COVID-19 feature detection with deep neural networks trained on simulated lung ultrasound B-mode images

Lingyi Zhao*, Tiffany Clair Fong†, and Muyinatu A. Lediju Bell*‡§
*Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD
†Department of Emergency Medicine, Johns Hopkins Medicine, Baltimore, MD
‡Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD
§Department of Computer Science, Johns Hopkins University, Baltimore, MD

***Abstract*—Deep learning has been implemented to detect COVID-19 features in lung ultrasound B-mode images. However, previous work primarily relied on *in vivo* images as the training data, which suffers from limited access to required manual labeling of thousands of training image examples. To avoid this manual labeling, which is tedious and time consuming, we propose the detection of *in vivo* COVID-19 features (i.e., A-line, B-line, consolidation) with deep neural networks (DNNs) trained on simulated B-mode images. The simulation-trained DNNs were tested on *in vivo* B-mode images from healthy subjects and COVID-19 patients. With data augmentation included during the training process, Dice similarity coefficients (DSCs) between ground truth and DNN predictions were maximized, producing mean ± standard deviatio values as high as 0.48 ± 0.29, 0.45 ± 0.25, and 0.46 ± 0.35 when segmenting *in vivo* A-line, B-line, and consolidation features, respectively. Results demonstrate that simulation-trained DNNs are a promising alternative to training with real patient data when segmenting *in vivo* COVID-19 features.**

***Index Terms*—deep learning, ultrasound imaging, COVID-19, segmentation**

I. Introduction

Multiple groups have introduced the potential of deep learning to aid COVID-19 diagnosis based on real-time lung ultrasound imaging [1]–[4]. Possible detectable features include A-lines [5], B-lines [6], and subpleural consolidations [7]. A-lines are commonly present in healthy lungs, appearing as horizontal reverberation artifacts of pleura caused by multiple reflections originating from the normal lung surface [5]. B-lines are characteristic of diseased lungs, appearing as laser-like vertical lines extending from the pleural line to the edge of the screen, obliterating A-lines in some cases [8]. Subpleural consolidations are similarly present in diseased lungs, appearing as echo-poor regions or tissue-like echo texture that extends to the pleural line and may have irregular margins [9].

While the first deep learning models implemented to detect COVID-19 features in lung ultrasound B-mode images primarily relied on *in vivo* labeled B-mode images as the training data [1], these datasets are difficult to obtain, and manually annotating *in vivo* data can be time consuming. In contrast to *in vivo* data, simulated data can be easily generated with known ground truths. Previous work demonstrated that a training set containing a mixture of simulated and *in vivo* B-mode images enabled deep neural networks (DNNs) to achieve better performance when segmenting *in vivo* bone surface features [10] and vessels [11]. In addition, when trained only on simulated raw ultrasound channel data, DNNs can detect cyst-like features in both phantom and *in vivo* B-mode images [12], [13].

Our group is the first to implement simulation-trained DNNs to identify *in vivo* B-line features in lung ultrasound images from COVID-19 patients [14]. The simulation-trained network found 39% more B-line features than a human observer, which is promising for training less experienced users and triaging the most problematic cases in an emergency setting. Despite this promise, the quantitative Dice similarity coefficient (DSC) scores measuring the overlap between network outputs and ground truths were low (i.e., $\leq$0.24, with a value of 1 indicating complete overlap). In addition, only B-line detection was previously investigated with our simulation-trained approach. We hypothesize that data augmentation will improve network performance when detecting B-lines. We also hypothesize that our simulation-trained approach is not limited to B-line detection.

This paper extends our simulation-trained approach to detect A-lines, B-lines, and subpleural consolidations after implementing data augmentation during the training process. We trained networks to identify these three features in simulated B-mode images. Network performance was then tested with B-mode lung ultrasound images from healthy individuals (to detect A-lines) and from COVID-19 patients (to detect B-lines and subpleural consolidations).

