Manuscript accepted by IEEE ICCE2026

Deep Sensor Array Decoding for Remote Health

Qingxue Zhang
Electrical and Computer Engineering
Temple University, Philadelphia, PA, USA
qingxue.zhang@temple.edu

***Abstract*—Deep sensor array decoding is essential for remote health monitoring and big data, leveraging the deep mining capability to reveal the subtle patterns in the decayed signal contactlessly captured. In this study, we propose a novel two-stage deep learning framework for deep array signal decoding, with the first stage for signal fidelity boosting through the latent-space transformation, and with the latter stage for medical insights generation through the convolutional neural network. The latent-space transformation with the deep autoencoder facilitates the suppression of the noise and interferences in the raw signal, thereby boosting signal fidelity crucially. We have further investigated the enhanced supervision for signal fidelity boosting by referring to the additionally captured near-body signal. Evaluated on the real-world application, acoustic sleep apnea detection, our novel deep learning framework has achieved a detection accuracy of 93.3%, superior to state-of-the-art. This study will greatly advance remote health big data with innovative deep sensor array mining.**

Keywords—Deep Array Decoding, Deep Fidelity Boosting, Remote Health Big Data

I. Introduction

Deep learning and vital sign perception have been advancing health practices recently [1, 2]. For instance, the wearable systems have been broadly reported for cardiac, brain, and biomechanical tracking, among others. In addition to wearables, the remote heath monitoring with noncontact systems has also been advanced quickly, leveraging the highly convenient settings.

We take a special interest in remote health monitoring with noncontact systems. There have been previously reported studies on remote health monitoring. The smartphone has been used for acoustic data acquisition and then deep learning has been designed for data analysis [3]. Smartphone has also been used in another study for breathing decoding, from Liu *et al.* [4]. Li *et al.* [5] reported the radar-based health monitoring. Husaini *et al.* [6] reported the Ultra-Wide-Band radar for breathing detection.

Nevertheless, the studies on deep learning and array signal decoding are still limited and urge the extensive efforts in this field. Notably, the array signal perception is expected to bring in spatial dynamics of the pattern of interest. In this study, we propose a novel array signal perception and decoding system. The deep learning of health signals has been researched for a while, but the previous studies usually directly target the inference. For instance, the Convolutional Neural Network (CNN) has been designed for remote acoustic decoding [7]. The reverse neural network has been reported for vital sign prediction [8]. Multi-task CNN [9] has been developed for remote vital sign processing. These studies are of great promise for the vital sign decoding but only directly perform the inference. Instead, in our study we further boost the signal fidelity before the inference, thereby boosting the performance of the system.

In this study, we propose a novel two-stage deep learning framework for deep array signal decoding, with the first stage for signal fidelity boosting through the latent-space transformation, and with the latter stage for medical insights generation through the convolutional neural network. The latent-space transformation with the deep autoencoder facilitates the suppression of the noise and interferences in the raw signal, thereby boosting signal fidelity crucially. We have further investigated the enhanced supervision for signal fidelity boosting by referring to the additionally captured near-body signal. This study will advance long-term continuous health monitoring, thereby facilitating the big data practices [10, 11].

Our major contributions are summarized as below:

(1) A novel deep sensor array decoding system with array signal perception, signal fidelity boosting, and medical insight inference;

(2) The latent-space encoding that suppresses the noise and interferences in the raw signal, for signal fidelity boosting before inference;

(3) Real-world experiments on sleep apnea detection to validate the two-stage deep learning framework that is composed of the fidelity boosting and inference steps.

II. Approach

A. System Diagram

The system diagram is given in Fig. 1. The breathing sound, while strong near the the human subject, decays rapidly when it propagates to the remote acoustic monitor. Therefore, we propose a novel two-stage deep learning framework, to firstly boost the array signal fidelity through the autoencoder that owns the encoder and decoder, and then generate the medical insight though the convolutional neural network.

B. Latent-Space Array Signal Fidelity Boosting

The latent-space array signal fidelity boosting aims to transform the raw signal to the sparse space for noise and

This work was supported in part by the U.S. NSF Grant 2317148 and NSF CAREER Award 2047849.

