Automatic quantification of pleural effusion volume in lung ultrasound: an AI-based approach

A. Bottino[1], C. Botrugno[2], E. Casciaro[2], F. Conversano[2], A. Lay-Ekuakille[1], N.I. Giannoccaro[1], R. Morello[2] and S. Casciaro[2]

***Abstract*— Pleural effusion is a clinical condition characterized by abnormal fluid collection within the pleural space. One of the most common treatments, especially in critical patients, is thoracentesis consisting of removing the excess fluid through a needle not only to restore a normal lung expansion, but also to analyze the pleural fluid to determine the underlying cause. Thanks to its high temporal resolution, Lung Ultrasound (LUS) not only offers a crucial real-time guidance for needle insertion during thoracentesis, but it allows a rapid and non-invasive evaluation of pleural effusion volume. However, while LUS enables for rapid assessment at the bedside, accurate and reproducible quantification of the volume of pleural effusion remains an unmet clinical need. In this work, effusion volumes were reconstructed by integrating segmented areas across longitudinal and transverse scans and compared with thoracentesis output. A modified ResNet-based architecture was trained on manually annotated 2D ultrasound images to segment pleural effusion areas. Automated segmentations were then blindly compared with expert manual delineations. The segmentation algorithm achieved a mean Dice Similarity Coefficient (DSC) of 0.86 ± 0.04, outperforming inter-operator agreement (DSC 0.82 ± 0.05). Starting from two longitudinal and one transverse 2D segmentations, a volume reconstruction and quantification was finally performed. Automated volume estimates showed good agreement with thoracentesis, with a MAE of 31.6 ± 14.2 cm³ and an RMSE of 35.9 cm³. These preliminary results demonstrated that deep learning–based analysis of LUS can provide fast, non-invasive and accurate quantification of pleural effusion volume. This can represent a crucial support in clinical decision-making.**

I. INTRODUCTION

Pleural effusion is a clinical condition characterized by an excessive accumulation of fluid in the pleural cavity. The main symptom is dyspnea, usually related to lung compression due to the presence of fluid in the pleural space. Other well-known symptoms include dry cough and chest pain. Among the main causes are congestive heart failure (CHF), malignancy, pneumonia, and pulmonary embolism. However, in approximately 20% of cases, its etiology remains unknown [1]. The main treatments for massive pleural effusion consist of removing the excess fluid either by needle aspiration (thoracentesis) or by inserting a chest tube (tube thoracostomy) to drain larger volumes. In both cases, the aim is to remove the fluid for diagnostic analysis in order to determine the underlying cause and guide classification. In particular, pleural fluid can be classified based on Light's criteria into transudate or exudate [2]. In this scenario, ultrasonography represents a crucial tool for the management of pleural effusion: on the one hand, it is considered the imaging modality of choice for detecting pleural effusion, due to its ability to detect very small volumes of fluid (potentially even below 20 mL under optimal conditions) compared with other techniques such as chest radiography (CXR), where a minimum of approximately 150–200 mL is generally required for detection in the upright position [3], [4]; on the other hand, thanks to well-known advantages of ultrasound, such as high temporal resolution and the absence of ionizing radiation, it provides essential real-time guidance for needle insertion during thoracentesis [5], [6]. As demonstrated in [7], image-guided thoracentesis reduces the risk of complications such as pneumothorax and increases diagnostic accuracy. Ultrasound also plays a key role in the evaluation of pleural thickening, helping to differentiate benign from malignant disease, as well as in identifying possible underlying etiologies [8]. Being a fluid with relatively homogeneous acoustic properties, pleural effusion appears as a hypoechoic or anechoic area adjacent to the diaphragm on ultrasound imaging. Moreover, unlike a well-aerated lung—where ultrasound waves are largely reflected at the air–tissue interface due to the marked difference in acoustic impedance—ultrasound waves are transmitted through the pleural fluid, enhancing the visualization of deeper thoracic structures, such as the spinal column [9]. These structures become key references for an automated assessment of pleural effusion. Depending on the clinical scenario, ultrasound examination allows the evaluation of pleural effusion through a qualitative analysis (classifying it as mild, moderate, or massive) [10], [11] or a quantitative assessment by estimating its volume [12], [13]. While for large volumes several studies have demonstrated a correlation between the pleural effusion volume (PEV) estimated by ultrasound and the actual volume

*This work was not supported by any organization.

