STConvSleepNet: A Spatiotemporal Convolutional Network for Sleep Posture Detection*

Dikun Hu, Weidong Gao, Kai Keng Ang, *Senior Member IEEE,* and Mengjiao Hu, *Member IEEE*

***Abstract*— Sleep posture, intricately connected to sleep health, has emerged as a crucial focus in sleep medicine. Studies have associated the supine posture with increased frequency and severity of obstructive sleep apnea (OSA), while lateral postures may mitigate these effects. For bedridden patients, regular posture adjustments are essential to prevent ulcers and bedsores, highlighting the need for precise sleep posture detection. In this work, we propose STConvSleepNet, a novel method for detecting sleep posture using piezoelectric sensor pressure data. It employs two shallow CNN2D networks to discriminate spatial features and two CNN1D networks to discriminate temporal features, with each network processing either the heart rate or the respiratory rate. These networks are trained to detect sleep postures from spatial features of the pressure distribution, and temporal features of heart rate and cardiopulmonary activities variability. We collected data from 22 participants with 300-450 samples each, for a total of 8583 samples using a 32-sensor array. We performed 5-fold cross-validation on the data using the proposed method. The results showed that the proposed STConvSleepNet yielded 91.11% recall, 92.89% precision, and 94.93% accuracy. This is comparable to the state-of-the-art method that needs a significantly increased number of sensors to achieve slightly increased accuracy of 96.90%. Hence these results showed promise of using the proposed STConvSleepNet for cost-effective home sleep monitoring using portable devices.**

***Clinical Relevance*— This sleep posture detection potentially suits diverse populations for long-term, at home settings.**

I. Introduction

Sleep posture, intricately connected to sleep health, has emerged as a crucial focus in sleep medicine. Research shows that the supine posture is closely linked to an increased frequency and severity of obstructive sleep apnea (OSA) events, while lateral posture may alleviate these issues [1], [2]. In addition, for bedridden patients, regular adjustments in sleep posture are essential to prevent ulcers and bedsores caused by prolonged immobility. These situations highlight the vital role of sleep posture detection in preventing or mitigating health conditions. Despite Polysomnography (PSG) being the gold standard for sleep posture assessment [3], its invasiveness and the necessity for professional oversight limit its practicality for continuous monitoring. Therefore, creating an innovative, portable, and non-invasive device for accurate sleep posture detection is of great significance.

Home sleep tests (HSTs), including both wearable and non-wearable devices, offer practical home monitoring solutions. Wearables like chest straps and smartbands offer posture detection and heart rate monitoring but may result in higher false positives and limitations for extended use [4], [5]. Non-wearables such as cameras, radar, and pressure-based systems [6]–[15] offer advanced tracking capabilities. However, they face several challenges: cameras raise privacy concerns [6], radar systems have difficulty distinguishing between left and right lateral [7]–[8], and traditional dense sensor arrays in pressure-based systems are often limited by their size and cost [9]–[15]. Recent sparse sensor arrays in mattresses show promise with fewer sensors but may compromise pressure image clarity, affecting model effectiveness [16]–[17]. In summary, these methods largely depend on static pressure distribution, limiting adaptability and generalizability.

In this study, we propose STConvSleepNet, a novel spatiotemporal convolutional network for sleep posture detection. It analyzes pressure signals from a limited array of piezoelectric sensors. To overcome sensor number limitations, our model employs feature extraction techniques, deriving critical indicators from multi-channel vibrational data. STConvSleepNet enhances adaptability by integrating spatial and temporal features through convolutional networks. Specifically, it utilizes two shallow CNN2D networks for spatial features within cardiorespiratory maps, and two CNN1D networks for temporal features in heart and respiratory rates. This integrated approach not only enables robust sleep posture detection but also improves generalization by capturing dynamic physiological changes.

