between the line and its neighborhood: it is assessed through the derivative of the intensity profile, both before (score $s_{2,1}$) and after (score $s_{2,2}$) the position of the candidate line. This derivative is calculated on the smoothed profile of intensity of the LUS image, and its value at five pixels before the position of the candidate line and at five pixels after are studied. In order to ensure that the line is bright enough compared to its surroundings, the higher in absolute value those two derivatives are, the higher the score associated to this parameter is. The function giving the score to one of the derivative d depending on its absolute value has been determined experimentally and is detailed in Algorithm 2;

• result of the ASF and opening operations: they are used to ensure that the line is continuous and has not been detected thanks to a single peak of brightness. Let ASF_{pres} and O_{pres} be Boolean values that indicate whether something has been detected after the ASF and the opening operations, respectively (note that $O_{pres} \leq ASF_{pres}$). Furthermore, we define L as the length in pixels of the detected element (in the opening if something has been detected there or in the ASF otherwise) and H its height. Considering these notations, the score s_3 given to this parameter is determined by Algorithm 3.

Algorithm 2 Scoring of the contrast of a candidate line

Input: d
Output: s_2
if $|d| < 0.15$ **then**
  $s_2 \leftarrow 0$
else if $|d| < 0.5$ **then**
  $s_2 \leftarrow 0.57 \times |d| - 0.085$
else if $|d| < 1.25$ **then**
  $s_2 \leftarrow 1.067 \times |d| - 0.33$
else
  $s_2 \leftarrow 1$
end if

Algorithm 3 Scoring of the result of the morphological operations for a given candidate line

Input: $O_{pres}, ASF_{pres}, L, H$
Output: s_3
if O_{pres} **then**
  $s_3 \leftarrow 0.6$
else if ASF_{pres} **then**
  $s_3 \leftarrow 0.5$
else
  $s_3 \leftarrow 0.4$
end if
if $L > 100$ **then**
  $s_3 \leftarrow s_3 + 0.4$
else if $L > 80$ **then**
  $s_3 \leftarrow s_3 + 0.3$
else if $L > 60$ **then**
  $s_3 \leftarrow s_3 + 0.2$
end if
if $L/H < 3$ **then**
  $s_3 \leftarrow \max(s_3([L/H] - 2), 0)$
end if

This algorithms means that if nothing has been detected by the ASF at the position of the line candidate (every pixel of the line in the image after the ASF operation is equal to 0) then the score is 0.4: this is because experience has shown that some A-lines are not horizontal enough to be highlighted by the ASF, and thus the absence of detection after the ASF operation does not mean that there is no A-line at the considered position. If something has been detected at the position of the line after the ASF operation but disappeared after the opening, then the score is a little higher; and if something has been detected after the ASF and is still perceptible after the opening, then the score increases once again. After this first score is given to the considered line, the shape of the detected element is taken into account: if it is long, like an A-line should be, then the score is increased, and it is decreased if it is too thick (because it means that a bright extensive region is located here, which does not match the profile of an A-line).

In order to conclude on the number and position of the A-lines in the frame, the scores are then aggregated by a weighted average operation that gives the confidence level c (Equation 1) that the line is an A-line. This operation has been chosen because some A-lines do not respect one of their theoretical characteristics (e.g. their position is slightly off from where it should be). In this case, this operation allows for compensation

if the other parameters are strong enough to make it clear that the detected line is an A-line in spite of that. The scores are aggregated as follows:

$$c = (0.5s_1 + s_{2,1} + s_{2,2} + s_3)/3.5 \quad (1)$$

This gives a lot of influence to whether the line is contrasted enough from its neighborhood; and a little less influence to the position in order to limit false detections which only have a truly accurate position. Furthermore, if the line is distant from the skin more than five times the distance between the pleural line and the skin, c is set to 0 as there is almost never an A-line at this position.

*5) **Validation**:* The final step to identify A-lines is to consider three regions (because there is almost never more than three A-lines in a LUS image). Each region is centered around the depth of a multiple of the distance between the skin and the pleural line. This is to ensure that there is at most one A-line detected in each region: thus, for each region, only the candidate line with the best score is kept. This line is considered to be an A-line if the confidence level which has been attributed to it is higher than a threshold τ (cf Figure 1g obtained for $\tau = 0.5$). This validation step, for $\tau = 0.5$, enables the eliminate 3.66 false detections per image in average, while only discarding 0.19 true detections per image.

