COVIDx-US: An Open Benchmark Ultrasound Imaging Dataset for AI-Driven COVID-19 Diagnosis

Anonymous authors
Paper under double-blind review

Abstract

The COVID-19 pandemic has not only gravely impacted the health and well-being of the global population but it has caused economic and social disruption. Effective screening, prognosis, and treatment planning are critical to combat the COVID-19 pandemic. As a highly mobile and accessible tool with non-invasive nature, point-of-care ultrasound imaging can simplify the screening process of COVID-19 patients, reducing the pressure on healthcare systems and healthcare providers. Motivated by this and the potential of artificial intelligence-driven tools to aid clinicians, we introduce COVIDx-US, an open-access benchmark dataset of COVID-19 related ultrasound imaging data. The COVIDx-US dataset was curated from multiple sources and its current version consists of 80 lung ultrasound videos and 8,549 processed images of patients with COVID-19 infection, non-COVID-19 infection, as well as normal control cases. The COVIDx-US is the largest open-access fully-curated dataset of its kind that has been systematically curated, processed, and validated to build and evaluate artificial intelligence algorithms and models specifically.

1 Introduction

The novel Coronavirus Disease 2019 (COVID-19) outbreak resulted in a shortage of medical equipment (Roy et al., 2020) and healthcare professionals. Effective screening of infected patients is of immense importance in controlling the COVID-19 spread (Wang et al., 2020a). The reverse transcription-polymerase chain reaction (RT-PCR) test, performed on biological samples taken from the patient, has been used as the main screening method for COVID-19 detection (Wang et al., 2020b). Despite the criticality of intensive testing and diagnostics to effectively control the pandemic, many countries could not implement it due to limited testing capacity (Roy et al., 2020), and/or lack of skillful healthcare professionals required to perform the test and interpret the results. Moreover, despite being the gold standard, RT-PCR testing is a time-consuming and complicated process with a highly variable sensitivity reported in the literature (West et al., 2020). For such reasons alternative/complementary solutions for COVID-19 screening that improve diagnostic precision while reducing user-dependence have attracted the attention of the scientific community.

Radiography examination is an alternative imaging method that has been utilized for COVID-19 screening and risk stratification where a radiologist visually inspects radiography images, e.g., chest X-ray (CXR) or computed tomography (CT) scans, to find indicators associated with SARS-CoV-2 viral infection as well as the severity of the infection. As a widely available and cost-effective imaging technique, ultrasound imaging is lately gaining attention such that point-of-care ultrasound (POCUS) is increasingly being used in (resource-limited) medical settings, e.g., in emergency rooms and intensive care units (ICUs), for detecting respiratory disorders (Mojoli et al., 2019; Raheja et al., 2019). Notably, ultrasound imaging has been reported to show a higher sensitivity in pneumonia detection compared with chest X-ray in some cases (Amatya et al., 2018). Moreover, recent findings suggest specific lung ultrasound (LUS) characteristics and imaging markers in patients infected with COVID-19 (Peng et al., 2020; Soldati et al., 2020).

Artificial intelligence (AI)-powered decision support systems, mostly based on deep neural network architectures, have shown exemplary performance in many computer vision problems in healthcare

Table 1: Ultrasound video data distribution per source and class in COVIDx-US v1.0.

DATA SOURCE	COVID-19	NON-COVID-19	NORMAL	COUNT
POCUS Atlas	18	9	5	32
GrepMed	8	9	3	20
Butterfly Network	20	0	2	22
LITFL	0	6	0	6

(Chilamkurthy et al., 2018; Brinker et al., 2019). By extracting complex hidden patterns in medical images, deep learning (DL) techniques may find relationships/patterns that are not instantly available to human analysis (Poplin et al., 2018). Compared to CXR and CT, lung ultrasound deep learning studies are comparably limited due to lack of well-established, organized, carefully labelled LUS data sets (Arntfield et al., 2020). Motivated by the recent open-source efforts of the research community in the fight against COVID-19 and to support alternative screening, risk stratification, and treatment planning solutions powered by artificial intelligence and advanced techniques, we introduce COVIDx-US, an open-access benchmark dataset of ultrasound imaging data that was carefully curated from multiple sources and integrated systematically, specifically for facilitating the building and evaluation of AI-driven analytics algorithms and models. The current version of the COVIDx-US dataset contains 80 videos and 8,549 processed ultrasound images of patients diagnosed with COVID-19, non-COVID-19 infection, as well as normal control patients. The COVIDx-US dataset was released as part of a large open-source initiative and will be continuously growing as more data sources become available. To the best of the authors' knowledge, COVIDx-US[1] is the first and largest open-access fully-curated benchmark lung ultrasound imaging dataset that is easy-to-use, reproducible, and easy-to-scale thanks to its modular well-documented design. In the rest of this paper, we introduce the COVIDx-US in detail.

