# Lung Ultrasound Imaging and Image Processing with Artificial Intelligence Methods for Bedside Diagnostic Examinations

**Gábor Orosz[1,2,*], Róbert Zsolt Szabó[1,3*], Tamás Ungi[4], Colton Barr[4], Chris Yeung[4], Gábor Fichtinger[4], János Gál[2], and Tamás Haidegger[1,5]**

[1]Antal Bejczy Center for Intelligent Robotics, Óbuda University, EKIK; Becsi ut 96/B, Budapest, 1034, Hungary, gabor.orosz@irob.uni-obuda.hu
[2]Semmelweis University, Department of Anesthesiology and Intensive Therapy; Ulloi ut 78, Budapest, 1083, Hungary, gal.janos@med.semmelweis-univ.hu
[3]John von Neumann Faculty of Informatics, Óbuda University, Budapest, Hungary, robert.szabo@irob.uni-obuda.hu
[4]Laboratory for Percutaneous Surgery, Queen's University; 2K7L 2N8, ON, Canada, {ungi,c.barr,chris.yeung,gabor}@queensu.ca
[5]Austrian Center for Medical Innovation and Technology; Viktor-Kalpan-str. 2., Wiener Neustadt, 2700, Austria, haidegger@acmit.at
[*]*These authors contributed equally*

*Abstract: Artificial Intelligence-assisted radiology has shown to offer significant benefits in clinical care. Physicians often face challenges in identifying the underlying causes of acute respiratory failure. One method employed by experts is the utilization of bedside lung ultrasound, although it has a significant learning curve. In our study, we explore the potential of a Machine Learning-based automated decision-support system to assist inexperienced practitioners in interpreting lung ultrasound scans. This system incorporates medical ultrasound, advanced data processing techniques, and a neural network implementation to achieve its objective. The article provides a comprehensive overview of the steps involved in data preparation and the implementation of the neural network. The accuracy and error rate of the most effective model are presented, accompanied by illustrative examples of their predictions. Furthermore, the paper concludes with an evaluation of the results, identification of limitations, and recommendations for future enhancements.*

*Keywords: AI-based image processing; Surgical Data Science; Applied medical imaging; Deep neural networks; Lung ultrasound*

## 1 Introduction

Over the past decade, lung ultrasound has emerged as a widely utilized diagnostic tool [1–6], particularly in bedside examinations [7, 8], owing to the advent

of portable and handheld devices, and even supporting robotic surgery applications [9–12]. Critical care [13], anesthesia [14], and emergency medicine [15] are among the primary medical specialties that have embraced this technique. Proficiency in recognizing lung ultrasound artifacts is essential for practitioners in these fields [16]. The critical care perspective played a pivotal role in demonstrating the utility of this technique to the medical community [17–19]. Unlike general ultrasonography, lung ultrasound examinations focus on identifying and analyzing patterns of artifacts [20]. Therefore, lung ultrasonography has become a distinct clinical modality within acute care specialties.

## 1.1 Sustainable Radiology

Ultrasound is also seen as an affordable, wide-spread diagnostic tool, subject to continuous innovation, able to bring sustainability to the imaging domain of modern medicine [21, 22].

Emergency ultrasonography has emerged as a widely researched application, supporting euqal access to care. This is aligned with the United Nations' Sustainable Development Goals (SDG) – especially with SDG 3 for promoting well-being across all age groups [23].

Notably, lung emergency ultrasound offers an eco-friendly and non-ionizing imaging option, particularly valuable in resource-limited regions and during pandemics like COVID-19.

In low- and middle-income countries (LMICs), diagnostic imaging is often inadequate [24], but clinician-performed, hand-carried, bedside ultrasound has gained popularity globally. Affordable, portable, and user-friendly machines have expanded its reach, bolstering diagnostic capabilities in rural hospitals.

The use of ultrasound in LMICs has gained recognition from health ministries, non-governmental organizations, and the World Health Organization (WHO). It significantly improves patient diagnosis and management, particularly in remote regions [25].

## 1.2 Automating Ultrasound Diagnostics

Our research aims to develop an automated ultrasound system, conserving resources and aiding in training and diagnostics [26]. This innovation holds the potential to transform healthcare delivery in resource-limited settings, empowering local practitioners and enhancing patient care. Emergency lung ultrasonography represents a powerful tool in achieving sustainable healthcare goals in underserved regions. Through our efforts to create an automated solution, we hope to strengthen healthcare infrastructure and reduce disparities, fostering improved health outcomes for all.

The utilization of ultrasound imaging in lung examinations presents distinct challenges. These difficulties arise from the presence of the rigid chest structure and the unsuitability of air as an ultrasound medium. As a result, the emergence of acoustic

shadows, a common occurrence due to the transducer's orientation perpendicular to the ribs, becomes noteworthy. Nevertheless, adept manipulation and positioning of the transducer can mitigate this artifact.

In the context of air-filled lungs, the ultrasound beam experiences complete reflection at the boundary between soft tissue and air or fluid. This interaction yields a tangible image of the chest's soft tissues and pleural layers, accompanied by distinctive artifact patterns. The crux of lung ultrasound examinations revolves around the detection and analysis of these artifacts. Among the array of artifacts, reverberation artifacts are prominent. In physiological conditions, they are referred to as A-lines, while in pathological contexts, they are predominantly known as B-lines. [14]. The nomenclature of these artifacts originates from the work of Lichtenstein et al. [19, 27], and consensus has been reached regarding their application [28, 29]. An atelectatic lung appears as a true tissue image, and the description of each significant pathological phenomenon is provided within its respective field of application.

