Application and Development of Embedded Brain-Computer Interface and Artificial Intelligence Deep Learning

Siyuan Yu[1] *

[1] Chaminade high school, Los Angeles, CA, 91324, USA

* Samyusiyuan@gmail.com

Abstract. With the development of artificial intelligence and wearable devices, brain-computer interface (BCI) has gradually shifted from basic research to practical applications. The data collected by traditional embedded brain-computer interfaces (BCIs) are often highly complex and uncertain, posing significant challenges for effective signal processing and interpretation. Moreover, the presence of large amounts of noise further obscures meaningful brainwave patterns, making it extremely difficult to isolate relevant signals. This is comparable to searching for a spotted mosquito deep within a dense forest—where the signal of interest is easily buried under layers of irrelevant or misleading data. In such a context, the ability to rapidly and accurately detect and extract key brainwave signals from massive, noisy datasets has become one of the most critical technical challenges in the current development of BCIs. Overcoming this issue is essential for advancing real-time neural decoding and improving the overall reliability and efficiency of brain-computer communication systems, laying the foundation for future applications in both clinical neuroscience and human-computer interaction.

Keywords: Brain-Computer Interface, Artificial Intelligence, Deep Learning

1 Introduction

With the rapid development of artificial intelligence technology and the popularity of wearable devices, brain-computer interface (BCI) has gradually shifted from basic neuroscience research to practical application scenarios, such as medical rehabilitation, auxiliary communication, emotion monitoring and even human-computer collaboration enhancement. The core task of BCI is to realize the direct information transmission between human intention and external devices through the acquisition and decoding of brain signals (such as electroencephalogram (EEG) or cortical neural signals). The data collected by traditional embedded brain-computer interfaces are complex and uncertain. At the same time, a large number of noise waves make the really important waves hidden deep, just like finding a spotted mosquito in the woods. Therefore, being able to quickly and accurately find important brain waves has become the biggest technical challenge at present.

A. J. Moshayedi (ed.), *Proceedings of the 2025 2nd International Conference on Electrical Engineering and Intelligent Control (EEIC 2025)*, Advances in Engineering Research 279,
https://doi.org/10.2991/978-94-6463-864-6_86

With the development of AI, the AI era, and the relationship between brain-computer interfaces and AI is getting closer and closer. Embedded BCI systems integrate acquisition, preprocessing, decoding and feedback modules into a small and low-power hardware platform, enabling BCI to be applied in actual environments such as wearable devices, auxiliary robots, and neural prostheses, thereby promoting "BCI from the laboratory to life" [1]. For example, Casson et al. pointed out that wearable EEG devices are gradually evolving towards miniaturization, high frequency, multi-channel and AI compatibility, making EEG-based portable neural interface systems possible [2].

At the same time, deep learning technology, especially convolutional neural networks (CNN), recurrent neural networks (RNN), graph neural networks (GNN), etc., have shown significant advantages in BCI tasks such as brain signal classification, motor imagery recognition, and emotion recognition. Their capabilities in feature extraction, nonlinear modeling, and end-to-end learning provide key support for the intelligent and automated development of BCI systems.

This review will focus on the integration of embedded BCI systems and deep learning AI, analyze system architecture and key technology trends (such as low-power deployment, multimodal fusion, adaptive decoding, etc.), and evaluate its current development level, core challenges, and prospects. By synthesizing cutting-edge technology papers, it provides ideas and directions for academic professionals.

2 Embedded Brain-Computer Interface System

As brain-computer interface technology moves from the laboratory to clinical and consumer applications, its system form is also changing rapidly. Traditional BCI systems usually rely on high-end experimental platforms, such as multi-channel EEG collectors and large PC computing clusters. Although they have high accuracy, they have high deployment costs, poor portability, and large response delays. In order to meet the needs of mobility, real-time performance, and energy consumption control, researchers have begun to explore the design and optimization of embedded brain-computer interface (Embedded BCI) systems.

2.1 System architecture overview

The basic architecture of embedded BCI includes four key modules: brain signal collector, data preprocessing and feature extraction module, decoding algorithm processing unit, and output feedback interface. Unlike traditional BCI systems, these modules are usually integrated on one or more small hardware platforms as show in Fig.1, which have stronger real-time and mobility.

