



Research on the Identification Method of Audiovisual Model of EEG Stimulation Source

Zhaoxiang Lu¹(✉), Mei Wang¹, and Wenbin Chai²

¹ Xi'an University of Science and Technology, Xi'an 710054, China
20206043038@stu.xust.edu.cn

² Xi'an Nuclear Instrument Co., Ltd., Xi'an 710000, China

Abstract. For the cloud center of the mining industry, one of the most important tasks is to obtain the special environment safety state information because the hidden dangers of the water and fire, as well as the gas, always threaten human life and the production. In addition, audiovisual model of EEG stimulation source is integrated by using the convolutional neural network model with the inception network. It fuses the EEG information with the additional environment sensor information to increase the environment safety classification accuracy. Through brain network visualization and brain connection, we can get the density and weight change of global network connection in different audio-visual stimulation stages, which makes the analysis of EEG signals more intuitive. The experimental results indicate that the environment safety recognition accuracy of the audiovisual model can respectively reach 87.98%, 88.4%, and 90.12% for the single visual, the single auditory, and the audiovisual stimulus. The audio-visual modality has the best performance for the audiovisual evocation than the single visual or auditory stimuli.

Keywords: Safety state information · Audiovisual model · EEG · Brain network · Improved CNN

1 Introduction

Neuromorphic computing science has the potential to lead the next generation of computing science because of its high computing efficiency [1]. The neuromorphic computing model draws from the human brain computation mechanism, which performs the critical functions of perception and cognition [2, 3]. Due to the 80% of external environment information is obtained through the visual neural mechanism, and the 90% of external environment information is obtained through the audiovisual neural mechanisms, then the neuromorphic computing model of audiovisual evocations is absolutely important for human perception and cognition mechanism. Meanwhile, the audiovisual neuromorphic computing model is equally significant for the patrol robot using the human perception and cognition model [4–6].

For the perception and cognition mechanism modeling by using human electroencephalograph (EEG) signals of audiovisual evocations, a specific filtering algorithm is

firstly needed to eliminate the noises from the raw EEG signals [7–9]. Researchers have used different filtering methods to remove the noises, such as the filter bank method to decompose the EEG signals into sub-band and have overlapped frequency cutoffs, and the coincidence filtering method to remove the different types of noises [10, 11]. Besides, scientists propose a denoising method of an empirical mode decomposition (EMD) fused with the principal component analysis for the seizure EEG processing [12]. However, the different descriptions of EEG signals of the different evocation patterns have particular characteristics, and different filtering methods are required [13–15].

The paper is proposed an improved convolutional neural network (CNN) parallel computation strategy for the convolutional layer and Brain source estimation. Results and comparisons are provided in conclusions.

2 Related Works

In this section, we introduce the existing studies related to our approach and describe the differences.

For neural computing, the researchers performed a simple fusion of EEG and neuro-computing [16–18]. Besides, the neuromorphic feature extraction is secondly required. There is usually the one-dimensional time-frequency feature of a single tunnel EEG feature [19], the two-dimensional EEG energy feature of the multiple tunnels [20–22], and the three-dimensional EEG connection feature of the brain network [23]. Furthermore, researchers propose the EEG feature of the composite multivariate multiscale fuzzy entropy to describe the motor imagery [24–26]. Besides, the EEG and the electrooculogram (EOG) are fused to represent the brain-muscle feature [27]. However, the evocation of EEG signals is a time process, the dynamic characteristic should be considered [28].

Moreover, the neuromorphic computation model is also an important part of modeling human perception and cognition mechanisms [29]. The hidden Markov models and recurrent neural networks are used for event detection. In addition, the multi-modal neural network is constructed for person verification using signatures and EEGs [30]. In addition, scientists give the brain informatics-guided systematic fusion method and propose a predictive model to build a bridge between brain computing and the application services. The self-regulated neuro-fuzzy framework and the hierarchical brain networks via volumetric sparse deep belief network are developed to conduct neuromorphic computation.

