

# Advancements in EEG-based BCI for cognitive decline diagnosis using deep learning and Transformers

**Yuanyuan Zhang**

School of Professional Study, New York University, New York, NY10003, United States of America

yz3732@nyu.edu

**Abstract.** Deep learning and Transformer models have revolutionized medical diagnostics, particularly in the advanced analysis of large datasets. Given the individual variability in cognitive decline status, the accuracy of diagnosing elderly cognitive decline diseases is crucial. Since it allows for the initiation of treatment and interventions at the earliest possible stage, potentially slowing the progression of the disease. In this paper, the author reviews and summarizes studies on the use of Transformers and deep learning in Electroencephalography (EEG) signal decoding, which is the most prevalent signal-detecting method in Brain-Computer Interface (BCI) systems for elderly cognitive decline. The review indicates that deep learning and transformer models outperform traditional methods in classifying cognitive decline, offering effective methods and improved accuracy by extracting and analyzing in-depth features from EEG data. Differently, deep learning models are proficient in single modality learning, but transformer models offer a unified approach to process and integrate information from diverse data sources and, with the increase in data volume, Transformers have the potential to further improve the accuracy of diagnosis.

**Keywords:** Electroencephalography, Transformer, Deep learning, Brain-Computer Interface.

## 1. Introduction

The global population aged 65 and older is estimated to increase to 1.6 billion by 2025, and cognitive decline as a common issue affecting the elderly will continue to be a significant issue that is worth focusing on. According to Alzheimer's Disease International, over 55 million people were living with dementia as of 2022. This number is projected to steadily increase, reaching 78 million by 2030 and 139 million by 2050. Cognitive decline encompasses various types of diseases, with mild cognitive impairment and Alzheimer's being two of the most frequently discussed conditions [1,2]. Mild cognitive impairment is often considered a precursor to Alzheimer's, the most severe form of dementia in the elderly [3,4]. The accuracy in detecting cognitive decline based on large datasets is crucial due to the variability of individual symptoms and the complexity of brain connectivity. With extensive usage of Brain-Computer Interface (BCI) technologies, BCI systems are used to aid in cognitive decline in the elderly [5].

BCI systems offer non-invasive methods leveraging neurofeedback mechanisms that leverage Electroencephalography (EEG) to capture biomarkers as indices for dementia assessment. EEG is prevalently used in BCI systems for dementia assessment or classification due to its capability to provide

real-time signals in relevantly various types of biomarkers, such as frequency bands, event-related potential P300, etc., that have been verified as significant biomarkers and processed to represent complex brain functional connectivity [6].

Traditional machine learning has shown its efficiency in analyzing EEG signals for cognitive decline assessment coordinating with the capabilities in processing and analysis of nonlinear and large data samples [4,7]. However, the brain structure is non-Euclid and the brain activities occur over time, traditional machine learning is inadequate for addressing the temporal and spatial features and is also limited in handling complex data. Therefore, the research aims to manifest the abilities of deep learning and Transformer models in processing and analyzing higher-dimension data paired with spatial and temporal features from EEG signals and demonstrate its capability of getting more accurate results for cognitive decline assessment than traditional machine learning.

## **2. Deep learning-based approaches for EEG-based BCI**

In this paper, a systematic literature review is done to explore the feasibility and identify the advantages of using deep learning and transformer methods to aid cognitive decline assessment with EEG-based BCI.

### *2.1. Deep learning-based EEG signal processing*

BCI iteration is driven by advanced computational capabilities, and the strength of deep learning and Transformer in handling intricate data in BCI systems is revealed in existing studies.

Inspired by the human neural network, deep learning models are capable of automatically extracting high-dimension features from raw data through learning by leveraging interconnected nodes, which exceeds traditional machine learning which requires manual efforts in extracting based on domain expertise [8]. Thus, deep learning is leveraged in brain disease diagnosis by decoding brain signals from neuroimaging tools working with BCI.

Systematic research was done by Khan et al. and Song et al. presented the studies using typical deep learning models including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM) to detect brain tumors or Alzheimer's in BCI working with various neuroimaging tools including EEG [9].

Elsayed et al. proposed a deep learning model that encompasses a Deep Briefly Network (DBN) and CNN to extract latent features related to audio, visual, and imagery stimuli from EEG signals and classify the signals based on corresponding stimuli type or subject characters [10].