He et al. pointed out in their review that the research focus of embedded BCI systems is shifting from functional integrity to efficient real-time and low-power integration, while emphasizing the trend of hardware-algorithm co-design [3]. For example, small processors based on embedded ARM or RISC architecture can complete EEG signal preprocessing, feature transformation, and even preliminary classification tasks, providing rapid support for subsequent intelligent feedback links.

Fig. 1. Head-mounted brain-computer interface [1]

2.2 Neural Acquisition Chips and SoC Integration

Similar systems on a chip (SoC) are being used in fields such as high-density signal channels, multimodal sensor fusion, and wireless transmission, and CMOS low-power processes are generally used to further ensure system reliability and wearable performance.

2.3 Wireless Systems and Low-Latency Communication

Traditional BCI systems often require cables to connect the EEG cap to the host, which limits the user's range of activities. The wireless BCI platform proposed by Zhou et al. integrates Bluetooth/LoRa communication modules on the chip side, which can achieve cortical motion control feedback with a delay of less than 30ms. Its application scenarios have been expanded to brain-controlled drones and brain-controlled wheelchairs [4].

This type of system not only breaks through the limitations of physical connections, but also poses new challenges, such as signal compression, encryption, and transmission robustness, which need to be optimized at the system level, as shown in Fig.2.

Fig. 2. Brain-computer interface sensors [4]

2.4 Open source platform and low-cost

Although high-precision neural chips and customized SoCs are very advanced, they are expensive and not conducive to academic research and early prototype testing. To this end, Kumar et al. proposed an embedded solution based on the combination of OpenBCI hardware and Raspberry Pi or Jetson Nano, which realizes real-time acquisition and edge AI processing of EEG signals. It is low-cost and highly modular, suitable for scientific research, teaching, and start-up prototypes [5].

These open source solutions greatly reduce the development threshold of BCI systems, and at the same time provide a training ground and actual testing environment for the deployment of deep learning algorithms on edge devices.

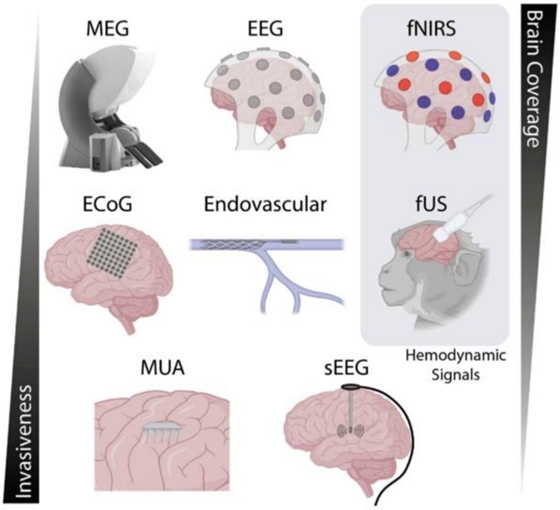

Fig. 3. Brain nerves and sensors [5]

In summary, embedded BCI systems are no longer auxiliary modules in the laboratory but have become the "nerve center" in the future BCI system structure as show in Fig.3. From high-performance neural acquisition chips to low-latency wireless communications, to open-source hardware combinations, embedded architectures are redefining the scalability, real-time performance, and wearability of brain-computer interface systems. The next chapter will discuss the neural decoding algorithm and explore how deep learning plays a key role in brain signal recognition.

3 Deep Learning Model and Brain Signal Decoding Technology

Brain signals are a type of highly complex, nonlinear, and time-varying bioelectric data. Traditional linear classifiers and manual feature extraction methods often have difficulty coping with the challenges of low signal-to-noise ratio, large individual differences, and high real-time requirements in actual tasks. In recent years, the introduction of deep learning models has completely changed the technical path of BCI signal decoding. Through end-to-end automatic feature extraction and nonlinear modeling capabilities, deep networks have greatly improved the accuracy and robustness of brain signal decoding.

3.1 Convolutional Neural Network (CNN)

Although EEG brain wave signals are one-dimensional time series, they can be constructed into a two-dimensional "spatial map" in a multi-channel system, which is very suitable for CNN modeling. The EEGNet network proposed by Lawhern et al. adopts a lightweight deep convolutional structure, combined with batch normalization and Dropout mechanism, which not only improves the classification accuracy, but also greatly reduces the number of parameters, making it suitable for deployment on edge computing devices [6].