## 1.3 Technology Against a Pandemic

During the COVID-19 pandemic, there was a rapid development and deployment of supporting technologies for diagnosis and treatment [30–33]. Lung ultrasound emerged as a reliable alternative to chest X-rays, exhibiting comparable accuracy to CT scans, which are considered the gold standard for lung imaging. According to a Cochrane review on thoracic imaging tests for diagnosing COVID-19, chest CT and lung ultrasound were found to be sensitive and moderately specific. Therefore, these modalities may be more valuable in ruling out COVID-19 than in differentiating SARS-CoV-2 infection from other respiratory illnesses [34].

## 1.4 Medical Background

Ultrasound stands as a preferred imaging modality in various scenarios owing to its array of advantages over other methods. Foremost, its cost-effectiveness renders it more accessible to patients. Furthermore, ultrasound furnishes real-time imaging capabilities, enabling instantaneous observation of anatomical structures and physiological dynamics. Distinct from alternative techniques like X-rays, ultrasound circumvents the use of ionizing radiation, thereby ensuring patient safety during recurrent examinations without the potential for adverse repercussions. [35].

The portability of ultrasound devices is another significant advantage, enabling their use at the bedside. This eliminates the need to transport patients within the hospital, reducing the risk of spreading infectious diseases and minimizing the challenges associated with handling critically ill and unstable patients [36]. In situations where other scanning methods may be impractical or unsafe, bedside lung ultrasound has emerged as a viable solution. Moreover, its diagnostic accuracy is comparable to or even superior to conventional radiographic measures, further validating its utility. However, a notable challenge with lung ultrasound lies in the interpretation of images, which demands experienced medical professionals. Diagnosis through ultrasound is inherently subjective and heavily reliant on the competency, experience, and mental state of the performing physicians. Factors like stress or exhaustion can

impact the accuracy of interpretations [20]. Given that lung ultrasound is frequently performed on critically ill patients, prompt and accurate diagnosis is of utmost importance. As not all physicians performing the examination possess expertise in interpreting lung ultrasound, the provision of a decision support system becomes crucial in expediting accurate diagnoses.

Moreover, despite the generally acknowledged safety and reproducibility of ultrasound assessments, the fundamental "ALARA" (As Low As Reasonably Achievable) principle remains pivotal in curtailing ultrasonography exposure. Hence, expediting the diagnostic process assumes significance from a safety standpoint. Regrettably, the absence of uniform standards hinders the seamless comparison of ultrasound investigations and their outcomes. Acknowledging this concern, authorities within the realm of lung ultrasound have advocated for the establishment of standardization guidelines encompassing machine configurations and scanned regions. [5, 31].

In addition to machine settings, the selection of the optimal scanning technique (such as longitudinal or transverse) remains a topic of ongoing discourse. This debate can give rise to potential misinterpretations or misdiagnoses in routine clinical settings. Such ambiguity may lead to uncertainty among medical personnel and elongate the duration of lung ultrasound assessments, ultimately extending patient contact time. It becomes imperative to address these challenges pertaining to standardization and scanning techniques, as doing so holds the key to enhancing the efficiency and dependability of lung ultrasound procedures. [37].

## 1.5 The BLUE protocol

The BLUE (Bedside Lung Ultrasound in Emergency) protocol, developed by Daniel Liechtenstein [19], is a clinical procedure commonly used in intensive care medicine to rapidly diagnose the underlying causes of acute respiratory failure (Fig. **??**). When administered by skilled practitioners, the BLUE protocol has demonstrated its efficacy and potency as a tool for assessing lung conditions in critically ill individuals. Furthermore, it is noteworthy to emphasize the exceptional adaptability of the BLUE protocol, which readily lends itself to a seamless evolution into the advanced semi-quantitative framework named BLUE-LUSS (Bedside Lung Ultrasound in Emergency-Lung UltraSound Score). This innovative concept, previously expounded upon by the authors in a prior publication, signifies a testament to the protocol's progressive potential.n [37]. This remarkable evolution empowers medical professionals to leverage the capabilities of numerical scoring, ushering in a paradigm shift towards enhanced precision and objectivity within the domain of diagnostic assessment.

Previously, the protocol relied on still frames of ultrasound images captured from standardized locations. However, recent advancements in the field have demonstrated the superiority of utilizing short ultrasound loops, typically ranging from 3 to 10 seconds in duration [31]. These dynamic ultrasound loops provide a more comprehensive view of the lung, allowing for a more accurate assessment of lung artifacts and abnormalities.

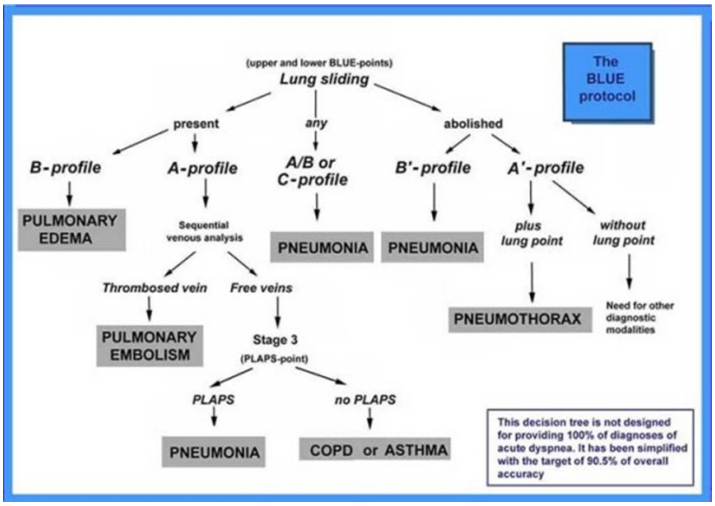


Figure 1
The core objective of the BLUE protocol revolves around streamlining the diagnostic timeline. This is achieved through the utilization of predefined points of analysis for ultrasound loop recording, coupled with a straightforward decision tree catering to the primary causes of acute respiratory failure. The protocol's design focuses on expediting the diagnostic process and enhancing its efficiency. Adopted from Lichtenstein et al.

