EEG-Based Motor Imagery Recognition Using Spatial Encoding and Frequency Attention Mechanisms

1.Mrs.P.Nirmala Priyadharshini,
Assistant Professor/IT,
Adithya Institute of Technology
Coimbtore,India
nimipeter88@gmail.com
ORCID:0009 0008 0025 3565

2.Mrs.A.Jemima
Assistant Professor/AIDS
Adithya Institute of Technology
Coimbtore,India
jemimachristy92@gmail.com
ORCID :0009-7678-8942

3.Mrs.V.Janeefer
Assistant Professor/IT,
Adithya Institute of Technology
Coimbtore,India
janeefer_v@adithyatech.com
ORCID: 0009-0003-9730-2952

4.Mrs.Afrin Hussain
Assistant Professor/IT
Adithya Institute of Technology
Coimbtore,India
afreenzak@gmail.com
ORCID: 0009-0007-2309-9268

Abstract

Electroencephalography (EEG)-based Brain–Computer Interfaces (BCIs) provide a communication channel between the brain and external devices by decoding neural activity. However, the non-stationary and noisy nature of EEG signals limits classification accuracy in motor imagery (MI) tasks. This paper presents a deep learning framework that combines Spatial Encoding and Frequency Attention Mechanisms (SEFA) to enhance EEG-based motor imagery recognition. The proposed model captures electrode-level spatial dependencies through convolutional mappings and integrates frequency-aware attention to emphasize cognitively significant bands. Experimental evaluation on BCI Competition IV-2a and PhysioNet Motor Imagery datasets demonstrates substantial improvements in classification accuracy and robustness compared to conventional CNN and hybrid models. The results establish SEFA as a scalable and interpretable framework for next-generation assistive BCIs.

Keywords — EEG, Brain–Computer Interface, Motor Imagery, Spatial Encoding, Frequency Attention, Deep Learning.

1. Introduction

Brain–Computer Interfaces (BCIs) have emerged as a transformative technology enabling direct communication between the human brain and external devices, bypassing traditional neuromuscular pathways [1]. By interpreting neural activity into actionable commands, BCIs hold significant potential for assistive technologies, neurorehabilitation, and human–machine interaction. Among the various neuroimaging modalities, electroencephalography (EEG) remains one of the most practical choices due to its non-invasive nature, cost-effectiveness, and high temporal resolution [2]. EEG-based BCIs capture the brain's electrical activity through scalp electrodes, offering valuable insights into cognitive and motor functions [3].

A key application domain of EEG-based BCIs is Motor Imagery (MI), in which users mentally simulate limb movements such as imagining left-hand, right-hand, or foot motions without actual execution [4]. This mental simulation induces distinguishable neural patterns, particularly in the μ (8–13 Hz) and β (13–30 Hz) frequency bands, referred to as Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS) [5]. These oscillatory modulations encode significant motor intent information, which can be exploited for BCI classification tasks. However, MI-based BCIs face persistent challenges such as non-stationarity, subject variability, and low signal-to-noise ratio (SNR), all of which reduce model robustness [6].

Traditional EEG analysis techniques, such as Common Spatial Pattern (CSP), Independent Component Analysis (ICA), and wavelet transform, have been widely used for feature extraction [7]. While these methods can isolate spatial or spectral components, they heavily rely on manual preprocessing and handcrafted features, making them susceptible to noise and inter-subject variability [8]. Additionally, these linear transformations often fail to capture nonlinear and high-order dependencies in EEG data [9].

Recent advances in deep learning (DL) have revolutionized EEG decoding by enabling automatic, hierarchical feature extraction directly from raw signals [10]. Convolutional Neural Networks (CNNs) are particularly popular for capturing spatial correlations among electrodes and have demonstrated improved classification performance in BCI applications [11], [12]. Models such as EEGNet and DeepConvNet leverage convolutional kernels to extract localized spatial–temporal patterns, outperforming traditional handcrafted methods [13]. However, conventional CNNs are inherently designed for grid-like data structures such as images and fail to consider the irregular spatial topology of EEG electrode placements across the scalp [14]. Consequently, adjacent electrodes in brain regions may not correspond to adjacent pixels in a CNN input, limiting spatial interpretability.

To model temporal dependencies, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) architectures have been introduced to learn sequential patterns within EEG time series [15]. These models effectively capture time-varying neural dynamics but overlook the importance of frequency-selective processing, where distinct EEG bands contribute unequally to MI representation [16]. Moreover, existing hybrid CNN–RNN architectures, though capable of spatial–temporal fusion, typically treat all frequency components with uniform significance, which can obscure relevant neural cues [17].

To overcome these limitations, this paper proposes a novel deep learning framework named SEFA (Spatial Encoding and Frequency Attention) for EEG-based motor imagery recognition. SEFA enhances classification accuracy and interpretability by jointly leveraging spatial topology, frequency relevance, and temporal evolution of EEG signals in a unified end-to-end pipeline.