Schirrmeister et al. further compared the performance of CNN and traditional filtering-feature engineering-classification pipeline in motor imagery tasks. The results showed that CNN can not only learn features end-to-end, but also has stronger generalization ability [7]. This structure has shown excellent performance in tasks such as target recognition, emotion discrimination, and attention monitoring.

3.2 Recurrent Neural Network (RNN) and long short-term memory network (LSTM)

Due to the strong dynamic characteristics of brain signals, long short-term memory network (LSTM) has become an effective tool for modeling the temporal structure of brain signals. RNN-type networks can capture the complex temporal relationship between the user's continuous thinking state and external stimuli through hidden state transmission.

The ChronoNet model designed by Roy et al. combines temporal convolution and recurrent units, which not only retains the local perception ability of CNN, but also has the long-term memory ability of LSTM, and exceeds the performance of traditional RNN in multiple brain signal classification tasks [8]. ChronoNet is particularly suitable for high sampling rate and high dimensional EEG signal processing.

3.3 Graph Neural Network (GNN)

Unlike images and speech, brain signals have significant inter-channel dependencies and non-Euclidean structures (such as prefrontal-occipital signal synergy). For this reason, GNN was introduced into the BCI field to model inter-channel topological graph structures.

The GNN structure proposed by Zhou et al. based on the emotion recognition task achieved significant improvement in EEG emotion classification by constructing a "brain function map" and extracting the correlation pattern between different brain regions through graph convolution [9]. This model architecture is particularly suitable for tasks such as emotion decoding and cognitive load assessment.

3.4 Transformer

The Transformer structure is gradually entering the field of EEG signal processing due to its excellent sequence modeling ability. Unlike RNN, Transformer can achieve modeling of long-range dependent signals using the self-attention mechanism, and is theoretically more suitable for processing long-term windows and multi-segment brain signal streams.

Although the application of Transformer in BCI is still in its infancy, studies have shown that it is superior to traditional recurrent networks in tasks such as attention state detection and language thinking prediction. In the future, with the emergence of lightweight Transformer structures, its potential in embedded BCI will be further released.

4 Trend of Low-power deployment and edge computing

The real-time, portability and user comfort of brain-computer interface systems determine that their core computing modules must have the characteristics of low latency, low power consumption and high reliability. Although deep learning models, especially CNN and LSTM in signal decoding, have excellent performance, their large amount of computation and resource consumption directly limit their deployment on traditional embedded platforms. Therefore, running deep models efficiently on edge devices is one of the key bottlenecks in the implementation of current BCI system engineering.

4.1 Compression and sparsification of deep neural networks

The EIE (Efficient Inference Engine) deep inference architecture proposed by Han et al. greatly compresses the model size by sparsifying, weight sharing and quantizing the neural network, achieving a compression rate of more than 10 times without sacrificing accuracy [10], as show in Fig.4. When EIE deploys convolutional neural networks in embedded systems, it can significantly reduce memory access frequency and energy consumption, allowing models originally running on GPUs to be executed on low-power platforms such as FPGAs or MCUs.

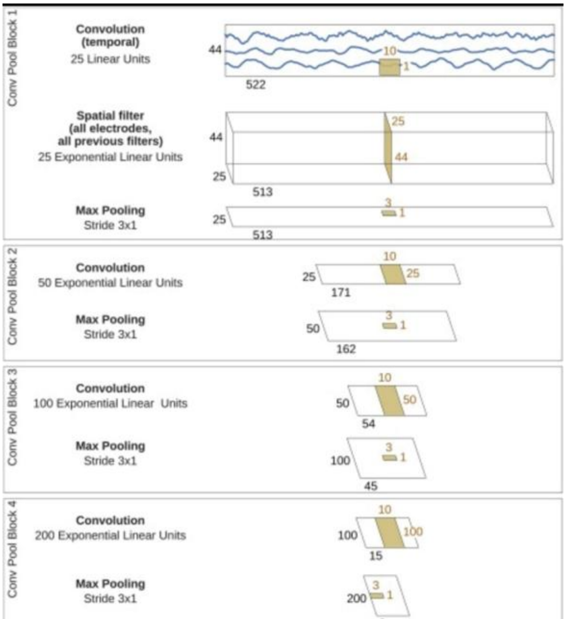

Fig. 4. Analyzing brain waves [10]

This type of compression technology is also applicable to BCI tasks, especially for structures such as EEGNet, which have relatively few parameters. After further sparse optimization, they can be inferred in real time on a milliwatt-level platform, greatly expanding the application scope of wearable brain-controlled devices.

