EEG-BASED FAST AUDITORY ATTENTION DETECTION IN REAL-LIFE SCENARIOS USING TIME-FREQUENCY ATTENTION MECHANISM

Zhuang Xie[1] Jianguo Wei[2] Wenhuan Lu[2] Zhongjie Li[2] Chunli Wang[3] Gaoyan Zhang[2]*

[1]*School of Future Technology, Tianjin University, Tianjin, China*

[2]*CCALab, College of Intelligence and Computing, Tianjin University, Tianjin, China*

[3]*School of Electrical and Information Engineering, LanzhouJiaotong University, Lanzhou, China*

[*] *corresponding to zhanggaoyan@tju.edu.cn*

Paper ID:3659

Background

Auditory attention detection (AAD) based on electroen cephalogram (EEG) has been extensively studied in laboratory settings, but its application in real-life scenarios is limited. Therefore, there is a need for improved AAD methods that can effectively detect auditory attention in real-life scenarios.

Proposed method

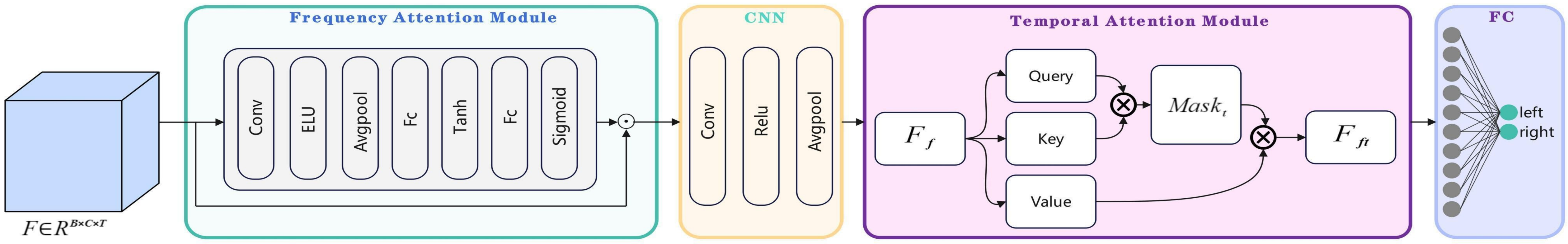

Fig. 1. Schematic diagram of time-frequency attention network.

This method consists of four steps:

(1) Using a frequency band attention module to assign weights to sub-bands of EEG signals.

(2) Employing a convolutional neural networks (CNN) to extract preliminary features.

(3) Using a temporal attention model to explore the temporal patterns of EEG signals.

(4) Using a fully connected layer to make classification decisions.

Dataset

EEG data were collected from 20 subjects. The participants were instructed to pay attention to one of two simultaneously presented speech streams. The speech direction was presented at 45 degrees to the left and right of the subject. All measurements were taken in a cafeteria. More details about the dataset can be found in [1].

Contrast experiment

We re-implemented four baseline AAD models, including stimulus reconstruc tion (SR)[2], canonical correlation analysis (CCA)[3], common spatial pattern (CSP)[4] method, and STA-net[5] model, and conducted experiments, as shown in Table 1.

Table 1. AAD performance (%) using different methods

Method	Decision Window(s)				
	0.1	0.2	0.5	1	2
SR	-	-	53.3	55.3	57.1
CCA	-	-	54.4	55.6	58.9
CSP	77.5	80.9	83.0	84.9	89.8
STA-net	83.4	88.3	92.5	93.1	95.3
CNN-FT(ours)	**91.8**	**93.1**	**95.2**	**96.7**	**98.1**

Ablation Study

To better study the impact of the frequency band and temporal attention module, we define four comparison models and conduct experiments on a 0.1 s window, as shown in Table 2.

Table 2. The results of the ablation experiment

Model	Accuracy (%)
CNN (without both types of attention)	84.1
CNN-T (without frequency attention)	86.2
CNN-F (without temporal attention)	88.9
CNN-FT (with time-frequency attention)	91.8

Visualization

To investigate the contributions of different sub-bands to AAD, we visualized the weights of each sub-band when the subjects were in different states on the 2-second decision window, as shown in Figure 2.

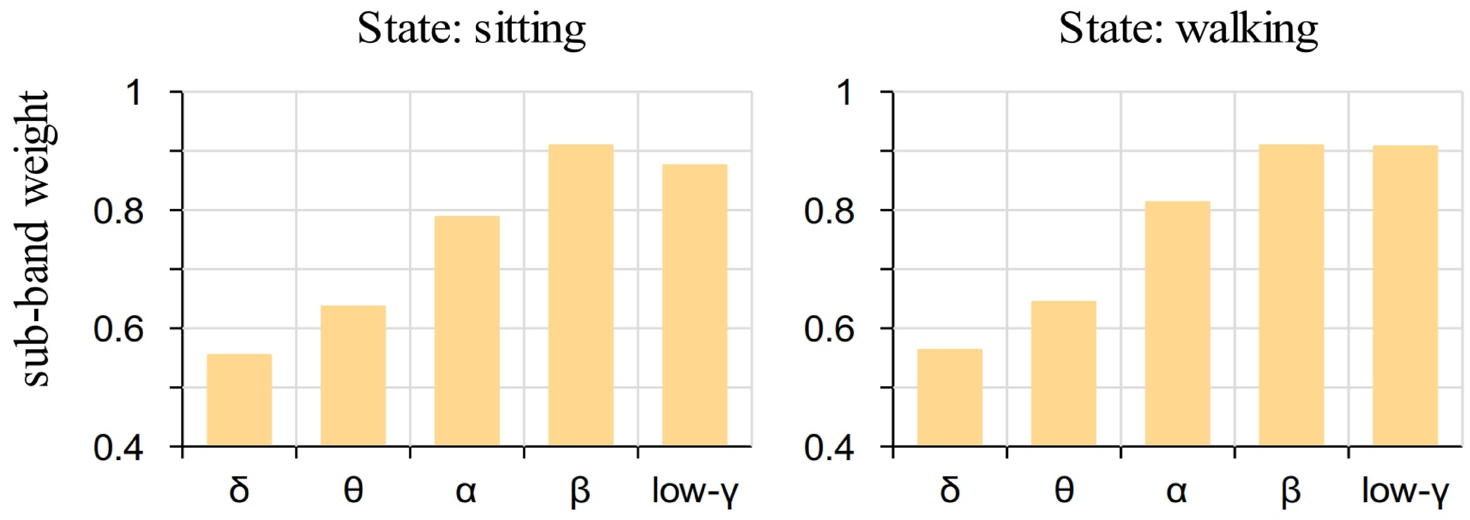

Fig. 2. The results of frequency attention visualization

Conclusion

In this study, the EEG-based AAD during walking state achieved 90% plus accuracy, showing the potential in real-life application. When using a decision window of 0.1s for AAD, our proposed model still has an accuracy of 91.8%, suggesting a fast decoding ability.

References

[1] L Straetmans, et al., "Neural tracking to go: auditory attention decoding and saliency detection with mobile eeg," 2022.

[2] James A O'sullivan, et al., "Attentional selection in a cocktail party environ ment can be decoded from single-trial eeg," 2015.

[3] Emina Alickovic, Thomas Lunner, et al., "A tutorial on auditory attention identification methods," 2019.

[4] Simon Geirnaert, et al., "Fast eeg-based decoding of the directional focus of auditory attention using common spatial pat terns," 2020.

[5] Enze Su, Siqi Cai, et al., "Stanet: A spatiotemporal attention network for decoding auditory spatial attention from eeg," 2022.