II. Methodology

A. Data collection

Considering the characteristics of flat shape, high sensitivity, low cost, excellent high-frequency response, and flexibility, we have elected to use piezoelectric ceramic materials for the construction of our monitoring mattress. Shown in Fig. 1(a), sensors consisting of a 3 cm piezoelectric element on a brass base are affixed to a 3cm by 5cm PCB, generating a microcurrent under pressure. Fig. 1(b) presents a 32-channel 8 by 4 matrix with electrodes spaced to match an adult's chest, ensuring precise vibration mapping. Fig. 1(c) displays the system's design, featuring an analog front end (AFE) with piezoelectric ceramics and non-inverting amplifiers (NIAs) to condition the pressure induced electrical signals. Post-conditioning, low pass filters (LPFs) refine these

D. Hu, and W. D. Gao are with Institute for Beijing University of Posts and Telecommunications (BUPT)， School of Communication Engineering, China. (Corresponding author to provide phone: +86 13522965162 e-mail: gaoweidong@bupt.edu.cn)

M. Hu, and K. K. Ang are with Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore.

K. K. Ang is also with School of Computer Science and Engineering, Nanyang Technological University, Singapore.

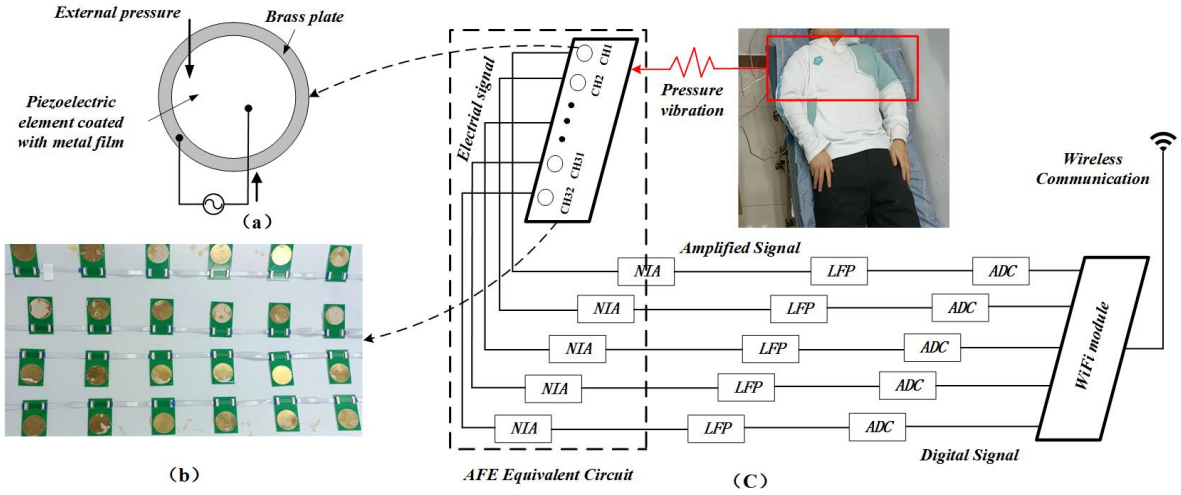


Fig.1 (a) Piezoelectric ceramic sensor unit; (b) Piezoelectric sensor array; (c) Circuit used to scan the pressure distribution.

signals, which are digitized by analog to digital converters (ADCs) and Wi-Fi module.

B. Signal prepossessing

The forces from the chest regions mainly derived from respiratory motion, heartbeat, and body movement. Respiration and heartbeat are high-frequency periodic activities within 0.1-0.8 Hz and 0.8-15 Hz frequency bands [18]. In contrast, body movement is characterized as slow, aperiodic activity concentrated in the low-frequency range. To eliminate motion artifacts from body movements, we employed sample entropy and energy entropy metrics to identify and eliminate affected non-periodic signal segments. For segments less impacted by motion artifacts, a Chebyshev type-I finite impulse response (FIR) band-pass filter is used to extract cardiac and respiratory signals, presenting harmonic-like features. Our circuit analysis confirmed a linear relationship between these periodic activities and their harmonic patterns, enabling further extraction of key physiological features from these signals.

