



OPEN

DATA DESCRIPTOR


Lung Ultrasound Imaging Dataset for Accurate Detection and Localization of LUS Vertical Artifact

Nixson Okila¹, Andrew Katumba^{2,3} , Joyce Nakatumba-Nabende^{1,3}, Sudi Murindanyi^{1,3} , Cosmas Mwikirize², Jonathan Serugunda², Samuel Bugeza⁴, Anthony Oriekot⁵, Juliet Bossa⁶ & Eva Nabawanuka⁶

Lung ultrasound (LUS) vertical artifacts are critical sonographic markers commonly used in evaluating pulmonary conditions such as pulmonary edema, interstitial lung disease, pneumonia, and COVID-19. Accurate detection and localization of these artifacts are vital for informed clinical decision-making. However, interpreting LUS images remains highly operator-dependent, leading to variability in diagnosis. While deep learning (DL) models offer promising potential to automate LUS interpretation, their development is limited by the scarcity of annotated datasets specifically focused on vertical artifacts. This study introduces a curated dataset of 401 high-resolution LUS images, each annotated with polygonal bounding boxes to indicate vertical artifact locations. The images were collected from 152 patients with pulmonary conditions at Mulago and Kiruddu National Referral Hospitals in Uganda. This dataset serves as a valuable resource for training and evaluating DL models designed to accurately detect and localize LUS vertical artifacts, contributing to the advancement of AI-driven diagnostic tools for early detection and monitoring of respiratory diseases.

Background & Summary

Lung ultrasound (LUS) has gained prominence as a valuable imaging modality in pulmonary and critical care medicine, owing to its portability, real-time imaging capabilities, and non-ionizing nature¹. Among the sonographic features identifiable in LUS, vertical artifacts, arising from resonance phenomena when ultrasound waves interact with fluid-filled alveoli or thickened interlobular septa in areas of increased pulmonary density, are particularly significant for evaluating pulmonary congestion^{2,3}. These artifacts serve a crucial role in the assessment of a broad spectrum of respiratory and systemic conditions, including pulmonary edema, decompensated heart failure, chronic kidney disease, pneumonia, interstitial lung disease, and COVID-19⁴. Despite the diagnostic utility of LUS, the modality inherently lacks anatomical imaging of the lung parenchyma¹. Consequently, clinical interpretation relies on the accurate recognition and analysis of artifact patterns, a process that is susceptible to considerable inter- and intra-observer variability, as demonstrated in recent studies^{5–7}. In a study involving LUS of COVID-19 patients, Lerchbaumer *et al.*⁵ evaluated agreement among ten observers who each rated 100 cine loops four times. The results indicated fair to moderate inter-observer agreement (Fleiss' $\kappa \approx 0.27$ for intermediate scores and ~ 0.59 for subpleural consolidations) and moderate to substantial intra-observer agreement (κ up to ~ 0.71 – 0.79 for distinct features). Similarly, Fatima *et al.*⁶ evaluated interrater agreement among six LUS experts for both video-level and prognostic-level scoring. At the video level, a kappa value of 0.404 was reported for overall scores, while at the prognostic level, the kappa value was 0.505, both reflecting moderate agreement among the experts. The findings underscore the inherent subjectivity associated with expert interpretation of LUS. In a related study⁷, interrater agreement was assessed between three LUS experts and AI-based solution at the video level. The mean kappa value for overall scores among the experts was 0.61, also indicating moderate agreement and further highlighting the subjective nature of human analysis.

¹Department of Computer Science, Makerere University, Kampala, Uganda. ²Department of Electrical and Computer Engineering, Makerere University, Kampala, Uganda. ³Emergent AI, Kampala, Uganda. ⁴Department of Radiology, Makerere University Hospital, Makerere University, Kampala, Uganda. ⁵Mulago Specialized Women and Neonatal Hospital, Kampala, Uganda. ⁶Mulago National Referral Hospital, Kampala, Uganda. e-mail: andrew.katumba@mak.ac.ug

To mitigate the challenges of LUS interpretation, recent studies^{8–12} have focused on leveraging deep learning (DL) techniques to automate LUS image analysis, with a particular focus on the detection and localization of LUS vertical artifacts. While these methods show considerable promise, their effectiveness remains constrained by the scarcity of annotated datasets designed explicitly for this task. In particular, the lack of datasets with detailed polygonal bounding box (PBB) annotations has hindered the development of models capable of accurately capturing the morphology and spatial distribution of LUS vertical artifacts. This limitation often results in models that exhibit suboptimal performance, such as imprecise artifact delineation and inclusion of extraneous background regions, thereby reducing diagnostic accuracy and increasing computational burden.