The main objective of the BLUE protocol is to categorize lung artifacts into a distinct "lung profile" based on anatomical signs and visual artifacts observed on the ultrasound images. By following a predefined decision tree outlined in the protocol, physicians can make on-the-spot diagnoses and identify the underlying causes of acute respiratory failure. The protocol has shown remarkable diagnostic accuracy, successfully identifying the six most common diseases associated with acute respiratory failure in 97 percent of cases, with an overall accuracy of 90.5 percent [35].

The profiling of the lung involves detecting specific signs in the ultrasound loops. The protocol defines ten basic signs, which can be further categorized into primary and secondary signs. Primary signs, such as the pleural line, A-line, quad sign, fractal sign, tissue-like sign, and B-lines or lung rockets, can be observed on single still frames without considering the entire ultrasound loop [19]. These signs serve as important indicators of lung pathology and assist in the diagnostic process.

In contrast, secondary signs, including lung sliding, sinusoid sign, and lung point, require a comprehensive analysis of the entire video loop. These signs often exhibit movement or dynamic patterns that are not discernible on still frames alone. Detecting these secondary signs necessitates a thorough examination of the ultrasound loop as a whole, considering the temporal progression of lung artifacts.

To streamline the identification of initial indicators, we've examined different methods, such as the utilization of 2D multiclass semantic segmentation methodologies. Through the segmentation of static frames, it becomes feasible to train deep learning

algorithms in precisely recognizing and categorizing various pulmonary anomalies. This automatic identification of primary indicators shows potential in accelerating and establishing a uniform diagnostic procedure, especially for medical practitioners with limited experience or within scenarios requiring prompt decisions.

Concurrently, we've undertaken the development of a classification framework that is specifically concentrated on discerning the presence or absence of lung sliding. This secondary indicator holds pivotal importance in distinguishing among diverse pulmonary conditions. Leveraging an extensive, custom-designed clinical dataset, we've educated convolutional neural networks (CNNs) to scrutinize B-mode (2D mode) image sequences and transform them into M-mode ("motion" mode) images. These M-mode images facilitate the extraction of sample lines by proficient annotators, thus facilitating the training and assessment of the CNNs. The most proficient CNN achieved an impressive precision rate of 93.0

In summary, the BLUE protocol has risen as a valuable instrument for swiftly diagnosing acute respiratory failure, especially within intensive care environments. The transition from static images to brief ultrasound sequences has notably elevated the precision and effectiveness of lung assessments. The continuous endeavors in automating the identification of initial indicators using deep learning methodologies demonstrate encouraging outcomes, laying the groundwork for more consistent and easily accessible diagnostic procedures. Furthermore, the creation of a classification network dedicated to identifying lung sliding enhances the comprehensive diagnostic prowess of the BLUE protocol. These innovations hold the capacity to transform the landscape of lung ultrasound analysis, empowering timelier and more precise diagnoses for critically unwell patients.

# 2 Methods

## 2.1 Deep Neural Networks

In the past few years, there has been a growing trend in employing deep neural networks for enhancing the precision and effectiveness of medical procedures within the realm of medical imaging. These sophisticated neural networks have found application across a range of medical imaging assignments, encompassing tasks like object identification, segmentation, and the reconstruction of images [38, 39].

Illustratively, convolutional neural networks have found utility in discerning and categorizing diverse irregularities present in radiological images. Noteworthy instances encompass the detection of tumors within MRI scans or the identification of lesions in dermatological images. As an illustration, a study conducted in 2020 focused on utilizing a convolutional neural network to classify chest X-ray images. The network was primed for binary categorization and successfully achieved a commendable accuracy rate of 94.6

In addition to its advantages, it's crucial to acknowledge the challenges associated with the application of deep learning networks in medical image analysis. Among

the foremost hurdles is the establishment of a dataset containing a requisite quantity and caliber of annotated medical data. The task of annotation involves the expertise of medical professionals, who furnish the reference truth necessary for training and appraising the network's performance. Given that annotation demands the involvement of these specialized medical experts, procuring a substantial volume of annotated medical images often proves to be arduous and complex.

Deep neural networks have demonstrated their effectiveness in the scrutiny of lung ultrasound images as well. Over the past years, a plethora of research findings have emerged in this domain, examining the efficacy of deep neural networks in identifying various irregularities within lung ultrasound images.

An additional instance is found in a study authored by Cheng and Lam in 2021. In this research, a customized U-Net architecture was employed for the binary segmentation of lung ultrasound images. Notably, the encoder component of the U-Net was substituted with a VGG16 network that had been pre-trained on the ImageNet database. This hybrid architecture was subsequently trained on a dataset comprising 400 ultrasound images tailored to the specific challenge. The outcome was a Dice score of 0.86, underscoring the success achieved in accurate binary segmentation.6 [40–42].

In a separate investigation documented in 2020, Roy et al. explored the segmentation of COVID biomarkers within lung ultrasound images. Their approach hinged on a U-Net-based model, which underwent training using a comprehensive set of 277 annotated ultrasound loops. This concerted effort yielded a noteworthy 96

These studies unequivocally showcase the practicality and relevance of employing deep neural networks for the realm of medical image processing.

## 2.2 U-Net

U-Net constitutes a fully convolutional neural network architecture that was initially devised specifically to address medical image segmentation assignments [42]. Since its original introduction in 2015, U-Net has persisted as the foundational framework for numerous cutting-edge neural network architectures, particularly within the sphere of medical image analysis, up until the present day. [43, 44]. The architecture of the network employs an encoder-decoder design, enriched by skip connections. The network's initial segment is the encoder, often referred to as the "contracting path," where each stage progressively diminishes the dimensions of the input image in terms of width and height, while simultaneously amplifying the extracted features. Subsequently, the decoder component (termed the "expanding path") methodically upscales the input and carries out additional convolutions.