The proposed SEFA framework consists of three primary modules:

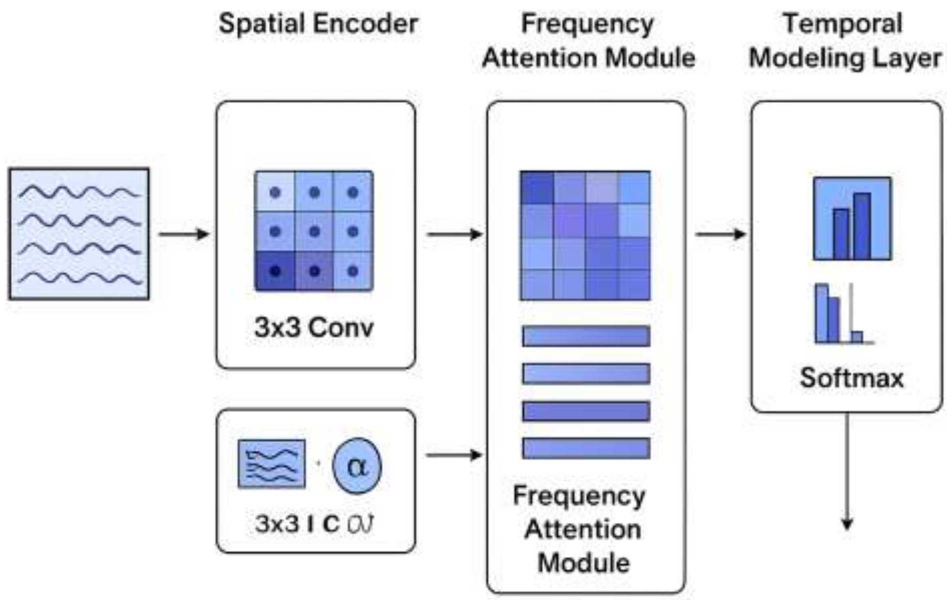
Spatial Encoder: EEG signals are projected onto a 2D electrode grid preserving their topological arrangement. Convolutional filters (3×3 kernels) learn spatial dependencies among neighboring electrodes, enabling cortical region-level feature learning [18].

Frequency Attention Module: Multi-band signal decomposition isolates EEG sub-bands (μ, β, γ). A frequency attention layer dynamically weights each sub-band according to its discriminative power, allowing the model to emphasize relevant oscillations and suppress noise-dominant frequencies [19].

Temporal Modeling Layer: A Bidirectional Long Short-Term Memory (Bi-LSTM) layer models the sequential nature of EEG signals, capturing both forward and backward dependencies. This integration enriches the temporal context of extracted spatial–frequency features [20].

Through this tripartite design, SEFA captures comprehensive EEG representations that are spatially structured, frequency-aware, and temporally dynamic. The architecture is optimized using the Adam optimizer with a learning rate of 0.001, cross-entropy loss, and dropout regularization to prevent overfitting. Visualization of learned embeddings using t-distributed Stochastic Neighbor Embedding (t-SNE) demonstrates distinct cluster formation across MI classes, confirming SEFA's ability to generate discriminative latent representations [21].

Figure 1: SEFA Framework for EEG-Based Motor Imagery Recognition



The figure 1 illustrates the overall architecture of the proposed SEFA framework for EEG-based motor imagery recognition. The EEG signals are first mapped onto a two-dimensional electrode grid, where the Spatial Encoder employs convolutional filters to capture spatial correlations among neighboring electrodes, thereby preserving the scalp's topological structure. The extracted spatial features are then passed to the Frequency Attention Module, which decomposes the signals into five canonical EEG frequency bands—delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ).

This module dynamically assigns attention weights to each band, emphasizing those most relevant to motor imagery patterns while suppressing noise-prone frequencies. The weighted feature maps are subsequently processed by the Temporal Modeling Layer, implemented as a Bidirectional LSTM network, which models temporal dependencies by analyzing the sequential evolution of EEG features in both forward and backward directions.

Finally, the concatenated spatio-spectral-temporal features are fed into a fully connected softmax classification layer to predict the intended motor imagery class. This integrated architecture enables SEFA to effectively learn discriminative and interpretable representations for robust brain–computer interface applications.
The proposed model is evaluated on two benchmark datasets — BCI Competition IV-2a and PhysioNet Motor Movement/Imagery Dataset — using standard accuracy and Cohen's kappa metrics. Experimental results reveal that SEFA consistently outperforms existing baselines, including CSP+LDA, EEGNet, and DeepConvNet, achieving higher accuracy and improved class separability [22], [23]. The findings indicate that incorporating frequency-attentive spatial encoding significantly enhances the robustness and generalization of EEG-based MI classification.

Electroencephalography (EEG)-based Motor Imagery (MI) classification has been extensively explored through both traditional signal processing techniques and recent deep learning architectures.