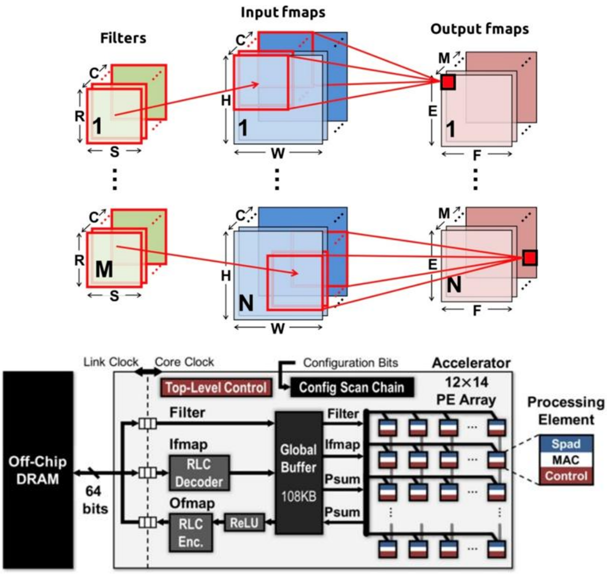

Fig. 5. Brainwave Algorithm [11]

4.2 Federated Learning: Individualized Model Training under Privacy Protection

Brain signals have significant individual differences. EEG waveforms between different users have large deviations in frequency, amplitude, and response delay, which makes the "general model" ineffective in practical applications. Li et al. proposed applying the federated learning mechanism to BCI model training, allowing each user to update the model locally, and the server only aggregates parameters without collecting raw data, thereby achieving individualized model optimization under the premise of ensuring privacy [11] as show in Fig.5.

This method has shown significant advantages in multiple BCI emotion recognition and motor imagery classification tasks, while effectively reducing network bandwidth consumption and security risks. Especially in medical BCI scenarios (such as Parkinson's disease and ALS patient intention recognition), federated learning provides a realistic and feasible path for modeling sensitive data.

4.3 Real-time and energy-efficient system optimization direction

In general, the realization of BCI edge deployment needs to start from three aspects:

- Model design side: Introduce lightweight structure, prunable network, neural structure search (NAS);
- Hardware design side: Use FPGA/ASIC for structure mapping and dynamic scheduling;
- System collaboration side: Build a heterogeneous computing architecture to achieve signal acquisition, decoding, and feedback integration.

These optimization solutions not only improve the system response speed, but also greatly reduce user fatigue and cognitive deviation caused by delays, providing a guarantee for high-quality human-computer interaction.

5 Cross-modal fusion and adaptive system

Although single-modal EEG signals have been widely used in brain-computer interfaces, their stability, noise resistance and information expression density are still limited. Especially in natural environments, EEG signals are easily affected by muscle movement, eye movement, and external electromagnetic interference, resulting in signal distortion or distortion. Therefore, integrating EEG signals with other physiological or behavioral data to build a multimodal BCI system has become a research hotspot in recent years. At the same time, considering the significant differences in individual brain signals, researchers are also paying more and more attention to the adaptive ability of BCI systems, hoping that it can continuously "learn" and adjust according to the user's state and habits.

5.1 Cross-Modal Fusion: Enhancing Signal Robustness and Recognition Accuracy

Xu et al. designed a hybrid brain-computer interface system based on the fusion of EEG and EOG (electro-oculogram) signals, combining the cognitive state information of EEG signals with the behavioral instructions of eye movement signals, which significantly improved the system's intention recognition accuracy and false trigger suppression ability [12]. The system achieved an accuracy rate of more than 90% in tasks such as text input and wheelchair navigation, and effectively reduced the risk of misrecognition caused by single channel fluctuations.