3 Estimation of Brain Network Source Model

Brain networks are novel methods to study the interactions between important brain regions, combined with brain inverse activity is possible to extract nonlinear information about the source of local electrical activity using nonlinear potential information generated directly from the activity of local organismal neural circuits in the cerebral cortex. The source model is calculated from the segmentation of Magnetic Resonance Image (MRI). Usually, the white matter-gray matter interface is chosen as the main source space generating region. The MRI anatomy and channel locations are used with the same anatomical landmarks (left and right ear anterior points and nasal tip).

According to dipole theory, the EEG signal $X(t)$ recorded from the M channel can be considered as a linear combination of P time-varying current dipole sources $S(t)$.

$$X(t) = \begin{pmatrix} x_1(t) \\ \dots \\ x_M(t) \end{pmatrix} = G \cdot \begin{pmatrix} s_1(t) \\ \dots \\ s_p(t) \end{pmatrix} + N(t) = G \cdot S(t) + N(t) \quad (1)$$

where G is the guide field matrix describing the deterministic quasi-instantaneous projection of the source onto the scalp electrodes. $N(t)$ is the measurement noise inherent in any acquisition process, and G is calculated from the head model and the position of the electrodes. If the source distribution is constrained to a current dipole field uniformly distributed across the cortex and normal to the cortical surface, the position of the source is defined and direction.

Using a validated mind model as well as an internally generated source model, each two elements within the network are connected to function as a whole network, and the effect of this network on EEG signals under audiovisual modal stimulation is then analyzed by the following experiments.

4 Experiments and Results

4.1 Experimental Environment

The EEG signals were collected by the 40 leads electrode cap of Neuroscan. The international 10–20 system electrode placement method is adopted to standardize the position of each lead electrode. And that distance between potential is moderate. According to the analysis of the visual and auditory experimental areas in previous studies, the electrode can accurately and effectively collect the original data of EEG signals, which can meet the requirement of this study. Moreover, the symmetry between potentials can guarantee the accuracy of the following brain source inverse estimation results. The placement of standard 10–20 system electrodes is shown in Fig. 1.

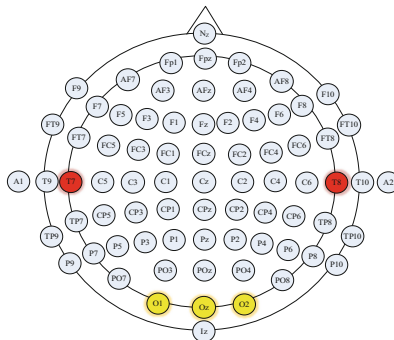


Fig. 1. Electrode distribution diagram of standard 10–20 system

The reference electrode is between Fz and Cz, and the ground electrode is between Fz and Pz. EEG signals are synchronously collected at a sampling frequency of 1 kHz,

and the impedance between scalp and electrode is less than $5\text{ k}\Omega$. Electrodes are placed on each lobe of the brain, and electrode labels identify their location. The odd number is left, the even number is right, and the middle electrode is “z”. For example, the F3 electrode comes from the left frontal lobe, and Cz is placed at the vertex on the top of the head.

In this experiment, five subjects were selected, the age distribution was about 25 years old, both sexes were male, without any history of brain diseases, and all of them had normal hearing and normal or corrected vision. Before the experiment, make sure that the subjects have enough rest time, and keep their hair clean within one hour before the experiment, to ensure less interference in the collection process. The experiment was conducted in a shielded room, including three kinds of stimuli: visual, auditory, and audio-visual stimuli.

4.2 Experiment and Analysis of Brain Source Estimation Under Audiovisual Stimulation

In this experiment, the 30s scalp EEG sequence of the subjects in the EEG acquisition experiment under audio-visual mixed stimulation was selected, and then the collected EEG signals were denoised by the CEEMDAN-FastICA algorithm, and then the source was estimated in different periods. Finally, the brain network connection model under audio-visual stimulation was constructed by the collected EEG signals.

As shown in Fig. 2, Fig. 3 and Fig. 4 from top to bottom, the audio-visual mixed modal stimuli collected in three different periods are used for brain source estimation. Figure 5. is a diagram of brain network connection under audio-visual mixed-mode simulation.