B. Feature Extraction

In this work, the features we extracted are divided into spatial and temporal features. The cardiac signal, originating from the heart's pumping action, includes five sub-waves known as IJKL [19], each corresponding to a specific stage of a heartbeat. The temporal features, reflecting dynamic characteristics, are derived from the cardiac signal's envelope, encompassing RR intervals and beat amplitude. Similarly, respiratory signals provide respiration intervals and peak intensity variations, showcasing dynamic sequential changes.

Spatial features represent cardiac and respiratory activity distribution across a 32-channel array. Sinusoidal wave decomposition refines the signal, preserving its active characteristics and enabling formulaic representation. Applying Maclaurin series expansion to our formulas eliminates time variables' influence, aiding in accurate activity calculations for cardiac and respiratory functions in each channel. Fig.2 shows activity intensities across sleep postures like supine, left lateral, and right lateral, revealing spatial information in these static features.

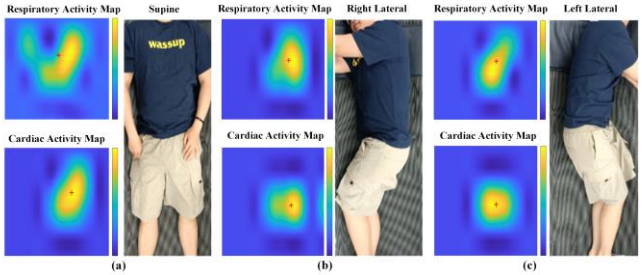


Fig.2 (a) Supine state; (b) Right lateral state; (c) Left lateral state.

C. STConvSleepNet Model

In setting up our experimental samples, we adopted a minute-based segmentation approach. Specifically, data from all 32 channels, collected continuously over a minute, were consolidated into a single sample. As previously described, the temporal features we extracted are one-dimensional attributes of heartbeat and respiration. These features include inter-beat intervals, heartbeat peaks, respiration intervals, and respiration peaks each minute. The spatial features associated with cardiopulmonary activity encompass the two-dimensional active distribution of cardiac and respiratory activities. To integrate these multimodal features, we developed STConvSleepNet, a novel network based on the Multi-Channel Convolutional Neural Network (MCNN) architecture.

STConvSleepNet, shown in Fig.3, uses four shallow CNNs: CNN1D-1 and CNN1D-2 for temporal heartbeat and respiration features, CNN2D-1 and CNN2D-2 for spatial cardiac and respiratory maps. This architecture effectively captures both temporal dynamics and spatial distributions, crucial for sleep posture detection. In the network, one-dimensional features undergo processing via CNN1D-1 and CNN1D-2, which are configured with diverse filter sizes and stride lengths. To maintain uniformity across inputs, we standardize the lengths of feature vectors through resampling: respiratory features are resampled to a consistent length of 90, while cardiac features are adjusted to 180. This resampling