2 COVIDx-US Design

The current version of the dataset, i.e., COVIDx-US v1.0., collects ultrasound video data from four data sources, i.e., Butterfly Network, the POCUS Atlas, GrepMed, and LITFL, and curates and integrates them under three categories, i.e., COVID-19 infection, non-COVID-19 infection (e.g., bacterial infection, non-SARS-CoV-2 viral infection, etc.), and normal. Table 1 shows the distribution of US video files per data source. The data is heterogeneous in nature and contains data of various characteristics. The COVIDx-US dataset contains US video recordings captured with either linear or convex US probes ($N = 15$ and 65, respectively). This variety not only enables higher generalizability of the models that are trained on the COVIDx-US but also provides users with higher flexibility to filter out video files based on the probe types, if required. Although the pandemic is recent, 57% of the US video data are from COVID-19 positive cases. Figure 1 shows the flow of processes and the steps taken to generate the COVIDx-US dataset. In the rest of this section, we explain the steps in more detail.

2.1 Data Curation

The data were curated from four different data sources, each with a different structure. We developed data curation engines (coded in Python), personalized for each of the data sources, to automatically curate lung POCUS video recordings and their metadata from each data source. The data are then integrated locally in a unified and organized structure. No original data is hosted in the COVIDx-US repository. The scripts to curate and integrate data are designed to be highly reproducible, easy to use, and extensible such that more data sources can be added to the pipeline in the future upon availability.

[1]The COVIDx-US dataset and all the scripts are available for the public at: *http://anonymousforreview* to enable researchers, clinical scientists, and etc. to drive innovation in this area.

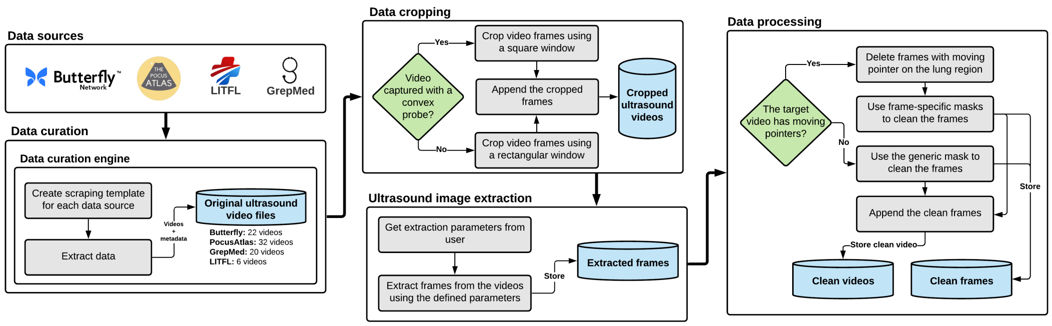


Figure 1: The conceptual flow of COVIDx-US dataset integration. The current version of dataset contains 80 ultrasound videos and 8,549 processed ultrasound images from the following four data sources: 1) Butterfly Network, 2) GrepMed, 2) The POCUS Atlas, and 4) LITFL. Original ultrasound videos are extracted from these data sources and are curated and integrated systematically in a unified and organized structure.