To address this gap, we introduce the LUS-BALD (Lung Ultrasound – LUS Vertical Artifact Localization and Detection) dataset, the first publicly available resource containing LUS images with corresponding PBB annotations for LUS vertical artifacts. The dataset comprises 401 images, partitioned at the patient level into training, validation, and test subsets, with annotations provided for the training and validation sets. The LUS-BALD dataset is intended to facilitate the development and benchmarking of deep learning algorithms for precise LUS vertical detection and localization, and it holds potential for practical applications such as AI-assisted educational platforms for training healthcare providers in LUS interpretation.

Methods

The methodological workflow for the systematic collection and curation of the LUS-BALD dataset is illustrated in Fig. 1.

Participants and ethical requirements. A total of 152 participants (113 males and 39 females; mean age: 40.91 ± 17.03 years) were recruited for a clinical study conducted at Mulago and Kiruddu National Referral Hospitals in Kampala, Uganda. The cohort comprised individuals with suspected COVID-19 infection, as well as patients presenting with other pulmonary pathologies, including cases of non-COVID-19 pneumonia. Ethical approval for the study was obtained from the Mulago Hospital Research and Ethics Committee (MHREC; Approval No. MHREC 2021-45, dated 24 August 2021) and the Uganda National Council for Science and Technology (UNCST; Approval No. SIR74ES, dated 22 October 2021).

Informed written consent was obtained from all participants prior to enrolment. The consent process included a comprehensive explanation of the study objectives, data collection procedures, and participant rights. Participants were informed of their right to withdraw from the study at any time without consequence. Consent also included permission for the use of anonymized data for research purposes, with explicit assurance that personal identifiers would remain confidential.

Data collection. The LUS imaging was conducted by three senior radiologists, each with over ten years of clinical experience in LUS. Two handheld curvilinear ultrasound probes were used: the Clarius C3 (Clarius Mobile Health Corp., Burnaby, BC, Canada) and the Philips Lumify C5-2 (Philips Ultrasound, Bothell, WA, USA). Acquisition settings were harmonized across devices: Clarius C3 operated at 2–6 MHz with a mechanical index (MI) ≤ 0.68 , and Philips Lumify C5-2 at 2–5 MHz with MI approximately 0.7. The focal depth was positioned at the pleural line (~3 cm), with a frame rate of 20 Hz, gain set to 53, and power output at –0.3 dB. Although a depth of 16 cm was uniformly applied in our acquisitions to capture the complete ultrasound field, the key LUS features (pleural line and vertical artifacts) were identifiable within the pleural line depth of 2–4 cm. This 2–4 cm depth range is consistent with previous reported findings^{13–15}.

During the LUS scanning, radiologists divided each patient's thorax into 12 regions—six on the left (upper posterior, lower posterior, upper anterior, lower anterior, upper lateral, lower lateral) and six on the right (upper posterior, lower posterior, upper anterior, lower anterior, upper lateral, lower lateral). Both longitudinal and transverse scans were performed in each region, generating a total of 24 image frames per patient, each stored in a unique folder labelled with the corresponding patient ID and uploaded to the data server.

Segmenting the thorax into these 12 distinct zones enabled radiologists to systematically evaluate different lung areas and detect localized abnormalities that could vary across regions¹⁶. Specifically, dividing the thorax into upper and lower, anterior, posterior, and lateral sections on both sides allowed for a more precise assessment of the distribution and severity of lung involvement¹⁷. Additionally, scanning both longitudinally and transversely within each region provided multiple perspectives, enhancing the detection of hidden pathologies that might otherwise be overlooked with a single scan orientation¹⁶.

