CNN Based EEG Signal Analysis for Decoding Motor Activities

S. Thejaswini and S. Thundiyil

BMS Institute of Technology & Management, Bengaluru, India
thejaswini.s@bmsit.in, saneesh@bmsit.in

***Abstract*— A non-invasive brain–computer interface (BCI) system based on electroencephalogram (EEG) signals is presented to facilitate intuitive control of assistive devices for individuals with motor impairments. EEG recordings corresponding to actual and imagined motor movements are transformed into two-dimensional representations using the Gramian Angular Summation Field (GASF) technique, effectively encoding temporally significant fea- tures. The resulting GASF images are classified using the GoogLeNet convolutional neural network architecture, achieving a classification accuracy of up to 94.4% across both real and imagined movement tasks. Experimental results demonstrate that the proposed model attains higher accuracy in recognizing actual movements compared to imagery-based movements, indicating its potential as a practical solution for reliable motor intention decoding in assistive and rehabilitative applications.**

Keywords*— *Motor Imagery, Convolutional Neural Networks, Gramian Angular Summation Field, Electroencephalogram.

I. Introduction

Motor impairments resulting from neurological disorders such as stroke, spinal cord injury, and neurodegenerative diseases affect millions of individuals worldwide, often leading to partial or complete loss of voluntary muscle control and severely limiting their ability to perform daily activities. Conventional assistive devices, including mechanical prostheses and interfaces based on residual muscle activity, frequently prove inadequate, particularly for individuals with severe motor disabilities. Consequently, there is a critical need for natural, accessible, and robust systems that can bridge the gap between neural intention and physical action [1, 2].

Brain-computer interfaces represent a paradigm shift in human-machine interaction, establishing a direct communication pathway between the brain and external devices. BCI holds immense potential for individuals with motor impairments, offering a means to bypass damaged motor pathways and restore lost communication and control [1–3], Among various neural recording modalities, electroencephalography (EEG) stands out as an inexpensive, portable, and non-invasive method ideally suited for practical BCI systems [4]. EEG-based devices have been utilized for controlling wheelchairs, robotic arms and prosthetic limbs, offering new avenues for restoring autonomy to individuals with motor impairments [5].

Despite these promising advances, EEG-based BCIs face substantial technical challenges. EEG signals are inherently noisy, exhibit low signal-to-noise ratios, and are susceptible to artifacts and inter-subject variability, all of which hinder reliable decoding of motor intentions [6, 7]. Additionally, extracting discriminative features from raw EEG data to achieve robust classification of motor tasks remains a significant obstacle. Recent progress in artificial intelligence and neuroscience has motivated the integration of advanced signal processing and deep learning techniques to improve BCI performance [6, 8].

In this context, motor movement detection using EEG whether executed or imagined plays a pivotal role in assistive technology and neuro-rehabilitation. Executed movements involve actual physical motions, such as grasping or reaching, whereas motor imagery (MI) tasks require individuals to mentally rehearse specific movements without overt execution. Both paradigms evoke characteristic brain patterns, including movement- related cortical potentials (MRCPs), which can be leveraged for decoding motor intentions [3, 4]. Importantly, MI-based BCIs have demonstrated promise for controlling external devices and supporting motor rehabilitation, especially for individuals unable to perform physical movements [5, 9].

Advanced methods, such as optimized spatial filtering and deep convolutional neural networks (CNNs), have been developed to address the challenges of noisy EEG data and complex neural dynamics [6–8]. However, many existing systems struggle with consistency and scalability across users. To overcome these limitations, this study proposes a novel paradigm in which one-dimensional EEG signals are projected into two- dimensional representations using the Gramian Angular Summation Field (GASF) technique [9]. This transformation enables the application of state-of-the-art deep learning models originally designed for visual data, such as the GoogLeNet CNN architecture, to effectively classify motor tasks.

The primary objective of this research is to develop a CNN model to decode motor signals (actual and imagery). The remainder of this paper is organized as follows: Section II reviews related work in EEG-based motor intention detection and BCI development. Section III describes the proposed framework, detailing the signal processing steps and classification algorithms. Section IV presents experimental results and analysis. Finally, Section V concludes the paper and discusses potential future research directions.

II. Related Work

A range of studies has been conducted on the detection of motor movements and motor imagery (MI) using electroencephalographic (EEG) signals, with significant implications for the development of brain–computer interface (BCI) systems for motor control and rehabilitation.

