Proceedings of the 3rd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/119/2025.21593

Research on Stroke Rehabilitation Technology Based on Brain-Computer Interface with Deep Learning

Jiale Cao[1,a,*]

[1]Shandong University, Jinan City, Shandong Province, 250100, China
a. 202300161055@mail.sdu.edu.cn
**corresponding author*

Abstract: Stroke has become the leading cause of disability in China and the second leading cause of death globally. Brain-Computer Interface (BCI) technology does not rely on the brain's normal output pathways but instead forms a connection between the brain and external machines based on the collection, output, transformation, and decoding of EEG signals. This provides a new solution for the field of neural rehabilitation. This paper focuses on the application of neural network models, particularly deep learning models (DLMs), in brain signal decoding. In the area of motor intent decoding for BCI, research on various models based on convolutional neural networks (CNNs) has addressed issues such as the accuracy of motor imagery EEG signal decoding. In the field of speech communication BCI decoding, deep learning contributes to the development of BCI spellers, especially in steady-state visual evoked potential (SSVEP) decoding and other speech-related signal decoding. By using various DLMs and optimization strategies, it improves decoding efficiency, accuracy, generalization ability, and robustness.

Keywords: Stroke, Brain-Computer Interface, Brain Signal Decoding, Deep Learning

1. Introduction

The China Cerebrovascular Disease Prevention and Treatment Report indicates that stroke has become the leading cause of disability in China [1]. According to the Global Burden of Disease Study, in 2021, there were approximately 7.3 million deaths due to stroke worldwide, accounting for 10.7% of all deaths. Each year, around 11.9 million new stroke cases are reported globally, making stroke the second leading cause of death, with this number continuing to rise annually [2]. One-third of stroke patients may face death, while another third may suffer permanent disabilities, including motor, speech, and cognitive impairments. Moreover, the irregular lifestyles and dietary habits of modern young people have led to a trend of stroke occurring at younger ages. Undoubtedly, stroke and similar neurological diseases cause significant disruptions to people's lives and impose a tremendous burden on countless families. Therefore, it is urgent to develop technologies through modern science that can help patients regain some motor abilities, improve cognitive function, and even support rehabilitation treatments.

The current development of Brain-Computer Interface (BCI) technology has generated significant expectations for its future. BCI is a novel human-machine information interaction technology that directly establishes a communication and control channel between the brain and external devices, capturing and converting brain signals into electrical signals to achieve information transmission and

DOI: 10.54254/2755-2721/119/2025.21593

control. This technology, especially in the medical neurorehabilitation field, has shown high feasibility. It can help restore and improve functions such as movement, sensation, and cognition, which are impaired due to neurological diseases, thus improving patients' daily lives. The key to implementing effective neurorehabilitation for patients lies in the precise connection between the brain and external devices, which depends on accurate decoding of brain signals. The brain signals used for BCI include electroencephalography (EEG), functional magnetic resonance imaging (fMRI), and functional near-infrared spectroscopy (fNIRS), among others. fMRI detects brain activity by measuring blood oxygen level-dependent (BOLD) signals, while fNIRS reflects brain function activity by measuring changes in the concentrations of oxygenated and deoxygenated hemoglobin in the cortical area of the brain using near-infrared light. However, these methods have limitations. For instance, fMRI has lower resolution and cannot reach the level of neurons or nerve fibers. Additionally, BOLD signals are easily affected by noise, which impacts decoding accuracy. fNIRS is limited by the transmission rate of light through brain tissue, which can affect detection depth and accuracy, and it also contains various physiological noises that require advanced signal processing techniques for separation [3]. EEG, with its millisecond-level time resolution, allows real-time capture of the instantaneous changes in neuronal activity, making it particularly advantageous in studying rapid brain processes such as cognition, perception, and emotion, which undergo dynamic changes within short time intervals. Machine learning, as the core technology for brain signal decoding, can extract valuable information from complex neural signal data, thereby improving decoding accuracy and stability. In particular, ensemble learning methods, by integrating the predictions of multiple models, have significantly enhanced the generalization ability of models [4]. However, when dealing with brain signals, which are unstructured data, machine learning faces challenges such as data complexity, noise interference, and scarcity, which limit the performance of traditional models like decision trees and support vector machines. Moreover, traditional machine learning (TML) primarily explores shallow learning structures, which makes it difficult to effectively extract deep features from brain signal data, thus affecting its performance in brain signal decoding tasks [5].

