Magnetoencephalography signal classification using Deep Learning approaches

Christos Papageorgiou Hugo Moura Panagiotis Andrikopoulos Stylianos Psara Tomás Carrilho

Abstract—**Magnetoencephalography (MEG) is a vital tool in neuroscience for exploring brain activity and understanding its correlation with various physiological and neurological conditions. This paper addresses the challenges of making use of MEG for medical diagnosis by proposing a Convolutional Neural Network (CNN) approach in classifying mental states based on MEG data. Two distinct strategies are employed: intra-subject classification, where deep learning is used to train and test models using the same subject, and cross-subject classification where a model is trained with a set of subjects and tested on new, unseen subjects. The goal will be to accurately classify whether the subject is in one of the following states: rest, math, memory, motor. The paper reviews related studies, identifies gaps in existing research, and discusses the applied dataset and preprocessing techniques. The results showcase a comprehensive qualitative and quantitative analysis. The intra-subject classification showcased outstanding accuracy results, highlighting the model's adeptness in capturing subject-specific neural characteristics. Although the cross-subject approach had a lower accuracy, it offers a practical evaluation of the model's ability to generalize across diverse individuals. The study contributes valuable insights into leveraging deep learning for further advancements in the neuroscience research domain highlighting the complex and individualistic nature of brain activity.**

I. Introduction

Magnetoencephalography (MEG) proves to be an indispensable method for investigating the dynamics and interconnections of extensive brain activity, exploring its relationship with the body and the surroundings across both healthy and impaired physiological and neurological conditions [1].

MEG enables the examination of brain function through the detection of magnetic fields produced by the electrical behavior of groups of neurons [1]. Additionally, MEG signals play a crucial role in diagnosing epilepsy, aiding in treatment strategies, monitoring of drug effects, and defining different neurological conditions like Parkinson's disease, Alzheimer's disease, autism, schizophrenia, stroke and head trauma [2].

Compared to other methods, a significant benefit of this approach is its capability to capture brain function directly and without intrusion, providing an exceptionally high temporal resolution [1].

The direct association between the captured magnetic field and the intrinsic neuronal currents implies that MEG remains unaffected by issues typically induced by intermediary processes (such as neurovascular coupling in fMRI [functional magnetic resonance imaging] or fNIRS [functional near-infrared spectroscopy]). Consequently, MEG can produce a dynamic and information-rich portrayal of extensive brain activity. These fundamental advantages of Magnetoencephalography lead neuroscientists to frequently employ it for investigating broad-scale brain dynamics in both healthy and pathological conditions [1].

Nonetheless, using MEG to diagnose medical conditions poses challenges for healthcare professionals. The extensive array of sensors, intricate preprocessing procedures required for extracting cortical signals, and the complexities involved in categorizing diverse waveform patterns demand considerable expertise. The application of deep learning in the classification of MEG signals aims to alleviate the workload on healthcare professionals and enhance the precision of neurological diagnoses [3].

In our paper, we will propose a Convolutional Neural Network (CNN) trained on MEG data that aims to classify the subjects in various mental states (rest, math, memory, and motor). Two approaches will be used. The first is intra-subject classification, where our model will be trained and tested using the same subjects. In the second approach (cross-subject classification), our model will be trained on a set of subjects but tested on new data from unseen subjects.

The rest of the paper is organized as follows. In the "Related work" section, relevant studies will be summarized. Then, some potential gaps and drawbacks of the previous research will be discussed. In the "Visualization and preprocessing" section, we will provide insights on the nature of our data as well as the preprocessing methods we applied on them before training our model. In The "Approach" section, we will first establish a baseline using machine learning and later describe the implementation details and the methodology used for developing our CNN model. Finally, a comparison between intra-subject classification and cross-subject classification will also be made using a statistical test. In the "Results" section, a qualitative and quantitative analysis of our results will be performed. Lastly, In the "Conclusion and Discussion" section, we will summarize the findings of our experiment, and we will discuss the observed key points of our study and the potential contribution of our findings to the field of neuroscience and MEG data.