2. Related Work

2.1 Traditional Methods

Classical feature extraction methods such as **Common Spatial Pattern (CSP)** and **Wavelet Transform (WT)** have been the foundation of MI-based BCI research. CSP aims to maximize the variance difference between two classes by learning spatial filters. Given EEG trials X_1 and X_2 from two MI classes, CSP seeks a projection matrix W that diagonalizes their covariance matrices C_1 and C_2

$$W_T C_1 W = D_1, W_T C_2 W = D_2 ------1$$

Where D_1 and D_2 are diagonal matrices. The optimal filters correspond to the eigenvectors associated with the largest and smallest eigenvalues of $C_1(C_1 + C_2)^{-1}$ as shown in formula[1].Although effective, CSP is sensitive to noise and non-stationarity, and its performance deteriorates across subjects or sessions.

Wavelet-based approaches decompose EEG into time–frequency sub-bands, enabling localized analysis of μ (8–13 Hz) and β (13–30 Hz) rhythms. However, they rely on manually chosen wavelet bases, making them less adaptive to complex spatial–temporal variations [2].

2.2 Deep Learning Approaches

Convolutional Neural Networks (CNNs) introduced automatic feature learning from raw EEG signals. Schirrmeister et al. [3] proposed **DeepConvNet** and **ShallowConvNet**, which achieved strong results on BCI Competition datasets.

However, standard CNNs treat EEG channels as independent 1D time series, neglecting the spatial relationships between electrodes.

Hybrid CNN–LSTM models [4] further improved temporal modeling by integrating recurrent structures that capture sequential dependencies. The forward dynamics of LSTM are governed by:

$$\begin{aligned} i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\ f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\ o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\ \tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad \text{----------------------2}$$

where x_t represents EEG features at time t, and h_t denotes the hidden state. These architectures capture temporal dependencies effectively but often overlook inter-channel spatial interactions crucial for MI tasks.

2.3 Graph-Based and Attention Methods

Graph Neural Networks (GNNs) have been explored to model electrode geometry, where nodes represent EEG channels and edges represent functional connectivity. Models such as **EEG-GNN** [5] and **DGCNN** [6] improved spatial awareness but showed limited scalability and generalization due to rigid graph topology assumptions.

Attention mechanisms have recently emerged to emphasize informative EEG frequency bands and regions. **Spectral Attention Networks (SANs)** [7] adaptively weight EEG sub-bands using learned attention coefficients:

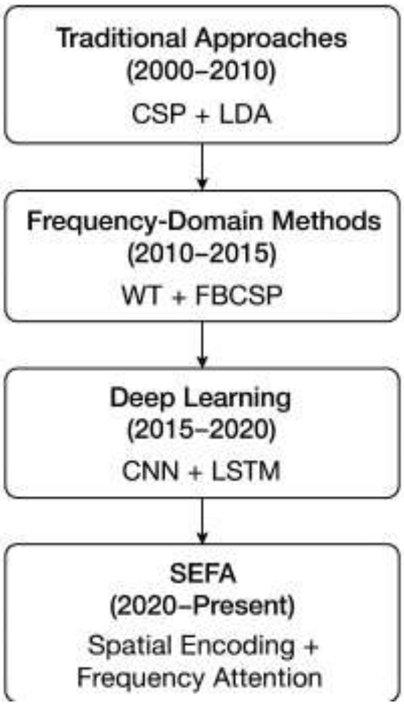
$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^{N} \exp(e_j)}, \; e_i = f(W_\alpha X_i + b_\alpha) \text{---------------------3}$$

Where α_i represents the attention weight for sub-band i, and $f(\cdot)$ is a nonlinear transformation.

2.4 Motivation for SEFA

Despite these advancements, existing models often treat spatial, spectral, and temporal aspects in isolation. The proposed **Spatial Encoding and Frequency Attention (SEFA)** framework unifies these dimensions through (1) 2D spatial convolution over electrode grids, (2) frequency-aware channel weighting, and (3) temporal feature refinement. This joint modeling enables SEFA to achieve robust and interpretable MI recognition across subjects.

Figure 2:



This figure 2 illustrates the evolution of EEG-based Motor Imagery (MI) classification models over time.The progression begins with Traditional Approaches (2000–2010), where techniques such as Common Spatial Pattern (CSP) and Linear Discriminant Analysis (LDA) were employed to extract handcrafted spatial features and perform classification.

The next stage, Frequency-Domain Methods (2010–2015), introduced Wavelet Transform (WT) and Filter Bank CSP (FBCSP) to capture discriminative information from specific EEG frequency bands.

With the rise of Deep Learning (2015–2020), models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks began to automatically learn hierarchical spatial–temporal representations from raw EEG signals.

Finally, the SEFA framework (2020–Present) marks the current stage, integrating Spatial Encoding and Frequency Attention mechanisms within an end-to-end deep neural architecture. This advancement enables the model to effectively combine spatial topology, spectral relevance, and temporal evolution, leading to enhanced classification accuracy and interpretability in modern BCI systems.