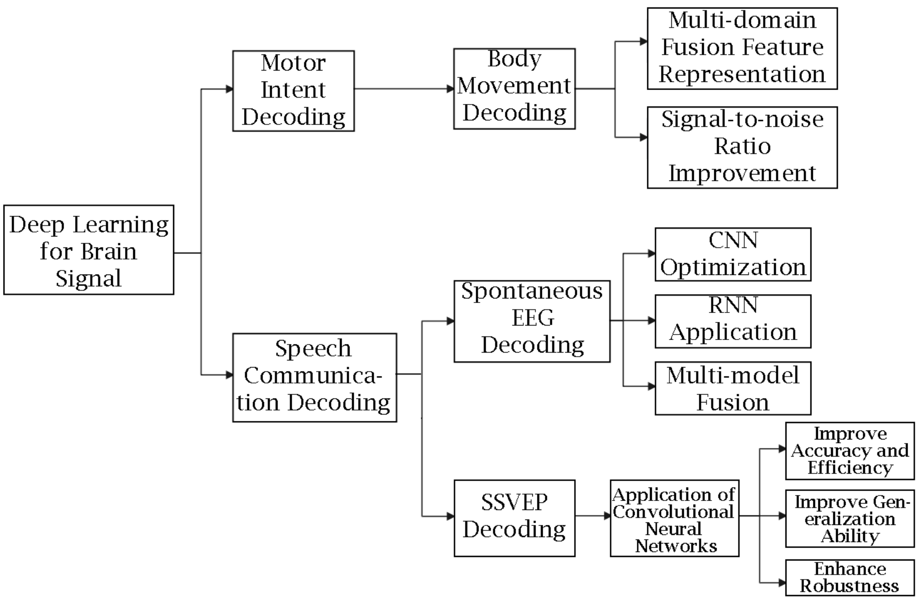

Figure 1: Application of Neural Network Models (NNMs) in Brain Signal Decoding.

This paper primarily introduces the application of NNMs, especially deep learning models (DLMs), in brain signal decoding (as shown in Figure 1). These models automatically learn complex features

DOI: 10.54254/2755-2721/119/2025.21593

and improve decoding accuracy and efficiency through training. To promote the role of BCI in neurorehabilitation for stroke patients, this paper focuses on the critical area of information decoding, specifically addressing the decoding of two types of neural signals: motor intent and speech communication using artificial neural networks [6].

2. Research on Brain-Computer Interface Motor Intent Decoding Algorithms

The Brain-Computer Interface (BCI) based on motor imagery (MI) uses sensorimotor rhythm (SMR) signals, including μ rhythm (frequency of 8–12 Hz) and β rhythm (frequency of 18–26 Hz), as the primary input [7]. Common motor imagery tasks include movements of the left hand, right hand, both feet, and tongue [8]. To decode rehabilitation training motion signals for upper limbs in stroke patients, Ren Zhiyang and colleagues designed a convolutional recurrent neural network (CRNN) model based on convolution, integrating an attention model and reconstructing array electromyographic (EMG) data, achieving accurate identification of relevant movements [9]. Yang Banghua and colleagues studied a binary classification motor imagery task for stroke patients (right-hand grasping and right-elbow swinging). They used convolutional neural networks (CNNs) to extract features from electroencephalography (EEG) signals and applied adaptive transfer learning methods to enhance the model's recognition ability for motor imagery tasks [10]. In research on lower-limb motor EEG signals, Lu Boyang designed three models for classification: a CNN model based on time-frequency images, a CNN model with improved common spatial patterns, and a recurrent CNN model with data augmentation. Experiments showed that these CNN models exhibited high accuracy in lower-limb motor imagery classification tasks [11]. In the future, multi-task learning, which expands to more complex body movements, will be a trend.