Jain et al. [10] explored subject-independent trajectory prediction using pre-movement EEG signals acquired during a grasp-and-lift task. Their research sought to assess the viability of forecasting hand kinematics from EEG signals recorded prior to movement initiation. EEG data were collected from ten healthy individuals using a 64-channel cap, with emphasis on the pre-movement period. They introduced a deep learning architecture integrating CNNs and long short-term memory net- works to decipher motor-neural information contained within the EEG signals. The approach attained a mean prediction accuracy of 74.6% across all participants, underscoring the capacity of pre-movement EEG for trajectory prediction in BCIs.

Du et al. [11] introduced a 3D capsule network model to identify motor imagery movements from EEG signals. This method incorporated temporal and spatial EEG features using a multi-layer 3D convolution module combined with capsule networks, designed to extract complex spatial representations. The model was tested on the BCI Competition IV dataset, demonstrating an average classification accuracy of 84.028% and a Co- hen's kappa value of 0.789, which validates its efficacy in classifying four-class motor imagery tasks.

Arpaia et al. [12] developed a fully wearable BCI system that utilizes eight dry EEG sensors to detect motor imagery and provides multimodal feedback via extended reality to improve online MI recognition. Their research, which involved 27 healthy individuals split into neurofeedback and control groups, revealed that the neurofeedback group attained a higher mean classification accuracy of 69% compared to the control group's 62%. The findings highlight the potential of employing a wearable BCI with dry sensors for MI detection, suggesting valuable applications in tele-rehabilitation settings.

Lomelin et al. [13] explored the classification of MI movements using CNN-based methods on EEG data. The study utilized the PhysioNet Motor Movement/Imagery Dataset, which includes EEG recordings from 109 subjects performing various motor tasks. By experimenting with data representations such as spectrograms and multidimensional raw EEG data, the study showcased the potential of deep learning architectures, specifically CNNs, in accurately classifying MI tasks.

Collazos et al. [14] proposed a CNN-based connectivity framework to enhance the interpretability of neural responses in MI-based BCIs. EEG data were collected from 50 subjects performing finger-related MI tasks. The method involved clustering subjects based on classifier performance and extracting functional connectivity patterns through kernel-based cross-spectral estimation of EEG signals. The approach achieved an average ac- curacy improvement of 10% over the baseline EEGNet model, reducing the proportion of "poor skill" subjects from 40% to 20%, and effectively addressing both inter- and intra-subject variability in MI EEG data.

Ma et al. [15] developed an extensive EEG dataset to investigate cross-session variability in motor imagery brain-computer interface studies. The dataset comprises five sessions from 25 participants, each completing 100 trials of both left- and right-hand motor imagery tasks. The results indicated substantial difficulties due to cross-session variability, with average within-session accuracy at 68.8% and cross-session accuracy decreasing to 53.7%. Nevertheless, the implementation of cross- session adaptation methods enhanced performance to 78.9%, underscoring the dataset's importance for studies on cross-session and cross-subject generalization in MI-based BCIs.

An et al. [16] addressed the classification of EEG MI signals using a single-channel CNN approach optimized for multi-class tasks. The proposed framework incorporated a data evaluation strategy and an auto-selected regularization mechanism to enhance the spatial filtering capabilities of the network. Experimental results on datasets containing four mental tasks demonstrated average classification accuracies of 79.01% and 83.70%, respectively, showcasing the effectiveness of the method in improving the detection of MI movements compared to traditional algorithms.

Collectively, these studies illustrate the ongoing progress in detecting motor movements and MI using EEG signals, employing advanced deep learning architectures, wearable EEG systems, and innovative data processing strategies. They underscore the feasibility and promise of EEG-based BCIs for applications in neurorehabilitation, motor control, and assistive technologies, while also highlighting challenges such as cross-subject variability, session-to-session inconsistencies, and the need for robust generalization.