A significant aspect of this architecture is the incorporation of skip connections. These connections link the output activation map of convolutional layers in the encoder segment to the input of convolutional layers in the decoder segment, preserving a match in the width and height dimensions between the encoder's output and the decoder's input. This strategic linkage allows the network to retain and integrate more information from preceding layers. This encompasses retaining the original

spatial characteristics of the input image, a pivotal attribute for segmentation undertakings.

# 3 Implementation

Although the present implementation is a considerable distance from direct clinical utilization, the chosen methodologies for its development and prototyping hold the potential to pave the way for a more robust and diagnosis-focused decision support system.

## 3.1 Approach

The central objective addressed within this paper revolves around the identification of fundamental indicators present in lung ultrasound images, bearing significance to the BLUE protocol. To address this objective, we employ a deep neural network tailored for multi-class semantic segmentation. This process is executed on every annotated frame contained within the captured ultrasound sequences. Essentially, this undertaking can be deconstructed into two primary components, each of which can be further subdivided into smaller constituent subtasks:

- Data Conditioning – in this step, we prepare and arrange our raw data into a format that is suitable for network training;
- Network Implementation – the actual implementation, training, and fine tuning of the neural network.

In the following sections we discuss both stages in detail.

## 3.2 Imaging protocol

Following the method developed by Daniel Lichtenstein, the creator of the BLUE and PINK protocols, we adhered to his specified approach of using three designated points on each side of the chest to gather imaging data. The imaging was integrated into the daily medical routine. Informed by prior expert analyses and considering the uneven and sporadic nature of COVID-19 pneumonia, we captured video loops in both longitudinal and transverse orientations. The specific locations and orientations were systematically labeled using predetermined codes. Our protocol enabled a single operator to conduct the assessments without the need for additional assistance. Given that all the examinations were carried out by lung ultrasound (LUS) specialists with over seven years of experience, we didn't document the precise duration of each examination, as it wasn't deemed significant. The assessment of the loops was generally done off-line. However, if the operator identified an urgent situation during the analysis, such as an emergency equivalent, they immediately communicated this to the clinical team. An example of this occurred during one examination when an acute pneumothorax on the right side was unexpectedly detected, leading to a timely intervention that ultimately saved the patient's life.

| Basic Settings | Obligatory value/range |
|---|---|
| Depth | Pleural line + 3–8 cm |
| Focus | Multifocus: OFF; Single focus: at pleural line |
| Gain | Optimized for main artifacts (A-lines, B-lines, consolidation) |
| Image-processing features | THI: OFF, XRes: OFF, CrossXBeam: OFF, SRI: OFF |
| MI | $< 0.7$ |
| TIs | $< 0.1$ |

Table 1
Abbreviations: MI: mechanical index; TI: soft tissue thermal index; THI: tissue harmonic imaging; XRes: speckle noise reduction; CrossXBeam: spatial compounding; SRI: speckle reduction

## 3.3 Ultrasound settings, data collection protocol

Our objective was to establish a consistent and comprehensive US examination procedure. Utilizing specific settings on our equipment, we successfully gathered a collection of images that effectively depicted the pleura and underlying artifacts, ensuring no critical observations were overlooked. Prioritizing patient safety, we carefully regulated the thermal index (TI) and mechanical index (MI) to adhere to global safety norms. The devices employed included a Philips CX50 (Philips Healthcare, The Netherlands) with a Philips C5-1 convex probe (1-5 MHz) and a GE Venue GO (GE Healthcare, IL, USA) with a C1-5-RS convex probe (1.4-5.7 MHz). Essential default settings are detailed in Table 1. Drawing from our clinical experience and aligning with global practices in lung ultrasound (LUS), we opted to capture video loops lasting 4–6 seconds instead of single images, for enhanced accuracy. The operator of the equipment was not informed about the patient's clinical progress, and did not participate in their care. These loops were recorded under a pseudo-anonymized system (using a unique patient code) and stored on the hard drive in DICOM (Digital Imaging and Communication in Medicine) format, fully compliant with GDPR (EU General Data Protection Regulation) guidelines. Subsequently, patient information was transferred to an encrypted, Microsoft Excel-based database, through a triple-layered security protocol.

## 3.4 Imaging protocol

Following the method developed by Lichtenstein, the creator of the BLUE and PINK protocols, we adhered to his specified approach of using three designated points on each side of the chest to gather imaging data. The imaging was integrated into the daily medical routine. Informed by prior expert analyses, and considering the uneven and sporadic nature of COVID-19 pneumonia, we captured video loops in both longitudinal and transverse orientations. The specific locations and orientations were systematically labeled using predetermined codes.

Our protocol enabled a single operator to conduct the assessments without the need

for additional assistance. Given that all the examinations were carried out by LUS specialists with over seven years of experience, we did not document the precise duration of each examination, as it was not deemed significant. The assessment of the loops was generally done off-line. However, if the operator identified an urgent situation during the analysis, such as an emergency equivalent, they immediately communicated this to the clinical team. An example of this occurred during one examination when an acute pneumothorax on the right side was unexpectedly detected, leading to a timely intervention that ultimately saved the patient's life.

## 3.5 Data Conditioning

The dataset harnessed for this endeavor has been meticulously curated by medical professionals affiliated with Semmelweis University, Budapest. In addition to supplying the requisite data, these experts contributed their invaluable medical acumen to the project. This encompassed the meticulous annotation of raw data and ensuring that the project's trajectory was aligned with a clinically coherent perspective.