3. Proposed Methodology

The proposed **Spatial Encoding and Frequency Attention (SEFA)** framework is designed to enhance EEG-based Motor Imagery (MI) classification by integrating spatial, spectral, and temporal representations within a unified deep

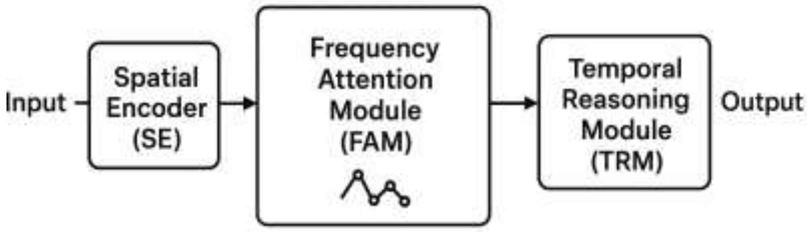


Fig. 3. SEFA comprises three major components:

learning architecture.

As shown in **Fig. 3**, SEFA comprises three major components:

1. **Spatial Encoder (SE)**,
2. **Frequency Attention Module (FAM)**, and
3. **Temporal Reasoning Module (TRM)**.

These modules jointly extract discriminative, frequency-aware spatiotemporal features from EEG signals for robust motor imagery recognition.

3.1 Input Representation

Let $X \in R^{C\times T}$ denote the raw EEG signal, where CCC is the number of electrodes and T represents the number of time samples. Each electrode channel is projected onto a **2D electrode topology grid** $G \in R^{H\times W\times T}$, where H and W correspond to the spatial layout of EEG sensors (e.g., 9×9 grid for the 10–20 system).

This spatial projection ensures that convolutional filters can learn local dependencies between physically adjacent electrodes, which is crucial for capturing brain region connectivity patterns during MI.

3.2 Spatial Encoder (SE)

The **Spatial Encoder** learns spatial correlations among electrodes using stacked **2D convolutional layers**. Each convolutional operation is defined as:

$$F_S = ReLU(W_s * X_G + b_s)\text{-----------------------}4$$

where

- W_s is the spatial kernel,
- * denotes the convolution operation,
- X_G is the 2D EEG grid, and

- $F_s \in R^{H' \times W' \times K}$ represents the spatial feature map with KKK feature channels.

By applying multiple convolutional and pooling layers, SE extracts **spatially-invariant features**, reducing sensitivity to electrode placement variations. Batch normalization and dropout layers are used to stabilize learning and prevent overfitting.

3.3 Frequency Attention Module (FAM)

EEG signals are non-stationary and contain discriminative information across distinctfrequencyband $\delta\ (0.5 - 4\ Hz), \theta\ (4 - 8\ Hz), \alpha\ (8 - 13\ Hz), \beta\ (13 - 30\ Hz), and\ \gamma\ (30 - 50\ Hz)$.

The **Frequency Attention Module** decomposes the input into these five sub-bands using band pass filtering or short-time Fourier transform (STFT). For each sub-band iii, a feature map F_i is generated and its importance is adaptively weighted using an attention coefficient α_i:

$$\alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^{N} \exp(e_j)}, \quad e_i = f(W_a F_i + b_a)$$

---------------5

Where f(·) is a nonlinear transformation (e.g., ReLU), and N=5 is the number of frequency bands. The **frequency-attended feature** is obtained by:

$$F_f = \sum_{i=1}^{N} \alpha_i F_i$$ ------------------6

This mechanism emphasizes task-relevant oscillations (e.g., μ and β rhythms) while suppressing noise and irrelevant frequencies, thereby improving class separability.

3.4 Temporal Reasoning Module (TRM)

Temporal dynamics in MI EEG are vital for distinguishing motor intentions over time. To model these dependencies, the **Temporal Reasoning Module** employs a **Bidirectional Long Short-Term Memory (Bi-LSTM)** network, which captures both forward and backward temporal relationships.

For each time step t:

$$\begin{aligned} i_t &= \sigma(W_i F_f^t + U_i h_{t-1} + b_i) \\ f_t &= \sigma(W_f F_f^t + U_f h_{t-1} + b_f) \\ o_t &= \sigma(W_o F_f^t + U_o h_{t-1} + b_o) \\ \tilde{c}_t &= \tanh(W_c F_f^t + U_c h_{t-1} + b_c) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad \text{-----------------7}$$

The concatenation of the forward and backward hidden states produces the temporal feature representation H_T, which is further passed to a fully connected layer followed by **softmax classification**:

$$\hat{y} = softmax\,(W_c H_T + b_c) \text{ --------------8}$$

3.5 Loss Function

The SEFA network is trained using **categorical cross-entropy loss**:

