Neuroendocrinology Letters Volume 45 No. 3 2024
ISSN: 0172-780X; ISSN-L: 0172-780X; Electronic/Online ISSN: 2354-4716
Web of Knowledge / Web of Science: Neuroendocrinol Lett
Pub Med / Medline: Neuro Endocrinol Lett

ORIGINAL ARTICLE

Hybrid Artificial Intelligence Solution Combining Convolutional Neural Network and Analytical Approach Showed Higher Accuracy in A-lines Detection on Lung Ultrasound in Thoracic Surgery Patients Compared with Radiology Resident

Martin Števík[1], Marek Malík[2], Štefánia Vetešková[1], Zuzana Trabalková[1], Maroš Hliboký[3], Michal Kolárik[3], Ján Magyar[3], Marek Bundzel[3], Martina Szabóová[3], František Babič[3], Marián Grendár[4], Kamil Zeleňák[1], Viktória Máčajová[1], Beáta Drobná Sániová[5], Anton Dzian[2]

1 Clinic of Radiology, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Kollárova 2, 036 59 Martin, Slovakia.
2 Department of Thoracic Surgery, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Kollárova 2, 036 59 Martin, Slovakia.
3 Department of Cybernetics and Artificial Intelligence, Technical University of Košice, Letná 9, 040 01 Košice, Slovakia.
4 Biomedical Centre Martin – BioMed Martin, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Malá hora 4, 036 01, Martin, Slovakia.
5 Department of Anesthesiology and Intensive Medicine, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Kollárova 2, 036 59 Martin, Slovakia.

Correspondence to: Marek Malík, MD, PhD.
Department of Thoracic Surgery, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, University Hospital Martin, Kollárova 4248/2, 036 59 Martin, Slovakia
TEL: 00421 43 4203 415, E-MAIL: marek.malik@uniba.sk

Submitted: 2024-06-13 *Accepted: 2024-06-29* *Published online: 2024-07-28*

Key words: **A-lines; artificial intelligence; convolutional neural network; analytical approach; lung ultrasonography; thoracic surgery**

Neuroendocrinol Lett 2024; **45**(3):229–237 **PMID:** 39146568 NEL450224L01 ©2024 Neuroendocrinology Letters • **www.nel.edu**

Abstract

OBJECTIVES: Lung ultrasound reduces the number of chest X-rays after thoracic surgery and thus the radiation. COVID-19 pandemic has accelerated research in lung ultrasound artifacts detection using artificial intelligence. This study evaluates the accuracy of artificial intelligence in A-lines detection in thoracic surgery patients using a novel hybrid solution that combines convolutional neural networks and analytical approach and compares it with a radiology resident and radiology experts' results.
DESIGN: Prospective observational study.
MATERIAL AND METHODS: Single-center study evaluates the accuracy of artificial intelligence and a radiology resident in A-line detection on lung ultrasound footages compared with the consensual opinion of two expert radiologists as the reference. After resident's first reading, the artificial intelligence results were presented to the resident and he was asked to revise the results based on artificial intelligence.

RESULTS: 82 consecutive patients underwent 82 ultrasound examinations. 328 ultrasound recordings were evaluated. Accuracy, sensitivity, specificity, positive and negative predictive values of artificial inelligence in A-line detection were 0.866, 0.928, 0.834, 0.741 and 0.958 respectively. The resident's values were 0.558, 0.973, 0.346, 0.432 and 0.962 respectively. The resident's values after correction based on artificial intelligence results were 0.854, 0.991, 0.783, 0.701 and 0.994 respectively.
CONCLUSION: Artificial intelligence showed high accuracy in A-line detection in thoracic surgery patients and was more accurate compared to a resident. Artificial intelligence could play important role in lung ultrasound artifact detection in thoracic surgery patients and in residents' education.

Abbreviations:

2D	- two dimensional
AI	- artificial intelligence
BLUE protocol	- Bedside Lung Ultrasound in Emergency protocol
BMI	- Body Mass Index
CNNs	- Convolutional Neural Networks
dB	- decibel
Hz	- Hertz
Inc.	- incorporated
LUS	- lung ultrasoound
MHz	- megahertz
mp4	- digital multimedia format
PLAPS	- Posterolateral Alveolar and/or Pleural Syndrome
Q1	- quartile 1
Q3	- quartile 3
USA	- The United States of America

INTRODUCTION

Lung ultrasonography (LUS) became an established method to diagnose and monitor pulmonary diseases (Born *et al.* 2021; Chavez *et al.* 2014; Lichtenstein *et al.* 2004). LUS has been demonstrated to be superior to chest X-ray (Bourcier *et al.* 2014; Dzian *et al.* 2021; Reali *et al.* 2014). A growing body of evidence for disease-specific patterns in LUS has led to advocacy for an amplified role of LUS in the research community (Buonsenso *et al.* 2020; Smith *et al.* 2020; Soldati *et al.* 2020). LUS can be utilized as a first-level examination technique among various medical fields (Brogi *et al.* 2017; Bourcier *et al.* 2016; Malík *et al.* 2021). LUS has some unique advantages compared to other imaging technologies. It is low cost, portable, rapid, radiation free and allows real time examination (Fatima *et al.* 2023; Marini *et al.* 2021). The main limitation of LUS is its high operator dependency, which opens room for subjective interpretation. Therefore, automatic detection of lung artifacts is highly relevant as it has been shown to reduce the time that physicians invest