### *2.2. Deep learning-based elderly cognitive decline diagnosis*

Traditional machine learning is limited in capturing in-depth features from intricate signals and is hard to balance calculation efficiency and prediction accuracy. Alvi et al. designed an LSTM-based framework that can exert the sequential feature of EEG signals to detect Mild Cognitive Impairment (MCI) at an early stage [11]. In this research, EEG data was collected from 27 participants, with 16 healthy and 11 MCI. Before the data is fed to the LSTM-based model for feature extraction and prediction, there is a pre-processing procedure of denoising with Butterworth filtering technique, segmenting (taking a 6-second recording from each 30-minute segment), and down-sampling via Average Filter to accelerate processing in the proposed LSTM-based framework.

The LSTM architecture was built with Python. Pre-processed data will be further processed in NumPy module to be converted into double-dimensional matrices, and randomly split into testing (25%) or training sets to develop the models using the Keras library with TensorFlow backend. 20 LSTM models were designed and tested in this research to find the most efficient one with the highest classification accuracy. These models vary in the number of hidden LSTM layers and the number of neurons per layer, and they share the same activation function loss function, optimizer, early stopping setup, batch size, and number of epochs. The model performance was evaluated by accuracy, sensitivity, false positive rate, precision, F1 score, Receiver Operating Characteristic (ROC) graph, and Area under

the ROC curve (AUC). Among all the tested models, model No.13 LSTM network, which compose of an input layer, 2 hidden LSTM layers, 1024 and 512 nodes, performed the best result.

Research of Nour et al. utilized a hybrid model, coalescing Deep Ensemble Learning (DEL) and 2D-CNN to accurately classify EEG signals to diagnose Alzheimer's (AD) and health subjects [12]. The EEG signals used in the research are from 2 databases including EEG signals of 104 AD patients and 36 healthy subjects in total. The EEG signals were denoised and filtered into 5 frequency bands by the band-filter technology then segmented into 1-second epochs of 128 x 19 2D array as inputs of the DEL model learning. The proposed model architecture contains 5 classifier models that are randomly generated by the 2D CNN models in this architecture and trained with EEG signal segmented in different bands. The randomly chosen feature is addressed by the internal 2D models in architecture in which the components of models and epochs change randomly for ensemble learning. Filtered EEG signals (certain frequency bands in 1s epochs and 128 x 19 array) were randomly fed into 5 2D-CNN classifier models. The prediction of multiple classifier models is combined together in the final layer of the Metamodel as confidence weight, and the Metamodel outputs the final classification based on average weights.

The design of multiple classier models and the final layer in the metamodel strength generation ability and lower the risk of overfitting. Moreover, 5 cross-fold validations are used to evaluate the model performance and help prevent overfitting, and McNamara's test was applied to compare the performance across different classifier models to further ensure the reliability of the classification results. The classification performance of the DEL model is measured with accuracy, precision, recall, specificity, F1-score, and AUC-ROC, and the overall accuracy of the DEL model is 97%. The test results of accuracy and F1 score of each frequency band show that the Gema frequency band seems to be the most distinctive in AD-Health classification since it shows the highest accuracy (0.95) and F1 Score (0.93).

### **3. Transformer-based approaches for EEG-based BCI**

#### *3.1. Transformer-based EEG signal processing*

Inheriting the merits of deep learning, the Transformer shows potential in disease diagnosis as it contains multiple layers of architecture of decoder/encoder and parallel blocks which empowers the model with advanced computational capabilities to handle high-dimension data focusing on spatial and temporal features and increase the prediction accuracy.

Natural Language Processing (NLP) makes the data analysis more independent of the format of the raw data, and Self-attention mechanism allows the model to weigh the importance of inputs relative to each and attend to the sequence of different parts which increases the learning efficiency and prediction accuracy. As mentioned, EEG is widely used as a brain signal-detecting method in BCIs, and the efficiency of the Transformer in processing EEG raw data, especially in capturing spatial and temporal features, has been demonstrated in existing studies.

Xie et al. proposed a transformer-based model encompassing 3 layers of architecture to improve the classification accuracy of motor imagery EEG signals in BCI. Spatial-Transformer (s-Trans) utilizes attention mechanisms to capture dependencies across different EEG channels. Temporal-Transformer (t-Trans) focuses on capturing long-range dependencies between time points in the EEG signals. Hybrid CNN-Transformer model handles spatial feature extraction with Transformers to learn temporal dependencies, thus ensuring the model capitalizes on both spatial and temporal information from raw EEG data [13].

