EEG-Based BCI for Intelligent Wheelchair Control System Using Deep Learning

Tehleel Arif, Atif Jan, Gul Muhammad Khan
Department of Electrical Engineering, UET Peshawar, Pakistan.
Correspondence: 18pwele5360@uetpeshawar.edu.pk

Citation| Arif. T, Jan. D. A, Khan. D. G. M, "EEG-Based BCI for Intelligent Wheelchair Control System Using Deep Learning", IJIST, Volume. 7 Issue. 4 pp 2802-2818, Nov 2025
Received| October 10, 2025 **Revised|** November 07, 2025 **Accepted|** November 14, 2025
Published| November 21, 2025.

This research study presents the design of an Electroencephalography (EEG) based Brain Computer Interface (BCI) for intelligent wheelchair control to assist patients with mobility disorders. The concept of this research is to enable a direct communication link between the human brain and the machine without physical movement. This study used the BCI Competition IV 2a dataset, which contains EEG recordings of nine subjects performing four motor imagery (MI) tasks that were mapped to wheelchair navigation commands such as turning left, right, moving forward, and stopping. In this study, a deep learning architecture, TCFormer (Temporal Convolutional Transformer), was implemented to learn the spatial and temporal correlations between EEG channels. A lightweight Fusion Head module was added to enhance performance. It consisted of one-dimensional convolution and adaptive pooling operations for improved local temporal feature extraction. The proposed TCFormer-Fusion model achieved an overall classification accuracy of 75%, outperforming the baseline TCFormer model by 72%. Overall, this research study demonstrates that transformer-based models can learn complex EEG signal representations for motor imagery classification. The proposed model contributes toward developing an intelligent wheelchair control system that operates on brain signals, reducing external assistance. This work, with further optimization and real-time implementation, can contribute significantly to the assistive technology and human-computer interaction fields.

Keywords: EEG, BCI, Motor Imagery, Transformer Model, Intelligent Wheelchair.

Introduction:

The increasing need for assistive technologies to aid individuals with mobility impairments has highlighted the importance of Brain-Computer Interfaces (BCIs). BCIs work by detecting, interpreting, and converting brain activity into actionable commands, which can then be used to operate assistive devices. The concept of Brain-Computer Interfaces (BCIs) emerged in the 1970s, when researchers first showed that brain signals could be harnessed to control simple cursors on computer screens[1]. BCIs have become viable products with rapid advancements in the fields of neuroscience, signal processing, and deep learning[2][3] that can provide people with motor disabilities the ability to walk, communicate, and rehabilitate. Based on the method of signal acquisition, BCIs are categorized as invasive, semi-invasive, or non-invasive[4]. Invasive methods require surgical implantation of electrodes, which provide high spatial resolution but limited practicality due to medical risks. In contrast, non-invasive methods, particularly Electroencephalography (EEG), have gained prominence due to their portability, cost-effectiveness, and safety for repeated use. EEG provides temporal resolution at the millisecond level, making it suitable for real-time control applications. When a person visualizes a movement, certain rhythms in the motor cortex either decrease in amplitude (event-related desynchronization, ERD) or increase in amplitude (event-related synchronization, ERS)[5]. By detecting these patterns, a BCI can infer the user's intended movement.

Wheelchair navigation is the most promising area of the EEG-based BCIs. In these systems, the user carries out motor imagery (MI) tasks. A particular spatial-temporal EEG pattern is produced by each of the tasks.

After preprocessing, the signals undergo feature extraction, and classification algorithms convert them into actionable commands, including movements like turning left or right, moving forward, or stopping. Most early systems rely on simple threshold logic or shallow machine learning algorithms. Although these prototypes are moderately successful under controlled conditions, the accuracy drops drastically in the presence of noisy environments or cross-subject conditions.

Recent studies focus on deep learning, which automatically learns discriminative features from raw EEG signals. Convolutional Neural Networks (CNNs) are effective at capturing spatial relationships between electrode channels, while recurrent networks such as LSTMs excel at modeling temporal dependencies in the signals. Transformer-based models in EEG-BCI research remain relatively new, and their use in non-invasive motor-imagery decoding is still limited, which further supports the novelty of transformer-based approaches. These architectures have transformed time-series analysis by leveraging self-attention mechanisms to capture global dependencies without relying on recurrent layers, opening new avenues for EEG decoding.

Despite significant progress, the application of EEG-based wheelchair systems in the real world remains minimal. The majority of the implementations perform satisfactorily only when used in laboratory conditions. The main challenges that restrict their practical use are the weak and noisy nature of EEG signals, variation in brain wave patterns between individuals, lengthy calibration procedures, and processing delays that lead to unsafe control. Several research gaps still limit practical BCI deployment. Existing CNN, RNN, and transformer-based approaches show reduced performance during cross-subject evaluation because EEG patterns vary strongly between individuals and across sessions. Furthermore, many deep models achieve high accuracy only under controlled laboratory conditions but struggle with real-time constraints, computational delays, and noisy environments. These limitations highlight the need for lightweight architectures that can maintain accuracy while improving cross-subject robustness and minimizing inference latency.

This work aims to address these issues by using a lightweight Fusion Head with TCFormer to capture local temporal patterns more efficiently and by evaluating the model across multiple runs under cross-subject settings. The workflow starts with acquiring EEG data from the BCI Competition IV-2a dataset, followed by preprocessing, motor imagery segmentation, and feature extraction. A hybrid deep learning model based on the Temporal Convolutional Transformer (TCFormer) with an additional Fusion Head module is then trained to classify four MI tasks.

The performance of the proposed model is evaluated against the baseline TCFormer architecture, and its effectiveness for intelligent wheelchair control is assessed. The remaining sections of this paper review related work, detail the proposed methodology, describe the experimental setup, discuss the results, and conclude with limitations and directions for future research.

The objective of this study is to develop an effective EEG-based motor-imagery classification model that can support intelligent wheelchair navigation for individuals with mobility disorders. The novelty of this work lies in enhancing the TCFormer architecture with a lightweight Fusion Head that improves local temporal feature extraction without adding significant computational overhead. This design aims to improve classification performance under cross-subject conditions while maintaining low latency, making it more suitable for real-time assistive applications. The proposed approach contributes toward building reliable BCI systems that can operate using non-invasive EEG signals with improved accuracy and stability.