As of the time of composing this article, our repository comprises data sourced from 22 RT-PCR confirmed SARS-CoV-2 infected patients, in addition to data extracted from 18 RT-PCR confirmed SARS-CoV-2 negative ("non-COVID") patients. This data has been meticulously gathered by Semmelweis University's Department of Anaesthesiology and Intensive Therapy. Notably, all the data acquisition procedures adhere rigorously to the BLUE protocol, a universally recognized and adopted standardized methodology. This meticulous adherence ensures the replicability of the procedure and the coherence of our data.

Each patient's dataset comprises multiple ultrasound loops, with each loop spanning 5 to 6 seconds in duration. These loops are captured from distinct points as delineated by the BLUE protocol. Given the aggregate of 40 patients, the dataset encompasses a cumulative count of approximately 630 loops, encompassing over 200,000 individual frames. Our efforts to augment this dataset are ongoing, involving the collaboration of additional hospitals and research establishments. To uphold the ethical considerations, all our data is pseudonymized, and we have procured explicit authorization from Semmelweis University's Research Ethics Committee to utilize this dataset. In general, most recent ethically aligned engineering design principles (IEEE 7000, IEEE 70007 standards) were followed during the project [45].

In the process of data preparation, we extensively leverage the capabilities of the 3D Slicer tool. [46]. Certainly, 3D Slicer stands out as an open-source, research-centric image processing tool, primarily geared towards medical imaging tasks. While its scope is not limited solely to medical imaging, it excels in this domain. A notable attribute of 3D Slicer is its robust segmentation capabilities. Moreover, the tool offers programmable extension capabilities and a conveniently embedded Python command line, which significantly facilitates data transformation and extraction procedures. These inherent functionalities align seamlessly with our objectives. We adeptly harness the built-in segmentation module to annotate our raw ultrasound loops, and the Python interface expedites the automated extraction of this annotated data, rendering 3D Slicer an ideal fit for our requirements.

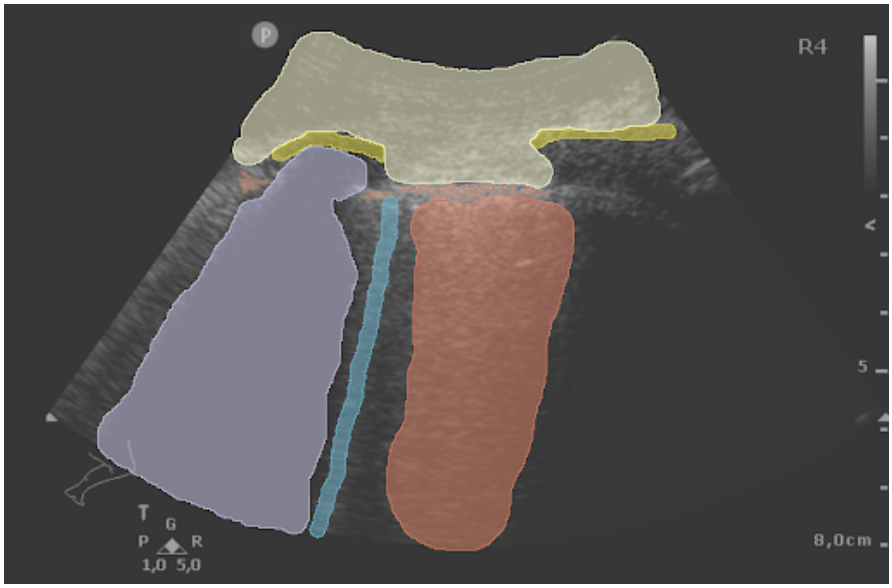


Figure 2
A manually segmented ultrasound frame. Visible segmented classes are: thoracic wall (pale yellow), rib (yellow), rib shadow (purple), B-line (blue), septal rockets (red), unhealthy pleural line (pale orange)

In the initial step, the raw data undergoes annotation under the supervision of medical experts within the 3D Slicer platform. This process involves the manual delineation of segments pertinent to our particular scenario. These encompass the primary indicators outlined earlier, accompanied by a select few supplementary landmark segments. The segmentation occurs on a frame-by-frame basis, with every fifth frame within each loop being subjected to this procedure. It's noteworthy that the entirety of the loop doesn't necessarily require annotation, as capturing a complete breathing cycle suffices for our purposes.

Upon completion of annotation, the subsequent phase involves the extraction of data from 3D Slicer. As not every individual frame undergoes segmentation, identification of frames with attached segmentation data becomes imperative. Moreover, a series of standardization and cleanup steps are necessary for each relevant frame. Certain extraneous details, originating from the ultrasound machine itself (such as machine settings and brand), are irrelevant to our analysis. To address this, a mask is applied to retain solely the pertinent ultrasound image, effectively eliminating the extraneous information. Additionally, the images are transformed into a uniform quadratic format, specifically 256x256 pixels.

To achieve this quadratic format, an initial resizing is performed, such that the larger dimension of the original image matches the desired 256 pixels, while maintaining the original aspect ratio. During this resizing, linear interpolation is employed for the ultrasound images, while nearest neighbor interpolation is used for the segmentation masks. Subsequently, the smaller dimension is extended to reach 256 pixels

using zero values, while preserving the central segment of the image.

Post-cleanup and extraction, these relevant frames are integrated to constitute a comprehensive dataset. This dataset is structured on a patient-specific basis, ensuring that data from a given patient remains exclusive either to the training or validation set. This segregation mimics real-world application scenarios, where the trained model is applied to new, unknown patients' BLUE ultrasound loops. This setup ensures more realistic outcomes in model evaluation.

At present, our dataset encompasses a total of 1097 manually segmented ultrasound frames, derived from the data of four patients.

