Benchmarking and Analysis of Deep Learning Methods to Classify EEG Motor Movement / Imagery Signals

Jayanth Jayakumar and **Sriram Acharya** and **Jack Praveen Raj Ilango** and
Sathiya Murthi Sankaran and **Pragatheeswaran Ravichandran** and **Neranjhana Ramesh**
University of Southern California, Los Angeles
{jjayakum, sriramac, jilango, sathiyam, pragathe, neranjha}@usc.edu

Abstract

The study aims to perform an analysis of the techniques and models available for the classification of Motor Movement / Imagery using Electroencephalogram (EEG) signals, testing general Deep Learning and Machine Learning model approaches to EEG classification, previously published models, and potential enhancements to these techniques. It aims to explore the various EEG Motor Imagery datasets available, different prepossessing and augmentation techniques, EEG signal embedding approaches, and test the model robustness to the addition of noise and artificially generated EEG data. The project focuses on the BCI Competition IV 2A datasets for Motor Imagery classification.

1 Introduction

The electroencephalogram (EEG) represents the brain's electrical activity commonly employed in clinical and research contexts. Classification of EEG data to determine various brain signal processes, such as motor movements or imagery, is essential for advances in Neurotechnology and brain-computer interfaces (BCIs). Classification of EEG signals can play a cardinal part in applications like Cognitive Load Monitoring, Medical Diagnosis of brain traits, Brain-Computer Interfaces, Neurofeedback, Biometric Authentication etc. The development of robust and efficient deep learning models for EEG analysis holds significant potential for advancing assistive technologies, rehabilitation, and human-machine interaction.

Machine Learning and Deep Learning have proven to be a valuable approach to perform EEG signal classification. Multiple techniques and models have been developed and tested on this task[9][13]. It becomes crucial to analyze the relative performance of these techniques and their robustness and provide a benchmark on the available approaches.

2 Background

2.1 EEG Signals

Electroencephalography (EEG) is a non-invasive technique widely used to monitor brain activity by measuring electrical signals generated by neurons. Electrodes are placed on the scalp which capture these signals, which represent the synchronized firing of neural populations. EEG recordings provide valuable insights into brain function, enabling researchers to study cognitive processes, neurological disorders, and brain-computer interfaces (BCIs). EEG's versatility allows for diverse applications, including diagnosing epilepsy, assessing sleep patterns, and investigating cognitive impairments. Moreover, its real-time monitoring capabilities make EEG instrumental in developing BCIs, enabling individuals to control devices using only their brain activity.

2.2 Datasets

There are various open corpus available used for EEG signal classification. Table 1 portrays several datasets which has comprehensive collection of EEG signal for Motor Imagery signals.

3 Current Progress

3.1 Literature Survey

- **[1] Vernon J Lawhern et al.** presents EEGNet, a compact convolutional neural network for EEG classification and evaluates it's accuracy and computational efficiency against other state-of-the-art methods in BCI tasks. It is also designed to be flexible and adaptable to different types of EEG-based BCI tasks. Its architecture can be easily adjusted to fit the specific requirements of different applications. Furthermore, it discusses the potential of EEGNet in practical BCI applications by highlighting its advantages in terms of speed,

Dataset Name	Subjects	Channels	Tasks/Description
Left/Right Hand MI	Includes 52 subjects	3D EEG electrodes	non-task related states
Motor Movement./Imagery Dataset	Includes 109 volunteers	64 electrodes	2 baseline tasks, motor movement, and imagery
Grasp and Lift EEG	12 subjects	32ch @ 500 Hz	for 6 grasp and lift events
Largest SCP Motor-Imagery	13 participants	BCI up to 6 states	60 hrs EEG recordings, 75 sessions
BCI Comp. IV-1	7 subjects	64 EEG ch at 1000 Hz	classes of left hand, right hand, foot
BCI Comp. IV-2a	9 subj., 2 sessions	22-electrode EEG	288 trials of imagined movements
BCI Comp. IV-2b	9 subj., 5 sessions	3-electrode EEG	imagined movements left/right hand
High-Gamma Dataset	14 subjects	128-electrode set	1000 four-second trials, 13 runs
L/R Hand 1D/2D movements	1 subject	19-electrode data	1D and 2D hand movements execution
Right-hand Thumb Movement	Single subject	8 electrodes at 256 Hz	5s epoch for imagined movement

Table 1: MI-EEG Datasets

accuracy, and computational requirements.