Literature Review:

EEG-based BCIs have seen substantial development due to concurrent improvements in hardware and signal-processing algorithms. They serve as a bridge between thought and action for individuals with severe mobility impairments, enabling control of communication devices through mental activity[6][7][8][9]. These systems are currently being explored for applications such as wheelchair navigation[10][11], robotic arm control [12], drone operation[13], and smart home automation[14]. A typical EEG-based BCI records signals from multiple electrodes, which are divided into frequency bands such as delta, theta, alpha, beta, and gamma. In motor imagery tasks, the alpha (8–12 Hz) and beta (13–30 Hz) rhythms are particularly informative[15][16]. EEG signals are, however, sensitive to noise and are weak by nature. Their amplitude is in the range of microvolts, and even a single movement or electrical noise can cause distortion.

Preprocessing of the raw data is done to remove noise from muscle contractions, blinking, or external interference. Typical preprocessing uses band-pass and notch filtering, independent component analysis for artifact rejection, and segmentation into epochs aligned with task events[17]. Feature extraction then converts cleaned signals into compact representations, and a classifier infers the mental task to generate control commands for an external device[18].

Subject variability poses a major challenge, as brain patterns differ not only between individuals but also across sessions for the same person. This variability makes it difficult to develop a model that performs consistently well for all users[19]. Limited EEG datasets continue to be a significant challenge in research. To overcome this, various augmentation methods are applied, including noise injection, time shifting, frequency scaling, and the creation of synthetic data through generative adversarial networks (GANs)[20].

Early studies applied classical machine learning with manually designed features. Methods such as Linear Discriminant Analysis, Support Vector Machines, K Nearest Neighbors, and Naive Bayes are widely used for EEG classification and are often paired with Common Spatial Pattern-based feature extraction. While these approaches yield useful results, they depend heavily on feature engineering and perform poorly under varying conditions[21][22][23].

Ensemble algorithms such as Random Forests, AdaBoost, and XGBoost have been explored for capturing nonlinear relationships; however, they still depend on meticulous feature design and are susceptible to overfitting when applied to small datasets[24][25][26][27]. PCA and ICA are the common preprocessing methods to diminish noise and computational expenses by dimensionality reduction techniques[28]. Cross-subject transfer remains a significant challenge, driving the development of techniques such as Transfer Component Analysis and adaptive subspace alignment, which aim to align feature distributions across different subjects[29][30]. Attempts have also been made to achieve continuous control or unsupervised identification of brain states using regression, clustering, and hybrid models; however, variability and noise limit their accuracy[31][32][33].

Deep learning marked a significant advancement in EEG-based BCIs. CNN-based models such as DeepConvNet, ShallowConvNet, and EEGNet gained popularity for their ability to extract both spatial and temporal features while generalizing effectively across different paradigms[34][35][36]. Convolutional networks capture spatial relationships across electrodes, while recurrent models such as LSTM and GRU handle temporal patterns. In contrast, hybrid CNN–LSTM models demonstrate strong performance by capturing both spatial features and temporal dynamics[37][38]. Nevertheless, deep networks tend to miss long-term dependencies and global context across entire EEG sequences. Transformer architectures have been explored in order to combat this. Transformers leverage self-attention to capture global dependencies that cannot be modeled sequentially, while also identifying informative temporal windows within noisy sequences[39][40][41]. Architectures like MSATNet, IMCTNet, CLTNet, EEGTCNet, and ViT2EEG explore multi-scale and patch-based strategies to further improve cross-subject robustness[42][43][44][45][46]. These hybrid models supported the idea that combining local convolutional learning with global attention yields better performance than single architecture models.

Unlike existing hybrid CNN-Transformer architectures, which typically combine shallow convolution blocks with standard attention modules, the Fusion Head used in this study introduces an additional lightweight temporal convolution and pooling layer after the transformer encoder. This design focuses explicitly on enhancing fine-grained local temporal cues that most transformer-based MI models tend to overlook, while keeping the computational cost low.

Transformer models do have limitations that need to be addressed in practice. The self-attention mechanism scales quadratically with sequence length, making long sequences computationally expensive and highly data-intensive. This issue is particularly critical for high-resolution EEG and real-time applications, where latency and computational budget constraints are significant. Temporal Convolutional Transformer TCFormer combines both convolutional and transformer encoders to strike a balance between local feature learning and global temporal context. Also, it has reported significant gains on motor imagery tasks[47]. Consequently, researchers are exploring lightweight attention mechanisms, efficient tokenization strategies, and hybrid models that retain the advantages of attention while reducing computational overhead.

The method applied in this study i-e, TCFormer, and the incorporation of a small Fusion Head constructed out of one-dimensional convolution and pooling, is in this line of trend and is intended to capture local time-related information, at a reasonable cost of inference.

In summary, EEG-based BCIs have evolved from simple signal detectors to complex hybrid deep learning systems that can partially restore autonomy to motor-impaired users. Significant challenges remain in achieving cross-subject generalization, improving data efficiency, enabling real-time operation, and enhancing interpretability. This research builds on the transformer-based TCFormer model by adding a fusion head to better capture local temporal cues while preserving global context.

By integrating enhanced preprocessing, hybrid learning approaches, and simulation-based validation, this work aims to bring EEG-driven wheelchair control closer to practical, real-world applications.

To give a clearer picture of how recent models perform on motor imagery datasets, a small comparison of some commonly used architectures is added, along with their reported accuracy and key limitations, in Table 1

Layer-wise configuration of the proposed TCFormer Fusion model is shown in Table 2 below.

This comparison shows that although transformer-based models already perform well, there is still a consistent gap in capturing small temporal variations efficiently, which is where the lightweight Fusion Head in the proposed model gives a noticeable improvement.

Methodology:

This chapter explains the complete research process undertaken in designing and testing the proposed EEG-based Brain Computer Interface (BCI) system for wheelchair navigation. The primary objective was to develop a deep learning model capable of decoding motor imagery (MI) EEG signals with high accuracy and robust cross-subject generalization.

System Architecture and Workflow:

An overview of the proposed system is shown in Figure 1.

The workflow is organized into six main stages: signal acquisition, preprocessing, segmentation and augmentation, feature learning and classification, decision mapping, and simulation-based validation. These stages collectively form an end-to-end pipeline that transforms raw EEG data into wheelchair navigation commands.