Tab. 1. Baseline demographic characteristics and basic surgical data

Patients (n)	82
Male, n (%)	41 (50)
Female, n (%)	41 (50)
Age (years), median (Q1[a]-Q3[b])	59 (51-68)
Weight (kg), median (Q1-Q3)	76.5 (65-90)
Height (cm), median (Q1-Q3)	168 (161-179)
BMI[c] (kg/m^2), median (Q1-Q3)	26 (23-30)
Operation side	
Right n (%)	42 (51%)
Left n (%)	40 (49%)
Surgical approach	
Thoracotomy n (%)	17 (21%)
Videotoracoscopy n (%)	65 (79%)
Type of surgical procedure	
Major Lung resection n (%)	18 (22%)
Wedge or atypical lung resection n (%)	42 (51%)
Procedure without lung resection n (%)	22 (27%)
Duration of the chest tube drainage (days), median (Q1-Q3)	3 (2-4)
Hospital length of stay (days), median (Q1-Q3)	5 (4-7)

[a] Q1 - quartile 1,
[b] Q3 – quartile 3
[c] BMI – Body Mass Index

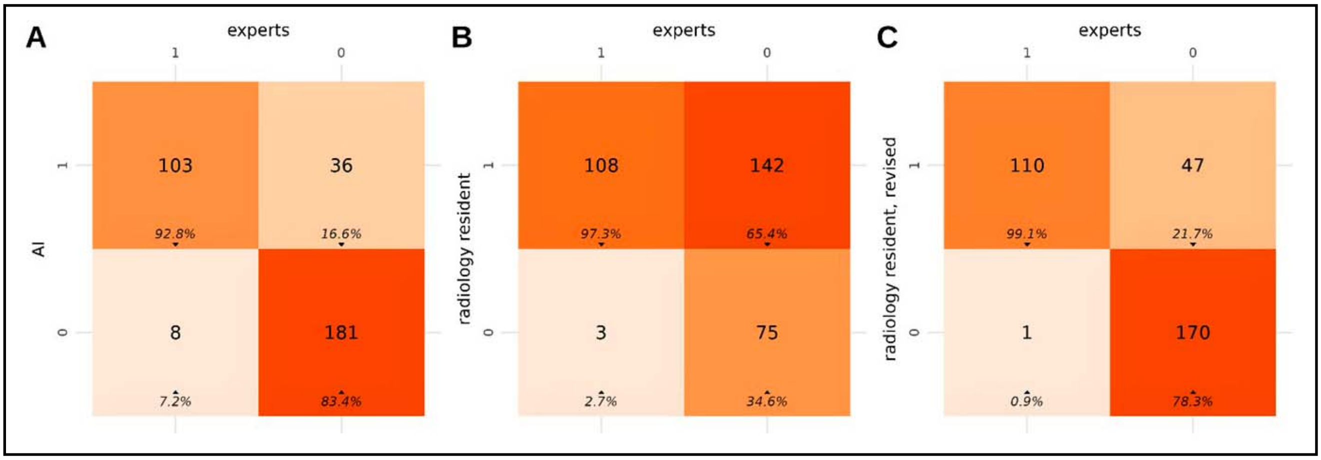

Fig. 1. Cross tabulations of A-lines detection. 1 - presence of A-lines, 0 - absence of A-lines. a artificial intelligence (AI) compared with expert radiologists' consensus (experts). b Radiology resident compared with expert radiologists' consensus. c Radiology resident revised results based on AI compared with expert radiologists' consensus. AI - artificial intelligence

to make a diagnosis (De Rosa *et al.* 2022). This has led to significant interest in developing artificial intelligence (AI) approaches for the interpretation of LUS imaging (Roy *et al.* 2020; Zhang *et al.* 2021). Deep learning, a foundational strategy within present-day AI techniques, has been shown to meet or exceed clinician performance across most visual fields of medicine (Chilamkurthy *et al.* 2018; Ohno *et al.* 2022; Ohno *et al.* 2024; Takamatsu *et al.* 2024).

Patients after thoracic surgery are monitored for several conditions, especially for surgically induced pneumothorax and pleural effusion. The importance of LUS after non-cardiac thoracic surgery in chest X-ray reduction is growing (Galata *et al.* 2022; Galetin *et al.* 2020; Malík *et al.* 2021). The accuracy of LUS in thoracic surgery depends on the complexity of the used LUS protocol and on examiner experience (Galata *et al.* 2022; Galetin *et al.* 2020; Jakobson *et al.* 2022). We assume an important role of AI in this field (Malík *et al.* 2023).