Lee et al. tested the accuracy of EEG signals classification of multiple transformed-based models, and a proposed fusion model, f-CTrans, with multiple branches combining temporal and spatial CNN with temporal-spatial transformers outperformed other proposed models in the classification. Moreover, this model utilized scaled dot-product attention (multi-attention heads to get the model to learn and integrate different aspects of data effectively [14].

In research by Hameed et al. Researchers investigated the efficiency of their proposed Temporal-Spatial Transformer (TST) in overcoming the limitation of traditional machine learning in processing long sequences with EEG data [15].

### 3.2. *Transformer-based elderly cognitive decline diagnosis*

Transformer Model architectures are able to integrate multiple layers of models to handle complex computations with high-dimension data and self-attention mechanisms increase the efficiency of learning, which means it may help increase the accuracy of cognitive decline assessment by analyzing more complex features in EEG signals.

Given the non-Euclid structure of the human brain, it requires more complex computation in feature extraction, while Transformer models can well adapt this to increase prediction accuracy with spatial and temporal features. This advantage was demonstrated in the brain network sequence-driven manifold-based transformer (BNMTrans) model, proposed by Qin et al, which used a brain network sequence-driven manifold-based transformer to analyze the brain synchronization decoded from EEG signals for cognitive impairment detection. The model is designed to simultaneously analyze and integrate both spatial and temporal patterns captured by the brain networks based on data from 46 MCI subjects and 43 healthy subjects to detect Mild Cognitive Impairment [16]. First, the EEG signals are processed to build brain networks. EEG signals are presented in matrices based on the corresponding recording channels and time points and then used to calculate the synchronization features through Pearson's Correlation Coefficient to build brain network sequences. In this step, the data is also prepared with Symmetric Positive Definite metrics for the manifold projection in the upcoming process of the model. After that, the matrix representations will be forwarded to the ST-extractor. In the extractor, the spatial relationships between EEG electrodes are captured by a Spatial 1D convolution layer, and temporal patterns of EEG signals will be captured by a temporal convolution layer, the features are further breakdown into small segments of each time point as the representation of the brain network at a certain time point, then the model will construct a new brain network, based on each segment and form the brain network sequence which will be projected to Riemannian manifold using SPD-transform in the Manifold-based Transform layer where self-attention mechanism is applied to analyze the time dependencies using the attention weight from the structure relationship between brain networks using bilinear mapping, and the output of this step will be flattened into vectors and forwarded to the final classifier to be labeled into MCI or Healthy group. The model performance was evaluated with AUROC, Precision, Recall, F1-Score, and accuracy of multiple EEG features combined with support vector machine (SVM) and k-nearest neighbors (KNN). Synchronization is found to be the most distinctive feature for classification with an accuracy of 0.81(0.09) in SVM and 0.80(0.14) in KNN. The researcher also compared the model performance with models with different techniques proposed in other research and BNMTrans model also outperformed other models with an overall 0.9 accuracy.

Compared with traditional machine learning approaches, the BNMTrans model is advanced in the following aspects: (1) The ST-extractor extracts both temporal and spatial feature patterns in EEG data while Traditional ML models often struggle to handle the complexity of time-series data while maintaining an understanding of spatial relationships. In addition, it uses Self-attention mechanism for temporal dependence to focus on important parts of the sequential brain network data and captures long-range dependencies over time. Unlike traditional ML, the End-to-End architecture learns the spatial-temporal patterns and relationships directly from the data, BNMTrans reduces the reliance on domain-specific knowledge and extensive data pre-processing. This leads to more accurate models with less bias from manual intervention. (2) BNMTrans also advanced in geometric features of the brain with the Manifold-based transformer, which incorporates geometric learning on a Riemannian manifold. (3) Changes in the frequency band can indicate the possibility of cognitive decline. These are not regular quantitative data for analysis where introducing a transformer-based model can effectively analyze and classify the signals.