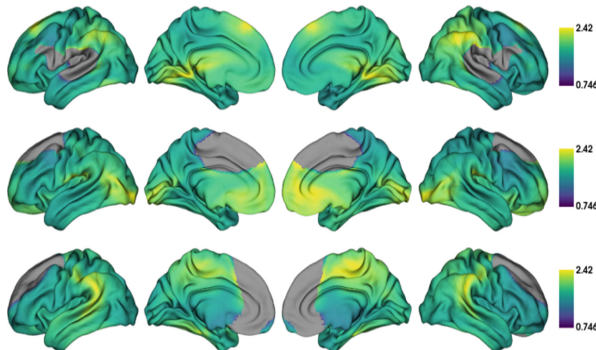


Fig. 2. Estimated images of brain sources under audio-visual stimulation in different periods(R0)

Through brain network visualization and brain connection, we can get the density and weight change of global network connection in different audio-visual stimulation stages, which makes the analysis of EEG signals more intuitive.

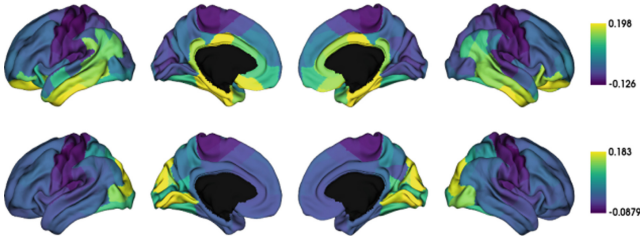


Fig. 3. Estimated images of brain sources under audio-visual stimulation in different periods(R1)

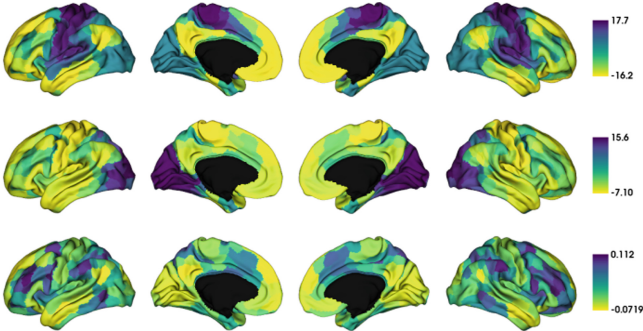


Fig. 4. Estimated images of brain sources under audio-visual stimulation in different periods(R2)

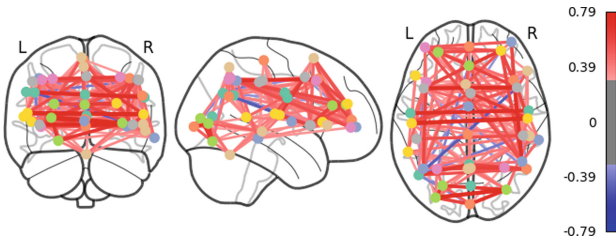


Fig. 5. Brain connection diagram under audio-visual stimulation

4.3 Experiment and Analysis of Improved CNN Audiovisual Model

In the experiment, binary cross-entropy is used as the cost function, L2 regularization is used, and the penalty term of the cost function is used to avoid overfitting. Visual stimulation is used as the evoked modality of EEG signals to obtain relevant EEG data. Taking the MSE value of the data as the input of the improved CNN, the curves of training accuracy and loss function are shown in Fig. 7.