We decided to evaluate the role of AI in LUS in thoracic surgery patients. LUS does not try to image the internal tissue of the lung, but rather focuses on artifacts that appear in the image below the pleural line. Artifacts allow the differentiation of lung and pleural pathologies (Soldati *et al.* 2016). Our research follows the 'Bedside Lung Ultrasound in Emergency (BLUE) protocol' (Lichtenstein, 2015). The LUS artifacts relevant in BLUE protocol are lung sliding (Lichtenstein & Menu, 1995), B-lines, A-lines (Lichtenstein *et al.* 2009), lung point (Lichtenstein *et al.* 2000), pleural effusion and lung consolidations (Lichtenstein, 2015). In the initial phase of our research, we decided to detect A-lines on LUS in thoracic surgery patients using AI. Healthy lung 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 (Carrer *et al.* 2020). The presence or absence of A-lines is not specific and can occur in various conditions but is important in the context of other LUS signs in complex LUS protocols (Erfanian Ebadi *et al.* 2021; Lichtenstein, 2015; Soldati *et al.* 2019). For applicability in clinical practice, in our future research we will train the AI to also recognize the rest of LUS signs relevant in BLUE protocol.

The primary aim of this study was to evaluate the accuracy of a trained AI model in A-line detection in LUS footages in thoracic surgery patients using a novel hybrid solution that combines convolutional neural networks (CNNs) and analytical approach. The secondary aim was to evaluate the accuracy of a radiology resident beginner in LUS in A-line detection in thoracic surgery patients and compare the results with AI. The tertiary aim was to evaluate the educational potential of AI in LUS. To evaluate this, the radiology resident was asked to revise her initial conclusions based on AI results and the accuracy of the revised results was evaluated.

MATERIALS AND METHODS

Study design and setting

A single-center prospective study was conducted in collaboration of Technical University of Košice and Jessenius Faculty of Medicine in Martin, Comenius University, Slovakia. All patients scheduled for the whole spectrum of non-cardiac thoracic surgery procedures at the Department of Thoracic Surgery Jessenius Faculty of Medicine in Martin, Comenius University, Slovakia and University Hospital Martin, Slovakia between October 2021, and April 2023 were consecutively enrolled in this study. Pediatric patients (under 18 years) and patients scheduled for pneumonectomy were excluded. Technically inadequate LUS studies were excluded as well.

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Slovakia on 29th June 2021, No. EK 44/2021. Written informed consent was obtained from all subjects involved in the study prior to examination. This study was supported by the Slovak Research and Development Agency, project No. APVV-20-0232.

LUS imaging protocol

The LUS examinations were performed by three examiners experienced in LUS. One of them was a skilled radiologist, two of them were thoracic surgeons. The patients' names were replaced by consecutive study numbers and another identification date was anonymized so that the stored videos were blinded to the patients' identity. Each patient underwent one examination in supine position. The LUS examination followed the 'Bedside Lung Ultrasound in Emergency (BLUE) protocol' (Lichtenstein, 2015). The probes were placed at three 'BLUE points' on each side: 'upper BLUE point', 'lower BLUE point' and 'Posterolateral Alveolar and/or Pleural Syndrome (PLAPS) point' as described by Lichtenstein and Meziere (2011). Six LUS videos were acquired from one LUS examination. Only videos from 'upper BLUE points' and from 'lower BLUE points' from both patient sides were evaluated for A-lines presence. The LUS videos from 'upper BLUE points' and from 'lower BLUE points' were performed by a Philips Lumify L12-4 (Lumify, Philips Ultrasound, Inc., Bothell, Washington, USA) portable linear array transducer (4-12 MHz) in lung preset (acoustic working frequency: 12-4 MHz, mechanical index: 0,7, soft-tissue thermal index: 0,1, preset: Lung, focal optimization: Gen, default penetration depth: 6cm, scan repetition rate: 30 Hz, power: -0,3dB, 2D Gain: 50). Videos from PLAPS points, taken with a convex probe, were not evaluated in this study. Probes were connected to handheld tablet devices (Samsung Galaxy S6 Lite Tab, Samsung, Suwon, South Korea) with an Android interface. Each scan consists of a ten second LUS video recording with a frame rate of 30 per second (300 frames for each clip) stored in mp4 format. Duplicate studies were discarded to avoid overfitting.

A-lines evaluation by radiologists

The LUS videos were reviewed and evaluated for the presence of the A-line by three radiologists. Two of them were expert radiologists experienced in LUS (2 board-certified general radiologist with 18-years experience and certified radiologist, 8 years of experience) and one was a radiology resident with limited LUS experience (2 years experience in radiology and LUS beginner). Observers were asked to identify A-lines. A-lines were defined as a hyperechoic horizontal lines, which occur when multiple reverberations are present below the pleural line running equidistant from each other equal to the distance between the skin and the pleura (Carrer *et al.* 2020; Erfanian Ebadi *et al.* 2021; Soldati, 2020). Each video was classified as (0) certainly no A-lines or (1) in case of A-lines presence. The reference standard was established by consensual agreement reading by two expert radiologists due to the lack of a corresponding sign or pathology to A-lines on any imaging modality that could serve as an objective reference. The discrepancies were resolved by consensus between expert radiologists. The clips were presented one at a time. Expert radiologists were blinded to the results of the the trained AI model and to the results of radiology resident. The videos were anonymized, so the readers did not have any information on the patients' history nor on any clinical or radiologic imaging results. Initially, the radiology resident was blinded to the expert radiologists and AI results and to clinical data. After initial reading, the AI results, in the meaning of verbal description (A-lines present or A-line absent), were presented to the radiology resident. Radiology resident was allowed to revise the initial conclusion regarding A-lines presence.