In the example in Figure 1, the confidence level of the two detected lines are both 0.89, and the confidence level of the line detected in the third region is 0.25, which means that this candidate line is not an A-line, as it can been seen in Figure 1a.

III. Dataset

The dataset used to test the proposed method includes LUS images from seven patients admitted to the ICU of the hospital La Pitié-Salpêtrière in Paris, which have undergone a successful spontaneous breathing trial and as such have been extubated afterwards. The clinical information of these patients (mean values) is detailed in Table I.

TABLE I: Patient data.

No. of patients	Sex	Age	BMI (kg/m²)	MV[a] (days)
7	M(4)-F(3)	52	24	7

[a]MV: Mechanical Ventilation

For each patient, LUS has been conducted just before the extubation as well as 2 hours after the extubation, 24 hours and 48 hours after the extubation if the patient was still at the ICU, and at the time of his discharge if he was discharged before 48 hours.

Lung Ultrasound was performed with a Philips Sparq, 2- to 4-MHz convex probe (cardiac probe) set to 40 frames/s. Each LUS sequence contains 81 frames and lasts for about 2 seconds. Each intercostal space of upper and lower parts of the anterior, lateral, and posterior regions of the left and right chest wall was carefully examined, for a total of 12 LUS sequences by patient and time of acquisition. This amount to a total of 244 sequences and 19763 LUS frames.

IV. Results and discussion

The proposed method has been applied on one image of every sequence of the aforementioned dataset, and has been judged on the score given to every candidate line. Thus, a true positive here is an A-line whose score is higher than τ and a false positive is a candidate line which is not an A-line but has a score higher than τ as well. With these definitions, the ROC (Receiver Operating Characteristic) curve representing the sensitivity of the method against its specificity shows an area under the curve of 0.95 (cf Figure 2), demonstrating that the proposed scoring method has a great capacity to distinguish between A-lines and other elements.

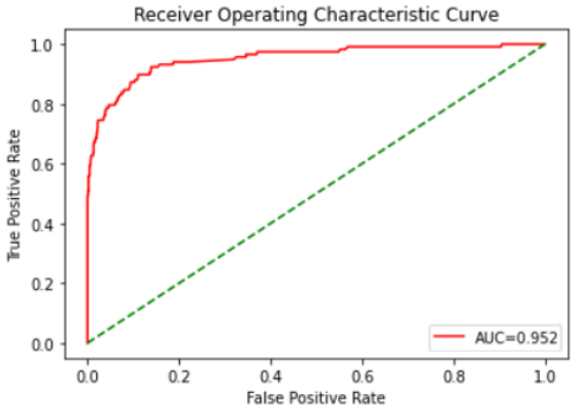

Fig. 2: ROC curve and its Area Under the Curve (AUC).

Because A-lines are patterns indicating a healthy lung, it is more important not to detect them wrongly than to detect all of them. Thus, to evaluate the choice of a threshold of 0.5 for the A-lines detection we use the F0.5-score, which corresponds to the harmonic mean between the precision and the recall by considering the precision to be two times as important as the recall. The performance of the method with $\tau = 0.5$ for this measure is 0.84, and its accuracy is 95%, proving this choice of threshold to be relevant for the given situation.

The proposed method is, to our knowledge, the only one that has been quantitatively evaluated on A-lines detection. It has been developed with the intention of being used for an automatic classification estimating the aeration of the lung, and to this end, is based on the scoring of the level of confidence that can be given to the fact that the detected line is an A-line. Indeed, for future works such as developing a classifier for the whole LUS scoring, having such an information may allow the classifier to estimate more accurately the score to give to the LUS image. Finally, the proposed method obtained great results on the task of A-lines scoring and detection, that we hope to confirm on an independent cohort of future patients.

Acknowledgment

The authors are thankful to Dr Vincent Joussellin and Dr Vincent Bonny who acquired the data at the hospital La Pitié-Salpêtrière (AP-HP) under the protocol NCT04180410 promoted by Bioserenity.

REFERENCES

[1] A. Soummer, S. Perbet, H. Brisson, C. Arbelot, J.-M. Constantin, Q. Lu, and J.-J. Rouby, "Ultrasound assessment of lung aeration loss during a successful weaning trial predicts postextubation distress," *Critical Care Medicine*, vol. 40, p. 2064–72, May 2012.

[2] C. McDermott, M. Lacki, B. Sainsbury, J. Henry, M. Filippov, and C. Rossa, "Sonographic diagnosis of COVID-19: A review of image processing for lung ultrasound," *Frontiers in Big Data*, vol. 4, p. 612561, Mar 2021.

