Semnan University

Journal of Modeling in Engineering

Journal homepage: https://modelling.semnan.ac.ir/

ISSN: 2783-2538

Motor Imagery Signal Recognition Using Deep Learning

Louiza Dehyadegari[1,*], **Razieh Rastgoo**[2]

[1]Department of Electronic Engineering, Sirjan University of Technology, Sirjan, Iran
[2]Electrical and Computer Engineering Faculty, Semnan University, Semnan, Iran
**Corresponding author: l.dehyadegari@sirjantech.ac.ir*

Received: 2025-06-1 Revised: 2025-08-22 Accepted: 2025-09-08

Abstract. Considerable individual differences in EEG patterns pose a significant challenge for designing accurate and generalizable models. Moreover, the ability to successfully recognize motor imagery from shorter signal durations has a direct impact on improving the efficiency and practical usability of these technologies. In this study, an innovative hybrid framework is proposed for classifying motor imagery EEG signals, introducing a two-stage architecture based on the combination of an enhanced Informer and the EEGNet model. In this architecture, the EEG signals, after initial frequency feature extraction, are first fed into the enhanced Informer module. This module, leveraging sparse attention mechanisms and adaptive frequency filters (FAA), effectively captures long-term temporal dependencies within the EEG data. The output of the Informer is then passed to the EEGNet model, which, through its specialized convolutional layers (spatial convolution, depthwise convolution, and separable temporal convolution), purposefully extracts spatial-temporal features from the EEG signals and generates a compact and discriminative representation for final classification. Experimental results demonstrate that the proposed model achieves 85.20% accuracy in cross-subject evaluation on the standard PhysioNet dataset with short 2-second trial durations. Comparative analyses with state-of-the-art models indicate that the proposed approach offers competitive and improved performance, particularly in handling shorter signal durations and participant diversity.

Keywords: Electroencephalography signals, Motor imagery, Informer, EEGNet, Convolution.

1. Introduction

Brain–Computer Interfaces (BCIs) have emerged from advancements in neuroscience and biomedical engineering, enabling direct communication between the brain and external devices to assist individuals with severe motor impairments [1–2]. A widely used approach in BCIs is motor-imagery electroencephalography (MI-EEG), where brain activity is recorded as users imagine movements without physically performing them [3–4]. Accurate MI-EEG classification is difficult because EEG signals are noisy, non-stationary, low-amplitude, and highly dependent on individual differences [5]. Traditional processing methods are computationally costly due to the nonlinear nature of EEG data and the presence of multiple noise sources [6]. Recent studies highlight the promise of AI and machine-learning algorithms, which can learn complex patterns and improve prediction, reasoning, and adaptability in EEG analysis [7]. Integrating neuroscience knowledge, such as brain-network information or anatomy-based pretraining, can further increase model interpretability.

Transformer networks have recently been applied to EEG for temporal-feature extraction and have outperformed many neural-network-based approaches, but they are inefficient for very long sequences [8–9]. Informer networks address this limitation by reducing computational complexity and better detecting long-term patterns in EEG, including voluntary-movement classification and neurological-disorder recognition [10]. The proposed study combines Graph Neural Networks, which model spatial relationships between EEG channels [11], with Informers to jointly capture spatial and temporal features, improving MI-EEG classification performance.

2. Methodology

The proposed architecture integrates an improved Informer model with EEGNet to enhance EEG-based classification. The Informer serves as the primary temporal-feature extractor, processing raw EEG signals with high efficiency using sparse attention and reduced complexity. Its output is fed into EEGNet, which captures spatial and spatiotemporal patterns and performs the final classification. The core idea is to combine the Informer's strength in modeling long-range temporal dependencies with the convolutional network's ability to extract local spatial features. A key innovation is adding frequency-aware attention to the sparse-attention mechanism, enabling simultaneous temporal and frequency-domain analysis of EEG signals for richer feature extraction.