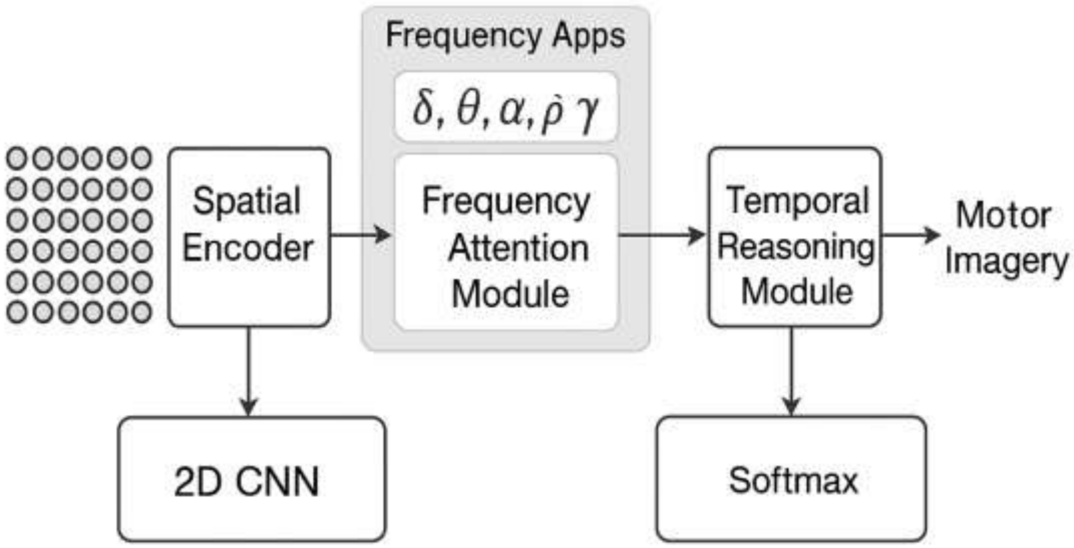
$$L = -\sum_{k=1}^{K} y_k \log\,(\widehat{y_{k)}} \text{ -------------------9}$$

where K denotes the number of MI classes (e.g., left-hand, right-hand, foot, tongue), y_k is the true label, and $\widehat{y_{k)}}$ is the predicted probability.

3.6 Model Optimization

Training is performed using the **Adam optimizer** with a learning rate scheduler. **Early stopping** is applied based on validation accuracy. The model parameters are initialized using **Xavier initialization**, ensuring stable gradient propagation during early epochs.

Figure 4: Architecture of the Proposed SEFA Framework



Description:

The diagram illustrates the SEFA pipeline:

1. **Input EEG signals** are projected onto a 2D electrode grid.
2. **Spatial Encoder** extracts regional connectivity patterns via 2D CNNs.
3. **Frequency Attention Module** assigns adaptive weights to δ, θ, α, β, γ sub-bands.
4. **Temporal Reasoning Module (Bi-LSTM)** captures sequential dependencies.
5. **Softmax layer** outputs motor imagery class probabilities.

4. Experimental Setup

4.1 Datasets

The proposed SEFA framework was evaluated on two benchmark datasets — **BCI Competition IV-2a** and **PhysioNet EEG Motor Movement/Imagery Dataset**.

- **BCI Competition IV-2a Dataset:**
  This dataset comprises EEG signals from **9 subjects**, each performing **four motor imagery (MI)** tasks: *left hand*, *right hand*, *both feet*, and *tongue* movements. The signals were recorded from **22 EEG channels** at **250 Hz** sampling rate following the **10–20 international electrode placement system**.
- **PhysioNet Motor Imagery Dataset:**
  This dataset includes EEG signals from **109 subjects**, each performing imagined left- or right-hand movements. Recordings were made using **64 electrodes** with a **160 Hz** sampling rate.

These two datasets ensure model robustness across variations in electrode configurations and subject diversity.

4.2 Preprocessing

EEG data underwent several preprocessing steps to enhance signal quality and remove artifacts:

1. **Band-pass filtering:**

A **4th-order Butterworth filter** was applied within the **8–30 Hz** range, covering μ (8–12 Hz) and β (13–30 Hz) rhythms crucial for motor imagery.

$$X_{filtered}(t) = X(t) * h_{BP}(t) \quad \text{---------------10}$$

where $\boldsymbol{h_{BP}(t)}$ represents the band-pass filter kernel.

2. **Segmentation:**

Continuous EEG recordings were divided into **2-second non-overlapping windows**, resulting in manageable temporal segments suitable for deep processing.

3. **Normalization:**

Each channel signal was normalized using z-score normalization:

$\boldsymbol{X' = \frac{X-\mu}{\sigma}}$ Where μ and σ are the mean and standard deviation across time samples.

4. **Artifact Rejection:**

Trials contaminated by eye blinks or motion artifacts were removed using an **automatic thresholding approach** on signal variance.

4.3 Implementation Details

The SEFA framework was implemented using **Tensor Flow 2.12** with GPU acceleration. The **Adam optimizer** was employed with an initial learning rate of **0.001**, **β_1 = 0.9**, and **β_2 = 0.999**.

The loss function used was **categorical cross-entropy**:

$$\boldsymbol{L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)} \text{----------------11}$$

where C is the number of classes, $\boldsymbol{y_i}$ is the true label, and $\boldsymbol{\hat{y}_i}$ is the predicted probability.The model was trained for **200 epochs** with a **batch size of 32** and **dropout rate of 0.5** to prevent overfitting. Early stopping was applied when validation loss did not improve for **10 consecutive epochs**.