II. Methods

A total of 30,000 lung phantoms were simulated with MATLAB based on publicly available *in vivo* lung ultrasound B-mode images with A-line, B-line, and consolidation features (10,000 phantoms per feature). The positions and the echogeneity of the features were changed to increase the data variability. Next, we simulated raw channel data with the MATLAB Ultrasound Toolbox [15] using these phantoms. The

2022 IEEE International Ultrasonics Symposium (IUS) | 978-1-6654-6657-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IUS54386.2022.9958899

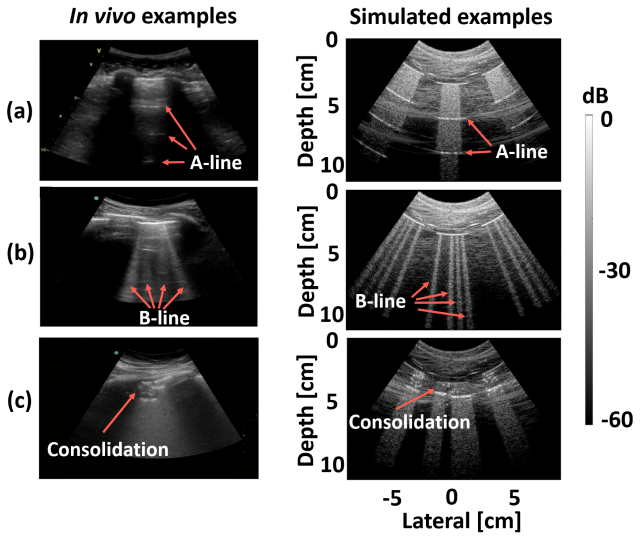

Fig. 1. *In vivo* (adapted from [16]) and simulated examples of (a) A-line, (b) B-line, and (c) consolidation features.

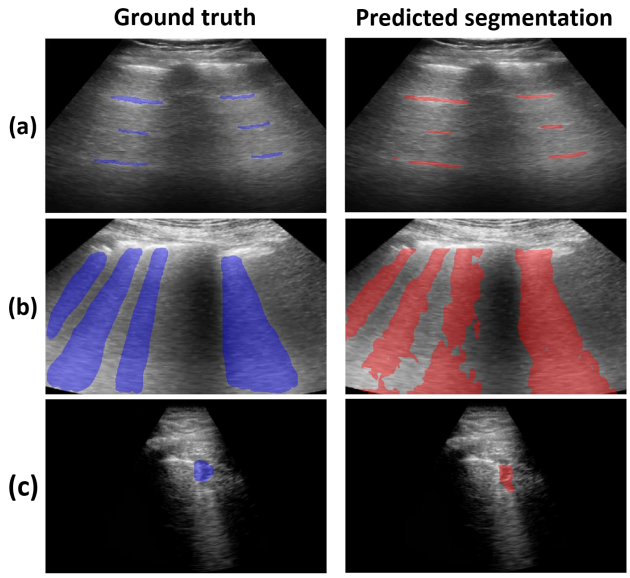

Fig. 2. Ground truth and predicted segmentations overlaid on lung ultrasound B-mode images from healthy and COVID-19 patients, containing (a) A-line (healthy), (b) B-line (COVID-19), and (c) consolidation (COVID-19) features.

simulated transducer was a convex probe with 192 elements, a field of view of 73°, and a central frequency of 4 MHz. The imaging depth was 10 cm, and the sampling frequency was 60 MHz. The simulated raw channel data were then processed with delay-and-sum beamforming, demodulation, envelope detection, and scan conversion to generate B-mode images with a dynamic range of 60 dB (i.e., a common dynamic range when displaying ultrasound images).

To demonstrate the similarity between simulated and real features, Fig. 1 shows real *in vivo* images (left) available in [16], alongside examples of simulated B-mode images (right). Simulated B-mode images were utilized to train and test DNNs based on the U-Net [17] architecture and a modified version of a previously reported deep learning architecture [12]. Each DNN was trained using the Adam optimizer [18] to detect one of the three features with an 80-20 training-testing split, 80 epochs, and a mini-batch size of 16. The training loss was the DSC loss, which is defined as:

$$\text{DSCLoss}(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(1 - 2\frac{|S_{p,i}(I_d;\theta) \cap S_{t,i}|}{|S_{p,i}(I_d;\theta) + S_{t,i}|}\right) \quad (1)$$

where $S_{p,i}$ and $S_{t,i}$ are the vectorized segmentation masks for each training example, and n is the total number of training examples in each mini-batch (i.e., the mini-batch size). The performance of each DNN was measured using the DSC score on a hold-out test set. Applied data augmentations included horizontal flipping with a 0.5 probability, cropping and resizing with a predefined region, contrast adjustment, and Gaussian blur with a kernal size ranging from 3 to 25 and with a 0.8 probability.