### 3.6 Network Training

Our approach involved implementing a standard U-Net model, enhanced by the incorporation of zero-padding within the convolutional layers. This modification was strategically employed to preserve the original dimensions of the image. In the testing phase, our network was configured with an input size of 256 by 256 pixels, and each convolutional layer was equipped with 16 filters, contributing to the model's feature extraction capabilities.

During the training process, we adopted a leave-one-out cross-validation technique rooted in patient-centric segmentation. This methodology entails setting aside the data from a single patient as the validation set in each round, while the model is simultaneously trained on the remaining dataset. This cross-validation scheme ensures comprehensive validation and helps mitigate potential overfitting.

In evaluating the performance of our model, we focused on two key metrics: validation accuracy and validation loss. To quantify the model's predictive accuracy and assess the fidelity of its output, we computed the validation accuracy. Meanwhile, the validation loss, measured using the sparse categorical cross-entropy loss function, provides insight into the dissimilarity between predicted and actual segmentation outcomes. This comprehensive approach offers a robust framework for effectively gauging the model's proficiency in the context of medical image analysis.

## 4 Results and Discussion

By employing the method outlined in the preceding section, we attained an average validation accuracy of 87.53

At this juncture, the most pronounced challenge lies in the scarcity of available annotated data. With the present dataset size, the model exhibits signs of overfitting within only a few training epochs, necessitating premature training termination. Our forthcoming efforts will be directed towards addressing this limitation. We intend to bolster our annotated data reservoir and explore diverse data augmentation techniques to alleviate this challenge effectively.

(a)
Ground truth

(b)
Model prediction

Figure 3
Illustrative Prediction Outcome of the Model: Noteworthy Segmentation Classes include the Thoracic Wall (highlighted in yellow), Ribs (displayed in orange), Rib Shadows (rendered in pale orange), and B-lines (depicted in purple).

Once a substantial annotated dataset is amassed, our focus will shift towards conducting a comprehensive hyperparameter sweep to determine the optimal configuration. While we've engaged in some level of hyperparameter optimization, the relatively modest dataset size has somewhat limited its impact. Once segmentation attains a satisfactory level of performance, our endeavors will extend towards identifying the secondary signs elucidated in the second chapter. Subsequently, we will be poised to embark on the comprehensive implementation of the entire decision tree integral to the BLUE protocol.

Another avenue for potential enhancement lies in revisiting the roster of segmentation classes currently utilized for model training. Streamlining the model by omitting or consolidating specific classes that do not contribute to relevant diagnostic insight could potentially simplify training and expedite the annotation process. An associated adjustment could involve modifying the loss function to facilitate the assignment of varying weights to individual segmentation classes. Notably, even in the existing segmentation schema, the prominence of the background class outweighs others. This concern could intensify with a reduced number of segments, and recalibrating the significance of the background class might yield improved outcomes.

Fine-tuning the model architecture and experimenting with diverse model types to facilitate performance comparison is another avenue for potential improvement. However, the dearth of sufficiently annotated data has thus far hindered the pursuit of this optimization. Given the present stage, a meaningful model architecture comparison might be elusive.

In addition to enhancing the current model, we intend to reevaluate our use case scenario to explore alternative approaches beyond segmentation. Our foremost challenge is the time-consuming process of manual annotation, particularly the meticulous contouring required for each frame. We're actively considering the prospect of leveraging object detection methods as an alternative. By employing this approach, annotating would involve simply drawing bounding boxes around objects rather than intricate contouring, thereby significantly expediting the annotation process. This transition alone could prove advantageous, even if the new approach doesn't yield immediate tangible benefits.

Another avenue involves the integration of the segmentation model with a medical ultrasound device, enabling real-time ultrasound image segmentation at the patient's bedside. While not directly leading to diagnosis, this real-time segmentation could still furnish valuable information to the attending physician during the examination, potentially aiding in timely decision-making.

These potential directions underscore our commitment to refining and innovating upon our current framework to overcome existing challenges and enhance the clinical utility of our approach.

Future research endeavors encompass the pursuit of full automation in image acquisition by integrating a collaborative robotic arm. This innovative approach holds the promise of enhancing data consistency, a crucial facet for accurate analysis. Moreover, such automation could prove particularly advantageous in pandemic scenarios, mitigating the risk of human exposure and facilitating safe and efficient medical

procedures. [33, 47].

# 5 Conclusion

Our research is dedicated to advancing computer-assisted support for the BLUE protocol. In the initial phase, we have successfully implemented a U-Net based deep neural network tailored for segmenting and identifying vital anatomical elements within lung ultrasound images.

The outcomes thus far reveal a highly promising conceptual foundation. Notwithstanding the modest count of annotated images and the brief training duration, the model's predictions exhibit discernible alignment with the ground truth. This alignment is particularly pronounced for segments that are more prevalent in nature, such as the thoracic wall or rib shadows. This initial progress underscores the potential of our approach to significantly enhance the diagnostic capabilities of the BLUE protocol through computer-assisted assistance.

**Future work**

Our current project is an essential component of a modular decision support system, incorporating artificial intelligence solutions. As per our vision, this system can be immensely beneficial to our clinical colleagues. The foundation of the system relies on well-validated and widely used emergency lung ultrasound protocols.

Our primary objective is to implement and enhance a framework that offers real-time support to physicians during bedside examinations. Additionally, this framework aims to facilitate the acquisition of lung ultrasound skills as part of structured medical training.

We plan to further advance the system by gathering more clinical data and expanding our annotated data bank. This expansion is expected to improve the system's metrics, making it suitable for real testing in clinical conditions.

Looking ahead, we also consider the possibility of automating and robotizing the system in various directions. Such advancements could lead to additional cost savings in terms of human resources for healthcare systems that operate within limited frameworks.

**Acknowledgment**

The authors thank the support of Roland Incze in the research.