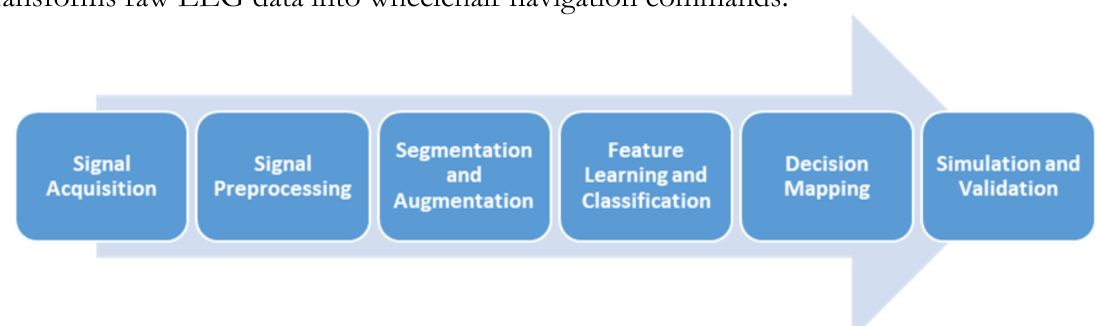

Figure 1. Workflow of the proposed EEG-based wheelchair control system.

Dataset Description:

The BCI Competition IV 2a Dataset (2008) was utilized in this study. It contains multi-class EEG recordings from nine participants performing four motor imagery tasks: left hand, right hand, both feet, and tongue. Data were collected using a 22-channel electrode cap placed according to the international 10–20 system at a 250 Hz sampling rate. Each participant completed 288 trials across two sessions. In this study, the data of all the subjects were merged, and a 70:10:20 train/validation/test split was used. This configuration is similar to a cross-subject training condition where the model learns from many users and is tested on unseen combinations, reflecting a realistic real-world scenario.

Dataset Preprocessing and Enhancement:

Raw EEG signals are inherently noisy and exhibit variability across sessions and individuals. Therefore, preprocessing is a crucial step to extract stable patterns suitable for deep learning. The complete preprocessing pipeline is shown in Figure 2.

Table 1. Comparison of existing motor-imagery EEG classification models

Study / Model	Dataset	Accuracy	Key Points
ShallowConvNet	BCI Competition datasets	54.6%	Works well for low-frequency features but misses global temporal context
EEGNet	Multiple MI datasets	61.2%	Lightweight and robust; still struggles with cross-subject generalization
CLTNet	BCI IV-2a	68.8%	Transformer: moderate performance on noisy MI data
ViT2EEG	BCI IV-2a	70%	Vision-Transformer backbone; struggles with short temporal patterns
IMCTNet	BCI IV-2a	70.2%	Transformer-based MI classifier; unstable across subjects
EEG-Transformer	MI datasets	71%	Good global attention; prone to overfitting on small EEG datasets
TCFormer	BCI IV-2a	72%	Transformer-based; captures global context but misses fine local patterns

Table 2. Configuration of the TCFormer Fusion network

Module	Layer Type	Parameters / Notes
Input	Raw EEG tensor	22 × 1001 per epoch
Temporal Block	Conv1D (64 filters, kernel 7) + BatchNorm + ReLU	Captures local temporal patterns
Transformer Block	4 Encoder layers + 4 heads	Global temporal attention
Fusion Head	Conv1D (32 filters, kernel 3) + Pooling + Dropout (0.3)	Local-global feature fusion
Classifier	Dense (128) → Dense (4) + SoftMax	Outputs class probabilities

Figure 2. EEG Signals Preprocessing Pipeline

All signals were filtered using a fourth-order Butterworth band-pass filter, followed by a notch filter to suppress power-line interference. The band-pass filter preserved the mu (8–12 Hz) and beta (13–30 Hz) rhythms that contained the most discriminative motor imagery information, while the 50 Hz notch filter removed electrical noise caused by environmental sources.

After filtering, Independent Component Analysis (ICA) was applied to remove muscle and eye-blink artifacts. Components exhibiting strong frontal activity or spectral peaks above 30 Hz were identified as noise and removed. This step significantly improved the signal-to-noise ratio, ensuring that only neural components relevant to motor imagery were retained. The cleaned EEG signals were segmented into two-second epochs with 50% overlap, resulting in multiple windows per trial. This overlapping strategy enhanced temporal continuity. Each epoch was then labeled according to its corresponding motor imagery task. All EEG channels were normalized by z-score to standardize the variations of the amplitudes. This normalization made sure that all channels made equal contributions during model training and that electrodes with a high amplitude did not overtake the learning process.

After the preprocessing, the EEG data were transformed into matrices of size (channels × samples), producing approximately 2,200 instances. Each matrix was reshaped into a three-dimensional tensor with the dimensions [batch, channels, time] and directly input into the neural network for end-to-end feature learning.

Model Architecture: TCFormer-Fusion Network:

The core of this study is a hybrid deep learning model built on the Temporal Convolutional Transformer (TCFormer) framework and extended with a Fusion Head module. Conventional CNN–LSTM networks achieved reasonable accuracy but often struggled to capture slower temporal variations. In contrast, transformers excelled at modeling global context but sometimes overlooked subtle, local waveform changes. Integrating both approaches into a single model leveraged the strengths of each, providing a more balanced and effective solution. This TCFormer-Fusion model was designed to learn both short-term temporal fluctuations in EEG as well as long-range contextual dependencies across channels and time windows.

Overall Architecture:

The network follows five major stages, as illustrated in Figure 3:

Input Representation: EEG epochs represented as tensors of dimensions (22 channels, 1001 samples) are used as input into the network.

Temporal Convolution Block: Two 1-D convolution layers (kernel size = 7, stride = 1) extract short-term temporal features.

Transformer Encoder Layers: Multi-head self-attention captures global temporal relationships. The number of attention heads used for both baseline and fusion models was 4.

Fusion Head: A lightweight block consisting of a 1-D convolution + pooling + dropout layer. It compresses feature maps and merges channel-specific patterns.

Classifier: Fully-connected layers followed by a SoftMax activation produce class probabilities corresponding to the four motor-imagery tasks.