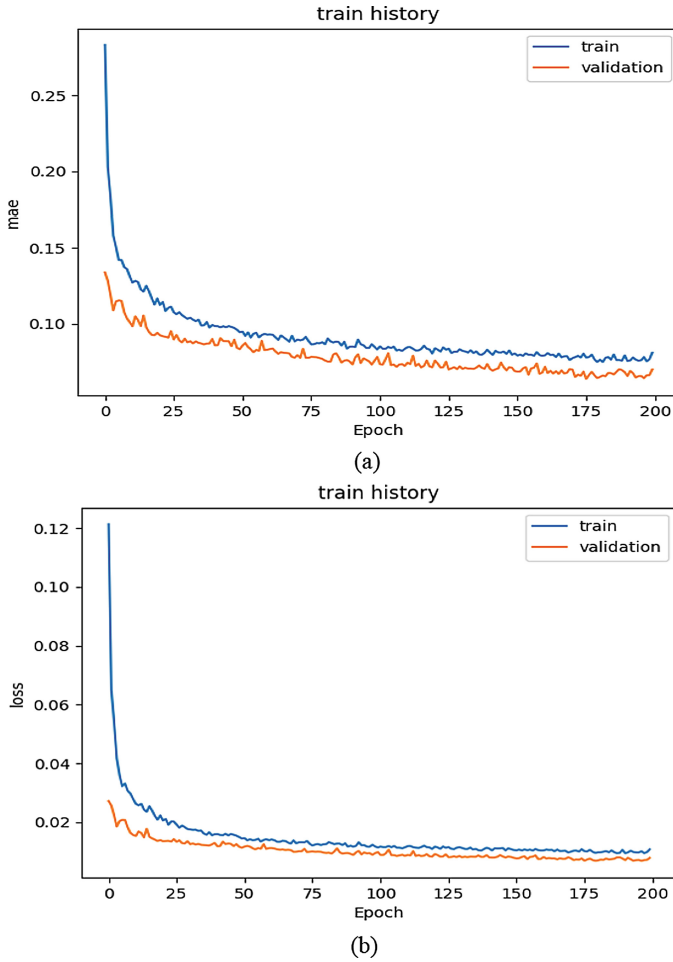


Fig. 6. Model performance under auditory stimuli: (a) Model accuracy under auditory stimuli; (b) Model loss value under the auditory stimulus

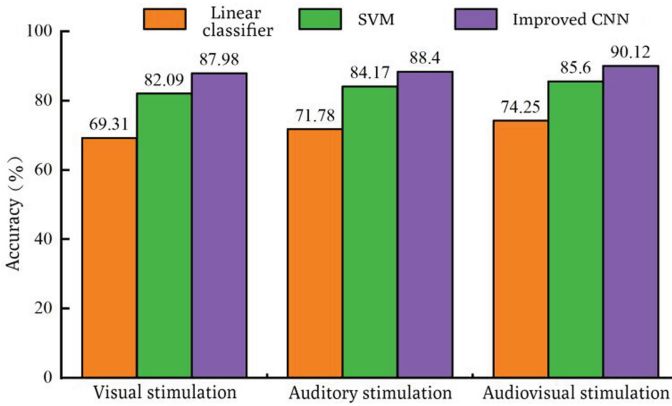
For visual stimulation, the best training effect is achieved at about the 200th time, when the training accuracy of the model is 82.34% and the loss value is 0.0117. For auditory stimulation, the best training effect is achieved at about the 870th time, and the training accuracy rate is 84.78%. The loss value is 0.0072. When the visual and auditory stimuli are mixed, the best training effect is achieved around the 200th time, when the training accuracy is 90.12% and the loss value is 0.0041 (Fig. 6).

By discussing the modes of visual stimulation, auditory stimulation, and audio-visual mixed stimulation, the following is the classification accuracy of the recognition model. The EEG signals denoised by the CEEMDAN-FastICA algorithm are finally used as the input of linear classifier, Support Vector Machine (SVM), and improved CNN. The classification accuracy is as shown in Table 1.

Table 1. Accuracy of EEG classification (%)

Stimulation mode	Linear classifier	SVM	Improved CNN
Visual stimulus	69.31	82.09	87.98
Acoustic stimuli	71.78	84.17	88.40
Audio-visual stimuli	74.25	85.60	90.12

For the feature vectors constructed with scale entropy values under the above three stimulation modes as the input of each classifier, the accuracy visualization results of the classifier are shown in Fig. 8.

**Fig. 7.** Comparison of results of EEG Signal classification by different classifiers

According to the classification accuracy results in the chart, among the linear classifier, the SVM, and the improved CNN, the classification results of the improved CNN are better than those of the other two. Meanwhile, for visual stimuli, auditory stimuli, and mixed audio-visual stimuli, the classification results of mixed audio-visual stimuli are better.

In this experiment we assess the significance of correlations between the first canonical gradient and data from other modalities (curvature, cortical thickness and T1w/T2w image intensity). A normal test of the significance of the correlation cannot be used, because the spatial auto-correlation in EEG data may bias the test statistic. In Fig. 9. we will show three approaches for null hypothesis testing: spin permutations, Moran spectral randomization, and autocorrelation-preserving surrogates based on variogram matching.