In BCIs, the accuracy of motor imagery EEG signal decoding is generally influenced by factors such as low signal-to-noise ratio (SNR) and individual differences. To address this issue, Shen Jianing proposed an improved method based on 3D-CNN, constructing a spatiotemporal filter set to perform multi-domain fusion feature representation on EEG signals, effectively improving the decoding accuracy of motor imagery EEG signals [12]. The SNR of EEG signals significantly affects classification accuracy. Song Shilin proposed a feature extraction algorithm that combines sample entropy with common spatial patterns (CSP), using wavelet packet decomposition and reconstruction to remove noise and improve SNR, followed by classification using a CNN [13]. The team led by Wu Xuejian first pre-processed the data and transformed it into a 3D spatial-frequency feature map dataset. They then built a specific 3D-CNN network and optimized the training. In addition to improving SNR, their approach also addressed the issue that traditional neural networks do not fully utilize frequency domain and spatial domain information, leading to low decoding recognition rates [14]. A similar solution was proposed by Hu Cunlin, who introduced a hybrid neural network pattern recognition method combining CNN and RNN (CNN-LSTM and CNN-GRU models). CNN was used to extract spatial features from EEG signals, while LSTM or GRU were employed to capture temporal features [15]. To overcome the limitations of traditional methods in feature extraction, where signals are prone to loss, Yan Heng proposed an EEG-fNIRS multimodal signal fusion model based on DGCNN and TCN. First, the EEG and fNIRS data were convoluted to extract features. After processing the EEG signals with phase-locked values, DGCNN and TCN were used to capture channel correlations and temporal features, and the features were then fused and classified using fully connected layers [16].

3. Research on Speech Communication Brain-Computer Interface Decoding Algorithms

Speech communication is a fundamental means of human interaction. BCI technology, by decoding brain signals, transforms brain activity into understandable language, offering new hope for individuals who have lost their ability to speak due to neurological diseases. In this field, the

DOI: 10.54254/2755-2721/119/2025.21593

development of BCI spelling devices has become a key focus of research. The application of deep learning technology has played an important role in improving the efficiency and accuracy of BCI spelling devices.

Deep learning can be applied to spontaneous electroencephalography (EEG) decoding and shows certain commonalities with natural language processing methods [17]. Specifically, DLMs such as CNNs, Recurrent Neural Networks, Long Short-Term Memory Networks, and Transformer can efficiently decode and recognize such EEG signals [18, 19].

Based on CNNs, the Cooney team optimized the CNN architecture and fine-tuned the input-output layers [20]. Tragoudaras adopted a data-driven offline optimization strategy and introduced transfer learning methods [21]. Xu-Geng designed a multi-input time-frequency-spatial hybrid CNN model to increase the EEG classification recognition rate [22]. RNNs, due to their recursive structure, are particularly suited for processing sequential data and can capture the temporal dependencies of EEG signals. Übeyli's research showed that power spectral density (PSD) estimates obtained via feature vector methods can well represent EEG signals, and RNNs trained on these features achieved high classification accuracy, validating the effectiveness of combining feature vector methods with RNNs for EEG signal classification [23]. Zhang Hang combined the method of partitioning filter banks and proposed a multi-scale feature fusion CNN-LSTM network (MFFCL), which is an effective end-to-end EEG decoding model. It demonstrates high classification accuracy while reducing parameter usage and computational resources, showcasing high practicality [24]. Song Y, Zheng Q, Liu B, and others proposed the EEG Conformer model by integrating CNN with Transformer architecture. This approach significantly improved the performance of EEG signal extraction for both local and global features and reached an advanced level in tasks like motor imagery and emotion recognition. It also enhanced the model's interpretability through a visualization strategy [25]. Moreover, Young-Eun Lee and colleagues utilized the self-attention mechanism of the Transformer architecture to improve the decoding technique of imagined speech EEG signals, enhancing decoding accuracy, reducing model complexity, and exploring the feasibility of decoding imagined speech using single-channel and ear EEG for practical BCI applications [26].