[1]A. Bottino is with the Department of Engineering for Innovation, University of Salento, Lecce, 73100, Italy. alberto.bottino@unisalento.it

[2]C. Botrugno is with the National Research Council – Institute of Clinical Physiology, Lecce, 73100, Italy. chiarabotrugno@cnr.it

[2]E. Casciaro is with the National Research Council–Institute of Clinical Physiology, Lecce, 73100, Italy. ernesto.casciaro@cnr.it

[2]F. Conversano is with the National Research Council–Institute of Clinical Physiology, Lecce, 73100, Italy. conversano@ifc.cnr.it

[1]A. Lay-Ekuakille is with the Department of Engineering for Innovation, University of Salento, Lecce, 73100, Italy. aime.lay.ekuakille@unisalento.it

[1]N.I. Giannoccaro, is with the Department of Engineering for Innovation, University of Salento, Lecce, 73100, Italy. ivan.giannoccaro@unisalento.it

[2]R. Morello is with the National Research Council–Institute of Clinical Physiology, Lecce, 73100, Italy. rocco.morello.ifc@gmail.com

[2]S. Casciaro is with the National Research Council–Institute of Clinical Physiology, Lecce, 73100, Italy. sergio.casciaro@cnr.it

drained [14], [15], [16], for smaller effusions (i.e., less than 250 mL), the evaluation becomes more challenging. In particular, measurement variability increases due to geometric assumptions and operator dependence. Remérand et al. [17] proposed a multiplane approach to improve accuracy even for small pleural effusions, thereby reducing estimation error. Various models have been proposed in the literature for PEV estimation from lung ultrasound images, and different formulas can be found in the literature. Tang et al. [18] compared three main techniques—single-section, two-section, and multi-section models—suggesting that the two-section model offers a favorable balance between accuracy and feasibility in routine clinical practice. Despite these advances, no universally accepted standard formula currently exists, and reported performance varies depending on patient positioning, probe orientation, and operator experience. For these reasons, we propose a deep learning–based framework for fast and accurate pleural effusion volume (PEV) quantification that integrates automated pixel-wise segmentation with multi-view ultrasound volumetric reconstruction and direct comparison with thoracentesis-derived drainage volumes, aiming to reduce dependence on manual measurements and simplified geometric estimation models.

II. MATERIAL AND METHODS

In this study, thirty-six patients with clinically significant pleural effusion underwent lung ultrasound examinations. Ultrasound acquisitions were performed using a convex probe (2–5 MHz) with patients in a seated or semirecumbent position, following a standardized scanning protocol that included both longitudinal and transverse intercostal views of the posterior and lateral thoracic regions. For each hemithorax, multiple frames were acquired at end-expiration to reduce respiratory motion artifacts and improve reproducibility. These acquisitions, yielded a dataset of a total of 108 ultrasound images, extracted from recorded clips and selected on the basis of image quality and clear visualization of the pleural space, where the pleural effusion areas were segmented using a Graphical User Interface (GUI) by an expert operator with more than ten years of experience in thoracic ultrasonography. Manual delineation was performed frame-by-frame, carefully excluding surrounding anatomical structures (e.g., diaphragm, lung parenchyma, ribs) to ensure accurate contouring of the effusion boundaries. The obtained masks represented the ground truth for training a ResNet-based convolutional neural network (ResNet-5 architecture) implemented in MATLAB using the Deep Learning Toolbox [19]. The reference standard for volumetric evaluation was the actual volume of pleural fluid drained during thoracentesis. The drained volume was recorded for each patient immediately after the procedure. It should be noted that, in routine clinical practice, complete evacuation of pleural fluid is not always achieved. For safety reasons (e.g., to reduce the risk of re-expansion pulmonary edema or patient discomfort), a residual volume is commonly left in situ. In our cohort, the estimated residual volume ranged approximately between 50 and 150 mL, depending on clinical conditions and operator judgment. This expected residual amount was considered during interpretation of discrepancies between ultrasound-derived and drained volume measurements.