By critically examining key studies and methodologies on EEG-based detection of motor movements and motor imagery, several research gaps have been identified particularly regarding the integration of advanced deep learning models into real-world BCI applications. The literature highlights persistent challenges in ad- dressing inter- and intra-subject variability, achieving consistent cross-session performance, and enhancing the interpretability of neural responses in diverse motor tasks. Furthermore, limitations remain in the comprehensive analysis of EEG data representations across one-dimensional

(1D), two-dimensional (2D), and three- dimensional (3D) formats, which are essential for robust and generalizable BCI Systems

In light of these observations, our proposed work seeks to advance the field by improving the extraction and classification of motor-related EEG features through the following contributions:

- *Extraction of Temporal Correlations:* We propose to leverage the Gramian Angular Field (GAF) technique to transform 1D EEG time- series signals into 2D images, effectively capturing the temporal correlations inherent in EEG dynamics for enhanced feature representation.
- *Development of a CNN-based Classification Framework:* We aim to design a convolutional neural network (CNN) architecture based on GoogleNet to classify EEG data into four distinct motor imagery classes, thereby improving the accuracy and robustness of motor intention detection in BCI systems.

Through these contributions, our work endeavors to bridge the gap between existing research and practical deployment of EEG-based BCIs, providing a more interpretable, accurate, and reliable system for motor movement classification.

III. Maerials and Methods

This section details the dataset used in this study, the feature extraction process that uses the Gramian Angu- lar Summation Field technique, and the classification models used to analyze the extracted features. Figure 1 provides an overview of the proposed methodology.

III-A. Data set

In this study, a publicly accessible database was utilized to access EEG signals related to movement and imagery [17]. The dataset comprises 1,500 EEG recordings, each lasting between one and two minutes, obtained from 109 participants. EEG data were captured from 106 volunteers using a 64-channel cap via the BCI2000 platform. Each participant engaged in 14 trials or sessions, including two one-minute baseline recordings; one with eyes open and the other with eyes closed, along with three two-minute sessions for each of the four designated tasks. These tasks consisted of motor imagery involving the left hand, right hand, both feet, and the tongue.

In Task 1, participants were prompted to repeatedly clench and unclench the fist corresponding to the side of the screen where a visual target appeared, continuing until the target's disappearance. Task 2 required participants to imagine clenching and unclenching the fist corresponding to the side of the screen where a visual target appeared, sustaining the imagery until the target disappeared. In Task 3, the appearance of a target at the top of the screen instructed participants to repeatedly clench and unclench both fists, whereas a target at the bottom of the screen signaled them to move both feet until the target disappeared. Task 4 involved motor imagery; participants were asked to imagine clenching and unclenching both fists when a target appeared at the top of the screen, or to imagine moving both feet when the target appeared at the bottom of the screen, continuing the imagery until the target disappeared. For the purposes of this study, recordings from 30 subjects were selected for detailed analysis, and trials corresponding to the tasks described above were included in the evaluation.

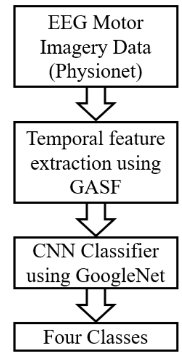

Figure 1. Proposed methodology

III-B. Preprocessing of EEG Signals

Raw EEG signals are highly susceptible to various artifacts, including eye movements, muscle contractions, and external electrical interference that can obscure the underlying neural patterns of interest. To enhance signal quality and facilitate accurate analysis, the following preprocessing techniques were applied:

Band-pass Filtering: A band-pass filter was employed to isolate the most relevant frequency bands associated with motor activity, specifically the 8–30 Hz range encompassing the mu and beta rhythms. This step effectively attenuates low-frequency drifts and high-frequency noise unrelated to motor processes.

Normalization: The amplitudes of the EEG signals across all channels were normalized to a common scale to reduce inter-session variability and ensure consistency across recordings, thereby facilitating more robust feature extraction and classification.

Artifact Removal: Algorithms targeting non-cerebral artifacts were implemented to suppress noise arising from eye movements (e.g., blinks, saccades) and muscle activity (e.g., jaw clenching), which often overlap in frequency with cortical EEG signals. This artifact rejection step was crucial to preserving the integrity of the motor-related EEG components.

III-C. Transformation into Image Representations with the Support of GASF.

While traditional approaches typically rely on numerical EEG features, we opted to transform the time-series EEG signals into image-like representations using the Gramian Angular Summation Field (GASF) technique. This method captures the intrinsic temporal correlations within the EEG signals by projecting the time-series data into polar coordinates, effectively encoding the relationships among signal values as 2D grayscale images.