Data filtering and splitting. We obtained 458 images with LUS vertical artifacts from the data server and, in consultation with radiologists, excluded 57 low-resolution images and those with unclear artifacts. To prepare the dataset for model training, we partitioned the images into train, val, and test splits at the patient level, ensuring that images from the same patient appeared in only one partition.

Data pre-processing. The LUS images included overlaid measurement scales and text labels. To eliminate these artifacts, a two-step pre-processing approach was employed: (i) image cropping to remove regions containing measurement scales, and (ii) inpainting techniques to remove overlaid text labels.

The image cropping procedure commenced by loading the input image using OpenCV and converting it to grayscale to reduce computational complexity while preserving essential structural information. Subsequently, Otsu's thresholding method was applied to binarize the image by automatically selecting an optimal threshold to differentiate foreground structures from the background. To enhance segmentation quality, morphological operations were performed. Specifically, morphological closing with a small rectangular kernel was employed to fill small gaps within detected structures, ensuring their continuity. This was followed by dilation to strengthen object boundaries and facilitate more accurate contour detection. Upon completion of pre-processing, contours were identified using OpenCV's contour detection algorithm. Each contour was evaluated based on its area, with

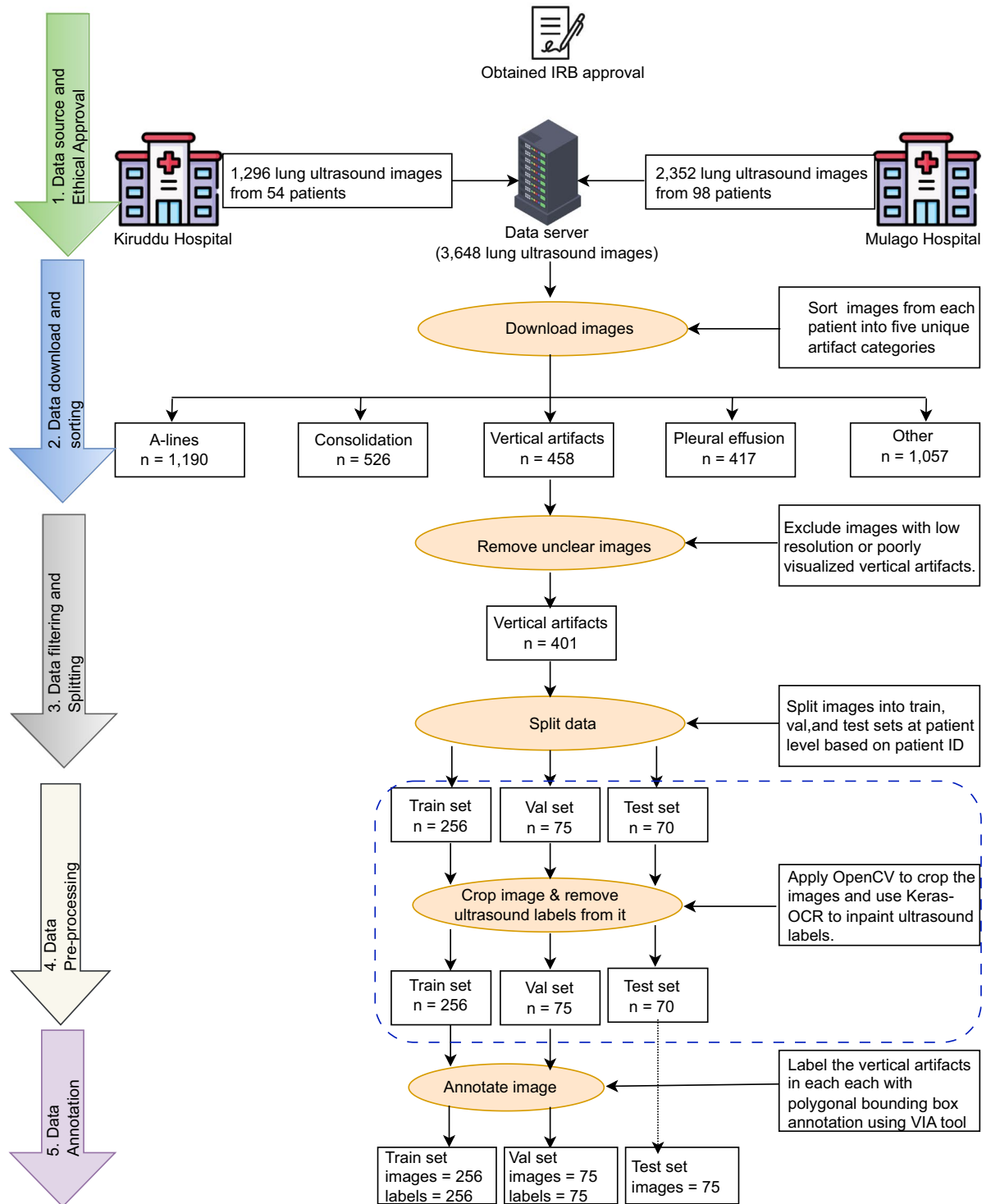


Fig. 1 Workflow for LUS-BALD dataset collection and curation. The process begins with ethical approval and data acquisition, followed by sequential stages of data filtering, image preprocessing, and detailed annotation.

the largest contour selected as the primary region of interest. A bounding box was then constructed around this region to delineate its spatial extent. The final step involved cropping the image using the computed bounding box, thereby isolating the principal structure from the background, as illustrated in Fig. 2.

To eliminate residual text labels from the ultrasound images, the Keras-OCR library was employed for text detection and recognition. The library identifies textual elements by extracting bounding box coordinates around detected text instances, each represented as a quadrilateral defined by four corner points. Following text detection, a binary mask was generated to localize the text regions. For each quadrilateral, two midpoints were calculated along its vertical edges, serving as anchor points for drawing a thick line across the text region. The

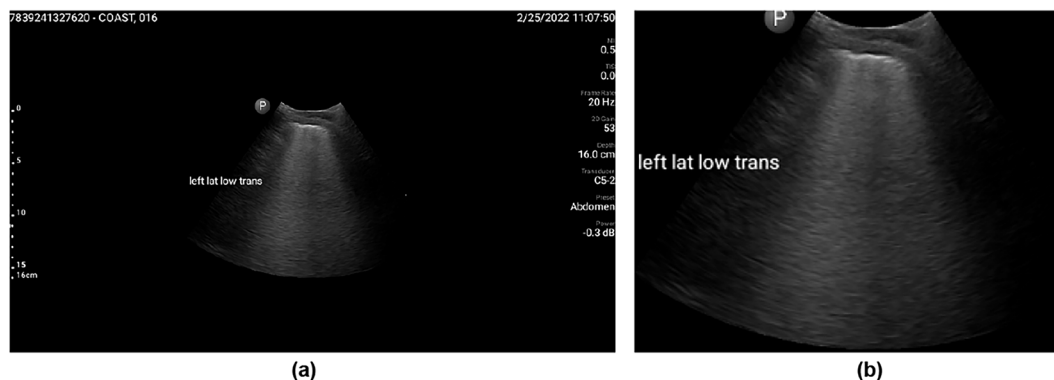


Fig. 2 Preprocessing step for removal of overlaid measurement scales and text labels from lung ultrasound images. **(a)** Input image with overlaid measurement scale and text labels. **(b)** Cropped image containing only the text labels, with the measurement scale removed.

