Improved Tracheal Sound Analysis by Transfer Learning from a Lung Sound Model

Fu-Shun Hsu[1], Shang-Ran Huang[1], Chien-Wen Huang[2], Chun-Chieh Chen[2], Yuan-Ren Cheng[1], and Feipei Lai[3]

[1] Heroic Faith Medical Science Co., Ltd., Taipei, Taiwan

[2] Avalanche Computing Inc., Taipei, Taiwan

[3] National Taiwan University, Taipei, Taiwan

Introduction

An automated tracheal sound analysis algorithm is required in a tracheal sound monitor. The algorithm can detect upper airway obstruction by identifying the patterns of continuous adventitious sound (CAS) and start an apnea alarm if breathing is not detected in a prolonged period, which is needed in anesthesia. Previously, we have established lung sound databases, HF_Lung_V1 [1] and HF_Lung_V2 [2], and proposed convolutional neural network (CNN)- bidirectional recurrent neural network (BiGRU) models for inhalation, exhalation, and CAS detection. However, lung sound and tracheal sound have different characteristics so that a lung sound model may failed in tracheal sound analysis. In this study, we collected a small size of tracheal sound, and trained the same CNN-BiGRU model for tracheal sound analysis by transfer learning.

Materials and Methods

A total 176 patients under intravenous general anesthesia were enrolled in this study. The protocol (case number: 19-006-A-2) for this study was approved by Joint Institutional Review Board, Taipei, Taiwan. We used an electronic stethoscope (AccurSound AS-101, Heroic Faith Medical Science Co., Ltd., Taipei, Taiwan) connected to a smartphone (Mi 9T pro, Xiaomi, Beijing, China) to record the tracheal sound. The obtained tracheal sound was truncated into 15 second recordings. The start and end time of inhalation, exhalation, and CAS events were labeled. The recordings and labels were used to train CNN-BiGRU models for segment detection of inhalation, exhalation, and CAS. The steps of preprocessing and postprocessing, model structure, task definition, and cross validation strategy can be found in [1]. We investigated the performance of the lung sound model in tracheal sound analysis. We also trained a tracheal model by using tracheal sound recordings only and three other tracheal sound models by using transfer learning. The learning rate was set at 1e-4.

Results

4772 of 15 second tracheal sound recordings were acquired. 9361 inhalation, 6376 exhalation, and 2334 CAS labels were created. The receiver operating characteristic (ROC) curves and the area under the ROC curves of all model are displayed in Fig 1. The lung sound model performed poorly in the tracheal sound analysis (black dotted lines and the corresponding AUCs in Fig 1). The models trained by tracheal sound performed worse than the lung sound model (red dashed lines and the corresponding AUCs in Fig 1). However, the models trained by transfer learning transcended the previous models by a large margin (green dashed line and the corresponding AUCs in Fig 1).

Discussion

The size of training dataset is an essential factor influencing whether a deep learning model can be trained successfully. With a small number of tracheal sound recordings, we failed to train a tracheal sound analysis model by using the tracheal sound alone, and the lung sound model was proved not competent for the tracheal sound analysis. However, by taking advantage of the previously built lung sound database, finetuning the lung sound model by transfer learning can greatly improve the performance of inhalation, exhalation, and CAS detection in tracheal sound.

References

[1] F.-S. Hsu *et al.*, "Benchmarking of eight recurrent neural network variants for breath phase and adventitious sound detection on a self-developed open-access lung sound database-HF_Lung_V1," *arXiv preprint arXiv:2102.03049,* 2021.

[2] F.-S. Hsu *et al.*, "An Update of a Progressively Expanded Database for Automated Lung Sound Analysis," *arXiv preprint arXiv:2102.04062,* 2021.