II. Related work

Due to the efficiency of the feature extraction process, CNN models are very popular among researchers who work on MEG signals. Antonio Giovannetti et al. [4] presented an ensemble classifier based on deep convolution neural network for the detection of Alzheimer disease in the early stages.

The data used for this experiment consists of image-based representation of MEG recordings and their corresponding magnetic resonance imaging scans. The proposed architecture achieved 89% and 87% accuracy respectively. In addition, Ahmad Hasasneh et al. [5] created a model that could classify ocular and cardiac artifacts in MEG recordings, without the need for additional electrocardiogram and electrooculogram signals. The data used for this study consists of 132 MEG signals obtained from 48 subjects, which are divided into task related experiments and rest experiments where the subjects were asked to have their eyes opened or closed. The author introduced a deep convolutional neural network and managed to achieve accuracy of 94.4%. Moreover, in order to address the complexity of the large and usually corrupted with noise MEG signals, Abdullah Caliskan et al. [6] suggested the use of the Riemannian feature extraction approach. The MEG data used for this experiment were produced as the brain output for visual stimuli based on faces and scrambled faces. Furthermore, the authors developed a deep neural network and achieved accuracy of 80.85% which outperformed other traditional classification methods such as SVM, KNN, NB and DT. Furthermore, Yifeng Bu et al. [7] proposed a MEG sensor-based CNN model to decode Rock-Paper-Scissors gestures involving only muscles from the fingers and hand performed by 12 subjects. In addition, both cross-subject and intra-subject approach was used in the classification process, with the intra approach achieving better performance. The results of the experiments show that the proposed model outperformed state-of-the-art deep learning architectures and an SVM algorithm with accuracy of 85.56%. Moreover, is important to note that the model achieved similar performance using sensor inputs from key brain regions instead of the whole-brain array.

On the other hand, some researchers decided to exploit the possibilities of RNN models in the MEG analysis field based on their ability to process sequence data. Hong Gi Yeom et al. [8] were the first who applied LSTM model to predict the arm movement using non-invasive MEG signals in order to control a robotic arm. The result achieved from the proposed model were promising and outperformed a conventional method such as Multiple linear regression. This study proved that is possible to control a robotic arm in real life using MEG signals. However, Jakub Orlinski er al. [9] followed a different approach by combining the efficiency of a CNN model to encode the spatial features with the ability of RNN model to track the changes over time. The purpose of this study was to use brain activity data for recognizing human emotions. The experiments were applied on the CiNet dataset which consists of MEG recordings obtained by 36 subjects. The subjects were asked to rate their emotional response to a white noise or song from 1 to 9 in the valence and arousal domain. The results show that the best performance was observed on the proposed model by achieving 56.5% on raw signal data.

To balance computational load and data quality, downsampling is essential, reducing the sampling rate, but caution is needed as lowering the rate might induce aliasing. Adhering to the Nyquist Sampling Theorem like H. Nyquist explained in the publication issued in ”Transactions of the American Institute of Electrical Engineers” when dealing with the ”Telegraph Transmission Theory”. Zhendong Zhang also introduced ”frequency pooling,” transforming features into the frequency domain, and restoring in the spatial domain. Proven shift-equivalent and anti-aliasing through Fourier transform, experiments demonstrate improved CNN accuracy and robustness. safeguards against aliasing issues [1], [2]. This precaution ensures accurate signal capture and avoid potential errors during subsequent processing steps. Improved schemes, DSD-FA and DDSD-FA, derived from existing methods, exhibit superior performance in experimental comparisons with conventional downsampling approaches, Lu Fang and other authors explored this approach in subpixel Downsampling via Frequency-Domain Analysis [3].

III. Visualization and Preprocessing

A. Data Visualization and Initial Assessment

Prior to implementing our deep learning methodology, a crucial step involved the initial visualization of our dataset. To achieve this, line graphs were employed, presenting a simplified yet informative view of the data. Specifically, these graphs depicted the raw values of random sensors, as well as the average values derived from all sensors at each distinct timepoint. This visualization strategy was pivotal in offering a clear understanding of the dataset's inherent structure and characteristics (Figure 1).