- **[2] Hermosilla et al.** present a study on the classification of motor imagery EEG for BCI applications using a Shallow Convolutional Network (SCN). Their research tests an end-to-end shallow architecture comprising two convolutional layers, hypothesizing that it could enhance EEG classification with fewer calibration stages. The system shows comparable and superior results to state-of-the-art on three public datasets, indicating promise in clinical applications for BCIs with minimal session-by-session calibration.

- **[3] Nijisha Shajil et al.** explore the use of transfer learning with CNN models such as AlexNet, ResNet50, and InceptionV3 for classifying EEG signals for motor imagery. Their findings highlight InceptionV3 as the most accurate model, indicating its potential in Brain-Computer Interface applications for individuals with neuromuscular disorders.

- **[4] Karel Roots et al.** introduced a novel multi-branch 2D convolutional neural network, EEGNet Fusion, for cross-subject EEG motor imagery classification. Their study uses the PhysioNet EEG Motor Movement/Imagery dataset and demonstrates that EEGNet Fusion achieves significantly higher accuracy (84.1% for executed movements and 83.8% for imagined movements) compared to other state-of-the-art models like EEGNet, ShallowConvNet, and DeepConvNet. However, the proposed model has a higher computational cost. This research contributes to the field by offering a more effective approach for cross-subject EEG classification, which can be crucial for brain-computer interface development.

- **[5] Radia Rayan Chowdhury, et al.** Developed for EEG-based brain-computer interface applications, EEGNet Fusion V2 is a multi-branch 2D convolutional neural network presented. With greater accuracy across many datasets—89.6% for actual motor activities and 87.8% for imagined motor activities on the eegmmidb dataset. With a higher computational cost, EEGNet Fusion V2 performs better than previous models such as EEGNet, ShallowConvNet, and DeepConvNet.

- **[6] Danilo Avola et al.** present a benchmark study on Machine and Deep Learning for EEG signal classification. They used four widely known models: Multilayer Perceptron (MLP), Convolutional neural network (CNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU), to determine which model could be a good starting point for developing EEG classification models. The study highlights the importance of choosing the right model for specific problems and data types in EEG signal analysis.

3.2 Identification of models and approaches

To perform the benchmark on the classification task, we have identified the following approaches:

- Basic Deep Learning Approaches for EEG Classification:
  - MLP, CNN, LSTM, CNN LSTM [6]
- Previous models fine-tuned for the task:
  - EEGNet [1]
  - ShallowConvNet [2]
  - DeepConvNet [3]
  - FusionNet [4]
  - Multi-Branch CNN [5]
- Preprocessing approaches for EEG signals:
  - Artifact removal
  - Bandpass filtering
  - Epoching
  - Baseline correction
  - Normalization
  - Spatial filtering
- Embedding approaches:
  - Triplet loss
  - Short Time Fourier Transform [10]
  - Graph Embeddings
  - Wavelet Transform [8][12]
- Testing robustness of the techniques:
  - Adding noise to the EEG signals
  - Testing performance on Artificial data

3.3 Dataset analysis

This section outlines the procedures and methods used in the analysis of the BCI Competition IV 2A dataset. The focus was on understanding the intricate details of the dataset and carrying out preliminary data processing steps that are essential for subsequent stages of EEG data analysis.

BCI Competition IV 2A - Graz Dataset

The BCI Competition IV 2A 2008 - Graz dataset [7] was meticulously studied. Initial efforts were dedicated to exploring the dataset structure, followed by performing a series of pre-processing tasks. These tasks were aimed at preparing the data for a thorough exploratory analysis, ensuring that the approaches can be benchmarked.

3.3.1 Dataset Structure and Description

The BCI Competition IV 2A dataset is a collection of EEG signal data from 9 different subjects, each participating in 2 sessions with 48 trials per class. This amounts to a total of 288 trials per session. The classes are categorized as follows:

- left hand (class 1)
- right hand (class 2)
- both feet (class 3)
- tongue (class 4)

Figure 1. shows the timing scheme of each trial. Each trial lasts for 7-8 seconds, sampled at 250Hz. The dataset is divided into two sets with a total of 5184 samples:

- Train data: 2592 trials (648 trials per class)
- Test data: 2592 trials (648 trials per class)

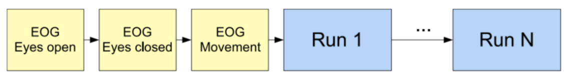

Figure 1: Timing scheme of each session

3.3.2 Dataset Format

The data is stored in the General Data Format (GDF), storing one file per subject and session. Only one session per subject contains class labels for all trials, while the other is used for classifier testing and performance evaluation. Each trial is associated with a target label indicating the intended motor imagery task (e.g., left-hand movement or right-hand movement).