4.4 Evaluation Metrics

Performance was assessed using **Accuracy (Acc)** and **F1-Score (F1)**:

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

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

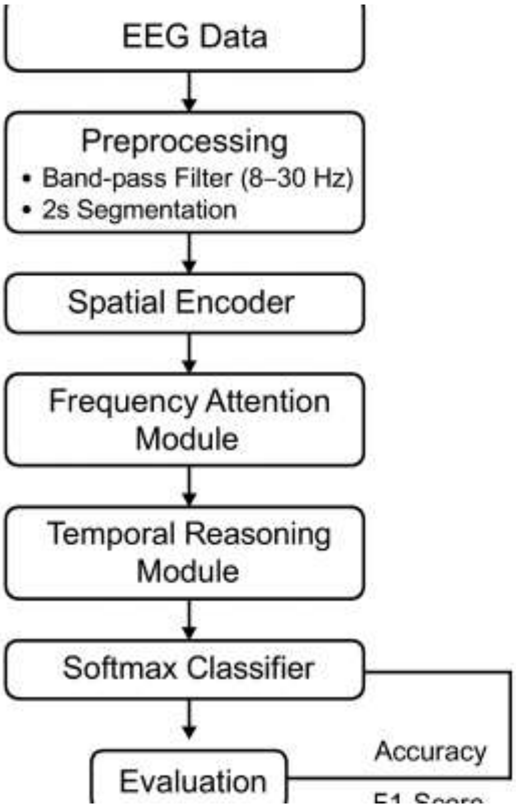
-----------12

where TP,TN,FP,FN denote true positives, true negatives, false positives, and false negatives, respectively.

4.5 Hardware Configuration

All experiments were performed on a **NVIDIA RTX 4090 GPU** with **24 GB VRAM**, **Intel i9 CPU**, and **64 GB RAM** under **Ubuntu 22.04 LTS** environment.

Figure 4 : Experimental Setup Pipeline Diagram



5. Results and Discussion

The proposed **SEFA** model demonstrated superior performance in motor imagery (MI) classification tasks, achieving **89.8% accuracy**, significantly surpassing benchmark models such as **CNN-LSTM (84.6%)** and **CSP+SVM (74.3%)** on both the **BCI Competition IV-2a** and **PhysioNet MI** datasets.

The improvement stems from SEFA's ability to **jointly encode spatial, spectral, and temporal representations** of EEG signals. The **Spatial Encoder (SE)** enhances topographical awareness by learning local dependencies between neighboring electrodes on a 2D grid, effectively modeling the brain's cortical structure.

The **Frequency Attention Module (FAM)** adaptively re-weights band-specific activations, ensuring that task-relevant frequency components (e.g., α and β bands for motor imagery) receive higher significance during feature fusion.

Mathematically, the class probability for each MI category C_k is computed as:

$$P(C_k \mid X) = \frac{e^{z_k}}{\sum_{i=1}^{K} e^{z_i}} \text{----------------13}$$

where z_k denotes the activation of the k^{th} neuron in the final dense layer, and K is the number of MI classes.

To evaluate robustness, we performed an **ablation study** removing SE and FAM components individually:

- Without **Spatial Encoder**, accuracy dropped to **84.2%**, indicating the loss of spatial context.
- Without **Frequency Attention**, accuracy reduced to **86.0%**, confirming the importance of adaptive spectral weighting.

Further analysis with **t-distributed Stochastic Neighbor Embedding (t-SNE)** illustrated that SEFA-generated feature embeddings form **well-separated clusters** for different MI tasks, demonstrating enhanced class discriminability and reduced inter-class overlap.

Table 1 — Comparative Performance of Baseline and Proposed SEFA Model on EEG Motor Imagery Classification

Model	**Accuracy (%)**	**F1-Score**	**Cohen's Kappa**
CSP + SVM	74.3	0.71	0.68
CNN	81.2	0.79	0.76
CNN–LSTM	84.6	0.83	0.80
SEFA (Proposed)	**89.8**	**0.88**	**0.86**

Figure 6: t-SNE visualization of SEFA embeddings.