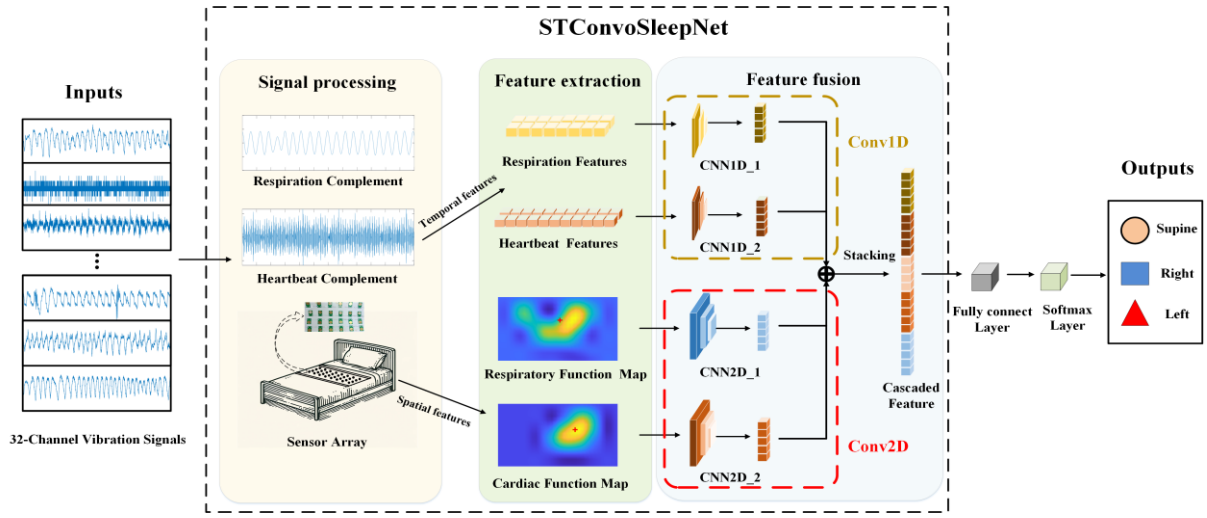


Fig.3 Architecture of STConvoSleepNet.

strategy not only ensures homogeneity of input dimensions for the CNN1D layers but also enhances the network's ability to discern intricate patterns within the temporal data.

Constrained by the number of sensors, the resolution of the cardiopulmonary activity map is only 4×8. To address this small size, we enlarge the original image 16 times using the bicubic interpolation. Subsequently, detailed features of cardiopulmonary activity are extracted by the CNN2D-1 and CNN2D-2. The input for CNN2D-1 is a 16×32 respiration activity map, processed by a shallow neural network with three 2-dimensional convolutional layers and two max-pooling layers. Similarly, CNN2D-2 processes a 16×32 heartbeat activity map, but it comprises a network with four 2-dimensional convolutional layers and three max-pooling layers. Finally, multi-modal features are fused into a cascaded feature to facilitate sleep posture recognition.

III. Result and Discussion

A. Dataset

This study was carried out at the Sleep Center of Peking Union Medical College Hospital, involved 22 participants (15 males and 7 females, aged 30-68 years). Most suffered from OSA of varying degrees, with diverse body types, including 8 overweight and 5 lean individuals. Nightly data were collected using a pressure-sensitive mattress for durations ranging from 4 hours 30 minutes to 7 hours, with subjects' sleep postures and body movements being simultaneously monitored using PSG. The recordings had a total duration of 8583 minutes, among which 4232 minutes (49.30%) were supine samples, 2934 minutes (34.18%) were right lateral samples, and 1417 minutes (12.06%) were left lateral samples. To eliminate the impact of device variability on data accuracy, calibration was conducted on different devices.

To evaluate the model's adaptability and generalizability, 22 nights of data were distributed into five subsets for a balanced cross-validation process. Systematic testing of the model's performance across these different data subsets was conducted through 5-fold cross-validation. The model's performance was evaluated using four metrics: F1 score, accuracy (Acc), recall (Rec), and precision (Pre). The model's generalization was further substantiated by analyzing its consistent performance across training datasets, ensuring its reliability under diverse environmental conditions.

B. Detection Performance

Enhancing this model required initializing key hyperparameters such as learning rate, batch size, and maximum epochs. The initial learning rate was 0.001, with training spanning up to 100 epochs. A batch size of 48 balanced performance and speed. Moreover, adding L1 regularization to the loss function helped mitigate overfitting.