AI model

In this study, a machine learning based software, the LUS AI solution, was used for automated detection and marking of A-lines in the LUS footages. The automated image processing software model was developed in collaboration with the Department of Cybernetics and Artificial Intelligence, Technical University of Košice (Hliboký *et al.* 2023). The details of the AI algorithm are vendor proprietary.

The presented research is based on real-life examinations' recordings. The dataset was not primarily gathered for supporting training AI models. The dataset for the development of the software consisted of LUS videos with frames manually labeled independently by two LUS experts to minimize bias. To reduce the workload of the experts, only every tenth frame of the videos was inspected (every one third of a second) and if A-lines were present, the frame was annotated. The data set was split into three subsets. One was used to train CNNs models, one to adjust the metaparameters and to determine when to stop training and one to generate the results. When preparing the models, 1718 frames were used for training (of which 227 frames contained A-lines), 430 frames for validation (46 positive examples), and 462 for testing (with 176 positives). For detecting A-lines, three separate models were trained on different subsets of the training data. The trained models do not provide direct predictions, instead they use heuristic rules that can be adjusted by experts based on evaluation results. The resulting prediction was obtained as an aggregate of the predictions of the models. This approach is known as bagging and improves the accuracy and generalization of the resulting aggregate model. The aggregation procedure was based on weighted voting and the validation set was used to determine how many votes are needed to decide whether the video has A-lines present.

The weight of the vote of the given model was based on its reliability, however, it could potentially be adjusted by human experts based on more detailed analysis and evaluation.

The training data were augmented by introducing random noise to the frames from standard distributions, as other data augmentation methods such as rotation and vertical flip would be impractical due to the nature of data. For A-line detection, we used a pre-trained ResNet-34 CNNs, what is called transfer learning. In that case the CNNs is already trained on a certain computer vision task and distributed to the public. Then the user re-trains only a portion of the CNNs using the target task data. This is done to reduce the training times, to improve generalization and accuracy of the resulting model. The ResNet-34 architecture was selected based on a preliminary comparison of performance of various neural network architectures, where ResNet was shown to be the most precise. CNNs and ResNet are among the most frequently architectures used in medical image processing and classification. To prevent overfitting, we used early stopping of the training process and other standard forms of regularization. We adjusted the loss function so that it penalized false negative predictions more than false positives with a ratio of 1:10. This was done to compensate the imbalance in the training set that contained more negative than positive samples and to reflect the higher medical cost of false negative prediction.

Statistical analysis

Baseline demographic characteristics such as age, sex, weight, height, BMI were recorded. Also, the side of the operation, surgical approach, type of surgical procedure, duration of the chest tube drainage and hospital length of stay were recorded.

Study of diagnostic accuracy was performed to compare the trained AI model in evaluation of A-lines presence in LUS videos with consensual agreement reading by two radiology experts as the reference. Also, the initial radiology resident conclusions in A-lines detection were compared with consensual agreement reading by two expert radiologists as the reference. The results of a trained AI model and radiology resident were compared. Finally, the revised radiology resident results based on AI were compared to expert radiologists' consensus. Categorical variables were summarized by counts and percentages. Continuous variables were expressed as median and range between first and third quartile (Q1-Q3). We assessed the model's performance by analyzing a confusion matrix. We determined the diagnostic power for sonographers and the trained AI model using accuracy, sensitivity, specificity, positive predictive value, and negative predictive value. For sensitivity and specificity, the 95% confidence interval was constructed by the bootstrap method.

RESULTS

82 consecutive patients were enrolled. Baseline demographic data and basic surgical data are shown in Table 1. 82 LUS examinations were performed. A total of 492 LUS videos were recorded, 6 videos were taken from each patient as described in methods. A total of 328 LUS videos from 'upper BLUE points' and 'lower

Tab. 2. Statistical evaluation of the artificial intelligence, the initial and the revised results of the radiology resident in A-line detection compared with radiology experts' consensus as the reference

	AI[a] vs. expert radiologists' consensus	Initial radiology resident results vs. expert radiologists' consensus	Revised radiology resident based on AI[a] results vs. expert radiologists' consensus
Sensitivity	0.928 (0.710 - 0.966)*	0.973 (0.881 – 0.993)*	0.991 (0.789 – 1.000)*
Specificity	0.834 (0.490 – 0.879)*	0.346 (0.148 – 0.410)*	0.783 (0.252 – 0.831)*
PPV[b]	0.741	0.432	0.701
NPV[c]	0.958	0.962	0.994
Accuracy	0.866	0.558	0.854
Balanced accuracy	0.881	0.659	0.887
Kappa	0.718	0.244	0.703
F-measure	0.824	0.598	0.821
Youden's index	0.762	0.319	0.774
Matthew's correlation coefficient	0.730	0.354	0.734

*95% confidence interval are given in brackets
[a] AI – artificial intelligence
[b] PPV – positive predictive value
[c] NPV – negative predictive value

Tab. 3. Review of articles on lung ultrasound artifacts detection by artificial intelligence in COVID-19 patients based on Wang et al. (2022)