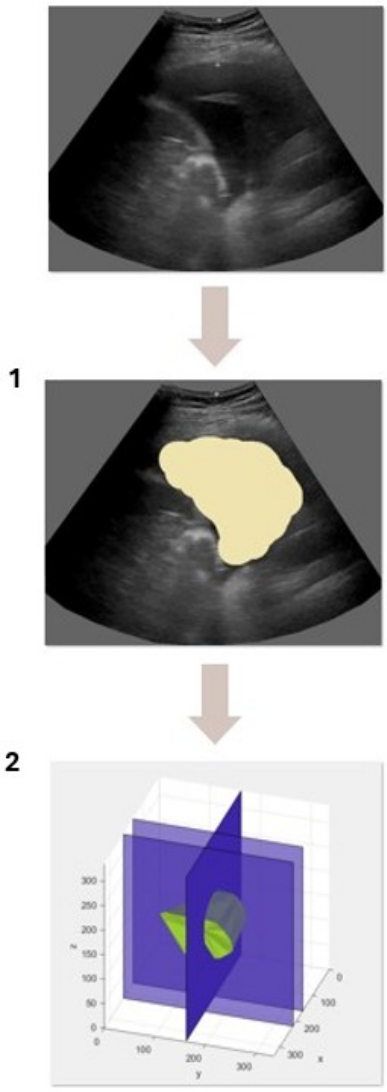

Fig. 1. After an initial dataset preparation step (1), for each patient, 2D segmentation (2) and volume reconstruction (3) were performed.

A. General workflow

Two main steps can be identified: automatic segmentation of pleural effusion on 2D frames and subsequent 3D reconstruction for volume quantification (Fig. 1).

1) *Automatic segmentation of 2D scans.* Preprocessed images from test set were given as input to the trained network and automatically segmented. To evaluate segmentation performance, the output masks were compared with those obtained by manual segmentation of an expert operator.
2) *3D pleural effusion reconstruction and volume quantification.* Starting from the 2D masks automatically obtained, pleural effusion volumes were reconstructed

by integrating segmented areas across longitudinal and transverse scans using a geometric reconstruction approach based on inter-slice distance estimation. Ultrasound-derived volumetric estimates were then compared with thoracentesis-derived drained volumes, considered the clinical reference standard.Agreement estimated volumes and drained volumes was assessed using absolute error metrics, and correlation analysis.

B. Segmentation Algorithm Description

The following workflow was implemented for automated 2D segmentation of pleural effusion:

1) *Dataset preparation.* Prior to training, poor-quality images were excluded. The selected 108 images were resized to match the neural network input dimensions and intensity-normalized. The dataset obtained was then divided into training, validation, and test subsets according to the following percentages: 75%, 15% and 10%, respectively.
2) *Data augmentation.* Data augmentation was applied to increase dataset variability and reduce overfitting. Specifically, random rotation, horizontal flipping, and mild scaling transformations were applied.
3) *Neural Network training.* A ResNet-5 was selected because residual connections improve optimization stability and gradient flow, which is particularly useful in small biomedical datasets. In preliminary experiments, this architecture provided the best trade-off between segmentation accuracy, robustness, and computational efficiency among the evaluated models. Training was performed using the Adam optimizer with an initial learning rate of 10^{-3} and step-wise reduction based on validation performance. A batch size of 8 and a maximum of 50 epochs were used. Early stopping was applied to prevent overfitting, together with L2 regularization and dropout. Hyperparameters were empirically optimized based on validation performance while maintaining consistent training protocols across models.
4) *Segmentation Evaluation.* The obtained automated segmentations were then blindly compared with the consensus manual delineations generated by the expert annotator, which served as ground truth (Fig. 2).Segmentation performance was evaluated using the Dice Similarity Coefficient (DSC), which quantifies spatial overlap between predicted and reference masks (range 0–1, where 1 indicates perfect agreement).