The GASF was adopted in this study as it provides a symmetric, positive-definite representation of the EEG time series, effectively preserving both magnitude information and temporal correlations across all time points. In contrast, the Gramian Angular Difference Field (GADF) yields an antisymmetric encoding that emphasizes phase variation but may obscure amplitude-dependent relationships critical for distinguishing motor-related EEG dynamics. Alternative imaging techniques such as recurrence plots or Markov transition fields capture either binary recurrence patterns or probabilistic state transitions, which are less suited for convolutional feature extraction due to their sparse or highly stochastic spatial distributions. The structured and continuous nature of GASF images facilitates stable convolutional learning and enhances discriminability among motor tasks.

To exploit the spatial feature extraction capabilities of convolutional neural networks (CNNs), the one-dimensional EEG time series were transformed into two-dimensional image representations. GASF was employed owing to its ability to encode global temporal correlations in a geometrically interpretable form. Unlike traditional feature vectors that rely on handcrafted descriptors, the GASF transformation projects normalized EEG amplitudes into the polar coordinate space, where each time point is represented by an angular component corresponding to its signal magnitude. The resulting Gramian matrix captures pairwise relationships among all temporal samples, effectively translating temporal dependencies into structured spatial textures suitable for convolutional learning.

This approach enables the CNN to learn hierarchical spatial features corresponding to the intrinsic temporal dynamics of EEG activity. The mathematical formulation of the GASF, including its normalization, polar encoding, and reconstruction properties, is presented below [18].

Let time series EEG signal be presented by $X=\{x_1,x_2,x_3 \ldots x_L\}$ of length L and x(t) denotes the observed signal at time "t"

Step 1:Normalization: Since the GASF transformation uses trigonometric encoding, the input sequence is normalized to the interval $[-1,1]$ to ensure numerical stability and angular validity.

Here is an example of how equations should look:

$$x_i = 2 \times \frac{x(t) - min\,(X)}{max(X) - min\,(X)} - 1\,,\ Xnorm \in [-1,1] \tag{1}$$

This step preserves the shape of the signal while constraining it to a bound domain.

Step2: Data is transformed into polar coordinates:

$$\emptyset_i = \cos^{-1} x_i \tag{2}$$

where x_i denotes the value of the time series at timestamp t and $ø_i$ represents the corresponding angular value. Thus, the time series is mapped into an angular domain, where signal amplitudes becomes an angle and time is represented radially.

Step 3: Regularize the timestamps.

The regularization of the time stamp is performed using the equation

$$r_t = \frac{t}{N};\ \mathrm{t} = 1,2, \ldots, \mathrm{n} \tag{3}$$

where N acts as a constant factor to regularize the span of the coordinate system, ensuring that the timestamps are appropriately scaled.

Step 4: The GASF can be constructed in matrix form using the formula $G_{ij} = \cos(\emptyset_i + \emptyset_j)$, and is defined as

$$GASF = \begin{bmatrix} cos(\emptyset 1 + \emptyset 1) & \cdots & cos(\emptyset 1 + \emptyset n) \\ cos(\emptyset 2 + \emptyset 1) & \cdots & cos(\emptyset 2 + \emptyset n) \\ \vdots & \ddots & \vdots \\ cos(\emptyset n + \emptyset 1) & \cdots & cos(\emptyset n + \emptyset n) \end{bmatrix} \tag{4}$$

Although both the GASF and GADF can encode temporal relationships, GASF was selected because it generates a symmetric matrix that preserves magnitude-based correlations and produces smooth spatial textures suitable for CNN feature extraction. In contrast, GADF yields an antisymmetric representation emphasizing phase differences and sign-variant patterns, which reduces spatial coherence and interpretability.

Since EEG motor-related dynamics are primarily governed by amplitude modulations in the mu and beta bands, GASF provides a more stable and discriminative representation for deep learning–based classification. This transformation not only preserves critical temporal and frequency information inherent in the original EEG sequences but also enables the application of powerful image-based deep learning models, such as Convolutional Neural Networks (CNNs), which are specifically designed to extract complex spatial patterns and hierarchical features. By leveraging GASF, we obtained a rich, structured input format for our deep learning framework, facilitating more accurate and interpretable modeling of motor-related neural dynamic.