Study	Patients/ Examinations	Techniques	Results
Barros *et al.* (2021)	185 videos	POCOVID-NET, DenseNET, RESNet, Xception, NASNet	Accuracy: 0.91-0.93 F1-score: 0.86-0.95
Demi *et al.* (2022)	220 patients	STN & U-Net	Accuracy: 0.8
Diaz-Escobar *et al.* (2021)	3326 images	VGG19, InceptionV3, Xception, RESNet5O	Accuracy: 0.89±0.02 F1-score: 0.88±0.03
Erfanian Ebadi et al. (2021)	300 patients	3D ConvNet	Accuracy: 0.9 F1 score: 0.87-0.94
Hu *et al.* (2021)	104 patients	ResNeXt	Accuracy: 0.94 Sensitivity: 0.76 Specificity: 0.96
La Salvia *et al.* (2021)	450 patients	ResNet18, ResNet50	Accuracy: 0.98-1.0 F1-score: 0.97-0.99

BLUE points' from both patients' sides were evaluated for A-lines presence. 164 LUS videos from PLAPS points from both patients' sides were not evaluated in this study.

Expert radiologists have consensually concluded the presence of the A-lines on 111 LUS videos and the absence of the A-lines on 217 LUS videos.

The trained AI model has detected the A-lines on 139 LUS scans and excluded the presence of A-lines on 189 LUS scans. Compared to expert radiologists' consensus, the AI evaluated 36 videos as false positive and 8 videos as false negative (Figure 1).

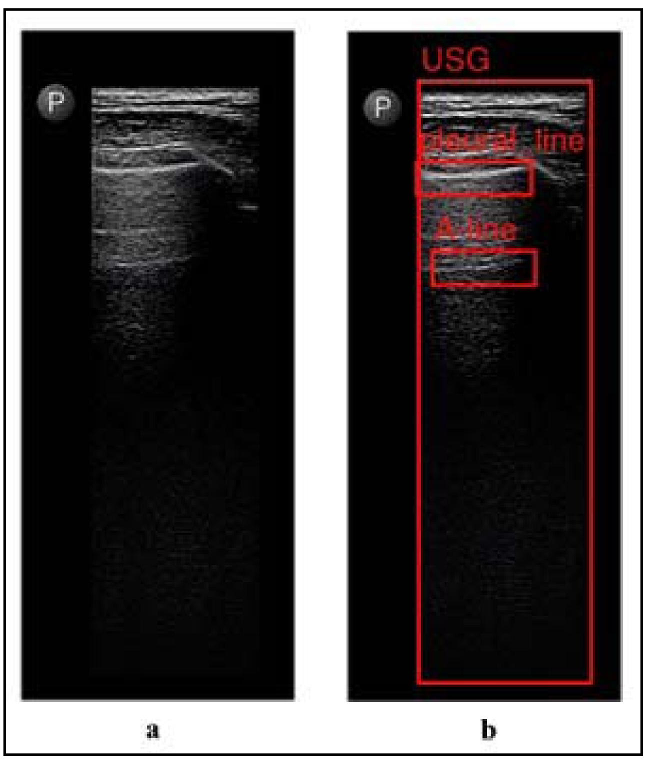

Fig. 2. Visual evaluation of the lung ultrasound video for the A-lines presence using artificial intelligence in a prepared application. **a)** A-lines on a lung ultrasound video prior to evaluation by artificial intelligence. **b)** Visualization of pleural line and A-line after evaluation of the same video by artificial intelligence

Radiology resident has detected the A-lines on 250 LUS videos and 78 LUS videos were concluded as not containing A-lines. 142 videos were false positive and 3 videos false negative when compared with expert radiologists' consensus (Figure 1).

Finally, the AI results were presented to radiology resident, and she was allowed to revise her conclusion in case of agreement with the AI result. After AI based revision, the radiology resident has concluded the presence of A lines on 157 LUS videos and 171 LUS videos were concluded as the absence of A-lines. After revision, 47 videos were false positive and 1 video false negative when compared to expert radiologists' consensus (Figure 1). The statistical evaluation of the AI, the initial and the revised results of the radiology resident in A-line detection compared to radiology experts' consensus as the reference are presented in Table 2.

Simoultaneously with our research, we are working on an application where the pleural line, A-lines and later also the other LUS artifacts will be visually marked (Figure 2).

DISCUSSION

Based on our results, hybrid AI solution combining CNNs and analytical approach can safely detect A-lines on LUS videos in thoracic surgery patients with an accuracy of 0.87, a balanced accuracy of 0.88, sensitivity 0.93 and F measure 0.82. AI achieved significantly higher accuracy compared to radiology resident, beginner in LUS (accuracy 0.56, balanced accuracy 0.66, sensitivity 0.97, F measure 0.60). Substantial agreement between AI and experts' radiologists as the reference (Kappa 72%) and fair agreement between expert radiologists and initial result of radiology resident (Kappa 24.4%) were observed. Re-evaluation of the radiology resident results based on AI model has significantly improved the accuracy and other statistical parameters of radiology resident (accuracy 0.854, balanced accuracy 0.88, sensitivity 0.99, F measure 0.821 and Kappa 70% - substantial agreement) compared with her initial conclusions. Our results showed the educational potential of AI in LUS artifact detection.