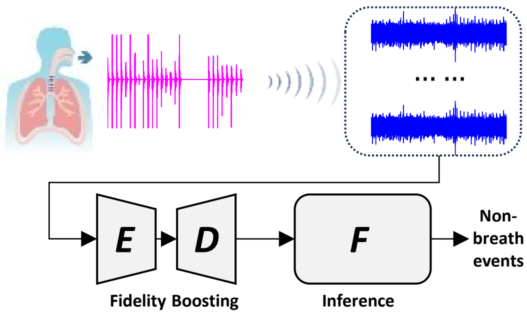

Fig. 1. The proposed novel two-stage deep learning framework, to firstly boost the array signal fidelity, and then generate the medical insight. (Notes) E: encoder; D: decoder; F: inference function.

inference suppression. It is achieved by an encoder D_{ψ}^{e} and a decoder D_{ω}^{d}, as (1), where, ψ and ω correspond to their parameters, respectively, x^{n} is the $n-th$ signal segment, and $\hat{x}^{n}$ is reconstructed signal from the latent space representation $D_{\psi}^{e}(x^{n})$.

$$\hat{x}^{n} = D_{\omega}^{d}(D_{\psi}^{e}(x^{n})) \tag{1}$$

The learning process is as (2), where, $\mathcal{L}^{n}(\cdot)$ is the mean square loss function, $x^{n,c}$ and $\hat{x}^{n,c}$ correspond to the $c-th$ channel in the input and output, respectively, C is the number of channels, and N is the number of instances.

$$\min_{\psi,\omega} \frac{1}{NC} \sum_{n=0}^{N-1} \sum_{c=0}^{C-1} \mathcal{L}^{n}(x^{n,c}, \hat{x}^{n,c}) \tag{2}$$

To further investigate the array signal fidelity boosting approach, we have considered two architectures, with the first named as 'AE + CNN' denoting autoencoder plus convolutional neural network, and with the latter named as 'Guided AE + CNN' that further introduces the nose-side acoustic signal as part of the ground truth during fidelity boosting. For the latter one, the loss function is thus as (3), where, $x_{t}^{n} + \mu g_{t}^{n}$ is the new ground truth that also considers the nose-side strong signal s_{t}^{n} regulated by the coefficient μ, and T is the number of samples in the signal segment. This will facilitate the understanding of whether additional training information can enhance detection performance.

$$\mathcal{L}^{n}(x^{n}, \hat{x}^{n}) = \frac{1}{T} \sum_{t=0}^{T-1} \|(x_{t}^{n} + \mu s_{t}^{n}) - \hat{x}_{t}^{n}\|^{2} \tag{3}$$

C. Inference and Evaluation

The boosted array signal is then fed into the inference stage, which is composed of the convolutional neural network and the fully connected neural network. To extensively evaluate the effectiveness of the proposed novel deep learning framework, we have compared it with the CNN-only method, and the CNN method with the boosted dataset (bstCNN) considering that the sleep apnea events in our experiments are minority cases [7]. Further, we have taken into account the Dynamic Time Warping (DTW) for comparison, which is a popular sequence matching technique for classification.

III. Results

A. Experiments

In out experiments, we have collected the human data with the IRB approval, from four human subjects. Each one performed two trials for training and testing, respectively, with each trial as 30-min. In each trial the participants were asked to have normal breath for 50min, and then repeat the process: hold the breath for about 10-sec and then have normal breath for about 50-sec. The customized array monitor was placed on the bedframe for noncontact acoustic monitoring.

B. Latent-Space Array Signal Fidelity Boosting

The latent-space array signal fidelity boosting has been illustrated in Fig. 2, where two boosting methods are given in a1-a2) and b1-b2), respectively. As shown, in a1) the 'AE + CNN' method can effectively suppress the signal distortions that may be induced by noise and interferences. One (channel No. 4) of the four channels is chosen for visualization. In a2), all four boosted channels are visualized together, indicating the effectiveness of the algorithm.

Further in b1), the 'Guided AE + CNN' shows enhanced signal fidelity boosting, since the reconstructed signal is more

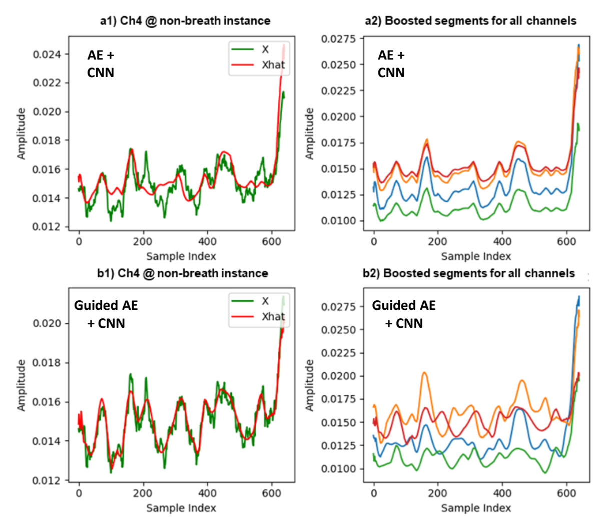

Fig. 2. The latent-space array signal fidelity boosting, where two boosting methods are given in a1-a2) and b1-b2), respectively.