C. Barr is supported by an NSERC CGS-D. G. Fichtinger is supported as an NSERC Canada Research Chair.

R. Szabó was supported by the MITACS Globalink Research Award.

T. Haidegger's work is partially supported by project no. 2019-1.3.1-KK-2019-00007, provided by the National Research, Development and Innovation Fund of

Hungary. T. Haidegger is a "Consolidator Researcher", supported by the Distinguished Researcher program of Óbuda University.

## References

[1] F. Mojoli, B. Bouhemad, S. Mongodi, and D. Lichtenstein. Lung ultrasound for critically ill patients. *American journal of respiratory and critical care medicine*, 199(6):701–714, 2019.

[2] S. Mongodi, D. De Luca, A. Colombo, A. Stella, E. Santangelo, F. Corradi, L. Gargani, S. Rovida, G. Volpicelli, B. Bouhemad, et al. Quantitative lung ultrasound: technical aspects and clinical applications. *Anesthesiology*, 134(6):949–965, 2021.

[3] R. Brat, N. Yousef, R. Klifa, S. Reynaud, S. S. Aguilera, and D. De Luca. Lung ultrasonography score to evaluate oxygenation and surfactant need in neonates treated with continuous positive airway pressure. *JAMA pediatrics*, 169(8):e151797–e151797, 2015.

[4] B. Bouhemad, S. Mongodi, G. Via, and I. Rouquette. Ultrasound for "lung monitoring" of ventilated patients. *Anesthesiology*, 122(2):437–447, 2015.

[5] G. Volpicelli, M. Elbarbary, M. Blaivas, D. A. Lichtenstein, G. Mathis, A. W. Kirkpatrick, L. Melniker, L. Gargani, V. E. Noble, G. Via, et al. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive care medicine*, 38:577–591, 2012.

[6] S. Mongodi, B. Bouhemad, A. Orlando, A. Stella, G. Tavazzi, G. Via, G. A. Iotti, A. Braschi, and F. Mojoli. Modified lung ultrasound score for assessing and monitoring pulmonary aeration. *Ultraschall in der Medizin-European Journal of Ultrasound*, 38(05):530–537, 2017.

[7] M. Szabó, A. Bozó, K. Darvas, S. Soós, M. Őzse, and Z. D. Iványi. The role of ultrasonographic lung aeration score in the prediction of postoperative pulmonary complications: an observational study. *BMC anesthesiology*, 21:1–10, 2021.

[8] A. Dargent, E. Chatelain, L. Kreitmann, J.-P. Quenot, M. Cour, L. Argaud, and C.-L. study group. Lung ultrasound score to monitor covid-19 pneumonia progression in patients with ards. *PLoS One*, 15(7):e0236312, 2020.

[9] R. Elek, T. D. Nagy, D. A. Nagy, B. Takács, P. Galambos, I. Rudas, and T. Haidegger. Robotic platforms for ultrasound diagnostics and treatment. In *2017 IEEE international conference on systems, man, and cybernetics (SMC)*, pages 1752–1757. IEEE, 2017.

[10] G. Fichtinger, J. Troccaz, and T. Haidegger. Image-guided interventional robotics: Lost in translation? *Proceedings of the IEEE*, 110, 2022.

[11] T. Haidegger, S. Speidel, D. Stoyanov, and R. Satava. Robot-assisted minimally invasive surgery—surgical robotics in the data age. *Proceedings of the IEEE*, 110:835–846, 2022.

[12] T. D. Nagy and T. Haidegger. Performance and capability assessment in surgical subtask automation. *Sensors*, 22:2501, 2022.

[13] R. Raheja, M. Brahmavar, D. Joshi, and D. Raman. Application of lung ultrasound in critical care setting: a review. *Cureus*, 11(7), 2019.

[14] M. Szabó, G. Orosz, Z. D. Iványi, and K. Darvas. Use of thoracic and lung ultrasound in general anesthesia. *Orvosi Hetilap*, 164(22):864–870, 2023.

[15] P. S. Pang, F. M. Russell, R. Ehrman, R. Ferre, L. Gargani, P. D. Levy, V. Noble, K. A. Lane, X. Li, and S. P. Collins. Lung ultrasound–guided emergency department management of acute heart failure (blushed-ahf) a randomized controlled pilot trial. *Heart Failure*, 9(9):638–648, 2021.

[16] C. Urbán, P. Galambos, G. Györök, and T. Haidegger. Simulated medical ultrasound trainers a review of solutions and applications. *Acta Polytechnica Hungarica*, 15(7):111–131, 2018.

[17] D. A. Lichtenstein and G. A. Meziere. Relevance of lung ultrasound in the diagnosis of acute respiratory failure*: the blue protocol. *Chest*, 134(1):117–125, 2008.

[18] D. A. Lichtenstein. Lung ultrasound in the critically ill. *Annals of intensive care*, 4(1):1–12, 2014.

[19] D. Lichtenstein. *Lung Ultrasound in the Critically Ill: The BLUE Protocol*. Springer International Publishing, 2015.

[20] G. Soldati, A. Smargiassi, L. Demi, and R. Inchingolo. Artifactual lung ultrasonography: It is a matter of traps, order, and disorder. *Applied sciences*, 10(5):1570, 2020.

[21] A. Khamis, H. Li, E. Prestes, and T. Haidegger. Ai: a key enabler of sustainable development goals, part 1. *IEEE Robotics & Automation Magazine*, 26(3):95–102, 2019.

[22] A. Khamis, H. Li, E. Prestes, and T. Haidegger. Ai: A key enabler for sustainable development goals: Part 2. *IEEE Robotics & Automation Magazine*, 26(4):122–127, 2019.