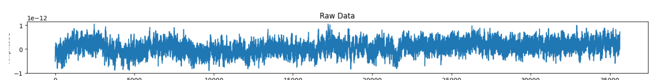

Fig. 1. Raw signal of a random sensor of a random observation.

B. Access and Initial Preprocessing of Data

The data for this project were conveniently stored on a cloud-based drive, enabling efficient access through Google Colab Python notebooks, the platform where our preprocessing scripts were developed and executed. A detailed inspection of the training data revealed its organization in a multidimensional structure, specifically a 32 X 248 X 35624 matrix for the Intra-subject dataset. This corresponds to 32 distinct files, each containing data from 248 sensors, with every sensor recording 35624 timepoints. Similarly the Cross-subject dataset corresponds to a 64 X 248 X 35624 matrix. Alongside the two datasets, their labels were also provided.

To standardize the data, z-score normalization was applied to the timepoints of each sensor. This process involved calculating the mean (μ) and standard deviation (σ) for each combination of sensor and timepoint across all files. The mean was then subtracted from each timepoint, and the result was divided by the standard deviation. Such normalization ensures that the data for each sensor-timepoint combination is centered around a mean of zero and has a standard deviation of one. This technique is crucial in facilitating more effective algorithmic processing and analysis, as it standardizes features to a uniform scale, enhancing the performance of many machine learning algorithms [10], [11]. The mathematical formulation of z-score normalization is given by:

$$\mathbf{z} = \frac{X - \mu}{\sigma}$$

where X represents the data point, μ is the mean, and σ is the standard deviation of the respective feature.

C. Addressing Data Size and Sampling Rate

Given the extensive size of the dataset, a reduction in the computational load was necessary. This was achieved through a process known as downsampling, a standard technique in image and signal processing. However, downsampling comes with the inherent risk of aliasing, a phenomenon where high-frequency components of a signal, sampled below their respective Nyquist rates, are misrepresented as lower-frequency components, potentially leading to inaccurate analysis [12].

To mitigate this risk, we adhered to the Nyquist Sampling Theorem [13], which posits that accurate signal representation requires a sampling rate at least twice the highest frequency present in the signal. To comply with this theorem and prevent aliasing [14], we first applied a lowpass filter to eliminate high-frequency components that could potentially cause aliasing during the downsampling process [15].

Subsequently, the data was downsampled by a factor of 2, effectively reducing the sampling rate from 2034Hz to 1017Hz. This strategic reduction in the sample rate significantly lowered computational demands while preserving the essential characteristics and integrity of the signal.

D. Statistical Insights from Sensor Data

In our preliminary analysis, we conducted a detailed statistical examination of the data on a per-file and per-sensor basis. This involved calculating the mean value across all sensors for each timepoint. These computations were visualized in a graph to ascertain if there were discernible differences between the various classes represented in the dataset. Our analysis revealed that certain classes, notably the "rest" and "motor" classes, exhibited more pronounced distinctions compared to others, with the former generally displaying the highest mean values and the latter the lowest. Conversely, classes such as "story math" and "working memory" showed a lack of distinguishability from one another (Figure 2).

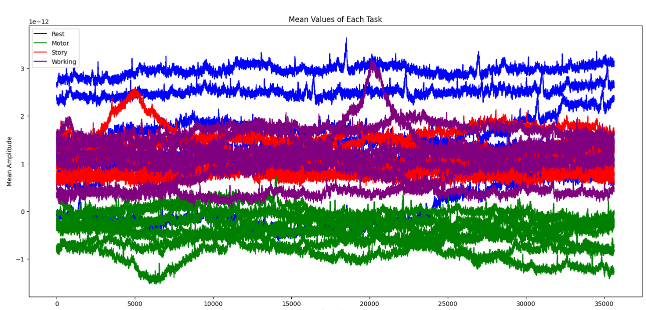

Fig. 2. Mean value across all sensors for each timepoint.