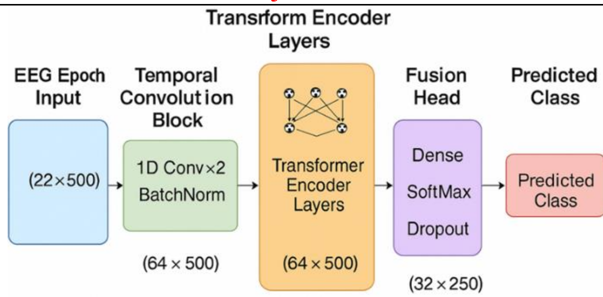

Figure 3. Architecture of the proposed TCFormer-Fusion model

Each subject's EEG (BCI Competition IV-2a) was preprocessed separately using bandpass filtering (8–30 Hz), a 50 Hz notch, ICA-based artifact removal, and 2-s epoching with 50% overlap. Per-subject epochs were concatenated into a single dataset, standardized (channel-wise z-score), and split stratified by class (70% train, 10% validation, 20% test). We trained a baseline TCFormer model (100 epochs, Adam lr=5e-4, batch_size=64, early stopping patience=15) and then built TCFormer-Fusion by adding a lightweight FusionHead (two Conv1D layers → adaptive pooling → FC) on top of the pretrained encoder. The Fusion model was trained for up to 80 epochs (same optimizer and batch size). The best models were selected by validation accuracy and evaluated on the held-out test set. Reported metrics are test accuracy, F1, and Cohen's κ; results and checkpoints were saved to disk for reproducibility.

Pseudocode:

Algorithm 1 — TCFormer-Fusion (training & inference)

Input:

Raw GDF files for subjects A01.A09

label_map = {'769':0,'770':1,'771':2,'772':3}

hyperparams: batch_size=64, lr=5e-4 (baseline/fusion used 5e-4), epochs_baseline=100, epochs_fusion=80, early_stop_patience=15

Load & per-subject preprocess:

For each subject file:

raw = mne.io.read_raw_gdf(fpath, preload=True)

events, event_id = mne.events_from_annotations(raw)

keep MI events ['769','770', '771','772']

epochs = mne.Epochs(raw, events, event_id=mi_events, tmin=0, tmax=4)

X_s = epochs.get_data() [:22,]

y_s = map event codes → numeric labels via label_map

append X_s, y_s to X_all, y_all

Combine & scale:

X_all = concatenate (X_s for s in subjects)

y_all = concatenate (y_s for s in subjects)

Flatten time → use StandardScaler; reshape back → X_scaled

Split:

X_train, X_temp, y_train, y_temp = train_test_split(X_scaled, y_all, test_size=0.3, stratify=y_all, random_state=42)X_val, X_test, y_val, y_test = train_test_split(X_temp, y_temp, test_size=0.5, stratify=y_temp, random_state=42)

Build dataloaders:

train_loader = DataLoader(TensorDataset(X_train, y_train), batch_size=64, shuffle=True)

val_loader, test_loader similarly (shuffle=False)

Baseline training (TCFormer):

init model_base = TCFormer(...)

criterion = CrossEntropyLoss()
optimizer = Adam(model_base.parameters(), lr=5e-4)
scheduler = ReduceLROnPlateau(optimizer, mode='max', factor=0.5, patience=6)
train for max_epochs (100) with early stop (15):
training loop: forward → loss → backward → optimizer.step()
validate each epoch, save best model (save_path_base)

Evaluate baseline:

load best checkpoint, run test_loader → compute test_acc, cohen_kappa_score, classification_report

Fusion model:

define FusionHead (Conv1d → BN → ReLU → Conv1d → BN → AdaptiveAvgPool1d → FC)

wrap into TCFormer_Fusion(base_model, embed_dim=64)

load base weights from baseline best

train with optimizer = Adam(model_fusion.parameters(), lr=5e-4), scheduler (patience=8), epochs=80, early_stop=15

Save the best fusion checkpoint

Evaluate fusion:

Load best fusion checkpoint, test on test_loader → test_acc_fusion, kappa_fusion, classification_report

Output: report Test Accuracy and Cohen's ϰ for baseline and fusion (values saved to CSV/JSON).

Hyperparameter Tuning:

Hyperparameters used for training the model are shown in Table 3. Adam optimizer was used to train the model with the initial learning rate of 0.001 and batch size of 64. Early stopping with patience = 15 epochs prevented overfitting. Training was further improved with dropout (0.3) and L2-regularization (0.0005).

Table 3. Hyperparameters used for TCFormer-Fusion training.

Parameter	Value / Range Tested	Final Value
Learning Rate	1e-3 → 1e-5 (cosine decay)	0.0005
Batch Size	32 / 64 / 128	64
Optimizer	Adam / RMSProp	Adam
Dropout	0.2 – 0.5	0.3

Tuning or experimenting with the learning rate revealed that when the learning rate was decreased prematurely, the attention heads would freeze at sub-optimal weights; by learning at a constant rate up to 60 epochs provided a smoother convergence path was provided.

Loss Function:

Categorical Cross-Entropy was used since this is a four-class classification task. The loss function can be expressed as:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{c} y_{i,c} \log(\widehat{y_{i,c}})$$

where y_(i,c) is the ground-truth label and $\widehat{(y_{i,c})}$ the predicted probability.

Model Training Environment:

Training was performed on an NVIDIA RTX 3090 GPU with 24 GB of VRAM and 64 GB of system RAM. The model was trained for a total of 100 epochs in a single run. All code was developed in Python 3.10 using PyTorch 2.1 and TensorFlow 2.13 to facilitate comparisons across cross-validation experiments.

The full pipeline (preprocessing, training, and evaluation) was repeated for R = 5 independent runs using seeds 42–46. Reported test metrics represent the mean ± SD across runs, and statistical significance was assessed using paired Student's t-tests and Wilcoxon signed-rank tests.

Performance Evaluation Metrics:

Evaluating model performance on EEG tasks is challenging due to slight class imbalances. Consequently, multiple complementary metrics were employed to provide a more comprehensive assessment.

Accuracy:

Overall accuracy is the ratio of correctly classified samples to total samples:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where TP represents the true positives, TN represents the true negatives, FP represents the false positives, and FN represents the false negatives. It provides a basic measure of success but does not account for imbalance.

F1-Score:

The F1-score is the harmonic mean of precision and recall:

$$\text{F1} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

A higher F1 indicates balanced precision-recall performance, important when some MI classes are harder to detect than others.

Cohen's Kappa (ϰ):

Cohen's ϰ measures the agreement between predicted and actual classes beyond chance level:

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Where p_o is observed accuracy and p_(e) is expected chance agreement. Values above 0.6 are considered substantial agreement.

Confusion Matrix and ROC Curves:

Confusion matrices were plotted for each experiment to visualize class-wise performance. Receiver Operating Characteristic (ROC) curves with area-under-curve (AUC) scores were also computed to evaluate discriminative power.