Automated pleural effusion volumes were reconstructed from the predicted segmentation masks by integrating the segmented areas across longitudinal and transverse intercostal views.

III. RESULTS

Segmentation performance was evaluated on the independent test set on a per-patient basis. Automated segmentations achieved a mean DSC of 0.86 ± 0.04 relative to the reference standard, while inter-operator agreement between two expert annotators yielded a mean DSC of 0.82 ± 0.05. Because DSC values are bounded between 0 and 1 and not assumed to follow a normal distribution, the comparison between automated segmentation performance and inter-operator agreement was performed using a two-sided Wilcoxon signed-rank test on paired per-patient DSC values. The difference was statistically significant ($p < 0.05$). Automated volume estimates were compared against thoracentesis-derived drained volumes, considered the reference standard. The mean absolute error (MAE) was 31.6 ± 14.2 cm³, and the root mean square error (RMSE) was 35.9 cm³. Pearson correlation analysis demonstrated a strong positive correlation between ultrasound-derived and drained volumes (r = 0.91, $p < 0.001$). Mean processing time per examination, including preprocessing, segmentation inference, and volumetric reconstruction, was below 3 seconds on a 13th Gen Intel(R) Core(TM) i5-1335U processor.

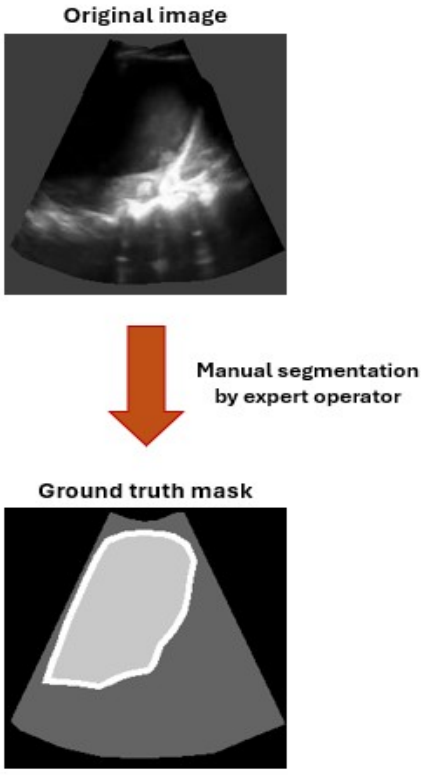

Fig. 2. Example of mask used for ResNet-5 training obtained through a GUI by an expert operator who manually segmented the pleural effusion.

TABLE I
PERFORMANCE METRICS OF THE PROPOSED AUTOMATED FRAMEWORK.

Metric	Value
Dice Similarity Coefficient (DSC)	0.86 ± 0.04
Inter-operator DSC	0.82 ± 0.05
Mean Absolute Error (MAE)	31.6 ± 14.2 cm^3
Root Mean Square Error (RMSE)	35.9 cm^3
Pearson Correlation (r)	0.91 ($p < 0.001$)
Mean Processing Time	< 3 s/exam

IV. DISCUSSION

Deep learning–based segmentation of pleural effusion in lung ultrasound images enabled accurate and reproducible volumetric estimation. The mean DSC (0.86 ± 0.04) indicates substantial overlap between algorithm output and expert manual segmentations. Moreover, the automated method showed lower variability compared with independently assessed inter-operator agreement (0.82 ± 0.05), suggesting improved reproducibility within the range of human annotation variability. The volumetric reconstruction approach yielded strong correlation with thoracentesis-derived measurements(r = 0.91, $p < 0.001$), with limited absolute error (MAE 31.6 ± 14.2 cm³). These findings support the feasibility of automated pleural effusion quantification from lung ultrasound imaging while reducing reliance on operator-dependent measurements or simplified geometric assumptions.