E. Effects of Scaling on Data Visualization

Interestingly, we observed that the application of z-score normalization, while beneficial for standardization, led to a convergence of these class distinctions in our visual representations. Post-normalization, the graphical distinction between classes was significantly reduced, resulting in overlapping lines in the plots. This phenomenon underscores the impact of scaling methods on data interpretation and visualization (Figure 3).

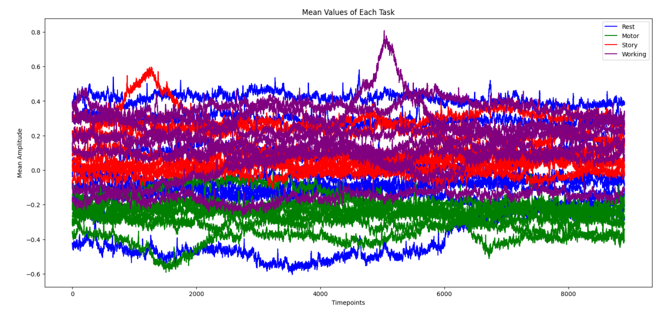

Fig. 3. Mean signals across all sensors after z-score scaling.

F. Exploration of Min-Max Scaling Per Observation

Further investigation into alternative scaling techniques revealed that min-max scaling per observation was particularly effective in enhancing the graphical distinction between different classes. This method amplified the disparities, making the class-specific trends more visually apparent (Figure 4).

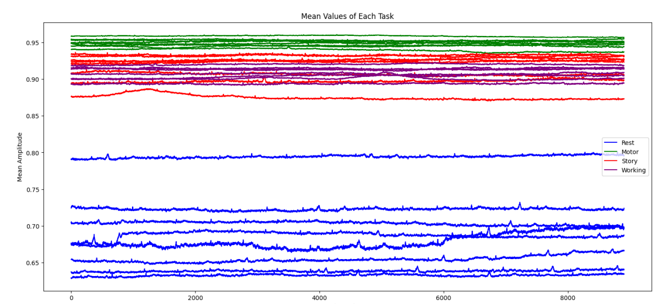

Fig. 4. Mean signals across all sensors after min-max scaling per observation.

IV. Approach

A. Establishing a Baseline with Machine Learning

Before delving into advanced deep learning methodologies, we deemed it crucial to establish a baseline using traditional machine learning techniques. To this end, we opted for a Support Vector Machine (SVM) model, utilizing the mean values of all 248 sensors for each timepoint as input features. SVM was chosen as the most widely used technique in EEG and MEG data classification [16]. These features were preprocessed using min-max scaling to accentuate class disparities. The SVM model underwent rigorous hyperparameter tuning, employing a grid search approach. This initial machine learning attempt yielded a promising accuracy of 75%. This result was particularly encouraging given the simplified nature of our training data, which was based on aggregated sensor readings.

B. *Intra-subject classification*

Upon completing the preprocessing of our dataset and establishing a machine learning baseline, we transitioned to employing deep learning techniques for more advanced analytical capabilities. For this purpose, we opted for a 2-dimensional Convolutional Neural Network (2D CNN). The choice of a 2D CNN [17] was primarily driven by its proficiency in capturing spatial relationships within the data. Unlike 1D CNNs that are suitable for time-series data, 2D CNNs excel in scenarios where the data contains spatial correlations [18], making them ideal for our multi-sensor MEG data which possesses spatial-temporal characteristics.

Our CNN architecture, detailed in the Figure 5, consists of three convolution layers. Given the substantial size of the dataset, exacerbated by a high sampling rate, we intentionally designed the model to be of low complexity with a moderate number of parameters. This approach strikes a balance between the model's capacity to learn from complex data and computational efficiency. Additionally, MaxPooling layers were incorporated into the architecture. These layers serve the crucial function of reducing the dimensionality of the data, thereby easing the computational load and helping to prevent overfitting, which is vital in a high-dimensional dataset.