3.3.3 Data Collection

In this dataset, the signals were originally sampled at 250 Hz. It was then bandpass-filtered between 0.5 Hz and 100 Hz. Further, an additional 50 Hz notch filter was applied to suppress line noise. The amplifier sensitivity determines the scale of the recorded signals. Here, it was set to 100 μV for EEG channels and 1 mV for EOG channels. To record this dataset, twenty-two Ag/AgCl electrodes were used with the left mastoid serving as the reference and the right mastoid as the ground. This was performed following the international 10-20 system.

3.3.4 Technical Specifications

- 25 electrodes are used, first 22 are EEG, last 3 are EOG
- Sampling frequency (f_s) is 250Hz
- 9 subjects

Time Duration

- 1 trial = 7-8s
- 1 run = 48 trials = 336-384s
- 1 session = 6 runs = 288 trials = 2016-2304s

3.3.5 Additional EOG Channels Data

In addition to the EEG channels, three monopolar EOG channels were recorded. Before the start of each session, EOG influence was sampled with the use of a 5-minute recording. This was further divided into three blocks:

- eyes open
- eyes closed
- eye movements

3.4 Data Loading and Preprocessing

3.4.1 Data Loading

GitHub link: https://github.com/sam189239/EEG-MI-Benchmarking (Repo is currently private)

Data analysis was performed on the BCI Competition IV 2A dataset.

To load the dataset and preprocess it, we used the dataset converted to .npz format. The dataset contains samples which were collected from 9 subjects. Figure 2. shows the loaded data.

- ‘s‘ contains continuous time-series recorded EEG signals are, shape of M x N array. Size may vary between subjects but N is fixed to 25, indicates 25 electrodes with 22 first EEG and 3 last EOG
- ‘etyp‘ stands for event type which indicate event occurrence, event code is described in Table 2
- ‘epos‘ stands for event position, denoting corresponding event begins at n-th sample at ‘s‘
- ‘edur‘ stands for event duration, denoting duration of corresponding event

Event type	Code	Description
276	0x0114	Idling EEG (eyes open)
277	0x0115	Idling EEG (eyes closed)
768	0x0300	Start of a trial
769	0x0301	Cue onset left (class 1)
770	0x0302	Cue onset right (class 2)
771	0x0303	Cue onset foot (class 3)
772	0x0304	Cue onset tongue (class 4)
773	0x030F	Cue unknown
1023	0x03FF	Rejected trial
1072	0x0430	Eye movements
32766	0x7FFE	Start of a new run

Table 2: Event types and their descriptions.

- ‘artifacts‘: size of 288 x 1, 288 comes form 6 x 48, 6 runs where @run has 48 trials, @class has 12 trials

The EEG signal from a specified subject and channel is extracted from the dataset, allowing for the isolation of specific neural activity of interest. This plotting allows for the examination of EEG activity patterns and dynamics during the event.

In the below example, we have extracted channel C3 from Subject1. ‘etyp‘ 768 corresponds to the start of the trial. We then extract the signal related to the event and Fig shows its plot in Figure 3.

3.4.2 Preprocessing

Bandpass Filtering

Bandpass filtering can be applied to retain only relevant frequency components associated with motor imagery. Bandpass filtering was performed on EEG signals using Butterworth bandpass filters across 9 frequency bands: 4-8Hz, 8-12Hz, 12-16Hz, 16-20Hz, 20-24Hz, 24-28Hz, 28-32Hz, 32-36Hz, and 36-40Hz. This filtering approach decomposes the EEG data into frequency-specific components, enhancing analysis by isolating neural activity within stable frequency ranges. Additionally, the filtering operation reshapes ’raw_EEG’ within each subject to an N x T matrix, facilitating electrode-wise analysis across samples.