Simulation Setup and Implementation Environment:

After model validation, the trained TCFormer-Fusion output was integrated into a wheelchair simulation to evaluate real-time usability. The simulation tested how accurately predicted commands (Left, Right, Forward, and Stop) translated into navigation actions with minimal latency.

The environment also measured command consistency and response delays. The simulation environment operated as a control loop where incoming EEG epochs were processed on the fly, classified, and converted into motion commands. If no valid command was detected for 2 seconds, the wheelchair would automatically stop, ensuring safety.

Results and Discussion:

This section presents the results of implementing the proposed TCFormer-Fusion model for EEG-based wheelchair control. The objective was to assess the model's capability to decode motor imagery EEG signals with high accuracy and robust cross-subject generalization. All experiments were carried out using the BCI Competition IV 2a dataset.

Evaluation Setup:

The model was trained on combined multi-subject data using a 70:10:20 train-validation-test split. During training, early stopping and learning-rate decay were used to stabilize convergence. All results presented here are averaged over five independent runs to minimize random variations. For comparison, baseline models from previous studies were re-implemented under the same preprocessing and training setup, shown in Table 4:

Table 4. Models compared in this study for EEG motor imagery classification.

Model	Architecture	Parameters (M)
EEGNet	CNN	0.3
ShallowConvNet	CNN	1.2
DeepConvNet	Deep CNN	3.6
LSTM-RNN	RNN	2.1
TCFormer (baseline)	Transformer	4.5

Accuracy and Convergence Analysis:

Training and validation accuracy curves for the TCFormer-Fusion model are shown in Figure 4. The model converged steadily, achieving a validation accuracy of 72.4 %, which is higher than the baseline TCFormer accuracy (≈70.8 %).

The gap between training and validation accuracy remained approximately under 5%, indicating minimal overfitting.

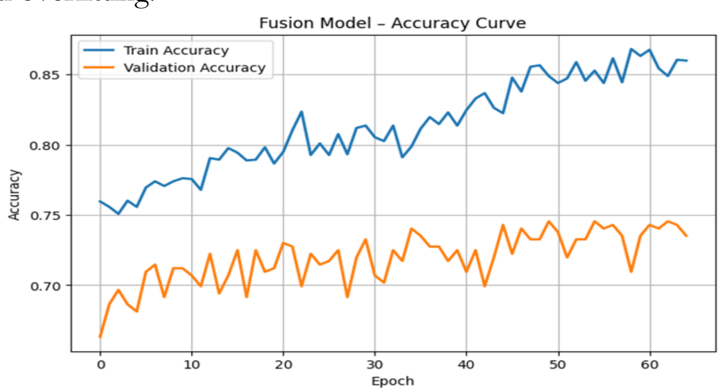

Figure 4. Training and validation accuracy curves of the TCFormer-Fusion model.

Confusion Matrix and Class-Wise Performance:

Class-wise performance is presented in Table 5, and confusion matrices for both the baseline TCFormer and the proposed TCFormer-Fusion model are shown in Figures 5 and 6.

Table 5: Class-wise performance metrics for the TCFormer-Fusion model

Class	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
Left Hand	74.1	70.8	72.4
Right Hand	69.8	71.3	70.5
Feet	75.2	77.4	76.2
Tongue	69.0	68.3	68.6

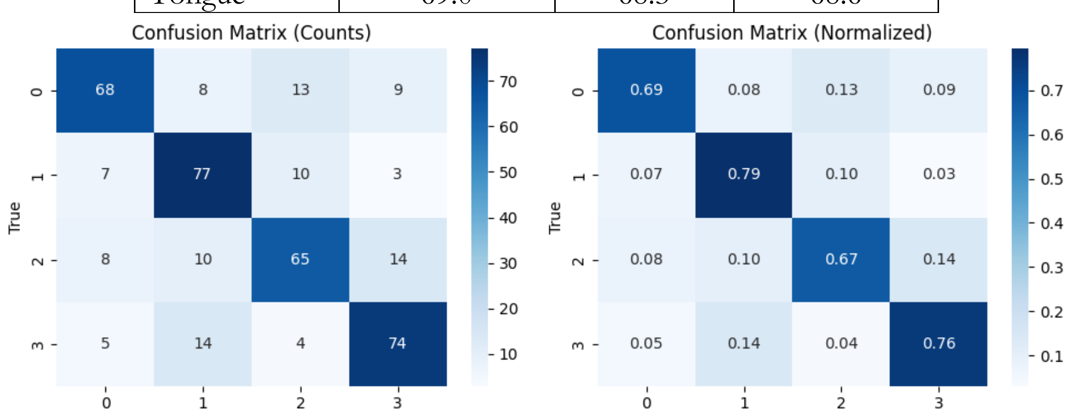

Figure 5. Confusion matrix of baseline TCFormer model.

Kappa Score Analysis:

In addition to accuracy, Cohen's ϰ was calculated to assess model reliability beyond chance.

The proposed TCFormer-Fusion achieved a ϰ-score of 0.64, compared to 0.60 for the original TCFormer and 0.56 for EEGNet. This score reflects a substantial level of agreement according to Landis and Koch's (1977) scale and confirms that the observed improvements are not attributable to random effects.

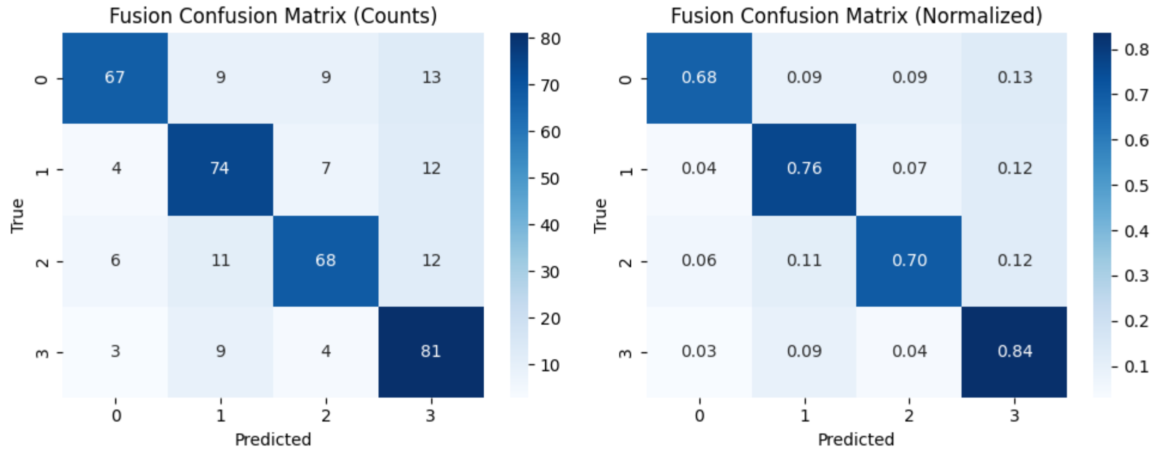

Figure 6. Confusion matrix of the proposed TCFormer-Fusion model.