Steady-state visual evoked potentials (SSVEP) have emerged as a predominant modality in brain-computer interface (BCI) research, surpassing spontaneously generated neural signals in practical applications. This neurophysiological phenomenon manifests when subjects maintain visual fixation on a rhythmic stimulus with specific flickering frequencies. Characterized by its distinctive spectral profile, SSVEP exhibits prominent peaks at both the fundamental frequency of stimulation and its harmonic components within the power spectrum, accompanied by consistent phase-locked responses in temporal analysis. The primary neural generators of this signal are distributed across the occipitofrontal cortical network.

SSVEP's ascendancy in BCI implementations stems from several inherent advantages: 1) Enhanced signal detectability due to superior signal-to-noise ratios (SNR $\geq$ 5 dB in typical paradigms); 2) Rapid information transfer capabilities (achieving ITR > 50 bits/min in optimized systems); 3) Minimal user acclimatization requirements; and 4) Robust inter-session reproducibility (ICC > 0.85 in longitudinal studies). These technical advantages establish SSVEP as a highly suitable choice for real-world BCI implementations that require robust and efficient interaction systems, such as neuroprosthetic devices, virtual reality navigation tools, and advanced communication interfaces. Such attributes underscore SSVEP's potential as a reliable signal for BCI applications, particularly in contexts that prioritize rapid, user-friendly, and consistent performance [27].

Among research on decoding SSVEP signals, CNNs have the most widespread application. N. K. N. Aznan and colleagues explored the application of CNN in decoding dry EEG signals based on SSVEP, verifying its efficiency and accuracy [28]. To address the performance bottleneck of TML methods in extracting and classifying SSVEP features, Li Chang trained a sub-band CNN (sbCNN)

DOI: 10.54254/2755-2721/119/2025.21593

model and applied transfer learning to achieve effective decoding of SSVEP signals, thus improving the decoding accuracy of SSVEP-BCI systems [29]. Zhu Y and others used the EEGNet model combined with ensemble learning methods [30]. Han Dan proposed a time-domain-based generalized filter bank convolution network (FBCNN-G) using generalized task-related component analysis (GTRCA) [31], while Oralhan Z, Oralhan B, Khayyat M M, and others employed a 3D-input CNN [32]. Both approaches significantly improved the efficiency and accuracy of signal decoding, providing new technical avenues for the design and optimization of brain-computer interface systems.

To achieve higher classification accuracy, Zhao X and colleagues proposed an augmented reality steady-state visual evoked potential (AR-SSVEP) multi-target rapid classification method based on CNN [33]. Du Yuxin combined a convolutional correlation analysis model with transfer learning and used transfer learning to enable cross-subject migration, thus improving the classification accuracy of SSVEP signals in the target domain and expanding the data domain to address the problem of insufficient training datasets [34].

To tackle the issues of high model complexity, large computational resource consumption, and insufficient generalization ability in EEG signal classification tasks, Lawhern V J and colleagues designed a compact CNN architecture to reduce the number of parameters, thereby lowering the risk of overfitting. They also employed regularization techniques and optimization algorithms, ultimately verifying the generalization ability of the EEGNet model on multiple BCI paradigms with limited training data [35]. On this basis, Yao H proposed the FB-EEGNet model, which fused EEG signals under multiple stimulus conditions and further optimized the CNN architecture and algorithms. This enhanced the model's generalization ability and reduced the complexity and computational resource consumption of EEG signal classification tasks [36].

Differences in EEG signals between individuals, along with the ability to maintain stable decoding performance under noise and artifact interference, are crucial to improving the reliability of the system. To address this, Cong Peichao designed the PLFC-DNN model, which deeply explores and utilizes phase and frequency features of SSVEP signals from both global and local perspectives, significantly improving the robustness of the algorithm [37]. Li Chang proposed a method that combines TRCA with sbCNN (TRCA-CNN), which helps reduce the overfitting risk of single models. By integrating the outputs of different models, it reduces performance fluctuations on different samples, thus enhancing the overall robustness of decoding.