To our best knowledge, this is the first study that has evaluated the accuracy of an AI model in LUS artifacts detection in non-cardiac thoracic surgery patients. In various medical fields, the body of evidence in LUS artifact detection using AI is growing. Tan *et al.* reported better results of doctors' examination combined with AI in A line detection than each alone in pneumonia diagnostics with LUS (Tan *et al.* 2020). Several pre-trained CNNs models were described in the literature. Most of these models have high computational complexity but have a very high accuracy in many applications (Muhammad & Shamim Hossain, 2021). Sloun *et al.* (2020) proposed a fully CNNs to identify and localize the artifacts in clinical LUS. Recently, various AI algorithms were built for the detection of A-lines. Camacho *et al.* (2022) automated detection algorithm for A-line detection in COVID patients had coincidence 70.7%, false positivity of 23.1% and false negativity of 6.2%. The authors concluded a good agreement between the algorithm and an experienced physician. Erfanian Ebadi et al. (2021) with their study including (475 scans) concluded the potential use of the automated analysis of the portable LUS to assist clinicians in screening and diagnosing patients, based on the results of their AI program (precision 0.87, recall 0.91, F1 score 0.88). Recent review of machine learning in LUS in COVID-19 patients by Wang *et al.* (2022) concluded that various machine learning architectures have been employed to evaluate LUS and showed high performance. Most AI models have been used to detect A-lines and B-lines for the classification and scoring of COVID-19 and bacterial pneumonia (Table 3).

Results of meta-analyses of AI LUS programs are needed. The heterogeneity of the datasets, the different methods of the individual AI programs, the use of different probes with a predominance of convex probes, the use of different ultrasound devices and settings, small number of training segmented datasets, small number of patients in the datasets make it difficult to compare the individual developed AI applications to each other. There is a lack of a comprehensive evaluation of AI models for various types of lung diseases. LUS seems to be an ideal modality for monitoring of the patient after thoracic surgery. It could be performed repetitively and thus allowed to assess the dynamics of possible postoperative conditions without radiation and at the bedside of the patient (Carrer *et al.* 2020; Dzian *et al.* 2021; Malík *et al.* 2023). User-dependent interpretation of LUS contributes to wide variation in disease classification, creating urgency for AI techniques that will improve diagnostic precision and reduce user dependence (Erfanian Ebadi *et al.* 2021). To confirm these assumptions, research using homogeneous, well-labeled, multicentre data and meta-analysis is indicated. Further studies are required to validate externally and elucidate the benefit of AI models for thoracic surgical patients. The detection of A-lines by AI on LUS videos is only the initial step of our research. Detection of A-lines only has poor clinical value in patients after thoracic surgery. As described above, A-lines need to be evaluated together with other important LUS artifacts. In our further research we are planning to continue with AI detection of the rest of LUS signs important in BLUE protocol. After the completion of this task a clinical trial will be performed to evaluate the accuracy of AI not only in LUS artifact detection but also for diagnosis of the important postoperative conditions after thoracic surgery. For the future, our effort is to develop a lightweight mobile-friendly efficient deep learning AI application.

Our study has several limitations. First, our results were based on single-centre recruitment of patients scheduled to thoracic surgery. Second, there was overall a relatively small number of patients recruited. Our presented AI program results still show some limitations, which is due to the limited amount of data and inability to use some data augmentation methods to address this issue. We also allow for the possibility of some inaccuracies in manual labeling, which could result in contrastive data that would hinder higher detection accuracy. Another aspect negatively impacting the results is the high variability in training data due to different ultrasonography devices used to gather data. While this is a desirable quality for higher generalization, it can also be a limiting factor with smaller amounts of data. Another important limitation is the reference. A-line is a LUS artifact without correlation with a specific radiological sign or a specific lung or pleural condition on other imaging modalities. Therefore, we have used the consensual opinion of two expert radiologists, which is not so objective.

In conclusion, in this study, high accuracy of the trained hybrid AI model was achieved in A-line detection on LUS in non-cardiac thoracic surgery patients with their specifics. In our study, the trained AI model had significantly better accuracy compared with radiology resident, beginner in LUS. The radiology resident accuracy improved after correction of the initial conclusions based on the given AI model results. Therefore, we assume the important role of AI in LUS artifacts education of beginner sonographists especially when the LUS artifacts are visualized. For the clinical application, further research is needed. AI should be trained to detect the rest of LUS artifacts important in complex LUS protocols and finally a clinical trial should be performed to evaluate the accuracy of AI in differential diagnosis of various lung and pleural pathologies in patient after thoracic surgery and not only in separate LUS artifact detection.

ACKNOWLEDGMENTS

Funding

This research is funded by the Slovak Research and Development Agency, grant number APVV 20-0232.

Competing Interest

The authors have no relevant financial or non-financial interests to disclose. The funders had no role in the design of the study, in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

REFERENCES

1 Barros B, Lacerda P, Albuquerque C, Conci A (2021). Pulmonary COVID-19: Learning Spatiotemporal Features Combining CNN and LSTM Networks for Lung Ultrasound Video Classification. Sensors. **21**(16): 5486.

2 Born J, Wiedemann N, Cossio M, Buhre C, Brändle G, Leidermann K, et al. (2021). Accelerating Detection of Lung Pathologies with Explainable Ultrasound Image Analysis. Appl Sci. **11**(2): 672.