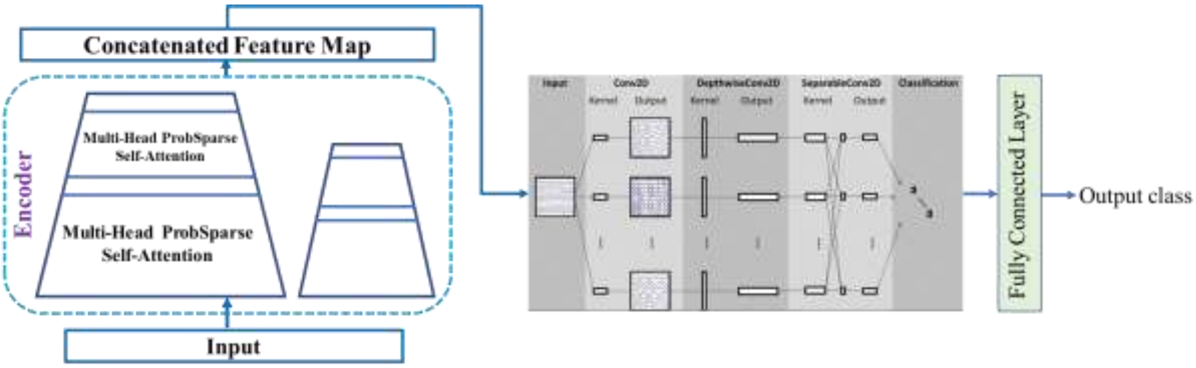

Figure 1. An overview of the proposed model.

Built on the Transformer architecture, the Informer is optimized for long time-series through intelligent sampling, avoiding the inefficiencies of full attention in classical Transformers. Only the encoder is used in this study, consisting of stacked blocks with sparse attention, feed-forward layers, normalization, dropout, and relative positional encoding. This design ensures efficient and deep temporal representation before EEGNet completes spatiotemporal processing and classification.

In this paper, a two-layer Informer is used as the temporal-feature extractor. To improve the accuracy of EEG feature extraction, a frequency-aware attention module is added to the sparse-attention pathway, operating as follows:

- First, a Fast Fourier Transform (FFT) is applied to the input signals to obtain their frequency-domain representation.
- The frequency features are then combined with the output of the sparse-attention mechanism through a weighted fusion module.
- This combination helps the model become more sensitive to dominant EEG frequency patterns associated with different mental activities. The proposed model therefore attends not only to temporal dependencies but also to critical EEG frequency features. This improves model accuracy, especially for noisy signals or those with subtle frequency variations.

The goal of the frequency-aware attention is to adaptively fuse frequency-domain features with the Informer's sparse attention output so the model captures both temporal and frequency-based EEG patterns. For each EEG input signal $X \in \mathbb{R}^{T\times C}$, where T is sequence length and C is the number of channels, an FFT is applied independently to each channel:

$$F_c = FFT(X_c), \quad \forall c = 1, \cdots, C \tag{1}$$

Here, X_c is the signal from channel c, and F_c is its frequency representation. To reduce dimensionality and focus on meaningful EEG frequency bands (e.g., alpha, beta, gamma), a frequency filter G is applied:

$$F_{filtered} = G(F) \in \mathbb{R}^{B\times C} \tag{2}$$

where F is the frequency-domain matrix for all channels, and B is the number of selected frequency bands. For temporal-feature extraction, the Informer's sparse-attention mechanism produces:

$$Z_{time} = Attention(Q, K, V) \in \mathbb{R}^{T\times d} \tag{3}$$

$$Q = XW_Q, \quad W_Q \in \mathbb{R}^{C\times d} \tag{4}$$

$$K = XW_K, \quad W_K \in \mathbb{R}^{C\times d} \tag{5}$$

$$V = XW_V, \quad W_V \in \mathbb{R}^{C\times d} \tag{6}$$

To adaptively combine temporal and frequency features, the frequency features are mapped through a fully connected layer of size $d \times T$:

$$F_{mapped} = \phi(F_{filtered}) \in \mathbb{R}^{T\times d} \tag{7}$$

where $\phi(\cdot)$ is a nonlinear mapping. The final combination is:

$$Z = \alpha \cdot Z_{time} + (1-\alpha) \cdot F_{mapped} \tag{8}$$

where $\alpha \in [0,1]$ is a learnable parameter. The combined output is normalized:

$$Z_{out} = LayerNorm(Z) \tag{9}$$

EEG data from all participants are first merged, randomly shuffled, and then divided into training, validation, and test sets. The Informer is trained on the training set and monitored using the validation set to prevent overfitting. After evaluation on the test set, if the results are satisfactory, the Informer's extracted features are saved and used to train the EEGNet model.