4. Conclusion

This paper reviews the current status of brain-computer interface (BCI) research based on deep learning, focusing on decoding motor intent and speech communication signals from the brain using deep learning methods. It analyzes the solutions to existing problems at this stage. In decoding motor intent, convolutional neural networks (CNNs) are mainly relied upon, and improvements and technology fusion have been explored based on this framework. Solutions to the common problems of motor imagery EEG signals, including those for upper and lower limbs, have been proposed with different models. To improve the decoding accuracy of motor imagery EEG signals, researchers have focused on enhancing signal-to-noise ratio (SNR) and multi-domain fusion features. In speech communication research, this paper first introduces various neural networks for decoding spontaneous EEG and then emphasizes various DLMs and optimization strategies. It highlights the significant advantages in SSVEP signal decoding from the perspectives of efficiency, accuracy, generalization performance, and robustness.

References

[1] Report on Stroke Prevention and Treatment in China Writing Group. (2023). Summary of the 2021 China Stroke Prevention and Treatment Report. Chinese Journal of Cerebrovascular Diseases, 20(11), 783–793.

[2] GBD 2019 Stroke Collaborators. (2021). Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. The Lancet Neurology, 20(10), 795.

[3] Hu, H. B. (2010). Research and application of functional near-infrared spectroscopy imaging (Master's thesis). University of Science and Technology of China.

[4] Chen, K., & Zhu, Y. (2007). A review of machine learning and its related algorithms. Statistics and Information Forum, 2007(5), 105–112.

[5] Yang, J. F., Qiao, P. R., Li, Y. M., et al. (2019). A review of machine learning classification problems and algorithms. Statistics and Decision, 35(6), 36–40. https://doi.org/10.13546/j.cnki.tjyjc.2019.06.008

[6] Chaudhary, U., Birbaumer, N., & Ramos-Murguialday, A. (2016). Brain–computer interfaces for communication and rehabilitation. Nature Reviews Neurology, 12(9), 513–525.

[7] Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. Clinical Neurophysiology, 110(11), 1842–1857.

[8] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. Clinical Neurophysiology, 113(6), 767–791.

[9] Ren, Z. Y. (2021). Research on upper limb rehabilitation training motion recognition for stroke based on convolutional neural networks (Master's thesis). University of Electronic Science and Technology of China. https://doi.org/10.27005/d.cnki.gdzku.2021.003493

[10] Yang, B. H., Zhang, Y. H. (2024). Feasibility study on improving the classification accuracy of stroke patients' binary motor imagery tasks based on convolutional neural networks. Shanghai Medical Journal, 47(4), 253–258. https://doi.org/10.19842/j.cnki.issn.0253-9934.2024.04.010

[11] Lu, B. Y. (2022). Research on lower limb motion decoding based on convolutional neural networks (Master's thesis). Southeast University. https://doi.org/10.27014/d.cnki.gdnau.2022.000303

[12] Shen, J. N. (2024). Research on motor imagery decoding based on 3D convolutional neural networks (Master's thesis). Xi'an University of Technology. https://doi.org/10.27391/d.cnki.gxagu.2024.000110

[13] Song, S. L. (2023). Classification algorithm of motor imagery EEG features based on deep learning (Master's thesis). Nanjing University of Posts and Telecommunications. https://doi.org/10.27251/d.cnki.gnjdc.2023.001142

[14] Wu, X. J., Chu, Y. Q., Zhao, X. G., et al. (2024). Motor imagery EEG decoding method based on space-frequency feature map learning of 3D convolutional neural networks. Journal of Biomedical Engineering, 41(6), 1145–1152.

[15] Hu, C. L., Ye, Y. (2024). Motor imagery EEG signal pattern recognition based on convolutional recurrent neural networks. Journal of Luoyang Institute of Technology (Natural Science Edition), 34(1), 50–55.

[16] Yan, H., Zhou, Z. K., He, X. S., et al. (2025). Dynamic graph-time convolutional neural network EEG-fNIRS multimodal motor imagery/execution decoding. Control Theory & Applications, 1-9.

[17] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30(2017).

[18] Zheng, Y. B. (2024). Research on natural language brain signal processing and decoding based on deep learning (Master's thesis). Beijing University of Posts and Telecommunications. https://doi.org/10.26969/d.cnki.gbydu.2024.000331

[19] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering, 16(5), 051001.