Converting to image

Another option to try is to convert the EEG signals as shown in Figure 5. Converting EEG time series data to an image format involves transforming the raw EEG signals into a visual representation that retains the relevant information for analysis or visualization. Spectrogram image representation

```
{'subject01': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A01T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject02': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A02T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject03': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A03T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject04': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A04T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject05': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A05T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject06': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A06T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject07': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A07T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject08': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A08T.npz' with keys: s, etyp, epos, edur, artifacts,
 'subject09': NpzFile '/Users/neranjhana/Downloads/datasets_npz/A09T.npz' with keys: s, etyp, epos, edur, artifacts}
```

```
pd.DataFrame(ori_data['subject01']['s']).head()
```

✓ 0.8s Python

	0	1	2	3	4	5	6	7	8	9	...	15	16	17	18	19
0	-21.142578	-23.681641	-21.484375	-25.146484	-25.732422	-27.929688	-14.550781	-22.509766	-25.439453	-28.710938	...	-30.615234	-29.638672	-26.660156	-20.947266	-25.439453
1	-21.923828	-23.925781	-24.316406	-25.341797	-26.074219	-25.244141	-16.064453	-22.753906	-28.320312	-29.248047	...	-31.054688	-30.712891	-26.123047	-22.998047	-27.099609
2	-15.625000	-19.726562	-18.847656	-21.582031	-20.751953	-21.728516	-12.451172	-20.410156	-23.730469	-25.927734	...	-28.759766	-28.417969	-24.462891	-21.386719	-25.097656
3	-16.699219	-13.720703	-20.361328	-20.849609	-24.169922	-22.021484	-10.986328	-16.357422	-25.585938	-24.951172	...	-27.783203	-28.857422	-25.537109	-19.335938	-24.316406
4	-19.335938	-17.626953	-20.410156	-26.464844	-27.392578	-29.882812	-9.277344	-18.359375	-24.609375	-30.517578	...	-31.884766	-33.740234	-32.666016	-18.554688	-26.611328

5 rows × 25 columns

Figure 2: Data loaded in python

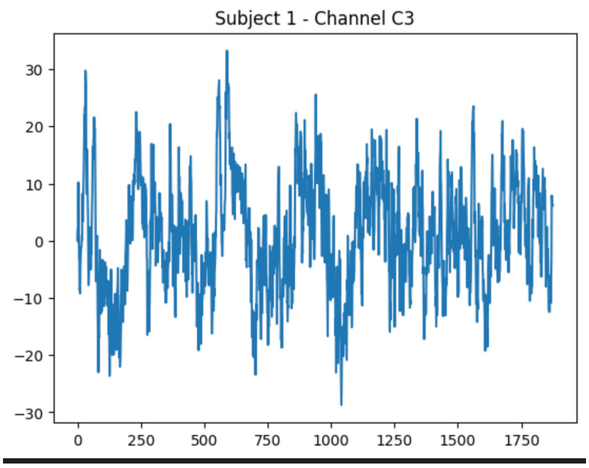

Figure 3: EEG signal for Subject 1 - Channel C3

can efficiently retain all the necessary EEG time series information for classification application. Each pixel in the spectrogram image corresponds to the intensity of a specific frequency at a particular time.

4 Noise and Synthetic Data Robustness

Changes in accuracy due to various external entities that could be present in the input data can be explored to benchmark the models for robustness. We explore options to add noise to input data and use synthetic data to test on the models which can be used for model evaluation.

Synthetic data generation techniques like GAN (General Adversarial Networks), VAE (Variable AutoEncoders) can be used to test the robustness of models used for EEG Classification and improve performance [11]. Robustness to Synthetic data can be used for benchmark comparison. Figure 6 shows the structural difference between raw data and synthetic data genrated by GAN.

A. G. Habash [14] proposes a general approach of using a GAN model to generate synthetic data based on the original BCI data. Figure 4 represents the GAN process for generating synthetic data. For generation of synthetic data, some pre-processing methods are applied to the original dataset. As represented in the dataset preprocessing, image representation of the time series data is applied with STFT (Short-time Fourier transform) and is converted to grayscale. GAN uses the preprocessed data for accurate generation of artificial data. Similar processes can be used to generate synthetic data and test the robustness of models. This robustness can be used to benchmark how different models vary with classification accuracy when tested with synthetic data.

5 Challenges and solutions

One of the major challenges that we would face would be the reproduction of results from large models. The large models often have millions of parameters which would require substantial memory and processing capacities to store and manipulate. Employing strategies such as distributed computing, model parallelism, and utilizing cloud services can help mitigate these constraints and facilitate the reproduction of results from large models to a certain extent.