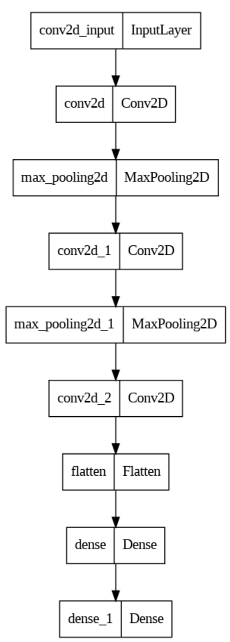

Fig. 5. Structure of our 2-D Convolutional Neural Network model.

For the optimization algorithm, we chose the 'Adam' optimizer. Adam, is known for its efficiency in handling large datasets and sparse gradients, making it a superior choice for our data. Furthermore, we used categorical crossentropy as our loss function. This choice is particularly appropriate for multi-class classification problems, as it measures the performance of the model whose output is a probability value between 0 and 1.

A key aspect of our training process was the use of a custom data generator. This generator facilitated efficient memory usage by iteratively providing batches of data and labels for training, rather than loading the entire dataset into memory at once. This approach is particularly advantageous for large datasets, allowing for dynamic adjustments during training.

The model was trained in small batches over multiple epochs. This incremental learning process enables the model to adjust and improve its classification accuracy progressively, making it more adept at handling the complexities of MEG data.

In the testing phase, the model was evaluated on a dataset derived from the same subject. This intra-subject testing approach is crucial for assessing the model's ability to generalize and accurately classify data within individual subjects, providing valuable insights into its performance and potential application in real-world scenarios.

C. *Cross-subject classification*

Building upon our intra-subject classification framework, we extended our deep learning approach to encompass cross-subject classification. This phase utilized the same 2-dimensional Convolutional Neural Network (2D CNN) model, proven effective in the intra-subject context. However, there was a significant difference at the training dataset's size and composition to adapt to the cross-subject paradigm.

For the cross-subject training, the dataset size was doubled, encompassing 64 files. This enlargement of the training set aimed to incorporate a broader spectrum of data variability, which is important for a model's ability to generalize across different subjects. By training the model on a more diverse set of MEG recordings, we sought to enhance its adaptability and predictive accuracy when exposed to data from subjects not included in the training phase.

The testing phase was structured to evaluate the model's performance across three distinct subjects, none of whom were part of the training dataset. This testing strategy was deliberate to rigorously assess the model's generalization capabilities in a cross-subject scenario. The subjects were tested sequentially allowing for an individualized assessment of the model's performance on each subject.

The model's effectiveness in cross-subject classification was assessed by aggregating results from tests on these three different subjects, providing insights into its robustness and generalization capabilities for MEG data analysis.

V. Results

A. *Quantitative analysis*

1) Intra-subject classification: As shown in Figure 6, the intra-subject classification approach resulted in a perfect prediction, reaching an accuracy of 100%, and therefore, precision and recall for all four classes also attained 100%. This can

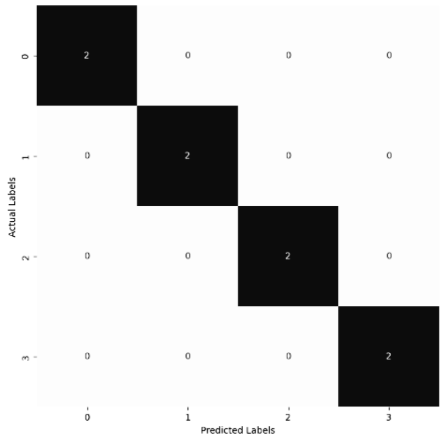

Fig. 6. Confusion matrix for the results of the classification on the intra-subject dataset. Classes 0 to 3 correspond to ”Resting Task”, ”Math & Story Task”, ”Working Memory Task” and ”Motor Task” in the order mentioned.