A. Comparison with other approaches

Ultrasound-based quantification of pleural effusion has traditionally relied on empirical formulas derived from linear measurements obtained in specific intercostal views. Most of these approaches approximate the effusion using simplified geometric structure and estimate volume using either single-distance [13] or two-distance models [20]. While such formulas are practical and fast, their accuracy may decrease in irregularly shaped or loculated effusions, and they are particularly sensitive to probe positioning and patient posture. More recent studies have explored multi-plane acquisition strategies to improve volumetric accuracy [20], especially for small-to-moderate effusions. Although these methods reduce geometric oversimplification, they still depend on manual measurements and operator-dependent contour selection. In parallel, deep learning methods for ultrasound image segmentation have demonstrated promising results in reducing operator variability and improving reproducibility. Encoder–decoder architectures such as U-Net and its variants have become common reference models in medical image segmentation, while attention-based mechanisms have been proposed to improve localization performance and boundary delineation [21]. The present approach differs in that as it performs pixel-wise segmentation of the effusion area, followed by numerical integration across orthogonal planes. By directly learning spatial features from annotated data, the proposed framework reduces reliance on predefined geometric models and may better accommodate anatomical variability. The strong correlation with drained volume observed in this study is consistent with the upper range of performance reported for multi-plane ultrasound estimation techniques. However, our findings should be interpreted in the context of previous automated pleural effusion segmentation studies. In particular, Vukovic et al. [22] reported an average DSC of 0.70 between the algorithm and expert annotations, compared with 0.61 between the two experts, on a dataset of 3,041 lung ultrasound images from 24 patients. Differences in segmentation performance between studies may reflect variability in dataset composition, image quality, annotation strategy, acquisition protocol, and reference standard generation. Therefore, although the present findings are encouraging, they should still be considered preliminary until externally validated on larger multicenter cohorts.

B. Clinical Implications

Operator dependency remains one of the main limitations of lung ultrasound in clinical practice, as variability in manual delineation can affect both qualitative and quantitative assessments. The proposed automated framework mitigates this limitation by enabling standardized segmentation and rapid volumetric reconstruction, requiring less than 3 seconds per examination. Such processing speed supports potential integration into bedside workflows, including emergency, intensive care, or post-procedural monitoring settings. Automated quantification could support therapeutic decision-making, longitudinal monitoring, and objective documentation of effusion progression or resolution. Furthermore, reducing subjectivity in segmentation may be particularly relevant in multicenter studies or telemedicine applications, where consistency across operators is essential. The segmentation performance observed in this study falls within the range generally considered indicative of strong agreement in medical image analysis. The volumetric reconstruction approach employed was based on integration of segmented areas across longitudinal and transverse views, balancing computational simplicity and anatomical realism while remaining compatible with real-time implementation. Although the error metrics (MAE and RMSE) were low relative to clinically significant effusion volumes, proportional error may increase for very small collections, where minor segmentation inaccuracies can produce larger relative deviations.