This research addresses the challenge of movement imagery detection using EEG signals, which are complex due to their non-stationary, noisy, and long-term dependent nature. Standard Transformers are computationally inefficient for high-resolution EEG data and struggle with sparse, long-range dependencies. The Informer model is proposed as a more scalable and robust solution, utilizing a probabilistic sparse attention mechanism and a distillation operation to efficiently model long-term temporal dependencies while preserving salient features. The extracted temporal features from the Informer are then processed by an EEGNet for deeper spatio-temporal analysis. EEGNet employs a multi-stage convolutional structure: 2D convolution captures spatial correlations across channels, depthwise convolution extracts channel-specific features efficiently, and temporal convolution precisely extracts fine-grained temporal patterns. Normalization and dropout are used to enhance training. A final fully connected layer performs the classification. The combination of the Informer's temporal feature extraction with EEGNet's spatio-temporal analysis allows for a more accurate understanding of complex brain signal dynamics, leading to improved movement imagery detection accuracy and efficiency. EEGNet's lightweight and efficient design also contributes to faster training and inference, reduced overfitting, and better performance even with limited data.

3. Discussion and Results

This section details the implementation and evaluation of a novel motor imagery EEG classification system. The proposed method combines an enhanced Informer model with EEGNet, leveraging the Informer's temporal dependency extraction and EEGNet's spatial pattern learning capabilities. Implemented using Keras and TensorFlow on Google Colab Pro, the model was trained and tested on the PhysioNet database, focusing on four motor imagery classes (right fist, left fist, both feet, rest) using only the first 2 seconds of the 6-second trials, with data from 105 participants.

The evaluation in Table 1 demonstrated that the combined model significantly outperformed individual models and achieved higher accuracy than several previous studies, even with shorter data segments. Notably, this improved performance was realized using only 2 seconds of data, compared to 4 seconds in reference studies, highlighting the model's efficiency for real-time Brain-Computer Interface applications. While some existing models showed higher accuracy, it was attributed to using fewer classes and longer trials. The study's group-level classification approach is deemed more practical for real-world environments despite potential minor accuracy trade-offs due to individual differences.

Table 1. Results of the proposed and comparative models.

Ref.	Year	Number of Participants	Classification Method	Classification Level	Number of Classes	Trial Duration	Input Data Type	Accuracy
[12]	2022	109	Transformer-CNN	Group	2	3 seconds	Raw Data	83.31
					3			74.44
					4			64.22
[13]	2023	103	EEGNet Fusion V2	Group	2	4 seconds	Processed Data	87.8
[14]	2020	105	EEGNet	Group Level / 5-fold Cross Validation	4	3 seconds	Raw Data	65.07
[15]	2024	20 participants with top PSD	Fractal dimension as discriminative feature + Machine Learning	Group Level / 5-fold Cross Validation	3	4 seconds	Processed Data	86.00
					2			81.00
[16]	2024	105	ConTraNet (CNN+Transformer)	Group Level / 5-fold Cross Validation	2	4 seconds	Raw Data	83.61
					3			74.38
					4			65.44
[17]	2025	105	Transformer + EEGNet	Group Level / 5-fold Cross Validation	4	2 seconds	Raw Data	66.90
Proposed	2026	105	Enhanced Informer + EEGNet	Group Level / 5-fold Cross Validation	4	2 seconds	Raw Data	85.20

4. Conclusion and Future Work

This research introduces an effective EEG-based motor imagery classification framework by combining an enhanced Informer architecture with EEGNet. The model excels at extracting temporal and spatial features from EEG signals, demonstrating superior performance even with shorter data segments compared to existing methods. This highlights its suitability for time-constrained, real-world applications.

However, the proposed model faces limitations, including high computational complexity, a potential decrease in performance with noisy signals, reduced generalizability due to signal non-stationarity, and the risk of losing subtle temporal patterns with short (2-second) data segments. Future research will focus on refining the model structure, testing it on diverse datasets, and addressing these limitations for improved real-time BCI applications.