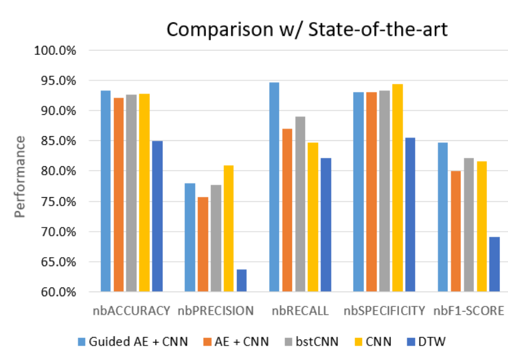

Fig. 3. The inference performance of the two methods, 'Guided AE + CNN' and 'AE + CNN', where state-of-the-art, bstCNN, CNN, and DTW, are also illustrated. (Note) nb: non-breath.

consistent with the trend of the raw signal. b2) further illustrates all channels boosted. The interesting results indicate that the introduction of the nose-side strong acoustic signal to the training process can facilitate the signal boosting essentially.

C. *Inference Performance vs State-of-the-art*

The inference performance of the two methods, 'Guided AE + CNN' and 'AE + CNN', is given in Fig. 3, where state-of-the-art methods [7], bstCNN, CNN, and DTW, are also illustrated.

The proposed 'Guided AE + CNN' approach has the highest accuracy (93.3%), as well as the highest recall and f1-score, regarding the non-breath events (nb). This indicates that the algorithm encourages the detection of the non-breath events, which is essential for sleep apnea monitoring.

IV. Conclusion

In this study, we have proposed a novel two-stage deep learning framework for array signal decoding, targeting the remote health big data. The array signal fidelity is firstly boosted in the first stage with the latent-space transformation through the deep autoencoder, thereby suppressing the noise and interferences. The enhanced supervision is further investigated through introducing the near-body signal as the additional reference. In the second stage, the enhanced signal is fed into the convolutional neural network for inference. Evaluated on the remote sleep apnea detection application, the proposed innovative framework can effectively boost the acoustic signal fidelity and detect the non-breath events with an accuracy up to 93.3%. This study will greatly advance remote health big data through deep signal array decoding innovations.

Acknowledgment

This study is in part supported by U.S. National Science Foundation Grant 2535766 and NSF CAREER Award 2535783. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] J. Zou and Q. Zhang, "EEG Prediction with Deep Recurrent Learning," in *2024 IEEE International Conference on Consumer Electronics (ICCE)*, 2024, pp. 1-3.

[2] J. Wong and Q. Zhang, "Deep Knowledge Distillation Learning for Efficient Wearable Data Mining on the Edge," in *2023 IEEE International Conference on Consumer Electronics (ICCE)*, 2023: IEEE, pp. 1-3.

[3] B. Islam *et al.*, "BreathTrack: detecting regular breathing phases from unannotated acoustic data captured by a smartphone," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,* vol. 5, no. 3, pp. 1-22, 2021.

[4] W. Liu, S. Chang, F. Li, Y. Xu, S. Yan, and Y. Liu, "Contactless breathing airflow detection on smartphone," *IEEE Internet of Things Journal,* vol. 10, no. 4, pp. 3428-3439, 2022.

[5] W. Li, B. Tan, and R. J. Piechocki, "Non-contact breathing detection using passive radar," in *2016 IEEE International Conference on Communications (ICC)*, 2016: IEEE, pp. 1-6.

[6] M. Husaini *et al.*, "Non-contact breathing monitoring using sleep breathing detection algorithm (SBDA) based on UWB radar sensors," *Sensors,* vol. 22, no. 14, p. 5249, 2022.

[7] J. Stauffer, Q. Zhang, and R. Boente, "DeepWave: Non-contact Acoustic Receiver Powered by Deep Learning to Detect Sleep Apnea," in *2020 IEEE 20th international conference on bioinformatics and bioengineering (BIBE)*, 2020, pp. 723-727.

[8] I. Jung, S. C. Park, and G.-N. Wang, "Two phase reverse neural network based human vital sign prediction system," in *2006 International Conference on Computational Inteligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*, 2006: IEEE, pp. 135-135.

[9] S. Chaichulee *et al.*, "Multi-task convolutional neural network for patient detection and skin segmentation in continuous non-contact vital sign monitoring," in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017: IEEE, pp. 266-272.

[10] J. Stauffer and Q. Zhang, "Big Data Edge on Consumer Devices for Precision Medicine," in *2022 IEEE International Conference on Consumer Electronics (ICCE)*, 2022: IEEE, pp. 1-3.

[11] J. Zou and Q. Zhang, "eyeSay: Eye Electrooculography Decoding with Deep Learning," in *2021 IEEE International Conference on Consumer Electronics (ICCE)*, 2021: IEEE, pp. 1-3.