[23] T. Haidegger, V. Mai, C. Mörch, D. Boesl, A. Jacobs, B. Rao R, A. Khamis, L. Lach, and B. Vanderborght. Robotics: Enabler and inhibitor of the sustainable development goals. *Sustainable Production and Consumption*, 43:422–434, 2023.

[24] H. Ostensen. Developing countries. *Ultrasound in Medicine and Biology*, 26:S159–S161, 2000.

[25] S. Sippel, K. Muruganandan, A. Levine, and S. Shah. Use of ultrasound in the developing world. *International journal of emergency medicine*, 4(1):1–11, 2011.

[26] R. Z. Szabó, G. Orosz, T. Ungi, C. Barr, C. Yeung, R. Incze, G. Fichtinger, J. Gál, and T. Haidegger. Automation of lung ultrasound imaging and image processing for bedside diagnostic examinations. In *2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, pages 779–784. IEEE, 2023.

[27] D. A. Lichtenstein. Current misconceptions in lung ultrasound: a short guide for experts. *Chest*, 156(1):21–25, 2019.

[28] S. Mongodi, E. Santangelo, D. De Luca, S. Rovida, F. Corradi, G. Volpicelli, L. Gargani, B. Bouhemad, and F. Mojoli. Quantitative lung ultrasound: time for a consensus? *Chest*, 158(2):469–470, 2020.

[29] G. Soldati, A. Smargiassi, R. Inchingolo, D. Buonsenso, T. Perrone, D. F. Briganti, S. Perlini, E. Torri, A. Mariani, E. E. Mossolani, et al. Time for

a new international evidence-based recommendations for point-of-care lung ultrasound. *Journal of ultrasound in medicine: official journal of the American Institute of Ultrasound in Medicine*, 40(2):433–434, 2021.

[30] F. Mento, T. Perrone, V. N. Macioce, F. Tursi, D. Buonsenso, E. Torri, A. Smargiassi, R. Inchingolo, G. Soldati, and L. Demi. On the impact of different lung ultrasound imaging protocols in the evaluation of patients affected by coronavirus disease 2019: how many acquisitions are needed? *Journal of Ultrasound in Medicine*, 40(10):2235–2238, 2021.

[31] G. Soldati, A. Smargiassi, R. Inchingolo, D. Buonsenso, T. Perrone, D. F. Briganti, S. Perlini, E. Torri, A. Mariani, E. E. Mossolani, et al. Proposal for international standardization of the use of lung ultrasound for patients with covid-19: a simple, quantitative, reproducible method. *Journal of Ultrasound in Medicine*, 39(7):1413–1419, 2020.

[32] L. Demi. Lung ultrasound: The future ahead and the lessons learned from covid-19. *The Journal of the Acoustical Society of America*, 148(4):2146–2150, 2020.

[33] A. Khamis, J. Meng, J. Wang, A. T. Azar, E. Prestes, Á. Takács, I. J. Rudas, and T. Haidegger. Robotics and intelligent systems against a pandemic. *Acta Polytechnica Hungarica*, 18(5):13–35, 2021.

[34] N. Islam, S. Ebrahimzadeh, J.-P. Salameh, S. Kazi, N. Fabiano, L. Treanor, M. Absi, Z. Hallgrimson, M. M. Leeflang, L. Hooft, et al. Thoracic imaging tests for the diagnosis of covid-19. *Cochrane Database of Systematic Reviews*, (3), 2021.

[35] D. A. Lichtenstein and D. A. Lichtenstein. The luci-flr project: Lung ultrasound in the critically ill–a bedside alternative to irradiating techniques, radiographs and ct. *Lung Ultrasound in the Critically Ill: The BLUE Protocol*, pages 217–225, 2016.

[36] B. Yousuf, K. S. Sujatha, H. Alfoudri, and V. Mansurov. Transport of critically ill covid-19 patients. *Intensive Care Medicine*, 46:1663–1664, 2020.

[37] G. Orosz, P. Gyombolai, J. T. Tóth, and M. Szabó. Reliability and clinical correlations of semi-quantitative lung ultrasound on blue points in covid-19 mechanically ventilated patients: The 'blue-luss'—a feasibility clinical study. *PLoS One*, 17(10):e0276213, 2022.

[38] S. K. Zhou, D. Rueckert, and G. Fichtinger. *Handbook of medical image computing and computer assisted intervention*. Academic Press, 2019.

[39] S. R. Chandaran, G. Muthusamy, L. R. Sevalaiappan, and N. Senthilkumaran. Deep learning-based transfer learning model in diagnosis of diseases with brain magnetic resonance imaging. *Acta Polytechnica Hungarica*, 19(5):127–147, 2022.

[40] D. Cheng and E. Lam. Transfer learning u-net deep learning for lung ultrasound segmentation. 10 2021.

[41] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. 2014.

[42] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[43] F. Isensee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S. Kohl, J. Wasserthal, G. Köhler, T. Norajitra, S. J. Wirkert, and K. H. Maier-Hein. nnu-net: Self-adapting framework for u-net-based medical image segmentation. *CoRR*, abs/1809.10486, 2018.

[44] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert. Attention u-net: Learning where to look for the pancreas, 2018.

[45] M. A. Houghtaling, S. R. Fiorini, N. Fabiano, P. J. Gonçalves, O. Ulgen, T. Haidegger, J. L. Carbonera, J. I. Olszewska, B. Page, Z. Murahwi, et al. Standardizing an ontology for ethically aligned robotic and autonomous systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023.

[46] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. M. Fennessy, M. Sonka, J. M. Buatti, S. R. Aylward, J. V. Miller, S. Pieper, and R. Kikinis. 3d slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9):1323–1341, 2012.

[47] T. Haidegger. Autonomy for Surgical Robots: Concepts and Paradigms. *IEEE Transactions on Medical Robotics and Bionics*, 1(2):65–76, 2019.