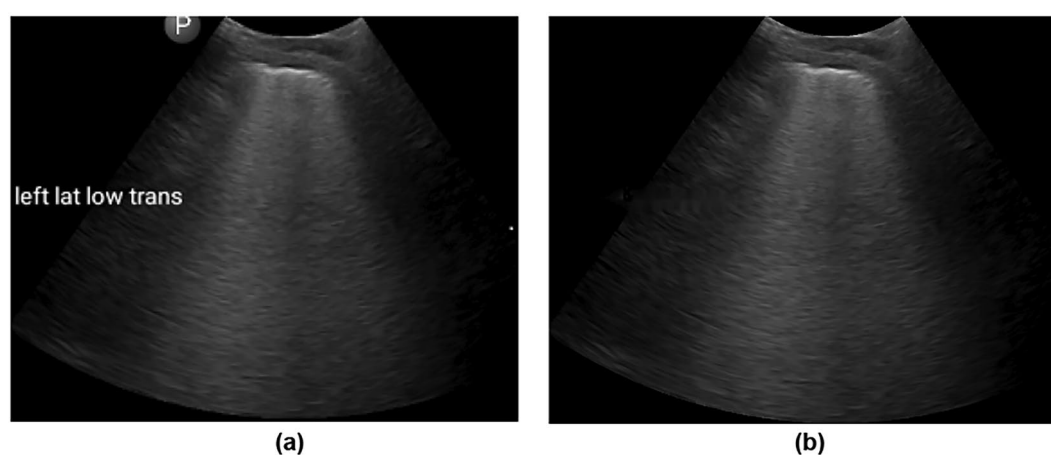


Fig. 3 Text label removal from lung ultrasound images. **(a)** Cropped image with overlaid text labels. **(b)** Processed image after text detection, masking, and inpainting, resulting in an image free of text labels.

line thickness was adaptively determined based on the Euclidean distance between adjacent text coordinates to ensure complete coverage of the text. The resulting mask was then processed using OpenCV's inpainting algorithm, which reconstructed the masked regions by interpolating pixel information from the surrounding areas. This procedure effectively removed text overlays while preserving the structural integrity of the underlying anatomical features. The processed image, free from textual annotations, is shown in Fig. 3(b).

Data annotation. The annotation of LUS vertical artifacts was performed by manually delineating bounding boxes around the regions of interest in the images. An open-source tool, the VGG Image Annotator (VIA)¹⁸, was utilized for this task. The VIA tool supports multiple annotation formats, including polygonal bounding boxes, which were selected for their flexibility in accurately capturing the irregular shapes of LUS vertical artifacts. The annotation process involved launching the VIA interface, uploading the collected images, and collaboratively annotating each image in the training and validation sets in consultation with experienced radiologists. Polygonal bounding boxes were drawn around each identified LUS vertical artifact, and the corresponding annotations were systematically saved. Figure 4 presents an example of the annotation process in VIA, illustrating two instances of LUS vertical artifacts annotated with polygonal bounding boxes.

Data Records

The dataset is available at Mendeley Data¹⁹. The root directory of the repository, named LUS-BALD, is organized into three principal subdirectories: train, val, and test, following a conventional partitioning strategy employed in supervised learning paradigms as shown in Fig. 5.

The train directory comprises two subfolders:

- images: Contains 256 lung ultrasound (LUS) image files in .png and .jpg formats.
- labels: Consists of 256 annotation files in YOLO (You Only Look Once) format, adapted to include four-vertex polygonal bounding boxes for the precise localization of LUS vertical artifacts. Each annotation file is paired with a corresponding image file.