[20] Cooney, C., Folli, R., & Coyle, D. (2019, October). Optimizing layers improves CNN generalization and transfer learning for imagined speech decoding from EEG. In 2019 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1311–1316). IEEE.

[21] Tragoudaras, A., Fanaras, K., Antoniadis, C., & Massoud, Y. (2023, May). Data-driven offline optimization of deep CNN models for EEG and ECoG decoding. In 2023 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1–5). IEEE.

[22] Xu, G. Y. (2024). Research on motor imagery EEG signal classification and recognition in brain-computer interfaces (Master's thesis). Xi'an University of Technology. https://doi.org/10.27391/d.cnki.gxagu.2024.000825

[23] Übeyli, E. D. (2009). Analysis of EEG signals by implementing eigenvector methods/recurrent neural networks. Digital Signal Processing, 19(1), 134–143.

[24] Zhang, H. (2024). Multi-class EEG signal recognition based on multi-scale CNN-LSTM networks (Master's thesis). Shandong Normal University. https://doi.org/10.27280/d.cnki.gsdsu.2024.000238

[25] Song, Y., Zheng, Q., Liu, B., & Gao, X. (2022). EEG conformer: Convolutional transformer for EEG decoding and visualization. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31, 710–719.

[26] Lee, Y. E., & Lee, S. H. (2022, February). EEG-transformer: Self-attention from transformer architecture for decoding EEG of imagined speech. In 2022 10th International Winter Conference on Brain-Computer Interface (BCI) (pp. 1–4). IEEE.

[27] Min, B. K., Dähne, S., Ahn, M. H., Noh, Y. K., & Müller, K. R. (2016). Decoding of top-down cognitive processing for SSVEP-controlled BMI. Scientific Reports, 6(1), 36267.

DOI: 10.54254/2755-2721/119/2025.21593

[28] Aznan, N. K. N., Bonner, S., Connolly, J., Al Moubayed, N., & Breckon, T. (2018, October). On the classification of SSVEP-based dry-EEG signals via convolutional neural networks. In 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3726–3731). IEEE.

[29] Li, C. (2024). Visual brain-computer interface decoding method research based on convolutional neural networks and task-related analysis (Master's thesis). Nanchang University. https://doi.org/10.27232/d.cnki.gnchu.2024.001361

[30] Zhu, Y., Li, Y., Lu, J., & Li, P. (2021). EEGNet with ensemble learning to improve the cross-session classification of SSVEP-based BCI from ear-EEG. IEEE Access, 9, 15295–15303.

[31] Han, D. (2023). SSVEP-BCI research based on correlation analysis and deep learning (Master's thesis). Taiyuan University of Technology. https://doi.org/10.27352/d.cnki.gylgu.2023.001114

[32] Oralhan, Z., Oralhan, B., Khayyat, M. M., Abdel-Khalek, S., & Mansour, R. F. (2022). [Retracted] 3D Input Convolutional Neural Network for SSVEP Classification in Design of Brain Computer Interface for Patient User. Computational and Mathematical Methods in Medicine, 2022(1), 8452002.

[33] Zhao, X., Du, Y., & Zhang, R. (2022). A CNN-based multi-target fast classification method for AR-SSVEP. Computers in Biology and Medicine, 141, 105042.

[34] Du, Y. X. (2022). Research on the convolutional correlation analysis model for the SSVEP brain–computer interface system (Unpublished master's thesis). Harbin Institute of Technology. https://doi.org/10.27061/d.cnki.ghgdu.2022.002472.

[35] Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. Journal of Neural Engineering, 15(5), 056013.

[36] Yao, H., Liu, K., Deng, X., Tang, X., & Yu, H. (2022). FB-EEGNet: A fusion neural network across multi-stimulus for SSVEP target detection. Journal of Neuroscience Methods, 379, 109674.

[37] Cong, P. C., Chen, X. L., Xiao, Y. X., et al. (2024). Research on the classification algorithm based on the phase-frequency characteristics of SSVEP signals. Electronic Measurement Technology, 47(5), 188–198. https://doi.org/10.19651/j.cnki.emt.2315255.