2.2 Data Cropping

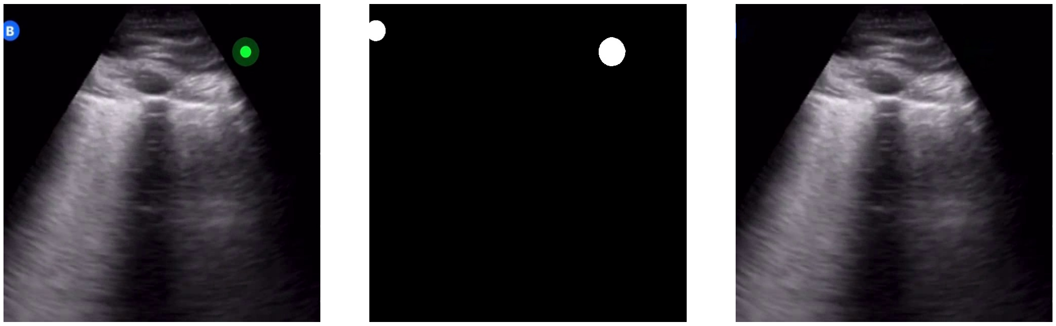
The original data contains different types of artifacts, e.g., measure bars and symbols, mostly on the sides of the videos. As an initial processing step, we cropped the collected video recordings to remove these peripheral artifacts. We defined and followed a separate cropping process for convex and linear US video files due to their different characteristics and with an aim to include a larger portion of the original file in the cropped video file. We used square and rectangular windows to crop the frames of the convex and linear US video files, respectively. The processing scripts are publicly available as part of the COVIDx-US release. Users can set different values for the parameters in the provided scripts, however, using the default parameters for the defined windows will remove artifacts such as bars and texts visible on the side or top of the collected US video files. The generated cropped files are stored locally.

2.3 Ultrasound Image Extraction

Our developed scripts are highly flexible and enable users to extract frames from the processed US video files based on their research goals and requirements, using a set of parameters such as: 1) the maximum number of frames to extract from each video, 2) target class(es) of the video files to extract frames from, 3) the data source(s), and 4) probe type(s). Using the default parameters all frames are extracted from all the video recordings. The extracted frames from the videos are stored locally.

2.4 Data Processing

The final processing step is performed on the extracted frames from the video files as follows: 1) videos with moving pointers are identified, 2) if the video contains a moving pointer, frames with a moving pointer on the lung region are deleted and frame-specific masks are generated and stored for the remaining frames, and 3) if the video does not contain a moving pointer, a generic mask that is suitable for all the extracted frames is generated and stored, and is used to remove artifacts from the frames. All the generated masks are provided as part of the COVIDx-US release. We used the inpainting technique introduced by Bertalmio et al. (2001) to remove the peripheral artifacts from the frames by replacing bad marks, i.e., pixels in the masked regions, with their neighboring pixels. Clean video files are then generated from the clean frames and both are stored locally on the user's device. Figure 2 shows a sample cropped ultrasound frame, the mask generated for this specific frame, and the final clean frame obtained by applying the mask to the original frame.



(a) A sample cropped frame (b) The respective generated mask (c) The resulted clean frame

Figure 2: Sample images in COVIDx-US dataset.

3 Data Quality Validation

As a crucial step to ensure the quality of images in the COVIDx-US dataset, our contributing clinician reviewed a randomly selected set of images. His findings and observations confirmed the existence of identifiers and indicators of disease in the COVIDx-US dataset. Such patterns and indicators can be exploited by AI-based analytics solutions for COVID-19 detection that will be built on the COVIDx-US dataset. Based on our contributing clinician's evolving experience, LUS has significant utility in the management of COVID-19 patients with respiratory symptoms.

4 Conclusion

Point-of-care ultrasound (LUS) is showing considerable promise as an alternative imaging solution to CXR as a first-line screening approach (Gazon et al., 2010). It may also have utility in predicting disease course and prognosis (Bonadia et al., 2020). This is due to its high portability, non-ionizing radiation nature, and the fact that it is being used as the preferred lung infection diagnosic method in resource-limited settings/environments, e.g., in emergency rooms or developing countries (Amatya et al., 2018). There are some limitations to LUS, however, as it requires significant operator training and experience before it can be used in the management of potentially unstable patients, or in those with suspected infectious syndromes. However, these limitations are mitigated by the potential of AI-driven solutions to aid clinicians with the screening process of COVID-19 patients, reducing the pressure on healthcare systems and healthcare providers. The COVIDx-US, to the best of our knowledge, is the largest open-access fully-curated benchmark LUS dataset of its kind that is systematically curated, highly reproducible, easy to use, and highly scalable. All the scripts are well-documented and modularly designed to ensure readability and scalability. The COVIDx-US is curated from multiple data sources and contains data of different types and characteristics. The scripts provided will perform the processes necessary to clean the collected POCUS videos, extract frames, and store them locally. All the scripts, metadata, and generated masks, necessary to reproduce the COVIDx-US data set, as described and explained in this manuscript, are available to the general public at *http://anonymousforreview*. The COVIDx-US dataset will be continuously growing as more data become available. We recommend that users check the COVIDx-US repository frequently, for the latest version of data and scripts.