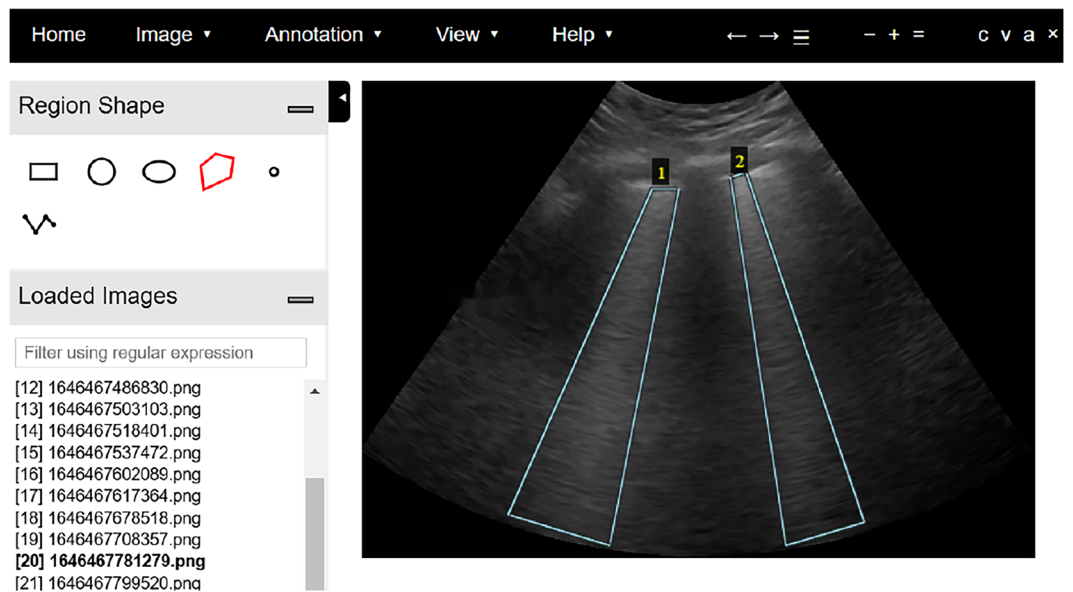


Fig. 4 Annotation using the VGG Image Annotator. The figure illustrates two instances of annotated lung ultrasound vertical artifacts.