Another challenge arises from the limitation of the dataset which contains samples from only 9 subjects. A small sample size may not fully capture the variability and complexity of real-world

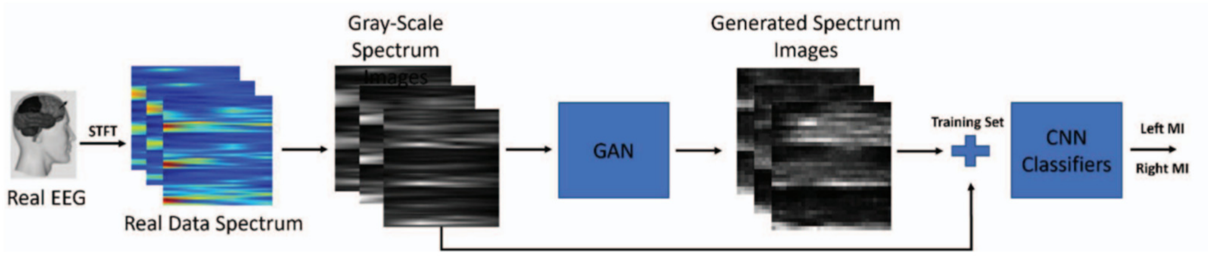

Figure 4: GAN for synthetic EEG signal generation

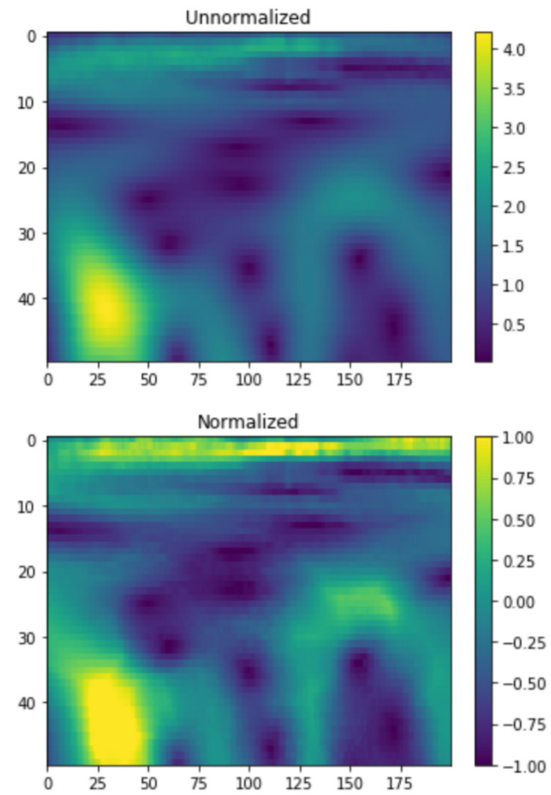

Figure 5: Image representation of EEG signals

scenarios, leading to concerns about the generalization and robustness of the trained models. Testing these models on synthetic data could provide us with valuable insights.

6 Conclusion

In conclusion, our project is focused on benchmarking various deep learning and machine learning models for EEG signal classification, specifically in the context of motor movement and imagery. Through our comprehensive literature review, we have identified several promising approaches for data processing and models. Our methodology encompasses a comprehensive approach to data

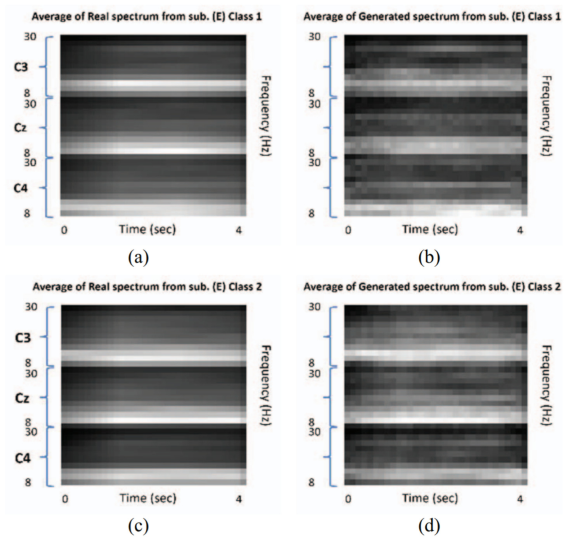

Figure 6: Spectrum of Real and Synthetically generated EEG signals

pre-processing, embedding, and model robustness evaluation. As we move forward, we will be implementing and testing these models to conduct a systematic benchmarking study. Our aim is to provide a clear and detailed comparison of the models' effectiveness in EEG signal classification, contributing to the ongoing development of reliable and efficient brain-computer interfaces. This project will serve as a valuable reference for researchers and practitioners in the field, guiding future advancements in EEG-based Neurotechnology.