References

Yogendra Amatya, Jordan Rupp, Frances M. Russell, Jason Saunders, Brian Bales, and Darlene R. House. Diagnostic use of lung ultrasound compared to chest radiograph for suspected pneumonia in a resource-limited setting. *International journal of emergency medicine*, 11:1–5, 2018.

Robert Arntfield, Blake VanBerlo, Thamer Alaifan, Nathan Phelps, Matthew White, Rushil Chaudhary, Jordan Ho, and Derek Wu. Development of a deep learning classifier to accurately distinguish covid-19 from look-a-like pathology on lung ultrasound. *medRxiv*, 2020.

Marcelo Bertalmio, Andrea L. Bertozzi, and Guillermo Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. *In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1*, 2001.

Nicola Bonadia, Annamaria Carnicelli, Alfonso Piano, Danilo Buonsenso, Emanuele Gilardi, Cristina Kadhim, Enrico Torelli, and et al. Lung ultrasound findings are associated with mortality and need for intensive care admission in covid-19 patients evaluated in the emergency department. *Ultrasound in Medicine & Biology 46, no. 11*, pp. 2927–2937, 2020.

Titus J. Brinker, Achim Hekler, Alexander H. Enk, Joachim Klode, Axel Hauschild, Carola Berking, Bastian Schilling, and et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, 113:47–54, 2019.

Sasank Chilamkurthy, Rohit Ghosh, Swetha Tanamala, Mustafa Biviji, Norbert G. Campeau, Vasantha Kumar Venugopal, Vidur Mahajan, Pooja Rao, and Prashant Warier. Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. *The Lancet 392, no. 10162*, pp. 2388–2396, 2018.

M. Gazon, N. Eboumbou, M.O. Robert, P. Branche, S. Duperret, and J.P. Viale. Agreement between lung ultrasonography and chest radiography in the intensive care unit. *In Annales francaises d'anesthesie et de reanimation (Vol. 30, No. 1, pp. 6-12)*, 2010.

Francesco Mojoli, Bélaid Bouhemad, Silvia Mongodi, and Daniel Lichtenstein. Lung ultrasound for critically ill patients. *American journal of respiratory and critical care medicine*, 199:701–714, 2019.

Qian-Yi Peng, Xiao-Ting Wang, Li-Na Zhang, and Chinese Critical Care Ultrasound Study Group. Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic. *Intensive care medicine*, 2020.

Ryan Poplin, Avinash V. Varadarajan, Katy Blumer, Yun Liu, Michael V. McConnell, Greg S. Corrado, Lily Peng, and Dale R. Webster. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering 2, no. 3*, pp. 158–164, 2018.

Ronak Raheja, Megha Brahmavar, Dhruv Joshi, and Dileep Raman. Application of lung ultrasound in critical care setting: a review. *Cureus*, 11:701–714, 2019.

Subhankar Roy, Willi Menapace, Sebastiaan Oei, Ben Luijten, Enrico Fini, Cristiano Saltori, Iris Huijben, and et al. Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound. *IEEE Transactions on Medical Imaging*, 8:2676–2687, 2020.

Gino Soldati, Andrea Smargiassi, Riccardo Inchingolo, Danilo Buonsenso, Tiziano Perrone, Domenica Federica Briganti, Stefano Perlini, and et al. Is there a role for lung ultrasound during the covid-19 pandemic? *Journal of Ultrasound in Medicine*, 2020.

Linda Wang, Zhong Qiu Lin, and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports 10*, 1: 1–12, 2020a.

Wenling Wang, Yanli Xu, Ruqin Gao, Roujian Lu, Kai Han, Guizhen Wu, and Wenjie Tan. Detection of sars-cov-2 in different types of clinical specimens. *Jama 323*, 18:1843–1844, 2020b.

Colin P. West, Victor M. Montori, and Priya Sampathkumar. Covid-19 testing: the threat of false-negative results. *In Mayo Clinic Proceedings*, 95:1127–1129, 2020.