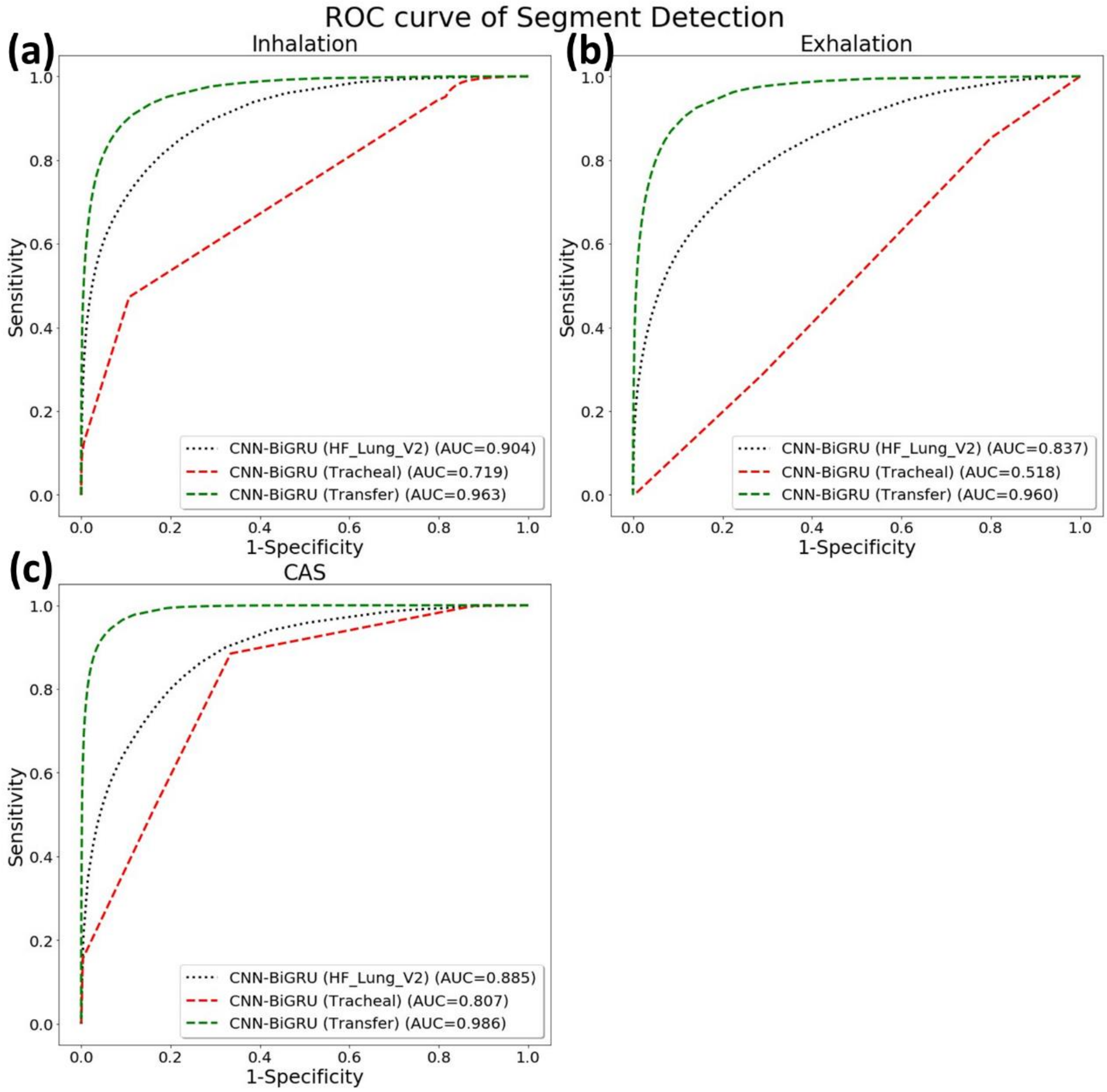

Figure 1. The ROC curves of (a) inhalation, (b) exhalation and (c) CAS detection. CNN-BiGRU (HF_Lung_V2) stands for the model trained by the training set of HF_Lung_V2. CNN-BiGRU (Tracheal) stands for the model trained by the collected tracheal sound. CNN-BiGRU (Transfer) stands for the model trained by transfer learning.