Similarly, researchers have also tried to fuse electromyography (EMG) and electrocardiogram (ECG) signals to build a cross-brain-body collaborative interface system. These solutions show that multi-source signal fusion can characterize the user's true intention state from multiple dimensions and enhance the stability and practicality of the BCI system.

5.2 Deep Autoencoder and Signal Reconstruction

In addition to fusing different modalities, there are also studies that use deep autoencoders to reduce noise and compress EEG signals. The autoencoder compression system proposed by Yin et al. can compress EEG data to 20% of the original while maintaining signal discrimination information, significantly improving wireless transmission efficiency and enhancing the system's robustness to electromagnetic interference [13].

This method is particularly suitable for wireless BCI systems and has significant value in scenarios such as cross-device communication and remote medical rehabilitation. By combining deep learning with knowledge of neural signal structure, the autoencoder not only improves signal quality, but also provides a "cleaner" input for subsequent decoding models.

5.3 Meta-learning and individual adaptive modeling

In BCI applications, the brain structures of different users are different.

The differences in structure, emotional response and usage habits between users lead to poor adaptability of the unified model. Therefore, how to make BCI "learn like a human" becomes a key issue. Park et al. introduced a meta-learning mechanism to train an initial model that can quickly adapt to new users with a small number of samples, so that the BCI system can still maintain high accuracy and fast tuning capabilities during "cold start" [14].

This method is based on the Model-Agnostic Meta-Learning (MAML) framework and can significantly improve the accuracy within 1-5 iterations, avoiding the problem of requiring a large number of samples and time in traditional individualized training. Meta-learning also enables the BCI system to have "transfer learning" capabilities, showing good flexibility in the same user's cross-task migration and cross-device migration.

5.4 Adaptive feedback mechanism and online update

In addition to the individual adaptation of the decoding model, the BCI system also begins to have an adaptive feedback mechanism, such as:

- Adjust the stimulus intensity based on the attention state;
- Adjust the task rhythm based on the cognitive load;
- Use online learning to update the classifier parameters to achieve "learning while using".

This type of mechanism enables BCI to gradually evolve from a "tool" to an "intelligent partner", truly realizing the closed-loop collaboration of brain-computer-environment.

6 Discussion and Outlook

With the rapid development of artificial intelligence and the continuous advancement of hardware platforms, brain-computer interfaces (BCIs) are becoming more practical and popular at an unprecedented speed. This paper systematically reviews the latest progress in embedded BCI systems and deep learning AI in structural design, algorithm implementation, low-power optimization, multimodal fusion and adaptive learning, and combines representative research results to show the multi-dimensional innovation path in this field. However, although embedded brain-computer interfaces have achieved breakthroughs in many directions, their large-scale implementation still faces a series of challenges and opportunities.

6.1 Main technical bottlenecks

First, signal stability and individual differences are still long-term problems in the field of BCI. Non-invasive EEG signals are easily affected by myoelectric and environmental interference, resulting in unstable model predictions; at the same time, the EEG patterns of different individuals vary greatly, making cross-population model migration difficult.

Secondly, model adaptation and energy efficiency scheduling of resource-constrained devices still need to be further optimized. Although model compression and edge accelerators have achieved initial results, existing embedded platforms still have difficulty supporting the real-time operation of complex deep models (such as Transformer and GNN) on high-channel, high-sampling EEG data, limiting their expansion in clinical rehabilitation and consumer markets.

In addition, data privacy and ethical issues have gradually become prominent. BCI devices collect the most private physiological and thinking data of users. How to protect user privacy while ensuring model personalization and how to formulate ethical standards for data collection and use have become issues that the entire industry must face.

6.2 Future development trends

Future embedded brain-computer interface systems are expected to make breakthroughs in the following directions:

- Neuromorphic computing: Brain-like hardware inspired by the human brain, such as Intel Loihi and IBM TrueNorth, is becoming a low-power alternative to deep networks. If combined with BCI, it will further promote the physiological integration of embedded intelligence.
- Lightweight Transformer and attention mechanism network: With the advent of structures such as MobileViT and TinyFormer, traditional high-power deep models are gradually being deployed in embedded systems, opening the door to more complex brain signal tasks.
- Unlabeled learning and small sample learning: Use self-supervised learning, meta-learning and other technologies to reduce data requirements and improve the BCI model's ability to quickly adapt to new users and new tasks.
- Neural-cognitive-environmental closed-loop system: Brain-computer interface will no longer be just a sensing tool between humans and machines, but a "thinking assistance system" that integrates cognitive modeling, behavior prediction, and environmental interaction.