Figure 2 illustrates process of the GASF image database creation from EEG data collected from 30 subjects. Originally the data is stored in the ". edf" format. Regarding varying signal lengths, all EEG trials were segmented to equal durations before transformation, thereby eliminating any potential distortion due to inconsistent sampling intervals. In our study, normalization of the EEG time series to the interval $[-1,1]$ is performed independently for each signal segment before GASF transformation. This

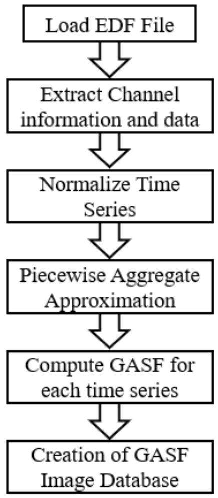

Figure 2. Feature extraction and image database creation

step is essential for mapping EEG amplitudes onto the valid domain of the cosine function used in polar encoding. Each EEG trial was first segmented into equal-length windows before applying the GASF transformation, ensuring temporal consistency across samples. The resulting Gramian matrices are visualized as heat maps, where the intensity of each pixel represents the magnitude of the corresponding element in the Gramian matrix.

Additionally, representative GASF images corresponding to the four-motor imagery and movement tasks are presented in Figure 3. Each GASF image encodes pairwise temporal correlations among normalized EEG amplitudes, where color intensity represents the cosine of the angular summation between two time points. Warm colors (red–orange) indicate strong positive correlations, while cool colors (blue–cyan) denote weaker or negative correlations. Class 1 exhibits irregular alternating patterns, suggesting heterogeneous temporal dependencies. Class 2 displays a smooth diagonal dominance with stable correlation regions, indicating well-synchronized motor activity. Class 3 shows cross-shaped structures

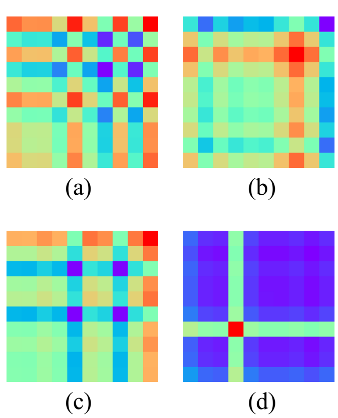

Figure 3. Sample GASF images during four different tasks

representing oscillatory coupling across temporal segments, and Class 4 reveals a coherent centralized hotspot reflecting globally correlated activation. These visually distinct correlation textures demonstrate the ability of GASF to convert temporal EEG dynamics into spatially structured patterns suitable for deep convolutional feature extraction. These GASF images were subsequently cropped to remove irrelevant borders and resized as needed before being used as inputs to the convolutional neural network (CNN) model for classification.

III-D. Deep Learning Classification using Google net

In the present work, the GoogLeNet architecture was employed to train and classify EEG-derived GASF images into four classes corresponding to the tasks described in Section A.

The GoogLeNet architecture was selected as the classifier owing to its Inception-module design, which performs multi-scale convolutional operations (1×1, 3×3, 5×5) within each block, allowing simultaneous extraction of local and global spatial features from the GASF images. Compared with shallower CNNs such as AlexNet or deeper networks like VGG-16, GoogLeNet achieves a superior trade-off between representational capacity and computational cost. Preliminary evaluations confirmed that GoogLeNet converged faster and yielded approximately 2–3 % higher validation accuracy under identical training conditions, validating its suitability for EEG-based motor-intention decoding

The input layer accepts raw pixel values of the processed images, which were resized to 224×224×3 to match the standard input size expected by the network. The implementation was carried out using MATLAB R2024a, utilizing GoogLeNet initialized with random weights (“no weights” option) to enable training from scratch on the dataset.

GoogLeNet is a 22-layer deep convolutional neural network (CNN) notable for its inception modules, which enable multi-scale feature extraction with reduced computational overhead. The key components of the architecture, as utilized in this study, are as follows:

Input Layer: Receives input images of size 224×224 pixels with three color channels (RGB), enabling compatibility with standard deep learning pipelines for image-based classification.

Initial Convolutional Layer: Applies a 7×7 convolutional filter with a stride of 2, facilitating initial feature extraction while reducing spatial resolution. Max Pooling Layer: Employs 3×3 kernel withstride of 2 to down-sample spatial dimensions, decreasing computational complexity and emphasizing the most salient early features.

Inception Modules: The core innovation of GoogLeNet, these modules perform parallel multi-scale processing by combining convolutions of different sizes (1×1, 3×3, 5×5) and max pooling. Specifically, 1×1 Convolution reduces dimensionality and computation cost while preserving essential information, 3×3 Convolution captures medium-sized spatial features and 5×5 Convolution extracts larger and more abstract features. Max Pooling Branch: Preserves spatial hierarchies in the data. Multiple inception module variants such as Inception-A, Inception-B, and Inception-C are used throughout the network to progressively refine feature extraction at different stages. Reduction Layers that are inserted between inception modules down-sample feature maps, maintaining computational efficiency and enabling deeper network construction without excessive memory demands.