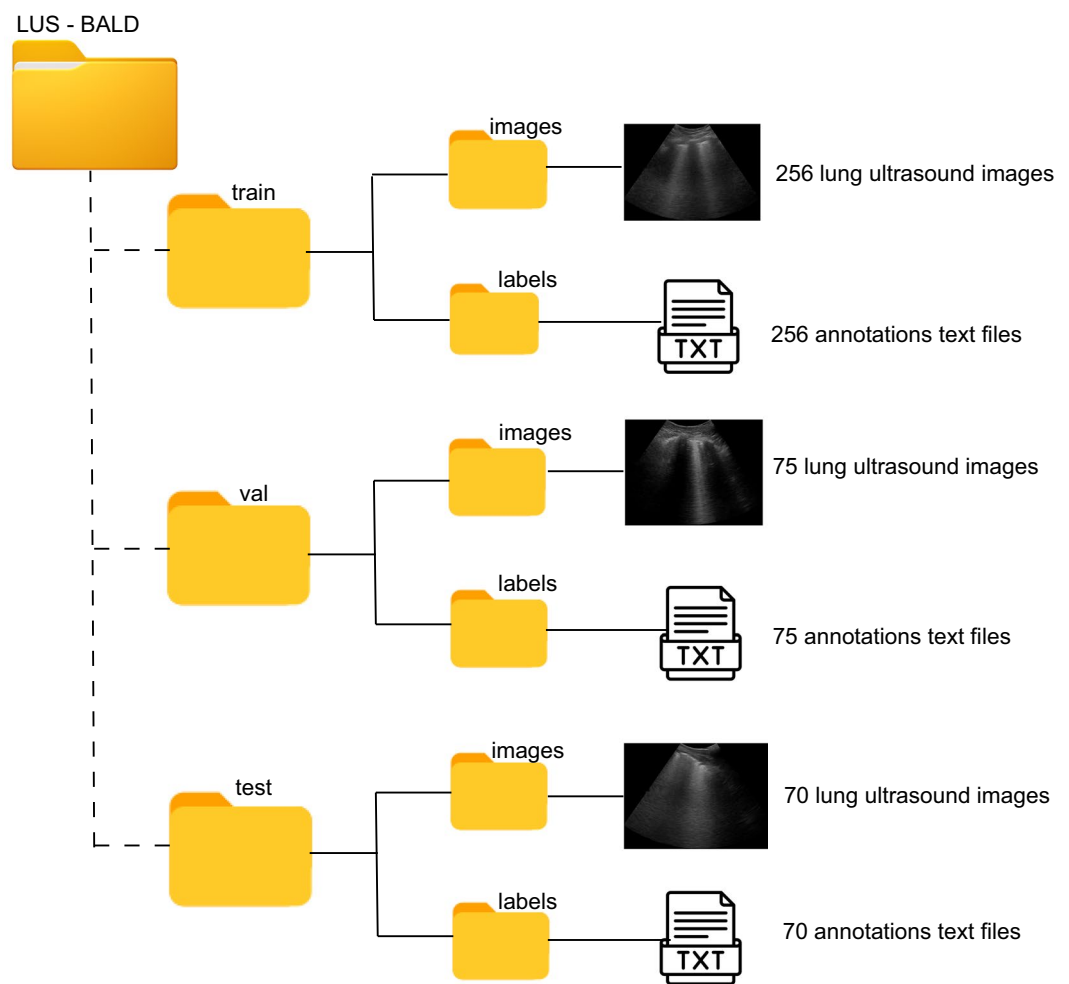


Fig. 5 Directory structure of the LUS-BALD dataset. The figure presents the hierarchical organization of files and folders within the LUS-BALD dataset.

The val directory mirrors the structure of the training set, containing 75 LUS image files and their corresponding annotation files, organized under images and labels subfolders, respectively.

The test directory contains a single images subfolder, comprising 70 LUS image files in .png and .jpg formats. Annotation files are not included for this subset, reflecting its intended use for model evaluation in a blind testing scenario.

Technical Validation

To ensure the reliability and utility of the dataset, multiple validation steps were incorporated throughout data acquisition, annotation, and quality control. Senior radiologists acquired lung ultrasound images following a standardized 12-zone scanning protocol. Imaging was performed using a convex probe in both longitudinal and transverse orientations, with each image systematically recorded and labelled according to the corresponding lung zone.

Following the acquisition, the images were reviewed and curated in collaboration with radiologists to ensure quality and clinical relevance. Images exhibiting motion artifacts, poor pleural line visibility, or incomplete anatomical views were excluded. Ultimately, a curated subset of 401 high-resolution LUS images containing LUS vertical artifact artifacts was selected based on imaging clarity and diagnostic significance.

The annotation of LUS vertical artifacts was conducted by the authors in collaboration with radiologists, utilizing a custom labelling interface to apply polygonal bounding boxes that precisely delineated the shape and orientation of each artifact. To ensure annotation consistency and reliability, a two-stage validation process was employed. A second expert radiologist independently reviewed the initial annotations, and any discrepancies were resolved through consensus adjudication. This rigorous validation approach minimized inter-observer variability and ensured accurate and reproducible representation of LUS vertical artifact features within the dataset.

A key limitation of the dataset lies in its reliance on high-resolution LUS images curated primarily for artifact clarity, which constitutes an idealized representation of lung ultrasound data. In clinical settings, the identification of vertical artifacts generally requires evaluation of full cine loops that often contain motion artifacts and considerable variability in image quality. Moreover, only a single representative frame was acquired per scanned region, rather than full video sequences. While this approach was practical for dataset curation, it fails to capture the temporal dynamics inherent to LUS and may introduce operator-dependent bias in frame selection. These constraints may limit the dataset's capacity to fully represent the variability and complexity of real-world LUS interpretation. Future dataset development should therefore emphasize acquisition of complete video sequences to mitigate selection bias and preserve the dynamic characteristics of LUS.

Code availability

The code for preprocessing LUS-BALD images is available at <https://github.com/Marconi-Lab/lus-blinc-artifact-prep/>.

Received: 4 May 2025; Accepted: 19 August 2025;