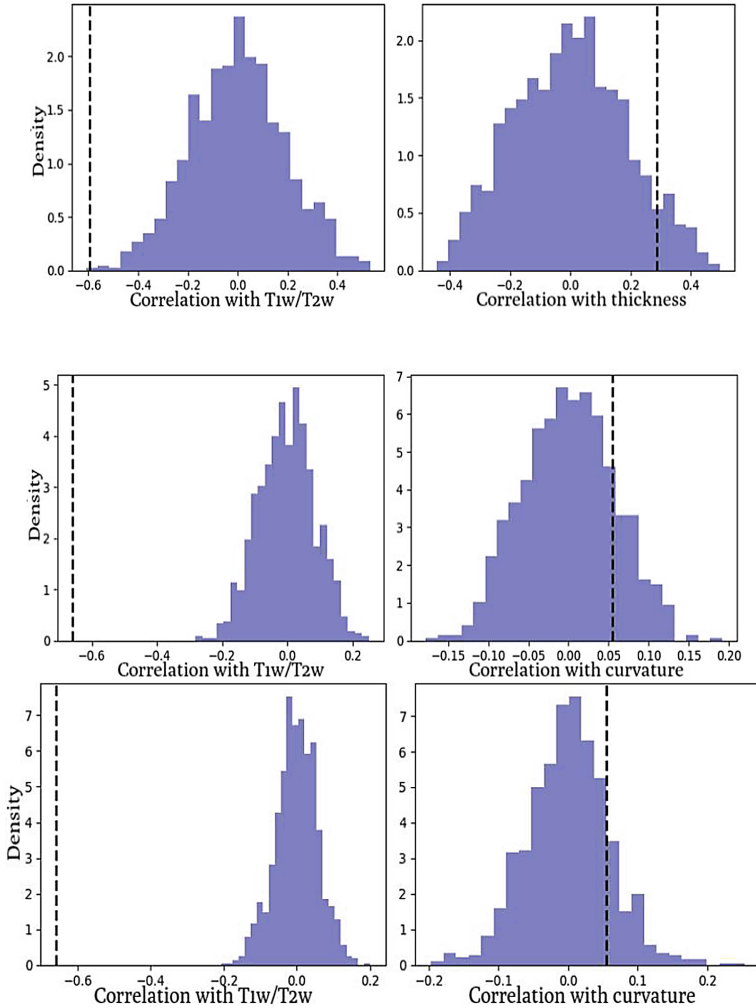


Fig. 8. Visualization of model evaluation metrics

5 Conclusions

EEG signals in different periods under audio-visual stimulation are analyzed by constructing a brain connection network and brain source map visualization, thereby providing a powerful guarantee for the processing of EEG signals of audio-visual stimuli.

Brain source estimation and model of audiovisual evocations, that is, the parallel computing strategy of the convolutional layer is proposed. The convolutional kernel in the convolutional layer is set as a vector to extract only spatial features, and the regularization operation is added to the network structure to prevent over-fitting. The audio-visual model is integrated by using the improved CNN model with the inception network.

Through experiments, the classification effects of EEG signals by traditional classification methods and improved CNN model are compared. The experimental results indicate that the environment safety recognition accuracy of the neuromorphic computing model can respectively reach 87.98%, 88.4%, and 90.12% for the single visual, the single auditory, and the audiovisual stimulus. The results show that the improved CNN model has better recognition accuracy, and the classification results under audio-visual mixed stimulus mode are better than those under single stimulus mode.

Acknowledgements. This work is supported by Yulin City Science and Technology Project under grants CXY-2020–026, The Chinese Society of Academic Degrees and Graduate Education under grant B-2017Y0002–170, and Shaanxi Province Key Research and Development Projects under grants 2016GY-040 and 2021GY-029.