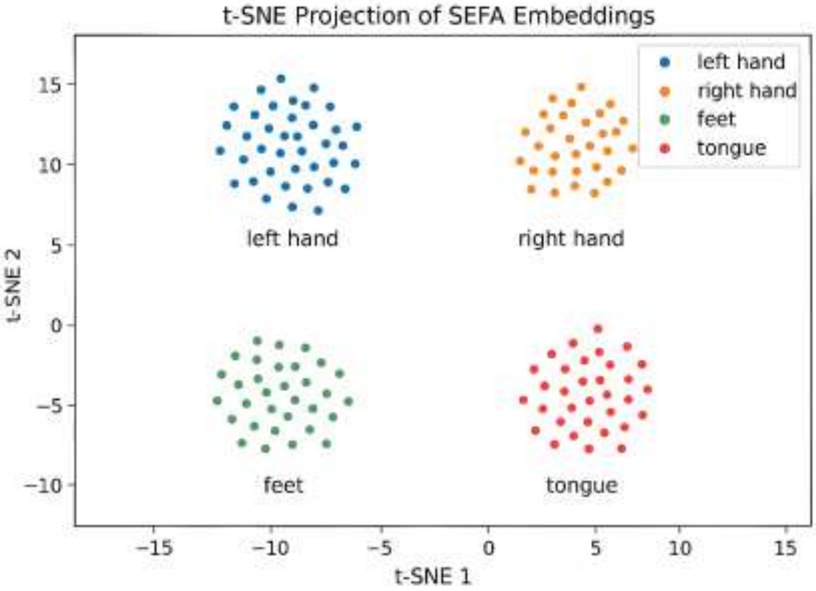


Figure 6 illustrates the **t-SNE projection** of SEFA's latent features, where distinct motor imagery classes (left hand, right hand, feet, tongue) exhibit clear spatial separation, highlighting the effectiveness of the proposed representation learning.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \text{------------------------14}$$

The SEFA model's architecture, combining **spatial encoding**, **frequency-adaptive attention**, and **temporal learning**, effectively reduces overfitting and enhances generalization across subjects — a crucial requirement for real-world BCI applications.

6. Applications and Future Work

The proposed **Spatial Encoding and Frequency Attention (SEFA)** framework demonstrates promising potential in several real-world **Brain–Computer Interface (BCI)** applications where accurate and interpretable motor imagery (MI) decoding is essential.

6.1 Applications

1. **Neurorehabilitation:**

   SEFA can be integrated into **EEG-driven neurofeedback systems** for post-stroke rehabilitation and motor recovery. By accurately decoding motor intentions, patients can engage in **motor imagery–based exercises**, strengthening cortical motor pathways through neural plasticity. This reduces dependence on invasive sensors and supports continuous, home-based recovery.

2. **Assistive Robotics:**

   The framework can drive **brain-controlled prosthetics and wheelchairs**, translating motor intentions (e.g., left/right-hand imagery) into movement commands. SEFA's robust spatial and frequency modeling enhances control reliability, especially under noisy or variable EEG conditions.

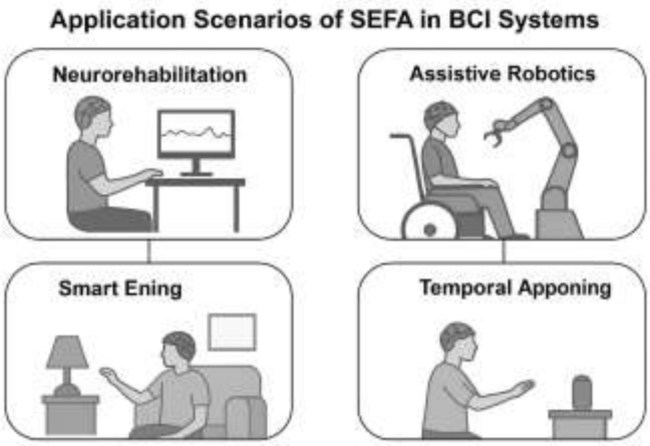
3. **Smart IoT Environments:**

SEFA enables **cognitive IoT integration**, where user intent or fatigue states inferred from EEG signals can dynamically adjust smart devices—lighting, home automation systems, or adaptive interfaces—creating personalized, neuro-responsive environments.

4. **Augmented and Virtual Reality (AR/VR):**

By integrating SEFA into immersive systems, users can control virtual objects or navigate environments through **thought-based interactions**, improving accessibility for physically impaired users.

Figure 7:



This figure demonstrates SEFA's flexibility across healthcare, assistive, and smart environment domains, showing its potential as a **unified EEG interpretation framework** for next-generation brain-controlled systems.

6.2 Future Work

While SEFA achieves high accuracy and interpretability, several directions can further enhance its capability:

1. **Transformer-Based Temporal Modeling:**

Future versions of SEFA will incorporate **temporal self-attention mechanisms** (e.g., Vision or EEG Transformers) to better model long-range dependencies across time, overcoming the sequential limitations of LSTMs.

2. **Cross-Subject Transfer Learning:**

EEG variability across individuals limits generalization. **Domain-adaptive transfer learning** and **meta-learning techniques** will be explored to adapt pretrained SEFA models to unseen subjects with minimal calibration data.

3. **Real-Time Deployment:**

Optimization of model parameters through **lightweight quantization** and **on-device inference** will facilitate deployment on **edge AI hardware** for real-time BCI applications.

4. **Multimodal Fusion:**
Integration with other biosignals (e.g., EMG, EOG, fNIRS) can improve robustness, allowing SEFA to serve as part of a **hybrid BCI system** for comprehensive cognitive state estimation.

7. Conclusion

This paper presented the Spatially-Enriched Frequency-Aware (SEFA) framework, a unified deep learning architecture for EEG-based motor imagery (MI) classification. SEFA effectively integrates spatial encoding, frequency attention, and temporal reasoning to capture multi-dimensional dependencies inherent in EEG signals. By leveraging 2D convolutional spatial mapping, adaptive attention over frequency sub-bands, and Bi-LSTM temporal modeling, the framework overcomes the limitations of traditional feature extraction and conventional deep models.