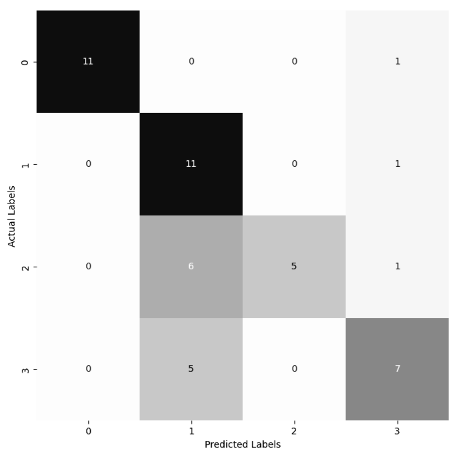

Fig. 7. Confusion matrix for the results of the classification on the cross-subject dataset.

possibly be explained by the fact that in both training and testing datasets there's only one subject, so it's reasonable to anticipate that the MEG signals are quite similar and therefore that the model can more accurately predict and distinct the different class labels.

2) Cross-subject classification: When it comes to the cross-subject classification approach not so accurate predictions were obtained, reaching a reasonably high mean accuracy value of 70.8% for the three test datasets.

Subject	1	2	3
Accuracy	93.8%	56.3%	62.5%
Recall	93.8%	56.3%	53.4%
Precision	95.0%	58.3%	58.3%
F1 score	94.4%	57.3%	55.8%

TABLE I
METRICS FOR EACH ONE OF THE CROSS-SUBJECTS.

It is evident that there's a significant difference between the accuracy value from intra-subject classification and the cross-subject. This was expected, as the model is trained and tested with only one subject in the intra-subject scenario, while in the cross-subject scenario, multiple subjects are involved in the testing phase.

In order to take a closer look into how well the model can distinct different classes, different metrics were computed for each one of the class labels and are presented in Table 1.

Metrics	Class 0	Class 1	Class 2	Class 3
Accuracy	97.9%	75.0%	85.4%	83.3%
Recall	91.7%	91.7%	41.7%	58.3%
Precision	100.0%	50.0%	100.0%	70.0%
F1 score	96.0%	65.0%	59.0%	64.0%

TABLE II
METRICS PER CLASS IN THE COMBINED CROSS-SUBJECT DATASET PER CLASS.

It's clear that the model can easily distinct classes 0 (resting) and 1 (math & story) from the others, but the same doesn't happen when it comes to classes 2 (working memory) and 3 (motor). From the confusion matrix showed in Figure 7 we can see that classes 2 and 3 are commonly mistakenly predicted as class 1, which shows that the models struggles differentiating Working Memory tasks and Motor tasks from Math & Story tasks, incorrectly predicting the latter one as the class label for 11 cases.

B. *Qualitative analysis*

As mentioned before the test for the cross-subject classification approach was conducted on three distinct subjects who were excluded from the training set. Our results revealed an anomaly related to the accuracy of the first subject in comparison with the accuracy of the remaining subjects.

The accuracy of the first subject reached 93.8%, contrasting with the accuracy of 56.3% and 62.5% for the other subjects, respectively. This may lead us to conclude that the results derived from MEG signals vary among different subjects and raises the possibilities that the first subject share more similarities with the subjects used in training.

According to Yifeng Bu et al. [7], the location of sensors can vary with respect to the underlying brain regions due to factors such as head sizes, head shapes, head orientations, body positioning, and stature. This variability could also explain the superiority of the intra-subject approach and the significant difference of the accuracy.

The noticeable variability in accuracy among subjects in the cross-subject classification can also be attributed to individual differences in neural architecture, cognitive processing, or even in the way MEG signals are captured across the three different subjects.

VI. CONCLUSION

In conclusion, our study offers important insights into the application of Convolutional Neural Networks in classifying motor and mental states based on MEG data. The distinct methodologies of intra-subject and cross-subject classification have revealed the challenges and potential in employing deep learning for neural pattern analysis.

While intra-subject classification demonstrated exceptional accuracy, indicative of the model's capability in capturing subject-specific neural characteristics, the cross-subject approach, despite its lower accuracy, provides a realistic assessment of the model's generalizability across different individuals.