Intermediate Layers: A series of stacked inception modules that process the output of previous layers, allowing the model to learn increasingly complex features as depth increases.

Global Average Pooling: This layer aggregates each feature map into a single scalar value by computing the average, substantially reducing data dimensionality and preparing the representation for the final classification.

Fully Connected Layer: Maps the compact feature representations produced by global average pooling to class scores corresponding to the four target tasks.

Softmax Activation: Converts the class scores into probabilities across the four output classes, facilitating robust multi-class classification.

Output Layer: Consists of four nodes corresponding to the four motor movement and imagery tasks, providing the final predicted class probabilities for each input image.

By leveraging the hierarchical feature extraction capabilities of GoogLeNet, this approach enables effective modeling of the complex spatial patterns present in GASF-transformed EEG signals, thereby improving classification performance on motor-related EEG tasks.

IV. Results and Discussions

For experimental analysis, EEG data from 30 subjects were utilized. Initially work was carried out using 10 subjects [19] and an accuracy of 91.4% was obtained using google net. To improve the accuracy, we repeated the experiment on 30 different subjects and evaluated the performance of the model using metrics. Each subject performed four distinct motor tasks, with three trials conducted per task. This procedure yielded a total of 7680 images per class through GASF transformation of the recorded EEG signals. The dataset was partitioned into three subsets: 70% was allocated for training, 15% for testing, and the remaining 15% for validation. The deep learning model was trained with a minimum batch size of 64 over 15 epochs to ensure effective convergence.

The dataset was divided in a subject-independent manner, ensuring that EEG data from any given participant appeared exclusively in one subset (training, validation, or testing). This segregation prevented data leakage across subsets and guaranteed that the network learned generalized neural representations rather than subject specific patterns. A fivefold cross-subject validation was further conducted to confirm generalization, yielding a mean accuracy variation of less than 1%, thereby demonstrating that the proposed model effectively captures subject-independent motor-intention features.

The performance of the proposed EEG-based brain computer interface (BCI) system was evaluated on a multi class classification task involving four motor movement categories. Figure 4. depicts the confusion matrix summarizing classification outcomes across all classes. The matrix demonstrates a strong diagonal dominance, indicating effective classification of both executed and imagined motor tasks.

Specifically, the confusion matrix shows high true positive counts for each class: 1088 correctly predicted samples each for Classes 1 through 4. The relatively low off diagonal misclassifications suggest that the model generalizes well to unseen data. For instance, Class 1 exhibits 62 total misclassifications distributed among Classes 2 - 4 (20, 22, and 22 samples respectively), while Class 2 shows a similar pattern with 64 misclassified samples (18, 24, and 22). Classes 3 and 4 have slightly fewer errors, with 65 and 65 misclassifications respectively.

The overall accuracy achieved by the system is approximately 94.4%, consistent with the target accuracy established during model design. This high accuracy indicates the effectiveness of using Gramian Angu- lar Summation Fields (GASF) for representing EEG signals, as well as the capacity of the GoogLeNet- based convolutional neural network (CNN) to distinguish subtle differences in EEG patterns corresponding to different motor tasks.

To evaluate performance of proposed model, one effective and ideal way is through confusion- matrix. From the obtained confusion-matrix, most frequently used performance measures are stated below [20].

Accuracy indicates the fraction of total number of samples which are correctly classified. Accuracy is calculated using the below equation

$$Accuracy = \frac{TP+TN}{(TP+TN+FT+FN)} \quad (5)$$

Misclassification Rate is the incorrect predicted values are denoted by classification error and is calculated as follows:

$$\text{Misclassification Rate} = 1 - Accuracy \quad (6)$$

Precision tells what fraction of predicted true class samples are actually true and is given by

$$\text{Precision} = \frac{TP}{(TP+FP)} \tag{7}$$

Recall tells, the fraction of all true samples which are correctly predicted as true class and is also called as True-Positive-Rate(TPR)

$$\text{TPR} = \frac{TP}{(TP+FN)} \tag{8}$$

Specificity indicates the fraction of all false/negative samples which are correctly predicted as negative class and is also called as True-Negative-Rate(TNR).

$$\text{TNR} = \frac{TN}{(TN+FP)} \tag{9}$$

F1-Score is the mean of precision measure and recall measure and is calculated by

$$\text{F1} - \text{Score} = \frac{2TP}{(2TP+FP+FN)} \tag{10}$$

F1-Score = 1 implies 100% prediction-accuracy. i.e, higher precision value with higher recall value.