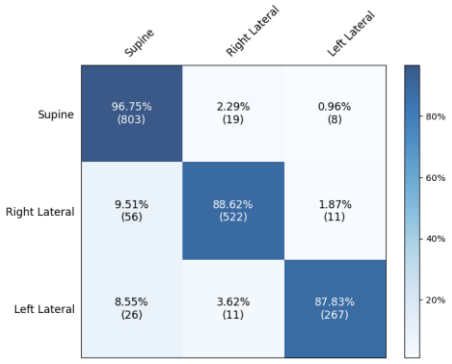


Fig.4 Confusion Matrix for Sleep Posture Detection.

Fig. 4 shows the confusion matrix of the best sleep posture detection achieved by the STConvSleepNet model. Under normal signal conditions, the accuracy, recall, precision, and F1 score of the proposed method in this study reached 94.93%, 91.11%, 92.89%, and 0.9188, respectively.

C. Ablation Study and Performance Comparison

To validate our model's performance and generalizability, we performed ablation experiments, shown in Table 1, focusing on our spatiotemporal feature integration strategy. STConvSleepNet, merging temporal and spatial features, surpasses networks using just one feature type. It shows minimal fluctuation in 5-fold cross-validation, indicating strong generalization.

TABLE I. DESCRIPTION OF THE DATASET

Metric	*CNN1D for Temporal*	*CNN2D for Spatial*	*STConvSleepNet*
Acc	0.8501 ± 0.0227	0.9124 ± 0.0113	0.9444 ± 0.0049
Rec	0.7567 ± 0.0261	0.8508 ± 0.0184	0.9086 ± 0.0072
Pre	0.7754 ± 0.0231	0.8657 ± 0.0147	0.9235 ± 0.0068
F1 score	0.7610 ± 0.0269	0.8625 ± 0.0172	0.9124 ± 0.0064

TABLE II. COMPARING WITH LEADING METHODS

Reference	*Number of sensors*	*Method*	*Accuracy*
Matar et al [9]	1728	FANN	97.1%
Enokibori et al [10]	3200	DNN	99.7%
Heydarzadeh et al [11]	2048	DNN	98.1%
Qisong et al [12]	1024	CNN+SVM	91.2%
Diao et al [13]	1024	RestNet	95.1%
Yousef et al [14]	2048	KNN	97.1%
Georges et al [15]	1728	FFANN	97.9%
Zhangjie et al [16]	768	HOG+PCA	89.0%
Jen et al [17]	220	CNN	96.9%
Our approach	**32**	**STConvSleepNet**	**94.9%**

Table 2 compares STConvSleepNet against leading sleep posture detection methods. While slightly less accurate than some deep neural networks, it needs far fewer sensors (95-99% less), making it ideal for embedded devices. Unlike typical lab data, we used real sleep data with significant motion artifacts, constituting about 10% of the total. We addressed this challenge with a model-based feature extraction method optimized for these artifacts, enhancing robustness. By combining the dynamic temporal and static spatial features, we greatly improved the generalization ability of the model. STConvSleepNet offers competitive performance in home and outdoor settings, outperforming methods that rely on extensive sensors or traditional techniques.

IV. CONCLUSION

In this study, we proposed STConvSleepNet, a hybrid network combining CNN1D and CNN2D to process temporal and spatial features extracted from a limited sensor array. Our approach showed comparable performance to the state-of-the-art method that needs significantly increased number of sensors, showing promises for cost-effective home sleep monitoring using portable devices. Yet, correlation between continuous sleep postures has not been investigated in this study. Given sleep's dynamic nature with frequent movements, future research will analyze postures' interrelations, including features like turning and movements, to enhance posture change detection.

ACKNOWLEDGMENT

The Research is supported by Project of Guangdong Province under grant No. 2022B1515130009 and the Special subject on Agriculture and Social Development under grant No. 2023B03J0172.

REFERENCES

[1] Scott Schecter and Junjie Liu, "411 rem predominance of osa: Associated with supine position, but not with cpap adherence," *Sleep*, 2021. Doi:10.1093/sleep/zsab072.410

[2] A. Menon and M. Kumar, "Influence of body position on severity of obstructive sleep apnea: A systematic review," *Int. Scholarly Res. Notices*, vol. 2013, no. 670381, pp. 1–7, Otc. 2013.