C. Limitations

Several limitations should be acknowledged. First, the study cohort was relatively small and derived from a single clinical center, potentially limiting variability in acquisition conditions and increasing the risk of overfitting. Consequently, the reported performance should be interpreted as an initial validation rather than definitive evidence of generalization. External validation on independent datasets acquired using different ultrasound systems, operators, and clinical environments remains necessary. Second, segmentation masks were generated by a single expert and served as ground truth. Although inter-operator agreement was assessed separately, the use of multi-expert consensus annotations as ground truth could further improve robustness and reduce annotation bias. Third, the relatively high segmentation performance observed in this study may be partially influenced by the controlled nature of the dataset. However, in real-world clinical practice, variations in posture, respiratory phase, and probe angulation may influence segmentation performance and volumetric estimation. In this context, differences in dataset composition, acquisition protocols, and annotation strategies may explain discrepancies with previously reported results in the literature. Therefore, the present findings, while promising, should be considered preliminary until validated

on independent external cohorts under heterogeneous acquisition conditions. Finally, although thoracentesis-derived volume was used as the reference standard, drained volume may not perfectly correspond to the total anatomical effusion volume due to incomplete drainage or procedural factors. In our cohort, a residual fluid volume was expected, which may introduce a systematic bias in the evaluation of volumetric accuracy, particularly for smaller effusions where such bias represents a larger proportion of the total volume.

D. Future Directions

Future research should focus on multicenter validation involving heterogeneous ultrasound systems, operators, and acquisition protocols to assess generalization across real-world clinical environments. Expanding the dataset to include a broader spectrum of effusion volumes, particularly small and loculated collections, would further clarify the limits of volumetric accuracy. Integration with real-time ultrasound platforms represents another relevant step. Embedding the segmentation model directly into ultrasound workstations could enable on-device inference and immediate quantitative feedback during bedside examinations. Moreover, combining automated pleural effusion quantification with AI-based assessment of pleural thickening, septations, or parenchymal abnormalities may enable more comprehensive thoracic ultrasound evaluation, supporting both diagnostic and prognostic applications. From a methodological perspective, the exploration of three-dimensional ultrasound acquisitions or spatiotemporal models incorporating respiratory dynamics may further improve volumetric precision. Additionally, incorporating uncertainty estimation mechanisms could enhance clinical interpretability by providing confidence intervals for automated measurements. Finally, since the present study focused on proof-of-concept validation of a single segmentation architecture rather than exhaustive benchmarking across multiple deep learning models, future studies should systematically compare the proposed framework with established architectures on larger multicenter datasets.

V. CONCLUSIONS

This study presents a deep learning–based framework for automated segmentation and volumetric quantification of pleural effusion from 2D lung ultrasound images. The proposed method achieved segmentation performance within the range of expert inter-operator variability and showed strong correlation with thoracentesis-derived volumes, with low absolute error and low computational time. By combining pixel-wise segmentation with multi-plane volumetric integration, the framework reduces operator dependency and minimizes reliance on simplified geometric assumptions. These characteristics suggesting potential for rapid, objective, and reproducible assessment of pleural effusion in clinical practice. Overall, the results support the feasibility of integrating artificial intelligence–based quantification tools into routine lung ultrasound workflows, with potential benefits in standardization, bedside decision-making, and longitudinal monitoring.

REFERENCES

[1] Karkhanis, V. S., & Joshi, J. M. (2012). Pleural effusion: diagnosis, treatment, and management. Open Access Emergency Medicine, 4, 31–52.

[2] Light, R. W., Macgregor, M. I., Luchsinger, P. C. & Ball, W. C. Jr. Pleural effusions: the diagnostic separation of transudates and exudates. Ann. Intern. Med. 77, 507–513.

[3] Colins JD, Burwell D, Furmanski S, Lorber P, Steckel RJ. Minimal detectable pleural effusions: a roentgen pathology model. Radiology. 1972; 105:51–53.

[4] U.K. Teichgräber, J. Hackbarth, et al. Sonographic bedside quantification of pleural effusion compared to computed tomographyvolumetry in ICU patients. Ultrasound Int Open, 4 (2018), pp. E131-5

[5] Carvajal Revuelta E, García Álvarez R (2020) Methods of estimation of pleural effusion by ecography. Rev Esp Anestesiol Reanim 67:521–526.