References

1. Zhang, Y., Qu, P., Ji, Y., et al.: A system hierarchy for brain-inspired computing. *Nature* **586**(7829), 378–384 (2020). <https://doi.org/10.1038/s41586-020-2782-y>
2. Kaushik, R., Akhilesh, J., Priyadarshini, P.: Towards spike-based machine intelligence with neuromorphic computing. *Nature* **575**(7784), 607–617 (2019). <https://doi.org/10.1038/s41586-019-1677-2>
3. Kuai, H., et al.: Multi-source brain computing with systematic fusion for smart health. *Information Fusion* **75**, 150–167 (Mar.2021). <https://doi.org/10.1016/j.inffus.2021.03.009>
4. Bhatti, M., Khan, J., Khan, M.U.G., Iqbal, R., Aloqaily, M., Jararweh, Y., Gupta, B.: Soft computing-based EEG classification by optimal feature selection and neural networks. *IEEE Trans. Industrial Informatics* **15**(10), 5747–5754 (2019). <https://doi.org/10.1109/TII.2019.2925624>
5. Lu, Y., Bi, L., Li, H.: Model predictive-based shared control for brain-controlled driving. *IEEE Trans. Intell. Transp. Syst.* **21**(2), 630–640 (Feb.2020)
6. Chakraborty, M., Mitra, D.: A novel automated seizure detection system from EMD-MSPCA denoised EEG: Refined composite multiscale sample, fuzzy and permutation entropies based scheme. *Biomedical Signal Processing and Control* **67**, 102514 (2021). <https://doi.org/10.1016/j.bspc.2021.102514>
7. Gao, Z., Li, Y., Yang, Y., Dong, N., Yang, X., Grebogi, C.: A coincidence filtering-based approach for CNNs in EEG-based recognition. *IEEE Trans. Industr. Inf.* **16**(11), 7159–7167 (Nov.2020). <https://doi.org/10.1109/TII.2019.2955447>
8. Saini, M., Satija, U., Upadhayay, M.D.: Wavelet-based waveform distortion measures for assessment of denoised EEG quality concerning noise-free EEG signal. *IEEE Signal Process. Lett.* **27**, 1260–1264 (Jul.2020)
9. Bhattacharyya, A., Ranta, R., Cam, S.L., et al.: A multi-channel approach for cortical stimulation artifact suppression in-depth EEG signals using time-frequency and spatial filtering. *IEEE Trans. Biomed. Eng.* **66**(7), 1915–1926 (Jul.2019)
10. Teng, T., Bi, L., Liu, Y.: EEG-Based detection of driver emergency braking intention for brain-controlled vehicles. *IEEE Trans. Intelligent Transportation Syst.* **19**(6), 1766–1773 (2018)
11. Tryon, J., Trejos, A.L.: Classification of task weight during dynamic motion using EEG–EMG fusion. *IEEE Sens. J.* **21**(4), 5012–5021 (Feb.2021)
12. Wu, W., Wu, Q.M.J., Sun, W., et al.: A regression method with subnetwork neurons for vigilance estimation using EOG and EEG. *IEEE Trans. Cognitive Developmental Syst.* **13**(1), 209–222 (Mar.2021)