The simulation-trained network was then tested on B-mode images from healthy and COVID-19 patients included in a public dataset (available at: https://github.com/BorgwardtLab/covid19_ultrasound) [19]. This public dataset is the largest publicly-available lung ultrasound dataset (202 videos + 59 images), comprising samples of COVID-19 patients, patients with bacterial pneumonia, (non-COVID-19) viral pneumonia, and healthy controls. For our *in vivo* test dataset, we included B-mode images acquired with convex probes from COVID-19 patients with B-line and subpleural consolidation features, and from healthy controls with A-line features. In total, the *in vivo* test dataset includes 32, 107, 27 images with A-line, B-line, and consolidation features, respectively.

III. Results

Fig. 2 shows example ground truth and predicted segmentations from *in vivo* B-mode images of each feature at training epochs where the highest averaged test DSCs were obtained. These images demonstrate that the predicted segmentation qualitatively achieves good agreement with the ground truth.

Fig. 3 shows the mean ± standard deviation DSC as a function of training epoch achieved with the *in vivo* test data when segmenting A-line, B-line, and consolidation features. In each case, the highest mean DSC was generally improved by employing data augmentation. Without data augmentation, the highest mean ± standard deviation DSCs were 0.40 ± 0.29 (epoch 1), 0.17 ± 0.15 (epoch 2), and 0.33 ± 0.35 (epoch 1) for A-line, B-line, and consolidation features, respectively. Employing data augmentation increased the highest mean ± standard deviation DSCs to 0.48 ± 0.29 (epoch 1), 0.45 ± 0.25 (epoch 20), and 0.46 ± 0.35 (epoch 75), representing 20%, 165%, and 39% improvement, respectively when detecting *in vivo* A-line, B-line, and consolidation features, respectively.

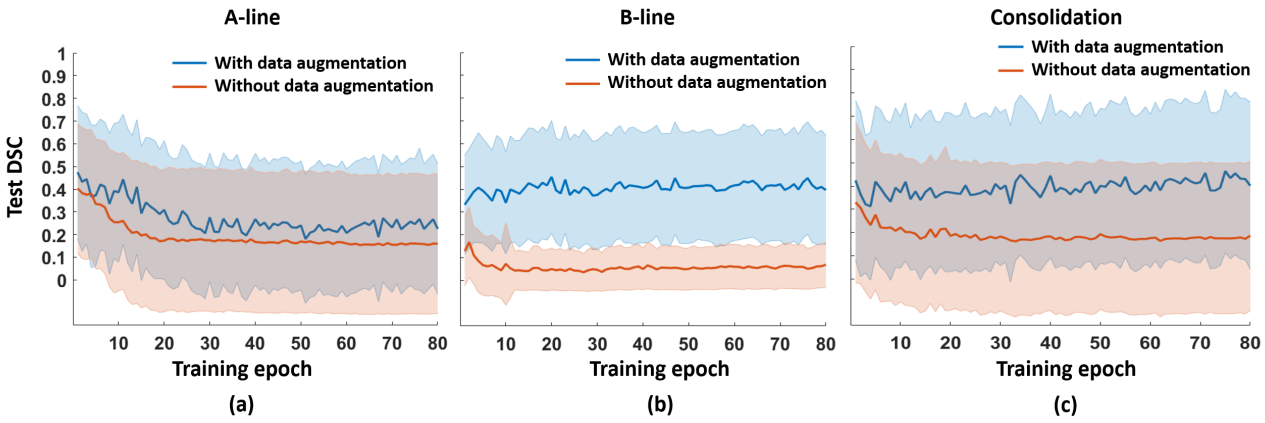

Fig. 3. Mean test DSC per training epoch ± one standard deviation shown as shaded error bars when segmenting *in vivo* (a) A-line, (b) B-line, and (c) consolidation features

IV. Discussion and Conclusion

This paper is the first to investigate the performance of DNNs trained with simulations and data augmentation to segment A-line, B-line, and consolidation features from *in vivo* lung ultrasound B-mode images from healthy individuals and COVID-19 patients. When compared with previously implemented simulation-trained DNNs [14], employing data augmentation improved DSC scores by 20% to 165%. In addition, the proposed simulation-trained approach is not only limited to LUS B-line detection, but also generalizes well to other LUS features including A-line and subpleural consolidation. Future work will investigate network designs that output multifeature segmentations.

Acknowledgment

This work was supported by NIH Trailblazer Award Supplement R21EB025621-03S1.