[6] Soni, N.J., Franco, R., Velez, M.I., Schnobrich, D., Dancel, R., Restrepo, M.I. and Mayo, P.H. (2015), Ultrasound in the diagnosis and management of pleural effusions. J. Hosp. Med., 10: 811-816.

[7] Andreas H. Diacon, Martin H. Brutsche, Markus Solèr, 'Accuracy of Pleural Puncture Sites: A Prospective Comparison of Clinical Examination With Ultrasound, Chest' Volume 123, Issue 2, 2003, Pages 436-441, ISSN 0012-3692.

[8] Sperandeo M, Filabozzi P, Varriale A, Carnevale V, Piattelli ML, Sperandeo G, Brunetti E, Decuzzi M. Role of thoracic ultrasound in the assessment of pleural and pulmonary diseases. J Ultrasound. 2008 Jun;11(2):39-46.

[9] Brant WE. Chest. In: McGahan JP, Goldberg BB, editors. Diagnostic ultrasound: a logical approach. 1st ed. Philadelphia: Lippincott-Raven, 1998;10631081.

[10] Tsai TH, Yang PC. Ultrasound in the diagnosis and management of pleural disease. Curr Opin Pulm Med. 2003;9:282–290.

[11] Ferreira AC, Filho FM, Braga T, Fanstone GD, Chodraui IC, Onari N. The role of ultrasound in the assessment of pleural effusion. Radiol Bras. 2006;39:145–150.

[12] Ibitoye BO, Idowu BM, Ogunrombi AB, Afolabi BI. Ultrasonographic quantification of pleural effusion: comparison of four formulae. Ultrasonography. 2018;37(3):254–60.

[13] Balik M, Plasil P, Waldauf P, Pazout J, Fric M, Otahal M, et al. Ultrasound estimation of volume of pleural fluid in mechanically ventilated patients. Intensive Care Med. 2006;32:318.

[14] Eibenberger KL, Dock WI, Ammann ME, Dorffner R, Hormann MF, Grabenwoger F. Quantification of pleural effusions: sonography versus radiography. Radiology.1994;191:681–684.

[15] Vignon P, Chastagner C, Berkane V, Chardac E, Francois B, Normand S, Bonnivard M, Clavel M, Pichon N, Preux PM, Maubon A, Gastinne H (2005) Quantitative assessment of pleural effusion in critically ill patients by means of ultrasonography. Crit Care Med 33:1757–1763.

[16] Roch A, Bojan M, Michelet P, Romain F, Bregeon F, Papazian L, Auffray JP (2005) Usefulness of ultrasonography in predicting pleural effusions $> 500ml$ in patients receiving mechanical ventilation. Chest 127:224–232.

[17] Remérand, F., Dellamonica, J., Mao, Z. et al. Multiplane ultrasound approach to quantify pleural effusion at the bedside. Intensive Care Med 36, 656–664 (2010).

[18] Tang, D., Yi, H. and Zhang, W. Ultrasound quantification of pleural effusion volume in supine position: comparison of three model formulae. BMC Pulm Med 24, 316 (2024).

[19] The MathWorks Inc. (2022). Deep Learning Toolbox version: 9.4 (R2022b), Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

[20] Scarlata S, Venditti I, Fulco G, Finamore P, Giua R, Antonelli Incalzi R, Silvestri S. Estimation of Pleural Effusion volume through chest ultrasound: validation of two Multiplanar models. Ultrasound Med Biol. 2020;46(8):1960–7.

[21] R. Azad et al., "Medical Image Segmentation Review: The Success of U-Net," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12, pp. 10076-10095, Dec. 2024.

[22] Vukovic, D., Wang, A., Antico, M. et al. Automatic deep learning-based pleural effusion segmentation in lung ultrasound images. BMC Med Inform Decis Mak 23, 274 (2023).