References

[1] Wolpaw, Jonathan, and Elizabeth W. Wolpaw. "Brain-Computer Interfaces: Principles and Practice." Oxford University Press, USA, 2012.

[2] Wolpaw, Jonathan R., Niels Birbaumer, Donald J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. "Brain–computer interfaces for communication and control." Clinical Neurophysiology 113, no. 6 (2002): 767-791.

[3] Pfurtscheller, Gert, and Christa Neuper. "Motor imagery and direct brain-computer communication." Proceedings of the IEEE 89, no. 7 (2001): 1123-1134.

[4] Hamidi, Arezoo, and Kourosh Kiani. "Motor Imagery EEG signals classification using a Transformer-GCN approach." Applied Soft Computing 170 (2025): 112686.

[5] Lotte, Fabien, Loïc Bougrain, Andrzej Cichocki, Maud Clerc, Alexandre Congedo, A. Rakotomamonjy, and F. Yger. "A review of classification algorithms for EEG-based brain–computer interfaces: A roadmap to algorithm design." Journal of Neural Engineering 15, no. 6 (2018): 1741-2552.

[6] Altaheri, Hassan. "Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: a review." Neural Comput. Appl. 35, no. 20 (2023): 14681–14722.

[7] Hamidi, Arezoo, and Kourosh Kiani. "Motor Imagery EEG signals classification using a Transformer-GCN approach." Applied Soft Computing 170, no. 10 (2025): 112686.

[8] Ahmadi, Amir Mohammad, Kourosh Kiani, and Razieh Rastgoo. "A Transformer-based model for abnormal activity recognition in video." Journal of Modeling in Engineering 22, no. 76 (2024): 213–221.

[9] Esfandiari, Nura, Kourosh Kiani, and Razieh Rastgoo. "A conditional generative chatbot using transformer model." Journal of Modeling in Engineering 23, no. 82 (2025): 99–113.

[10] Ma, Zhen. "Transformed common spatial pattern for motor imagery-based brain-computer interfaces". Front. Neurosci 17, no. 3 (2023): 1116721.

[11] Hou, Yanlong. "GCNs-net: a graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals". IEEE Trans. Neural Netw. Learn. Syst. 35, no. 6 (2022): 7312 – 7323.

[12] Xie, Jin, Jie Zhang, Jiayao Sun, Zheng Ma, Liuni Qin, Guanglin Li , Senior Member, IEEE, Huihui Zhou , and Yang Zhan. "A transformer-based approach combining deep learning network and spatial-temporal information for raw EEG classification." IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022. 30: pp. 2126-2136.

[13] Wang, Xiaying, Michael Hersche, Batuhan Tömekce, Burak Kaya, Michele Magno, and Luca Benini. "An accurate eegnet-based motor-imagery brain–computer interface for low-power edge computing." In 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA), pp. 1-6. IEEE, 2020.

[14] Chowdhury, Radia Rayan, Yar Muhammad, and Usman Adeel. "Enhancing cross-subject motor imagery classification in EEG-based brain–computer interfaces by using multi-branch CNN." Sensors 23, no. 18 (2023): 7908.

[15] Moaveninejad, Sadaf, Valentina D'Onofrio, Franca Tecchio, Francesco Ferracuti, Sabrina Iarlori, Andrea Monteriù, and Camillo Porcaro. "Fractal Dimension as a discriminative feature for high accuracy classification in motor imagery EEG-based brain-computer interface." Computer Methods and Programs in Biomedicine 244 (2024): 107944.

[16] Ali, Omair, Muhammad Saif-ur-Rehman, Tobias Glasmachers, Ioannis Iossifidis, Christian Klaes. "ConTraNet: A hybrid network for improving the classification of EEG and EMG signals with limited training data." Compusters in Biology and Medicine, 2024. 168: p. 107649.

[17] Hmaidi, Arezoo, Kourosh Kiani, Razieh Rastgoo. " A Fusion of Transformer and EEGnet Models for Motor Imagery EEG Decoding." Journal of Modeling in Engineering 23 (2025): 191-202.