References

[1] L. Zhao and M. A. L. Bell, "A review of deep learning applications in lung ultrasound imaging of COVID-19 patients," *BME Frontiers*, vol. 2022, Article ID 9780173, 2022.

[2] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli *et al.*, "Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2676–2687, 2020.

[3] R. Arntfield, B. VanBerlo, T. Alaifan, N. Phelps, M. White, R. Chaudhary, J. Ho, and D. Wu, "Development of a convolutional neural network to differentiate among the etiology of similar appearing pathological B lines on lung ultrasound: A deep learning study," *BMJ Open*, vol. 11, no. 3, p. e045120, 2021.

[4] M. J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, and N. Shukla, "COVID-19 detection through transfer learning using multimodal imaging data," *IEEE Access*, vol. 8, pp. 149 808–149 824, 2020.

[5] J. Liu, R. Copetti, E. Sorantin, J. Lovrenski, J. Rodriguez-Fanjul, D. Kurepa, X. Feng, L. Cattaross, H. Zhang, M. Hwang *et al.*, "Protocol and guidelines for point-of-care lung ultrasound in diagnosing neonatal pulmonary diseases based on international expert consensus," *Journal of Visualized Experiments*, no. 145, p. e58990, 2019.

[6] C. Xing, Q. Li, H. Du, W. Kang, J. Lian, and L. Yuan, "Lung ultrasound findings in patients with COVID-19 pneumonia," *Critical Care*, vol. 24, no. 1, pp. 1–3, 2020.

[7] Z. I. Bitar, M. Shamsah, O. Maadarani, O. M. Bamasood, A. Z. Bitar, and H. Alfoudri, "Lung ultrasound and sonographic subpleural consolidation in COVID-19 pneumonia correlate with disease severity," *Critical Care Research and Practice*, vol. 2021, 2021.

[8] B. Bouhemad, S. Mongodi, G. Via, and I. Rouquette, "Ultrasound for "lung monitoring" of ventilated patients," *Anesthesiology*, vol. 122, no. 2, pp. 437–447, 2015.

[9] A. Saraogi, "Lung ultrasound: Present and future," *Lung India: Official Organ of Indian Chest Society*, vol. 32, no. 3, p. 250, 2015.

[10] H. Patel and I. Hacihaliloglu, "Improved automatic bone segmentation using large-scale simulated ultrasound data to segment real ultrasound bone surface data," in *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2020, pp. 288–294.

[11] L. van Knippenberg, R. J. van Sloun, M. Mischi, J. de Ruijter, R. Lopata, and R. A. Bouwman, "Unsupervised domain adaptation method for segmenting cross-sectional CCA images," *Computer Methods and Programs in Biomedicine*, p. 107037, 2022.

[12] A. A. Nair, K. N. Washington, T. D. Tran, A. Reiter, and M. A. L. Bell, "Deep learning to obtain simultaneous image and segmentation outputs from a single input of raw ultrasound channel data," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 12, pp. 2493–2509, 2020.

[13] S. Seoni, G. Matrone, N. Casali, E. Spairani, and K. M. Meiburger, "Generalization of a deep learning network for beamforming and segmentation of ultrasound images," in *2021 IEEE International Ultrasonics Symposium (IUS)*. IEEE, 2021, pp. 1–4.

[14] B. Frey, L. Zhao, T. C. Fong, and M. A. L. Bell, "Multi-stage investigation of deep neural networks for COVID-19 B-line feature detection in simulated and in vivo ultrasound images," in *Medical Imaging 2022: Computer-Aided Diagnosis*, vol. 12033. SPIE, 2022, pp. 38–45.

[15] S. Shahriari and D. Garcia, "Meshfree simulations of ultrasound vector flow imaging using smoothed particle hydrodynamics," *Physics in Medicine and Biology*, vol. 63, no. 20, p. 205011, 2018.

[16] A. Miller, "Practical approach to lung ultrasound," *BJA Education*, vol. 16, no. 2, pp. 39–45, 2016.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer Assisted Intervention*. Springer, 2015, pp. 234–241.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[19] J. Born, N. Wiedemann, M. Cossio, C. Buhre, G. Brändle, K. Leidermann, A. Aujayeb, M. Moor, B. Rieck, and K. Borgwardt, "Accelerating detection of lung pathologies with explainable ultrasound image analysis," *Applied Sciences*, vol. 11, no. 2, p. 672, 2021.