3 Bourcier JE, Paquet J, Seinger M, Gallard E, Redonnet JP, Cheddadi F, et al. (2014). Performance comparison of lung ultrasound and chest x-ray for the diagnosis of pneumonia in the ED. Am J Emerg Med. **32**(2): 115–118.

4 Bourcier JE, Braga S, Garnier D (2016). Lung ultrasound will soon replace chest radiography in the diagnosis of acute community-acquired pneumonia. Curr Infect Dis Rep. **18**(12): 43.

5 Brogi E, Bignami E, Sidoti A, Shawar M, Gargani L, Vetrugno L, et al. (2017). Could the use of bedside lung ultrasound reduce the number of chest x-rays in the intensive care unit? Cardiovasc Ultrasound. **15**(1): 23.

6 Buonsenso D, Pata D, Chiaretti A (2020). Covid-19 outbreak: less stethoscope, more ultrasound. Lancet Respir Med. **8**(5): e27.

7 Camacho J, Muñoz M, Genovés V, Herraiz JL, Ortega I, Belarra A, et al. (2022). Artificial Intelligence and Democratization of the Use of Lung Ultrasound in COVID-19: On the Feasibility of Automatic Calculation of Lung Ultrasound Score. Int. J. Transl. Med. **2**(1): 17–25.

8 Carrer L, Donini E, Marinelli D, Zanetti M, Mento F, Torri E, et al. (2020). Automatic Pleural Line Extraction and COVID-19 Scoring from Lung Ultrasound Data. IEEE Trans Ultrason Ferroelectr Freq Control. **67**(11): 2207–2217.

9 Chavez MA, Shams N, Ellington LE, Naithani N, Gilman RH, Steinhoff MC, et al. (2014). Lung ultrasound for the diagnosis of pneumonia in adults: a systematic review and meta-analysis. Respir Res. **15**(1): 50.

10 Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. (2018). Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet. **392**(10162): 2388–2396.

11 Demi L, Mento F, Di Sabatino A, Fiengo A, Sabatini U, Macioce VN, et al. (2022). Lung Ultrasound in COVID-19 and Post-COVID-19 Patients, an Evidence-Based Approach. J. Ultrasound Med. **41**(9): 2203–2215.

12 De Rosa L, L'Abbate S, Kusmic C, Faita F (2023). Applications of artificial intelligence in lung ultrasound: Review of deep learning methods for COVID-19 fighting. Artif Intell Med Imaging. **3**(2): 42–54.

13 Diaz-Escobar J, Ordóñez-Guillén NE, Villarreal-Reyes S, Galaviz-Mosqueda A, Kober V, Rivera-Rodriguez R, et al. (2021). Deep-learning based detection of COVID-19 using lung ultrasound imagery. PLoS One. **16**(8): e0255886.

14 Dzian A, Malík M, Hamada L, Skaličanová M, Zeleňák, K, Števík M, et al. (2021). Lung ultrasound could reduce X-ray after major lung resection. Bratisl Lek Listy. **122**(12): 871–875.

15 Erfanian Ebadi S, Krishnaswamy D, Bolouri SES, Zonoobi D, Greiner R, Meuser-Herr N, et al. (2021). Automated detection of pneumonia in lung ultrasound using deep video classification for COVID-19. Inform Med Unlocked. **25**: 100687.

16 Fatima N, Mento F, Zanforlin A, Smargiassi A, Torri E, Perrone T, et al. (2023). Human-to-AI interrater agreement for lung ultrasound scoring in COVID-19 patients. J Ultrasound Med. **42**(4): 843–851.

17 Galata C, Cascant Ortolano L, Shafiei S, Hetjens S, Müller L, Stauber RH, et al. (2022). Are Routine Chest X-rays Necessary Following Thoracic Surgery? A Systematic Literature Review and Meta-Analysis. Cancers. **14**(18): 4361.

18 Galetin T, Defosse J, Schieren M, Marks B, Lopez-Pastorini A, Koryllos A, et al. (2020). Sensitivity of chest ultrasound for postoperative pneumothorax in comparison to chest X-ray after lung resecting surgery. Eur J Cardiothorac Surg. **57**(5): 846–853.

19 Hliboký M, Magyar J, Bundzel M, Malík M, Števík M, Vetešková Š, et al. (2023). Artifact Detection in Lung Ultrasound: An Analytical Approach. Electronics. **12**(7): 1551.

20 Hu Z, Liu Z, Dong Y, Liu J, Huang B, Liu A, et al. (2021). Evaluation of lung involvement in COVID-19 pneumonia based on ultrasound images. Biomed Eng. **20**(1): 27.

21 Jakobson D, Cohen O, Cherniavsky E, Batumsky M, Fuchs L, Yellin A (2022). Ultrasonography can replace chest X-rays in the postoperative care of thoracic surgical patients. PLoS One. **17**(10): e0276502.

22 La Salvia M, Secco G, Torti E, Florimbi G, Guido L, Lago P, et al. (2021). Deep learning and lung ultrasound for Covid-19 pneumonia detection and severity classification. Comput Biol Med. **136**: 104742.

23 Lichtenstein DA, Menu Y (1995). A bedside ultrasound sign ruling out pneumothorax in the critically ill. Lung sliding. Chest. **108**(5): 1345–1348.