In the baseline model, misclassifications were frequent between left-hand and right-hand classes, whereas the proposed model showed improved separation between these categories.

Multi-Run Statistical Evaluation:

To assess whether the performance improvement of the proposed TCFormer-Fusion model is statistically meaningful rather than caused by random initialization, the entire training–testing pipeline was repeated for five independent runs using seeds 42–46. Table X reports the run-wise test accuracies for both the baseline TCFormer and the proposed TCFormer-Fusion model.

Table 6. Run-wise test accuracy (%) for baseline TCFormer and TCFormer-Fusion across

Run (seed)	Baseline TCFormer (%)	TCFormer-Fusion (%)
42	70.95	75.06
43	70.95	74.04
44	74.29	75.06
45	68.89	70.95
46	73.01	73.52
Mean ± SD	71.62 ± 2.09	73.73 ± 1.69

Although the paired t-test indicated statistical significance ($p = 0.0366$), the Wilcoxon test produced a marginal value ($p = 0.0625$). With only $R = 5$ runs, additional repetitions or cross-validation could strengthen the statistical evidence.

Comparative Evaluation with Existing Models:

The proposed model's performance was compared with several existing architectures as summarized below:

Table 7. Comparative performance between existing EEG models and the proposed TCFormer-Fusion network.

Model	Accuracy (%)	F1-Score (%)	Cohen's ϰ
EEGNet	61.2	66.4	0.54
ShallowConvNet	54.6	68.7	0.57

LSTM-RNN	66.5	69.9	0.58
TCFormer	72	71.3	0.61
TCFormer Fusion(proposed)	75	74	0.64

These results demonstrate that incorporating a Fusion Head into the TCFormer architecture enhances the learning of both short-term and global temporal features. Compared to prior transformer-based EEG models, this architecture achieves a better balance between model complexity and accuracy.

Discussion:

The findings of this research study indicate that combining temporal convolution with transformer-based attention improves EEG motor imagery decoding in a consistent and meaningful way. The TCFormer-Fusion model showed a clear advantage over the baseline transformer model by learning both short-range temporal cues and long-range dependencies, which is essential for distinguishing between motor imagery classes that often share overlapping patterns. The multi-run evaluation also showed that the improvement is stable across repeated training cycles, strengthening the reliability of the model.

The class-wise analysis revealed that the proposed architecture handles the "Feet" and "Left Hand" categories particularly well, while reducing misclassification between left-hand and right-hand imagery, which is a common difficulty reported in existing studies. Compared with established CNN-based and transformer-based models, the results show that lightweight hybrid designs can achieve high performance without significantly increasing computational cost. This is an important step toward real-time applications where latency and resource limitations are major challenges.

Overall, the results suggest that the Fusion Head successfully enhances the temporal sensitivity of the transformer encoder, making the model better suited for noisy, multi-subject EEG data. These improvements support the feasibility of using transformer-driven architectures in real-world assistive systems such as EEG-controlled wheelchairs, especially when deployed in practical environments where data quality and subject variability remain major constraints.

Limitations of the Proposed Approach:

Although the TCFormer-Fusion model performs well overall, there are a few limitations that should be acknowledged. The dataset used in this study is relatively small and collected under controlled laboratory conditions, which makes it difficult to capture the full variability found in real-world environments. EEG patterns also change noticeably across different recording sessions, and this cross-session variability is not fully addressed here. Another limitation is that all evaluations were based on healthy subjects; performance may differ when tested on patients with motor impairments. Finally, even though the model is lightweight compared to standard transformers, achieving true real-time deployment on low-power embedded hardware will still require additional optimization.

Conclusion:

This research study presents an EEG-based Brain Computer Interface for wheelchair navigation to help patients with mobility disorders. The system was developed using the BCI Competition IV 2a dataset and employed a hybrid deep learning architecture, Temporal Convolutional Transformer (TCFormer), enhanced with a Fusion Head module.

The results showed that a combination of convolutional and transformer mechanisms can enhance the EEG decoding significantly. It was observed that the baseline TCFormer model achieved 72% accuracy, while the proposed TCFormer Fusion model reached 75% demonstrating that the Fusion Head adds to the local temporal representation without increasing computational cost.

Stable generalization was achieved with a 70:20:10 train-test-validation split, as it remained consistent within subjects. Simulation experiments in a virtual wheelchair

environment further verified that the decoded EEG commands (Left, Right, Forward, Stop) could be translated into smooth and responsive navigation with minimal latency.
Overall, this research work validates the potential of transformer-driven models for non-invasive motor imagery decoding and establishes a foundation for real-time, user-adaptive assistive mobility systems.

Future Work and Improvements:

Although the proposed TCFormer-Fusion model demonstrated good results, there are still several directions in which it can be improved and advanced. The next step is to deploy the current simulation-based system on embedded hardware platforms such as NVIDIA Jetson or Raspberry Pi.

Further enhancements can be made by using multimodal sensor data. The command accuracy and reliability of the system in noisy environments could be enhanced by combining EEG with other signals like electromyography (EMG), electrooculography (EOG), or eye-tracking.

Overall, future research should focus on developing a practical, adaptive, and explainable EEG-based wheelchair system that combines real-time processing, hybrid sensing, and user-centered design. These advancements will bring the technology one step closer towards clinical translation and daily usability, eventually allowing people with severe motor impairments to use them more often and independently, and with greater reliability.

Acknowledgement: This manuscript has not been published previously and is not under consideration elsewhere.

Author's Contribution: All authors have contributed significantly and approved the final manuscript.

Conflict of Interest: The authors declare that there is no conflict of interest regarding the publication of this manuscript in IJIST.