Miltiadous et al. proposed DICE-Net comprises a Hybrid Convolution-Transformer Architecture that combines depthwise convolution layers with Transformer encoders in a parallel block structure to

decode and analyze Relative Band Power (RBP) and Spectral Coherence Connectivity (SCC) from EEG and complete the classification for Alzheimer's diagnosis. The research is based on data from 88 participants [17]. The EEG signals are filtered into 5 frequency band waves and segmented by 30-second time windows with 15-second overlaps, and Relative Band Power (RBP) and Spectral Coherence Connectivity (SCC) were extracted from each segment. RBP was calculated with Power Spectral Density (PSD) of each frequency band across time, channels, and bands, which was estimated using the Welch method. The RBP is presented in a 3-dimensional matrix. SCC is used to quantify brain signal synchronization, which is also represented in a 3-dimensional matrix, calculated with Morlet Wavelet Transform in a real-time domain by comparing PSD of each signal with the cross-spectral density between pairs of signals, and then the processed features are fed to the DICE model for classification. The depthwise convolution layers allow independence of the data at a given dimension and extract the spatial and temporal feature to make it more effective for the computation in the next layer, after that the data will be encoded with transformer encoders. Further, the processed features will be fed into the Feedback-Forward Network to complete classification. The combination of its Hybrid Convolution - Transformer architecture is distinguished from traditional machine learning methods such as Support Vector Machines (SVM), k-nearest Neighbors (k-NN), or simple neural networks. The depthwise convolution layers extract local spectro-temporal patterns in EEG data, while the Transformer encoders focus on global relationships between features through self-attention mechanisms. This hybrid approach allows DICE-net to effectively capture both localized and long-range dependencies in the EEG signals, which traditional ML models struggle to handle without explicit feature engineering.

Multiple algorithms, such as k-Nearest Neighbors with Principal Component Analysis, Support Vector Machines with PCA (PCA-SVM), Multilayer Perceptron, etc, were applied for performance validation from the corresponding perspective. Leave-One-Subject-Out is also applied to validate the performance of the proposed model. The overall performance of DICE-model significantly outperformed in classifying AD and Healthy subjects with an accuracy of 83.28% compared with other benchmark algorithms.

#### 4. Discussion

Based on the research in existing studies, some areas may be worthy of further optimization in future studies.

**Gap Between Laboratory and Clinical Practice.** All existing studies are based on open-resource data from AD or MCI subjects or collaborating hospitals to get reliable experimental samples. However, how to smoothly transfer these diagnosis technologies to clinical practice still needs to be discussed. All models were trained and tested with a manageable number of subjects, but the clinical practice needs to encounter a more complex environment. Participants of the studies were from single regions, so the performance of the model with a more diverse demographic and larger population needs further validation.

**Data Integration.** Experiments were conducted in different labs using different datasets, even though the validation of model performance is solid using typical verification methods, the biomarkers extracted from EEG signals were not uniform. Some are analyzed with frequency bands, and some are analyzed based on the synchronization directly extracted from EEG signals. Using various signal formats can enhance the diagnosis but it may need a solid method to integrate the analysis based on all types of EEG signals to reduce inconsistencies when applying the technologies to clinical practice.

**User Acceptance.** For tools leveraged for signal recording, EEG requires a long detecting time and is relatively easy to be impacted by noise which may also influence the result of prediction, consequently lead bias on the cognitive decline assessment. Additionally, the EEG devices aren't portable and are more suitable in the lab environment. Cognitive decline is not perceptible at early stages and it may deteriorate rapidly. Current applications may need optimization to better fit in the clinical diagnosis or home application to ensure a prompt diagnosis and early detection.

## 5. Conclusion

Deep learning and Transformers are leveraged in processing complex features latent in the EEG signals, and the efficacy is proved in existing studies, which indicates the feasibility of EEG-based BCI working with Deep Learning and Transformer approaches for cognitive decline diagnosis. According to the research, Deep Learning and Transformer show a promising capability to increase the efficiency of assessment of cognitive decline in the elderly. All proposed models using Deep Learning or Transformer architectures yield good accuracy of cognitive decline classification based on solid performance validation. Compared with the traditional machine learning approach, both deep learning and Transformer models exceed in computational capabilities of high-dimension EEG signal data to help get accurate classification results of cognitive decline diseases. Transformer models manifest more potential in cognitive decline diagnosis, as the architecture with multiple components works better in extracting both spatial and temporal features to provide a more holistic analysis based on brain function connectivity and geometry features thus increasing the accuracy of cognitive decline diagnosis. In the existing research, the performances of the proposed Transformer models outperformed other benchmark algorithms in the classification accuracy and other performance metrics, such as AUROC, Precision, Recall, and F1-Score.