24 Lichtenstein DA, Mezière GA, Biderman P, Gepner A (2000). The "lung point": An ultrasound sign specific to pneumothorax. Intensive Care Med. **26**(10): 1434–1440.

25 Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby JJ (2004). Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. Anesthesiology. **100**(1): 9–15.

26 Lichtenstein DA, Mezière GA, Lagoueyte JF, Biderman P, Goldstein I, Gepner A (2009). A-lines and B-lines: Lung ultrasound as a bedside tool for predicting pulmonary artery occlusion pressure in the critically ill. Chest. **136**(4): 1014–1020.

27 Lichtenstein DA, Mezière GA (2011). The BLUE-points: three standardized points used in the BLUE-protocol for ultrasound assessment of the lung in acute respiratory failure. Crit Ultrasound J. **3**(2): 109–110.

28 Lichtenstein DA (2015). BLUE-protocol and FALLS-protocol: two applications of lung ultrasound in the critically ill. Chest. **147**(6): 1659–1670.

29 Malík M, Dzian A, Skaličanová M, Hamada L, Zeleňák K, Grendár M (2021). Chest Ultrasound Can Reduce the Use of Roentgenograms in Postoperative Care After Thoracic Surgery. Ann Thorac Surg. **112**(3): 897–904.

30 Malík M, Dzian A, Števík M, Vetešková Š, Al Hakim A, Hliboký M, et al. Lung Ultrasound Reduces Chest X-rays in Postoperative Care after Thoracic Surgery: Is There a Role for Artificial Intelligence? - Systematic Review (2023). Diagnostics. **13**(18): 2995.

31 Marini TJ, Rubens DJ, Zhao YT, Weis J, O'Connor TP, Novak WH, et al. (2021). Lung Ultrasound: The Essentials. Radiol Cardiothorac Imaging. **3**(2): e200564.

32 Muhammad G, Shamim Hossain M (2021). COVID-19 and non-COVID-19 Classification using Multi-layers Fusion From Lung Ultrasound Images. Inf Fusion. **72**: 80–88.

33 Ohno Y, Aoyagi K, Arakita K, Doi Y, Kondo M, Banno S, et al. (2022). Newly developed artificial intelligence algorithm for COVID-19 pneumonia: utility of quantitative CT texture analysis for prediction of favipiravir treatment effect. Jpn J Radiol. **40**(8): 800–813.

34 Ohno Y, Aoki T, Endo M, Koyama H, Moriya H, Okada F, et al. (2024). Machine learning-based computer-aided simple triage (CAST) for COVID-19 pneumonia as compared with triage by board-certified chest radiologists. Jpn J Radiol. **42**(3): 276–290.

35 Reali F, Sferrazza Papa GF, Carlucci P, Fracasso P, Di Marco F, Mandelli M, et al. (2014). Can lung ultrasound replace chest radiography for the diagnosis of pneumonia in hospitalized children? Respiration. **88**(2): 112–115.
36 Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, et al. (2020). Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. IEEE Trans Med Imaging. **39**(8): 2676–2687.
37 Sloun RJG, Cohen R, Eldar YC (2020). Deep Learning in Ultrasound Imaging. Proceedings of the IEEE. **108**(1): 11–29.
38 Smith MJ, Hayward SA, Innes SM, Miller ASC (2020). Point-of-care lung ultrasound in patients with COVID-19 - a narrative review. Anaesthesia. **75**(8): 1096–1104.
39 Soldati G, Demi M, Inchingolo R, Smargiassi A, Demi L (2016). On the Physical Basis of Pulmonary Sonographic Interstitial Syndrome. J Ultrasound Med. **35**(10): 2075–2086.
40 Soldati G, Demi M, Smargiassi A, Inchingolo R, Demi, L (2019). The role of ultrasound lung artifacts in the diagnosis of respiratory diseases. Expert Rev Respir Med. **13**(2): 163–172.
41 Soldati G, Smargiassi A, Inchingolo R, Buonsenso D, Perrone T, Briganti DF, et al. (2020). Is There a Role for Lung Ultrasound During the COVID-19 Pandemic? J Ultrasound Med. **39**(7): 1459–1462.
42 Takamatsu A, Ueno M, Yoshida K, Kobayashi T, Kobayashi S, Gabata T (2024). Performance of artificial intelligence-based software for the automatic detection of lung lesions on chest radiographs of patients with suspected lung cancer. Jpn J Radiol. **42**(3): 291–299.
43 Tan G, Lian X, Zhu Z, Wang Z, Huang F, Zhang Y, et al. (2020). Use of Lung Ultrasound to Differentiate Coronavirus Disease 2019 (COVID-19) Pneumonia from Community-Acquired Pneumonia. Ultrasound Med Biol. **46**(10): 2651–2658.
44 Wang J, Yang X, Zhou B, Sohn JJ, Zhou J, Jacob JT, et al. (2022). Review of Machine Learning in Lung Ultrasound in COVID-19 Pandemic. J Imaging. **8**(3): 65.
45 Zhang Y, Xue H, Wang M, He N, Lv Z, Cui L (2021). Lung Ultrasound Findings in Patients with Coronavirus Disease (COVID-19). AJR Am J Roentgenol. **216**(1): 80–84.