Published online: 16 October 2025

References

1. Demi, L. *et al.* New international guidelines and consensus on the use of lung ultrasound. *J. Ultrasound Med.* **42**, 309–344 (2023).
2. Corradi, F., Via, G., Forfori, F., Brusasco, C. & Tavazzi, G. Lung ultrasound and LUS vertical artifacts quantification inaccuracy: Be sure to have the right solution. *Intensive Care Med.* **46**, 1081–1083 (2020).
3. Mento, F., Soldati, G., Prediletto, R., Demi, M. & Demi, L. Quantitative lung ultrasound spectroscopy applied to the diagnosis of pulmonary fibrosis: The first clinical study. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**, 2265–2273 (2020).
4. Baloescu, C. *et al.* Automated lung ultrasound B-line assessment using a deep learning algorithm. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**, 2312–2320 (2020).
5. Lerchbaumer, M. H. *et al.* Lung ultrasound for early diagnosis and severity assessment of acute respiratory distress syndrome in the critically ill. *Sci. Rep.* **11**, 10678 (2021).
6. Fatima, N. *et al.* Human-to-AI Interrater Agreement for Lung Ultrasound Scoring in COVID-19 Patients. *J. Ultrasound Med.* **42**, <https://doi.org/10.1002/jum.16052> (2022).
7. Fatima, N. *et al.* Deep learning approaches for automated classification of neonatal lung ultrasound with assessment of human-to-AI interrater agreement. *Comput. Biol. Med.* **183**, 109315 (2024).
8. Van Sloun, R. J. G. & Demi, L. Localizing B-lines in lung ultrasonography by weakly supervised deep learning, *in-vivo* results. *IEEE J. Biomed. Health Inform.* **24**, 957–964 (2019).
9. Kerdegari, H. *et al.* Automatic detection of B-lines in lung ultrasound videos from severe dengue patients. In *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, 989–993 (2021).
10. Kulhare, S. *et al.* Ultrasound-based detection of lung abnormalities using single shot detection convolutional neural networks. In *Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation: Proc. POCUS 2018, BIVPCS 2018, CuRIOUS 2018, and CPM 2018, Held with MICCAI 2018, Granada, Spain*, 65–73 (Springer, 2018).
11. Roy, S. *et al.* Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. *IEEE Trans. Med. Imaging* **39**, 2676–2687 (2020).
12. Howell, L., Ingram, N., Lapham, R., Morrell, A. & McLaughlan, J. R. Deep learning for real-time multi-class segmentation of artefacts in lung ultrasound. *Ultrasonics* **140**, 107251 (2024).
13. Mento, F. & Demi, L. Dependence of lung ultrasound vertical artifacts on frequency, bandwidth, focus and angle of incidence: An *in vitro* study. *J. Acoust. Soc. Am.* **150**, 4075–4082 (2021).
14. Demi, L., Van Hove, W., Van Sloun, R. J. G., Soldati, G. & Demi, M. Determination of a potential quantitative measure of the state of the lung using lung ultrasound spectroscopy. *Sci. Rep.* **7**, 12746 (2017).
15. Mento, F. *et al.* State of the art in lung ultrasound, shifting from qualitative to quantitative analyses. *Ultrasound Med. Biol.* **48**, 2398–2416 (2022).
16. Duggan, N. M., Shokoohi, H., Liteplo, A. S., Huang, C. & Goldsmith, A. J. Best practice recommendations for point-of-care lung ultrasound in patients with suspected COVID-19. *J. Emerg. Med.* **59**, 515–520 (2020).
17. Marini, T. J. *et al.* Lung ultrasound: the essentials. *Radiol. Cardiothorac. Imaging* **3**, e200564 (2021).
18. Dutta, A. & Zisserman, A. The VIA annotation software for images, audio and video. In *Proc. 27th ACM Int. Conf. on Multimedia*, 2276–2279 (2019).
19. Katumba, A. *et al.* Lung Ultrasound Imaging Dataset for Accurate Detection and Localization of LUS Vertical Artifacts, *Mendeley Data V2*, <https://doi.org/10.17632/h923m366sf.2> (2025).

Acknowledgements

This research was supported by funding from the Government of Uganda through the Makerere University Research and Innovation Fund (RIF), the COAST project, and the Swedish International Development Cooperation Agency. Additional financial support was provided by UK International Development and the International Development Research Centre (IDRC), Ottawa, Canada, under the AI for Development: Responsible AI, Empowering People Program (AI4D). The authors gratefully acknowledge the support provided by the Departments of Radiology at Mulago National Referral Hospital and Kiruddu Referral Hospital, which facilitated this study.

Author contributions

This manuscript represents a collaborative effort among all authors. N.O. and S.M. were responsible for conceptualization, conducting the data analysis, and drafting the initial manuscript. A.K., J.N.-N., C.M., and J.S. contributed to the conceptualization of the study, provided supervision, secured funding, oversaw project administration, and critically reviewed the manuscript. S.B., A.O., J.B., and E.N. were responsible for data collection.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025