[3] Diane Lim, and Kate Sutherland. "Reinventing polysomnography in the age of precision medicine." *Sleep Medicine Reviews*, 52, 2020. Doi:10.1016/j.smrv.2020.10131

[4] Kaichen Xu, and Satoko Honda, "A wearable body condition sensor system with wireless feedback alarm functions," *Advanced Materials*, 21, 2021. Doi:10.1002/adma.202008701.

[5] Mustafa Radha, and Marco Ross, "A deep transfer learning approach for wearable sleep stage classification with photoplethysmography," *npj Digital Medicine*, 135, 2021.Doi:10.1038/s41746-021-00510-8

[6] Shi Lan, and Chen Yan, "Sleep classification with artificial synthetic imaging data using convolutional neural networks," *IEEE Journal of Biomedicaland Health Informatics*, 27(1): 421–432,2023. Doi10.1109 /JBHI.2022.3210485

[7] Ye Guanhua, Yin Hongzhi, and Chen, "Fenet: A frequency extraction network for obstructive sleep apnea detection," *IEEE Journal of Biomedical and Health Informatics*, 25(8):2848–2856, 2021. Doi: 10.1 109/JBHI.2021.3050113.

[8] Bing Luo,"Human Sleep Posture Recognition Method Based on Interactive Learning of Ultra-Long Short-Term Information," *IEEE Sensors Journal*, VOL. 23, NO. 12, 15 JUNE 2023

[9] Georges Matar, and Georges Kaddoum, "Artificial neural network for in-bed posture classification using bed-sheet pressure sensors," *IEEE Journal of Biomedical and Health Informatics*, 24:101–110, 2020

[10] Yu Enokibori and Kenji Mase, "Data augmentation to build high per-formance Dnn for in-bed posture classifification," Journal of Information Processing, 26:718–727, 2018

[11] Heydarzadeh Mehrdad, and Nourani Mehrdad, "In-bed posture classifification using deep autoencoders," In 2016 38th Annual International Conference of the IEEE Engineer

[12] Hu Qisong, and Tang Wei, "A real-time patient-specifific sleeping posture recognition system using pressure sensitive conductive sheet and transfer learning," *IEEE Sensors Journal*, 21(5):6869–6879, 2020

[13] Diao Haikang,"Deep residual networks for sleep posture recognition with unobtrusive miniature scale smart mat system," IEEE Transactions on Biomedical Circuits and Systems, 15:111–121, 2021

[14] Yousefi Ostadabbas, and Farshbaf, "Bed posture classification for pressure ulcer prevention," *In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 7175–7178, 2011. Doi:10.1109/IEMBS.2011.6091813

[15] Georges Matar, and Georges Kaddoum, "Artificial neural network for in-bed posture classification using bed-sheet pressure sensors," *IEEE Journal of Biomedical and Health Informatics*, 24:101–110, 2020.

[16] Chen Zhangjie and Ya Wang, "Remote recognition of in-bed postures using a thermopile array sensor with machine learning," *IEEE Sensors Journal*, 21:10428–10436, 2021.

[17] Jen Kau," Pressure-sensor-based sleep status and quality evaluation system," *IEEE Sensors Journal*, vol. 23, no. 9, 1 May. 2023

[18] Mai Yaozong, and Yu Baoxian, "Non-contact heartbeat detection based on ballistocardiogram using unet and bidirectional long short-term memory," *IEEE Journal of Biomedical and Health Informatics*,26(2):3720–3730, 2022. Doi:10.1109/JBHI.2022.3162396

[19] Aydin Leonhardt, "An algorithm for heart rate detection from the bcg based on continuous spline wavelet transform," *IEEE Transactions on Instrumentation and Measurement*, 69:3297–3307, 2020.