6.3 Application prospects

From disability assistance to game entertainment, from immersive learning to military control systems, the application boundaries of embedded BCI systems are constantly expanding. It can be foreseen that in the next 10 years, with the maturity of technology, cost reduction and improvement of regulations, BCI is expected to become the "third generation of human-computer interaction interface" after the mouse and touch screen.

7 Conclusion

This paper presents a comprehensive review of recent advances in embedded brain–computer interface (BCI) systems, focusing on deep learning-driven innovations in structural design, algorithm implementation, energy efficiency, multimodal fusion, and adaptive learning. Despite significant progress, challenges such as EEG signal instability, individual variability, computational constraints, and data privacy remain critical barriers to large-scale deployment. The paper explores future trends including neuromorphic computing, lightweight attention-based models, self-supervised learning, and closed-loop cognitive systems. It also outlines expanding applications from healthcare and rehabilitation to gaming and defense, suggesting that BCI may become the next generation of human–computer interaction technology.

The essence of brain-computer interface (BCI) lies in its revolutionary potential to redefine the boundaries between thinking, behavior, and technology. It transcends traditional engineering by engaging in a profound exploration of the fusion of human consciousness with intelligent systems. This fusion is not merely functional—it is philosophical, challenging our understanding of what it means to be human in an age of machines. The development of embedded AI, capable of operating in close physical and cognitive proximity to the human body, is a vital springboard in this evolution. Only when AI systems can truly “understand the brain,” respond with human-level intelligence, and integrate seamlessly with biological structures, can BCI unleash its transformative power. This convergence promises to reshape medicine, cognition, and the very fabric of human–machine interaction, ultimately enabling technologies that not only assist but intuitively amplify human potential.

References

1. He H, Wu D, et al.: A review of embedded brain–computer interface systems, Journal of Neural Engineering, 17(5), 051001(2020).
2. Casson, AJ.: Wearable EEG and beyond, IEEE Reviews in Biomedical Engineering, 12, 204–212(2019).
3. Muller R, et al. A miniaturized 64-channel 225 μW neural recording IC, IEEE Journal of Solid-State Circuits, 50(1), 344–359(2015).
4. Zhou A, Santacruz SR, et al. A wireless BCI system for real-time cortical control, Nature Biomedical Engineering, 3(12), 1036–1044(2019).
5. Kumar S, et al. Low-cost open-source EEG system for portable BCI, Sensors, 21(6), 2169(2021).
6. Lawhern VJ, et al. EEGNet: A compact CNN for EEG-based brain–computer interfaces, Journal of Neural Engineering, 15(5), 056013(2018).
7. Schirrmeister RT, et al. Deep learning with convolutional neural networks for EEG decoding and visualization, Journal of Neural Engineering, 14(1), 016008(2017).
8. Roy Y, et al. ChronoNet: A deep recurrent neural network for brain signal classification, 2019 IEEE EMBC, 4352–4356(2019).
9. Zhou H, et al. Graph neural networks for EEG-based emotion recognition, IEEE Transactions on Affective Computing. 2022.
10. Han S, et al. EIE: Efficient inference engine on compressed deep neural network, ISCA 16: Proceedings of the 43rd International Symposium on Computer Architecture, 243–254(2016).
11. Li T, et al. Federated learning on non-IID data silos: BCI applications and beyond, IEEE Internet of Things Journal, 7(6), 5159–5169(2020).
12. Xu M, et al. A multimodal hybrid BCI with EEG and EOG for enhanced intention detection, IEEE Transactions on Biomedical Circuits and Systems, 15(1), 119–130(2021).
13. Yin Z, et al. EEG signal compression using deep autoencoders, IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(6), 1298–1307(2020).
14. Park J, et al. Meta-learning for user-adaptive BCI decoding, Neural Computation, 34(7), 1572–1601(2022).