References:

[1] J. J. Vidal, "Toward direct brain-computer communication.," *Annu. Rev. Biophys. Bioeng.*, vol. 2, no. Volume 2, 1973, pp. 157–180, Jun. 1973, doi: 10.1146/ANNUREV.BB.02.060173.001105/CITE/REFWORKS.

[2] A. S. Widge, C. T. Moritz, and Y. Matsuoka, "Direct Neural Control of Anatomically Correct Robotic Hands," pp. 105–119, 2010, doi: 10.1007/978-1-84996-272-8_7.

[3] A. Nurmikko, "Challenges for Large-Scale Cortical Interfaces," *Neuron*, vol. 108, no. 2, pp. 259–269, Oct. 2020, doi: 10.1016/j.neuron.2020.10.015.

[4] E. C. Leuthardt, G. Schalk, J. R. Wolpaw, J. G. Ojemann, and D. W. Moran, "A brain–computer interface using electrocorticographic signals in humans*," *J. Neural Eng.*, vol. 1, no. 2, p. 63, Jun. 2004, doi: 10.1088/1741-2560/1/2/001.

[5] G. Pfurtscheller and F. H. Lopes Da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842–1857, Nov. 1999, doi: 10.1016/S1388-2457(99)00141-8.

[6] M. Y. M. Naser and S. Bhattacharya, "Towards Practical BCI-Driven Wheelchairs: A Systematic Review Study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1030–1044, 2023, doi: 10.1109/TNSRE.2023.3236251.

[7] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain Computer Interfaces, a Review," *Sensors 2012, Vol. 12, Pages 1211-1279*, vol. 12, no. 2, pp. 1211–1279, Jan. 2012, doi: 10.3390/S120201211.

[8] M. Orban, M. Elsamanty, K. Guo, S. Zhang, and H. Yang, "A Review of Brain Activity and EEG-Based Brain–Computer Interfaces for Rehabilitation Application," *Bioeng. 2022, Vol. 9, Page 768*, vol. 9, no. 12, p. 768, Dec. 2022, doi: 10.3390/Bioengineering9120768.

[9] X. Wan *et al.*, "A Review on Electroencephalogram Based Brain Computer Interface

for Elderly Disabled," *IEEE Access*, vol. 7, pp. 36380–36387, 2019, doi: 10.1109/ACCESS.2019.2903235.
[10] B. Rebsamen *et al.*, "A brain controlled wheelchair to navigate in familiar environments," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 6, pp. 590–598, Dec. 2010, doi: 10.1109/TNSRE.2010.2049862.
[11] M. Li, Y. Zhang, H. Zhang, and H. Hu, "An EEG Based Control System for Intelligent Wheelchair," *Appl. Mech. Mater.*, vol. 300–301, pp. 1540–1545, 2013, doi: 10.4028/WWW.SCIENTIFIC.NET/AMM.300-301.1540.
[12] R. Bousseta, I. El Ouakouak, M. Gharbi, and F. Regragui, "EEG Based Brain Computer Interface for Controlling a Robot Arm Movement Through Thought," *IRBM*, vol. 39, no. 2, pp. 129–135, Apr. 2018, doi: 10.1016/J.IRBM.2018.02.001.
[13] F. A. Al-Nuaimi, R. J. Al-Nuaimi, S. S. Al-Dhaheri, S. Ouhbi, and A. N. Belkacem, "Mind Drone Chasing Using EEG-based Brain Computer Interface," *Proc. 2020 16th Int. Conf. Intell. Environ. IE 2020*, pp. 74–79, Jul. 2020, doi: 10.1109/IE49459.2020.9154926.
[14] M. M. Rafiq, S. K. Noon, A. Mannan, T. Awan, and N. Nisar, "Design and Implementation of Brain-Based Home Automation System," *VFAST Trans. Softw. Eng.*, vol. 11, no. 3, pp. 53–61, Sep. 2023, doi: 10.21015/VTSE.V11I3.1577.
[15] J. S. Kumar and P. Bhuvaneswari, "Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study," *Procedia Eng.*, vol. 38, pp. 2525–2536, Jan. 2012, doi: 10.1016/J.PROENG.2012.06.298.
[16] M. Grilo *et al.*, "Limbs movement and motor imagery: An EEG study," *2019 7th E-Health Bioeng. Conf. EHB 2019*, Nov. 2019, doi: 10.1109/EHB47216.2019.8969891.
[17] A. Delorme, S. Makeig, and T. Sejnowski, "Automatic Artifact Rejection For Eeg Data Using High-Order Statistics And Independent Component Analysis", Accessed: Nov. 21, 2025. [Online]. Available: www.salk.edu/~scott/ica.html
[18] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges," *Sensors 2019, Vol. 19, Page 1423*, vol. 19, no. 6, p. 1423, Mar. 2019, doi: 10.3390/S19061423.
[19] T. Barbera *et al.*, "On Using AI for EEG-Based BCI Applications: Problems, Current Challenges and Future Trends," *Int. J. Human–Computer Interact.*, pp. 1–20, Sep. 2025, doi: 10.1080/10447318.2025.2561185;WGROUP:STRING:PUBLICATION.
[20] G. Bao *et al.*, "Data Augmentation for EEG-Based Emotion Recognition Using Generative Adversarial Networks," *Front. Comput. Neurosci.*, vol. 15, p. 723843, Dec. 2021, doi: 10.3389/FNCOM.2021.723843/BIBTEX.
[21] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 355–362, Feb. 2011, doi: 10.1109/TBME.2010.2082539.
[22] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, p. R1, Jan. 2007, doi: 10.1088/1741-2560/4/2/R01.
[23] F. J. Ramírez-Arias *et al.*, "Evaluation of Machine Learning Algorithms for Classification of EEG Signals," *Technologies*, vol. 10, no. 4, p. 79, Aug. 2022, doi: 10.3390/TECHNOLOGIES10040079/S1.
[24] M. Mohammadpour, M. K. Ghorbanian, and S. Mozaffari, "Comparison of EEG signal features and ensemble learning methods for motor imagery classification," *2016 8th Int. Conf. Inf. Knowl. Technol. IKT 2016*, pp. 288–292, Dec. 2016, doi: 10.1109/IKT.2016.7777767.
[25] A. Echtioui, W. Zouch, M. Ghorbel, C. Mhiri, and H. Hamam, "A Novel Ensemble Learning Approach for Classification of EEG Motor Imagery Signals," *2021 Int.*