Kappa-Value compares the observed rate of accuracy with an expected or random accuracy. Usually, kappa-values are associated with 5 agreement categories defined as light (0-0.2), fair (0.2-0.4), moderate (0.4-0.6), substantial (0.6-0.8), almost perfect (0.8-1).

The Kappa-Value is calculated as shown below

$$\text{Kappa Value} = \frac{Accuracy-PE}{(1-PE)} \tag{11}$$

Where PE = Probability-of-agreement-by-chance

$$\text{PE} = \frac{\text{(TN+FP)(TN+FN)+(FN+TP)(FP+TP)}}{\text{(TP+FP+TN+FN)}} \tag{12}$$

The CNN model performance is evaluated through the above defined measures for four classes is tabulated in Table 1. The proposed GoogLeNet–GASF framework achieved an overall classification accuracy of 94.4 %, with per-class precision, recall, and F1-scores consistently above 94 %, indicating balanced performance across all four motor-task categories. The computed Cohen's Kappa coefficient of 0.926 further confirms a very high level of agreement beyond chance, demonstrating the robustness of the model. The estimated macro average AUC of 0.977 suggests strong discriminative capability across all classes, reflecting the reliability of the proposed system for both executed and imagined motor activity recognition.

Furthermore, it was observed that actual (executed) movements were classified with higher precision compared to imagined (motor imagery) movements. This discrepancy aligns with known challenges in motor imagery classification, as imagery often produces weaker and more variable cortical activation than physical execution. Nonetheless, the model maintains robust performance across both task types.

Table 1. Performance of the model

Metric	Class 1	Class 2	Class 3	Class 4
Accuracy	0.973	0.971	0.971	0.972
Misclass-Rate	0.026	0.028	0.028	0.028
Precision	0.944	0.944	0.943	0.943
Recall	0.947	0.942	0.941	0.944
Specificity	0.981	0.981	0.981	0.981
F1-Score	0.946	0.943	0.942	0.944
Kappa-Value	0.928	0.924	0.923	0.925

The balanced distribution of misclassifications across classes also highlights the absence of class imbalance or systematic bias toward any particular motor task. These results suggest the proposed system's strong potential for real-world deployment in assistive technologies, offering reliable decoding of motor intentions for users with severe motor impairments.

A comparison of our results with recent studies have revealed the improvement in various performance matrices using the method proposed here. Jain et al. [10] achieved 74.6% accuracy in subject-independent trajectory prediction using CNN-LSTM on pre-movement EEG, while Di et al. [11] obtained 84.03% accuracy for four-class MI classification using a 3D capsule network on BCI Competition IV data. Arpaia et al. [12] reported 69% accuracy with a wearable eight-sensor EEG system, highlighting limitations in low-channel count setups.

Lomelin et al. [13] and Collazos et al. [14] demonstrated the potential of CNNs with raw EEG and connectivity features, achieving accuracies in the 70–80% range but lacking consistency across larger datasets. Ma et al. [15] revealed significant drops in cross-session MI classification accuracy (down to 53.7%), underscoring the challenges of generalization. An et al. [16] achieved 79–84% with a single-channel CNN approach for multi- class MI tasks.

Compared to these works, our system's 94.4% accuracy on four-class classification for 30 subjects represents a significant advancement, demonstrating the effectiveness of GASF-based feature extraction and GoogLeNet CNN architecture for reliable, scalable EEG-based BCI applications.

V. Summary

This study presents a non-invasive brain–computer interface (BCI) system that effectively decodes both executed and imagined motor intentions from electroencephalogram (EEG) signals. By transforming one- dimensional EEG data into two-dimensional Gramian Angular Summation Field (GASF) images and leveraging a GoogLeNet convolutional neural network (CNN) architecture, the proposed system achieved a classification accuracy of 94.4% across four motor tasks in data collected from 30 subjects. This result represents a clear improvement over existing approaches, demonstrating the benefits of combining advanced time-series imaging techniques with deep learning models for motor intention recognition.