## References

- [1] Lau, S. Y. J., & Agius, H. (2021). A framework and immersive serious game for mild cognitive impairment. *Multimedia Tools and Applications*, 80(20), 31183-31237.
- [2] Kantayeva, G., Lima, J., & Pereira, A. I. (2023). Application of machine learning in dementia diagnosis: A systematic literature review. *Heliyon*, 1-13.
- [3] Sibilano, E., Brunetti, A., Buongiorno, D., Lassi, M., Grippo, A., Bessi, V., ... & Bevilacqua, V. (2023). An attention-based deep learning approach for the classification of subjective cognitive decline and mild cognitive impairment using resting-state EEG. *Journal of Neural Engineering*, 20(1), 016048.
- [4] Ieracitano, C., Mammone, N., Hussain, A., & Morabito, F. C. (2020). A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Networks*, 123, 176-190.
- [5] Zhang, X., Yao, L., Wang, X., Monaghan, J., Mcalpine, D., & Zhang, Y. (2019). A survey on deep learning based brain computer interface: Recent advances and new frontiers. *arXiv preprint arXiv:1905.04149*, 66.
- [6] Fukushima, A., Morooka, R., Tanaka, H., Kentaro, H., Tugawa, A., & Hanyu, H. (2021). Classification of dementia type using the brain-computer interface. *Artificial Life and Robotics*, 26, 216-221.
- [7] AlSharabi, K., Salamah, Y. B., Abdurraqeeb, A. M., Aljalal, M., & Alturki, F. A. (2022). EEG signal processing for Alzheimer's disorders using discrete wavelet transform and machine learning approaches. *IEEE Access*, 10, 89781-89797.
- [8] Abibullaev, B., Keutayeva, A., & Zollanvari, A. (2023). Deep learning in EEG-based BCIs: a comprehensive review of transformer models, advantages, challenges, and applications. *IEEE Access*, 11, 127271-127301.
- [9] Khan, P., Kader, M. F., Islam, S. R., Rahman, A. B., Kamal, M. S., Toha, M. U., & Kwak, K. S. (2021). Machine learning and deep learning approaches for brain disease diagnosis: principles and recent advances. *IEEE Access*, 9, 37622-37655.
- [10] Elsayed, N. E., Tolba, A. S., Rashad, M. Z., Belal, T., & Sarhan, S. (2021). A deep learning approach for brain computer interaction-motor execution EEG signal classification. *IEEE Access*, 9, 101513-101529.
- [11] Alvi, A. M., Siuly, S., & Wang, H. (2022). A long short-term memory based framework for early detection of mild cognitive impairment from EEG signals. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(2), 375-388.

- [12] Nour, M., Senturk, U., & Polat, K. (2024). A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach. *Biomedical Signal Processing and Control*, 89, 105751.
- [13] Xie, J., Zhang, J., Sun, J., Ma, Z., Qin, L., et. al. (2022). A transformer-based approach combining deep learning network and spatial-temporal information for raw EEG classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30, 2126-2136.
- [14] Lee, Y. E., & Lee, S. H. (2022). EEG-transformer: Self-attention from transformer architecture for decoding EEG of imagined speech. In *2022 10th International winter conference on brain-computer interface*, 1-4.
- [15] Hameed, A., Fourati, R., Ammar, B., Ksibi, A., Alluhaidan, A. S., Ayed, M. B., & Khleaf, H. K. (2024). Temporal–spatial transformer based motor imagery classification for BCI using independent component analysis. *Biomedical Signal Processing and Control*, 87, 105359.
- [16] Qin, R., Song, Z., Ren, H., Pei, Z., Zhu, L., et. al. (2024). Bnmtrans: A brain network sequence-driven manifold-based transformer for cognitive impairment detection using eeg. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016-2020.
- [17] Miltiadous, A., Gionanidis, E., Tzimourta, K. D., Giannakeas, N., & Tzallas, A. T. (2023). DICE-net: a novel convolution-transformer architecture for Alzheimer detection in EEG signals. *IEEE Access*, 11, 71840-71858.