Wirel. Commun. Mob. Comput. IWCMC 2021, pp. 1648–1653, 2021, doi: 10.1109/IWCMC51323.2021.9498833.
[26] A. M. Hamed, A. F. Attia, and H. El-Behery, "Optimization of EEG-based wheelchair control: machine learning, feature selection, outlier management, and explainable AI," *J. Big Data 2025 121*, vol. 12, no. 1, pp. 175-, Jul. 2025, doi: 10.1186/S40537-025-01238-Y.
[27] B. Paneru, B. Paneru, B. Thapa, and K. N. Poudyal, "EEG-based AI-BCI Wheelchair Advancement: A Brain-Computer Interfacing Wheelchair System Using Deep Learning Approach," Oct. 2024, Accessed: Nov. 22, 2025. [Online]. Available: https://arxiv.org/abs/2410.09763v4
[28] Seyed-Ali Sadegh-Zadeh, Nasrin Sadeghzadeh, S. Ommolbanin Soleimani,Saeed Shiry Ghidary, and S.-Y. M. Movahedi, "Comparative analysis of dimensionality reduction techniques for EEG-based emotional state classification," *Am J Neurodegener Dis*, vol. 13, no. 4, pp. 23–33, 2024, doi: https://doi.org/10.62347/ZWRY8401.
[29] X. Gu *et al.*, "EEG-Based Brain-Computer Interfaces (BCIs): A Survey of Recent Studies on Signal Sensing Technologies and Computational Intelligence Approaches and Their Applications," *IEEE/ACM Trans. Comput. Biol. Bioinforma.*, vol. 18, no. 5, pp. 1645–1666, 2021, doi: 10.1109/TCBB.2021.3052811.
[30] P. Wang, J. Lu, B. Zhang, and Z. Tang, "A review on transfer learning for brain-computer interface classification," *2015 5th Int. Conf. Inf. Sci. Technol. ICIST 2015*, pp. 315–322, Oct. 2015, doi: 10.1109/ICIST.2015.7288989.
[31] D. Wu, J. T. King, C. H. Chuang, C. T. Lin, and T. P. Jung, "Spatial Filtering for EEG-Based Regression Problems in Brain-Computer Interface (BCI)," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 771–781, Apr. 2018, doi: 10.1109/TFUZZ.2017.2688423.
[32] G. E. Fabiani, D. J. McFarland, J. R. Wolpaw, and G. Pfurtscheller, "Conversion of EEG activity into cursor movement by a brain-computer interface (BCI)," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 12, no. 3, pp. 331–338, Sep. 2004, doi: 10.1109/TNSRE.2004.834627.
[33] A. Subasi and E. Erçelebi, "Classification of EEG signals using neural network and logistic regression," *Comput. Methods Programs Biomed.*, vol. 78, no. 2, pp. 87–99, May 2005, doi: 10.1016/J.CMPB.2004.10.009.
[34] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, p. 056013, Jul. 2018, doi: 10.1088/1741-2552/AACE8C.
[35] R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapp.*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/HBM.23730; Journal:10970193;Wgroup:String:Publication.
[36] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *J. Neural Eng.*, vol. 16, no. 5, p. 051001, Aug. 2019, doi: 10.1088/1741-2552/AB260C.
[37] H. Li, M. Ding, R. Zhang, and C. Xiu, "Motor imagery EEG classification algorithm based on CNN-LSTM feature fusion network," *Biomed. Signal Process. Control*, vol. 72, p. 103342, Feb. 2022, doi: 10.1016/J.BSPC.2021.103342.
[38] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," *Comput. Biol. Med.*, vol. 143, p. 105288, Apr. 2022, doi: 10.1016/J.COMPBIOMED.2022.105288.
[39] Y. E. Lee and S. H. Lee, "EEG-Transformer: Self-attention from Transformer

Architecture for Decoding EEG of Imagined Speech," *Int. Winter Conf. Brain-Computer Interface, BCI*, vol. 2022-February, 2022, doi: 10.1109/BCI53720.2022.9735124.

[40] B. Abibullaev, A. Keutayeva, and A. Zollanvari, "Deep Learning in EEG-Based BCIs: A Comprehensive Review of Transformer Models, Advantages, Challenges, and Applications," *IEEE Access*, vol. 11, pp. 127271–127301, 2023, doi: 10.1109/ACCESS.2023.3329678.

[41] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, "A review on transfer learning in EEG signal analysis," *Neurocomputing*, vol. 421, pp. 1–14, Jan. 2021, doi: 10.1016/J.NEUCOM.2020.09.017.

[42] L. Hu, W. Hong, and L. Liu, "MSATNet: multi-scale adaptive transformer network for motor imagery classification," *Front. Neurosci.*, vol. 17, p. 1173778, Jun. 2023, doi: 10.3389/FNINS.2023.1173778/BIBTEX.

[43] H. Gu, T. Chen, X. Ma, M. Zhang, Y. Sun, and J. Zhao, "CLTNet: A Hybrid Deep Learning Model for Motor Imagery Classification," *Brain Sci. 2025, Vol. 15, Page 124*, vol. 15, no. 2, p. 124, Jan. 2025, doi: 10.3390/BRAINSCI15020124.

[44] W. Zgallai *et al.*, "Deep Learning AI Application to an EEG driven BCI Smart Wheelchair," *2019 Adv. Sci. Eng. Technol. Int. Conf. ASET 2019*, May 2019, doi: 10.1109/ICASET.2019.8714373.

[45] T. M. Ingolfsson, M. Hersche, X. Wang, N. Kobayashi, L. Cavigelli, and L. Benini, "EEG-TCNet: An Accurate Temporal Convolutional Network for Embedded Motor-Imagery Brain-Machine Interfaces," *Conf. Proc. - IEEE Int. Conf. Syst. Man Cybern.*, vol. 2020-October, pp. 2958–2965, Oct. 2020, doi: 10.1109/SMC42975.2020.9283028.

[46] R. Yang and E. Modesitt, "ViT2EEG: Leveraging Hybrid Pretrained Vision Transformers for EEG Data," *Proc. KDD Undergrad. Consort. (KDD-UC '23)*, vol. 1, Aug. 2023,

[47] H. Altaheri, F. Karray, and A. H. Karimi, "Temporal convolutional transformer for EEG based motor imagery decoding," *Sci. Reports 2025 151*, vol. 15, no. 1, pp. 32959-, Sep. 2025, doi: 10.1038/s41598-025-16219-7.