References

[1] A. Al-Saegh, S. A. Dawwd, J. M. Abdul-Jabbar, "Deep learning for motor imagery EEG-based classification: A review," Microprocessors and Microsystems, vol. 80, Article ID 103311, Oct. 2020, doi: 10.1016/j.micpro.2020.103311.

[2] D. Milanés Hermosilla et al., "Shallow Convolutional Network Excel for Classifying Motor Imagery EEG in BCI Applications," IEEE Access, vol. 9, June 2021, doi: 10.1109/ACCESS.2021.3091399.

[3] N. Shajil, M. Sasikala, and A. M. Arunnagiri, "Deep Learning Classification of two-class Motor Imagery EEG signals using Transfer Learning," in Proc. 8th IEEE International Conference on E-Health and Bioengineering (EHB), Grigore T. Popa University of Medicine and Pharmacy, Romania, Oct. 2020.

[4] K. Roots, Y. Muhammad, and N. Muhammad, "Fusion Convolutional Neural Network for Cross-Subject EEG Motor Imagery Classification," in Computers, vol. 9, no. 72, pp. 1-9, Sep. 2020, doi:10.3390/computers9030072.

[5] R.R. Chowdhury, Y. Muhammad, and U. Adeel, "Enhancing Cross-Subject Motor Imagery Classification in EEG-Based Brain–Computer Interfaces by Using Multi-Branch CNN," Sensors, vol. 23, no. 7908, 2023, doi:10.3390/s23187908.

[6] D. Avola, M. Cascio, L. Cinque, A. Fagioli, G. L. Foresti, M. R. Marini, and D. Pannone, "Analyzing EEG Data with Machine and Deep Learning: A Benchmark," arXiv:2203.10009v1 [cs.LG], Mar. 2022.

[7] C. Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008 – Graz Data Set A," in Institute for Knowledge Discovery Graz University of Technology, Austria, and Institute for Human-Computer Interfaces Graz University of Technology, Austria, [Year of Publication].

[8] T. D. Pham, "Classification of Motor-Imagery Tasks Using a Large EEG Dataset by Fusing Classifiers Learning on Wavelet-Scattering Features," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, 2023. DOI: 10.1109/TNSRE.2023.3241241.

[9] H. Alwasiti, M. Z. Yusoff, and K. Raza, "Motor Imagery Classification for Brain Computer Interface Using Deep Metric Learning," IEEE Access, vol. 8, Jun. 2020, doi: 10.1109/ACCESS.2020.3002459.

[10] S. K. Bashar, A. R. Hassan, and M. I. H. Bhuiyan, "Motor Imagery Movements Classification Using Multivariate EMD and Short Time Fourier Transform," in Proceedings of IEEE INDICON, 2015, doi: 10.1109/INDICON.2015.7443449.

[11] Z. Zhang, F. Duan, J. Solé-Casals, J. Dinarès-Ferran, A. Cichocki, Z. Yang, and Z. Sun, "A Novel Deep Learning Approach With Data Augmentation to Classify Motor Imagery Signals," IEEE Access, vol. 7, Feb. 2019, doi: 10.1109/ACCESS.2019.2895133.

[12] S.K. Bashar, A.R. Hassan, and M.I.H. Bhuiyan, "Identification of Motor Imagery Movements from EEG Signals Using Dual Tree Complex Wavelet Transform," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 290-296, 2015, doi: 10.1109/ICACCI.2015.7275679.

[13] R.T. Schirrmeister, J.T. Springenberg, L.D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," Human Brain Mapping, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/hbm.23730.

[14] A. G. Habashi, A. M. Azab, S. Eldawlatly, and G. M. Aly, "Motor Imagery Classification Enhancement using Generative Adversarial Networks for EEG Spectrum Image Generation," in Proceedings of the 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), 2023, doi: 10.1109/CBMS58004.2023.00243.