13. Around, A., Mirkovic, B., De Vos, M., Doclo, S.: Impact of different acoustic components on EEG-based auditory attention decoding in noisy and reverberant conditions. *IEEE Trans. Neural Syst. Rehabil. Eng.* **27**(4), 652–663 (Apr.2019)
14. Wang, M., Huang, Z., Li, Y., Dong, L., Pan, H.: Maximum weight multi-modal information fusion algorithm of electroencephalographs and face images for emotion recognition. *Computers and Electrical Eng.* **94**, 107319 (2021). <https://doi.org/10.1016/j.compeleceng.2021.107319>
15. Li, P., Liu, H., Si, Y., Li, C., et al.: EEG based emotion recognition by combining functional connectivity network and local activations. *IEEE Trans. Biomed. Eng.* **66**(10), 2869–2881 (Oct.2019)
16. Cai, J., Wang, Y., Liu, A., McKeown, M.J., Wang, Z.J.: Novel regional activity representation with constrained canonical correlation analysis for brain connectivity network estimation. *IEEE Trans. Med. Imaging* **39**(7), 2363–2373 (Jul.2020)
17. Wang, M., Ma, C., Li, Z., Zhang, S., Li, Y.: Alertness estimation using connection parameters of the brain network. *IEEE Trans. Intelligent Transportation Syst.* (2021) <https://doi.org/10.1109/TITS.2021.3124372>
18. Ting, C.M., Sandin, S.B., Tang, M., Ombao, H.: Detecting dynamic community structure in functional brain networks across individuals: a multilayer approach. *IEEE Trans. Med. Imaging* **40**(2), 468–480 (Feb.2021)
19. Mammone, N., et al.: Brain network analysis of compressively sensed high-density EEG signals in AD and MCI subjects. *IEEE Trans. Industr. Inf.* **15**(1), 527–536 (Feb.2019)
20. Masulli, P., Masulli, F., Rovetta, S., Lintas, A., Villa, A.E.P.: Fuzzy clustering for exploratory analysis of EEG event-related potentials. *IEEE Trans. Fuzzy Syst.* **28**(1), 28–38 (Feb.2020)
21. Li, M., Wang, R., Yang, J., Duan, L.: An improved refined composite multivariate multi-scale fuzzy entropy method for MI-EEG feature extraction. *Computational Intelligence and Neuroscience* (2019). <https://doi.org/10.1155/2019/7529572>
22. Bhattacharyya, A., Tripathy, R.K., Garg, L., Pachori, R.B.: A novel multivariate-multiscale approach for computing EEG spectral and temporal complexity for human emotion recognition. *IEEE Sens. J.* **21**(3), 3579–3591 (Jun.2021)
23. Zhang, G., Etemad, A.: Capsule attention for multimodal EEG-EOG representation learning with application to driver vigilance estimation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **29**, 1138–1149 (Jun.2021)
24. Zhang, G., Cai, B., Zhang, A.Y., Stephen, J., et al.: Estimating dynamic functional brain connectivity with a sparse hidden Markov model. *IEEE Transactions on Medical Imaging* **39**(2), 488–498 (2020)
25. Khalifa, Y., Mandic, D., Sejdić, E.: A review of hidden Markov models and recurrent neural networks for event detection and localization in biomedical signals. *Information Fusion* **69**, 52–72 (May2021)
26. Chakladar, D.D., Kumar, P., Roy, P.P., Dogra, D.P., Scheme, E., Chang, V.: A multimodal-siamese neural network (MSN) for person verification using signatures and EEG. *Information Fusion* **71**, 17–27 (Jul.2021). <https://doi.org/10.1016/j.inffus.2021.01.004>
27. Jafarifarmand, A., Badamchizadeh, M.A., Khanmohammadi, S., Nazari, M.A., Tazehkand, B.M.: A new self-regulated neuro-fuzzy framework for classification of EEG signals in motor imagery BCI. *IEEE Trans. Fuzzy Syst.* **26**(3), 1485–1497 (Jun.2018)
28. Dong, Q., Ge, F., Ning, Q., Zhao, Y., et al.: Modeling hierarchical brain networks via volumetric sparse deep belief network. *IEEE Trans. Biomed. Eng.* **67**(6), 1739–1748 (Jun.2020)
29. Wang, Y., Song, W., Tao, W., et al.: A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion* (2022)
30. Kumar, S., Yadava, M., Roy, P.P.: Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Information Fusion* **52**, 41–52 (2019)

31. Lu, H., Zhang, M., Xu, X.: Deep fuzzy hashing network for efficient image retrieval. *IEEE Trans. Fuzzy Systems* **29**(1), 166176 (2020). <https://doi.org/10.1109/TFUZZ.2020.2984991>
32. Huimin, L., Li, Y., Chen, M., et al.: Brain Intelligence: go beyond artificial intelligence. *Mobile Networks Appl.* **23**, 368–375 (2018)
33. Huimin, L., Li, Y., Shenglin, M., et al.: Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. *IEEE Internet Things J.* **5**(4), 2315–2322 (2018)
34. Huimin, L., Qin, M., Zhang, F., et al.: RSCNN: A CNN-based method to enhance low-light remote-sensing images. *Remote Sensing* **13**(1), 62, 2020
35. Huimin, L., Zhang, Y., Li, Y., et al.: User-oriented virtual mobile network resource management for vehicle communications. *IEEE Trans. Intell. Transp. Syst.* **22**(6), 3521–3532 (2021)