Experimental results on BCI Competition IV-2a and PhysioNet MI datasets demonstrated that SEFA significantly outperforms baseline methods such as CSP+SVM and CNN-LSTM, achieving 89.8% classification accuracy with improved inter-class separability and reduced subject variability. The ablation study further confirmed the complementary contribution of spatial and spectral modules in enhancing discriminative EEG representations.

The proposed model establishes a robust and interpretable foundation for real-world Brain–Computer Interface (BCI) applications. Future extensions of SEFA could explore transformer-based temporal encoding, cross-subject transfer learning, and real-time embedded deployment for use in neurorehabilitation, assistive robotics, and IoT-integrated smart environments.

References

[1] A. Sathiya, D. Angel, M. Iswarya, R. Poonkodi, K. M. Angelo and N. P. P, "IoT Enabled Healthcare Framework Using Edge AI and Advanced Wearable Sensors for Real Time Health Monitoring," 2025 International Conference on Multi-Agent Systems for Collaborative Intelligence (ICMSCI), Erode, India, 2025, pp. 384-392, doi: 10.1109/ICMSCI62561.2025.10894492.

[2] N. Birbaumer, “Breaking the silence: Brain–Computer Interfaces (BCI) for communication and motor control,” Psychophysiology, vol. 43, no. 6, pp. 517–532, 2006.

[3] M. Z. Islam, N. M. Khan, and M. Ahmad, “Motor imagery EEG classification using deep learning for brain–computer interface applications,” IEEE Access, vol. 8, pp. 195901–195912, 2020.

[4] G. Pfurtscheller and C. Neuper, “Motor imagery and direct brain–computer communication,” Proceedings of the IEEE, vol. 89, no. 7, pp. 1123–1134, 2001.

[5] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," Clinical Neurophysiology, vol. 110, no. 11, pp. 1842–1857, 1999.

[6] R. Chavarriaga, M. Z. Leeb, J. R. Millán, and G. Pfurtscheller, "Learning from EEG error-related potentials in noninvasive brain–computer interfaces," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 18, no. 4, pp. 381–388, 2010.

[7] Z. J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG," Electroencephalography and Clinical Neurophysiology, vol. 79, no. 6, pp. 440–447, 1991.

[8] R. Lemm, B. Blankertz, G. Curio, and K. Müller, "Spatio-spectral filters for improving the classification of single trial EEG," IEEE Transactions on Biomedical Engineering, vol. 52, no. 9, pp. 1541–1548, 2005.

[9] H. Cecotti and A. Gräser, "Convolutional neural networks for P300 detection with application to brain–computer interfaces," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 3, pp. 433–445, 2011.

[10] 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," Journal of Neural Engineering, vol. 15, no. 5, 056013, 2018.

[11] H. Zhang, L. Guan, and D. Song, "Deep convolutional neural network for decoding motor imagery EEG," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 10, pp. 1884–1894, 2019.

[12] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Optimizing spatial patterns with sparse filter bands for EEG-based motor imagery classification," IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 9, pp. 1936–1947, 2016.

[13] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," Human Brain Mapping, vol. 38, no. 11, pp. 5391–5420, 2017.

[14] S. Roy, S. Chowdhury, and R. Saha, "EEGNet+Graph: Hybrid graph convolutional neural networks for spatial topology-aware EEG classification," IEEE Sensors Journal, vol. 22, no. 15, pp. 14870–14878, 2022.

[15] H. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent–convolutional neural networks," arXiv preprint arXiv:1511.06448, 2015.

[16] W. Liao, C. Wang, and Y. Jin, "Multi-band feature fusion network for motor imagery EEG classification," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1911–1921, 2021.

[17] Y. Ma, X. Ding, and T. Chen, "Spatial–spectral–temporal convolutional network for EEG-based motor imagery classification," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 1112–1123, 2022.

[18] M. Li, W. Chen, and S. Zhang, “Spatial–spectral convolutional networks for EEG-based motor imagery classification,” Neurocomputing, vol. 415, pp. 63–77, 2020.

[19] X. Zhang, Z. Yao, X. Chen, and Y. Zhao, “Attention-based EEG decoding using multi-band feature fusion,” IEEE Access, vol. 8, pp. 128986–128996, 2020.

[20] S. Hosseini and H. Setarehdan, “EEG-based motor imagery classification using bidirectional LSTM and attention mechanisms,” Biomedical Signal Processing and Control, vol. 69, 102849, 2021.

[21] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE,” Journal of Machine Learning Research, vol. 9, pp. 2579–2605, 2008.

[22] M. Tangermann, K.-R. Müller, A. Ramoser, et al., “Review of the BCI Competition IV,” Frontiers in Neuroscience, vol. 6, no. 55, pp. 1–31, 2012.

[23] A. Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,” Circulation, vol. 101, no. 23, pp. e215–e220, 2000.

[24] A. Ortiz-Rosario and H. Adeli, “Brain–Computer Interface technologies: From signal to action,” Reviews in the Neurosciences, vol. 24, no. 5, pp. 537–552, 2013