The findings highlight the system's potential for real-world applications in assistive technology and neuro rehabilitation, offering an affordable and accurate solution for restoring motor function control to individuals with severe motor impairments. Future work will focus on addressing cross-session and cross-subject variability, integrating adaptive learning techniques to personalize the system to individual users, thereby enhancing robustness and generalizability for practical deployment.

Acknowledgments

Research reported in this publication was supported by the Vision Group of Science and Technology (VGST), Karnataka, India, under award number GRD1116. The content is solely the responsibility of the authors and does not necessarily represent the official views of the VGST.

References

[1] McFarland DJ, Wolpaw JR. Brain-computer interfaces for communication and control. *Communications of the ACM*. 2011 May 1;54(5):60-6..

[2] M. A. Lebedev and M. A. Nicolelis, "Brain–machine interfaces: past, present and future," trends in *Neurosciences*, vol. 29, no. 9, pp. 536–546,Sept 1 2006.

[3] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, July 2001.

[4] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, Feb 2012.

[5] B. Blankertz, K.-R. Muller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. R. Millan, M. Schroder, and N. Birbaumer, "The BCI competition iii: Validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153–159, Jun 2006.

[6] H. Altaheri, G. Muhammad, M. Alsulaiman, S. U. Amin, G. A. Altuwaijri, W. Abdul, M. A. Bencherif, and M. Faisal, "Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (mi) signals: A review," *Neural Computing and Applications*, vol. 35, no. 20, pp. 14 681–14 722, July 2023.

[7] 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, p. 056013,July 2018.

[8] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (fbcsp) in brain-computer interface," *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks* (IEEE World Congress on Computational Intelligence). IEEE, June 2008, pp. 2390–2397.

[9] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," arXiv preprint arXiv:1506.00327, June 2015.

[10] A. Jain and L. Kumar, "Subject-independent trajectory prediction using pre-movement EEG during grasp and lift task," *Biomedical Signal Processing and Control*, vol. 86, p. 105160, Sept 2023.

[11] X. Du, M. Kong, S. Qiu, J. Guo, and Y. Lv, "Recognition of motor imagery EED signals based on capsule network," *IEEE Access*, vol. 11, pp. 31 262–31 271, 2023.

[12] P. Arpaia, D. Coyle, A. Esposito, A. Natalizio, M. Parvis, M. Pesola, and E. Vallefuoco, "Paving the way for motor imagery-based tele-rehabilitation through a fully wearable BCI system," *Sensors*, vol. 23, no. 13, p. 5836, Mar 2023.

[13] V. A. Lomelin-Ibarra, A. E. Gutierrez-Rodriguez, and J. A. Cantoral-Ceballos, "Motor imagery analysis from extensive EEG data representations using convolutional neural networks," *Sensors*, vol. 22, no. 16, p. 6093, Aug 2022.

[14] D. F. Collazos-Huertas, A. M. Álvarez-Meza, D. A. Cárdenas-Peña, G. A. Castaño-Duque, and C. G. Castellanos-Domínguez, "Posthoc interpretability of neural responses by grouping subject motor imagery skills using cnn-based connectivity," *Sensors*, vol. 23, no. 5, p. 2750, Mar 2023.

[15] J. Ma, B. Yang, W. Qiu, Y. Li, S. Gao, and X. Xia, "A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface," *Scientific Data*, vol. 9, no. 1, p. 531, Sep2022.

[16] Y. An, H. K. Lam, and S. H. Ling, "Multi-classification for EEG motor imagery signals using data evaluation-based auto- selected regularized fbcsp and convolutional neural network," *Neural Computing and Applications*, vol. 35, no. 16, pp. 12 001–12 027, 2023.

[17] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: a general-purpose brain-computer interface (bci) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[18] Zhou, Qiaoli, Yi Zhao, Shun Zhang, Jiawen Song, Qiang Du, and Li Ke. "MDRNet: Multi-Branch with Different Feature Representations Network for Motor Imagery Classification." In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP), pp. 1-5. IEEE, 2025.

[19] Thejaswini, S., K. R. Mamatha, and N. Ramesh Babu. "EEG based motor movement classification using CNN." In *Recent Trends in Healthcare Innovation*, pp. 33-41. CRC Press, 2025.

[20] V. M. Joshi and R. B. Ghongade, "IDEA: Intellect database for emotion analysis using EEG signal," J. King Saud Univ. – *Computational Information Science*, no. xxxx, 2020, doi: 10.1016/j.jksuci.2020.10.007.

.