[3] L. Zhao and M. A. Lediju Bell, "A review of deep learning applications in lung ultrasound imaging of COVID-19 patients," *BME Frontiers*, vol. 2022, Feb 2022.

[4] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli, E. Peschiera, R. Trevisan, G. Maschietto, E. Torri, R. Inchingolo, A. Smargiassi, G. Soldati, P. Rota, A. Passerini, and L. Demi, "Deep learning for classification and localization of Covid-19 markers in Point-of-Care lung ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, pp. 2676–2687, May 2020.

[5] Z. Hu, Z. Liu, Y. Dong, J. Liu, B. Huang, A. Liu, J. Huang, X. Pu, X. Shi, J. Yu, Y. Xiao, H. Zhang, and J. Zhou, "Evaluation of lung involvement in Covid-19 pneumonia based on ultrasound images," *BioMedical Engineering OnLine*, vol. 20, p. 27, Mar 2021.

[6] F. Mento, T. Perrone, A. Fiengo, A. Smargiassi, R. Inchingolo, G. Soldati, and L. Demi, "Deep learning applied to lung ultrasound videos for scoring Covid-19 patients: A multicenter study," *The Journal of the Acoustical Society of America*, vol. 149, p. 3626–3634, May 2021.

[7] L. Carrer, E. Donini, D. Marinelli, M. Zanetti, F. Mento, E. Torri, A. Smargiassi, R. Inchingolo, G. Soldati, L. Demi, F. Bovolo, and L. Bruzzone, "Automatic pleural line extraction and Covid-19 scoring from lung ultrasound data," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, p. 2207–2217, Nov 2020.

[8] C. Baloescu, G. Toporek, S. Kim, K. McNamara, R. Liu, M. M. Shaw, R. L. McNamara, B. I. Raju, and C. L. Moore, "Automated lung ultrasound B-line assessment using a deep learning algorithm," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, p. 2312–2320, Nov 2020.

[9] R. Moshavegh, K. L. Hansen, H. M. Sørensen, M. C. Hemmsen, C. E. M.D., M. B. Nielsen, and J. A. Jensen, "Novel automatic detection of pleura and B-lines (comet-tail artifacts) on in vivo lung ultrasound scans," in *Medical Imaging 2016: Ultrasonic Imaging and Tomography* (N. Duric and B. Heyde, eds.), vol. 9790, pp. 157 – 163, International Society for Optics and Photonics, SPIE, 2016.

[10] 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, p. 309–317, Feb 2019.

[11] O. Karakuş, N. Anantrasirichai, A. Aguersif, S. Silva, A. Basarab, and A. Achim, "Detection of line artefacts in lung ultrasound images of Covid-19 patients via non-convex regularization," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, p. 2218–2229, Nov 2020.

[12] N. Anantrasirichai, M. Allinovi, W. Hayes, and A. Achim, "Automatic B-line detection in paediatric lung ultrasound," in *2016 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4, 2016.

[13] 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, p. 2045–2056, Oct 2017.

[14] C. Brusasco, G. Santori, E. Bruzzo, R. Tro, C. Robba, G. Tavazzi, F. Guarracino, F. Forfori, P. Boccacci, and F. Corradi, "Quantitative lung ultrasonography: A putative new algorithm for automatic detection and quantification of B-lines," *Critical Care*, vol. 23, Dec 2019.

[15] W. F. Weitzel, J. Hamilton, X. Wang, J. L. Bull, A. Vollmer, A. Bowman, J. Rubin, G. H. Kruger, J. Gao, M. Heung, and P. Rao, "Quantitative lung ultrasound comet measurement: Method and initial clinical results," *Blood Purification*, vol. 39, no. 1–3, p. 37–44, 2015.

[16] L. J. Brattain, B. A. Telfer, A. S. Liteplo, and V. E. Noble, "Automated B-line scoring on thoracic sonography," *Journal of Ultrasound in Medicine*, vol. 32, no. 12, p. 2185–2190, 2013.

[17] R. J. G. 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*, vol. 24, p. 957–964, Apr 2020.

[18] X. Wang, J. S. Burzynski, J. Hamilton, P. S. Rao, W. F. Weitzel, and J. L. Bull, "Quantifying lung ultrasound comets with a convolutional neural network: Initial clinical results," *Computers in Biology and Medicine*, vol. 107, p. 39–46, 2019.

[19] H. Susanti and Suprijanto, "Image processing framework for pleural line (A-line) detection in video lung ultrasonography," in *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, p. 99–102, Mar 2021.