This contrast highlights the complex and individualistic nature of brain activity, underscoring the importance of personalized approaches in neural diagnostics.

Looking ahead, future work could explore refining these models to better accommodate individual variances in MEG data, potentially opening new pathways in personalized neurological research and treatment.

References

[1] J. Gross, "Magnetoencephalography in Cognitive Neuroscience: A Primer," Neuron, vol. 104, no. 2, pp. 189-204, 2019.

[2] F. A. L. Fred et al., "A Brief Introduction to Magnetoencephalography (MEG) and Its Clinical Applications," Brain Sciences, vol. 12, no. 6, Art. no. 788, 2022.

[3] J. Aoe, R. Fukuma, T. Yanagisawa et al., "Automatic diagnosis of neurological diseases using MEG signals with a deep neural network," Sci Rep, vol. 9, Art. no. 5057, 2019.

[4] A. Caliskan, M. Yuksel, H. Badem, and A. Basturk, "A Deep Neural Network Classifier for Decoding Human Brain Activity Based on Magnetoencephalography," Elektronika ir Elektrotechnika, vol. 23, Apr. 25, 2017.

[5] A. Hasasneh, N. Kampel, P. Sripad, N. Shah, and J. Dammers, "Deep Learning Approach for Automatic Classification of Ocular and Cardiac Artifacts in MEG Data," Journal of Engineering, vol. 2018, pp. 1-10, May 2, 2018.

[6] A. Giovannetti et al., "Deep-MEG: Spatiotemporal CNN Features and Multiband Ensemble Classification for Predicting the Early Signs of Alzheimer's Disease with Magnetoencephalography," Neural Computing and Applications, vol. 33, pp. 1-17, May 19, 2021.

[7] Y. Bu et al., "Magnetoencephalogram-based Brain–Computer Interface for Hand-Gesture Decoding Using Deep Learning," Cerebral Cortex, vol. 33, May 13, 2023.

[8] H. G. Yeom, J. Kim, and C. K. Chung, "LSTM Improves Accuracy of Reaching Trajectory Prediction From Magnetoencephalography Signals," IEEE Access, vol. 8, pp. 1-1, Jan. 27, 2020.

[9] J. Orlinski, "MEG-Driven Emotion Classification Using Convolutional and Recurrent Neural Networks," Jan. 1, 1970.

[10] H. Nyquist, "Certain Topics in Telegraph Transmission Theory," Transactions of the American Institute of Electrical Engineers, vol. 47, pp. 617-644, 1928.

[11] O. Akbulut, "Feature Normalization Effect in Emotion Classification based on EEG Signals," Sakarya University Journal of Science, vol. 24, no. 1, pp. 60-66, 2020.

[12] A. Apicella, F. Isgrò, A. Pollastro, R. Prevete, "On the effects of data normalization for domain adaptation on EEG data," Engineering Applications of Artificial Intelligence, vol. 123, Part A, 106205, 2023.

[13] S. M. Nacy, S. N. Kbah, H. A. Jafer, and I. Al-Shaalan, "Controlling a Servo Motor Using EEG Signals from the Primary Motor Cortex," American Journal of Biomedical Engineering, vol. 6, no. 5, pp. 139-146, 2016.

[14] L. Fang, O. C. Au, K. Tang, and A. K. Katsaggelos, "Antialiasing Filter Design for Subpixel Downsampling via Frequency-Domain Analysis," IEEE Transactions on Image Processing, vol. 21, no. 3, pp. 1391-1405, March 2012.

[15] Z. Zhang, "Frequency Pooling: Shift-Equivalent and Anti-Aliasing Downsampling," submitted to arXiv on 24 Sep 2021.

[16] Richhariya, B., Tanveer, M. (2018). EEG signal classification using universum support vector machine. Expert Systems with Applications, 106, 169-182. ISSN 0957-4174

[17] L. Meng and J. Xiang, "Brain Network Analysis and Classification Based on Convolutional Neural Network," Front. Comput. Neurosci., vol. 12, Dec. 10, 2018.

[18] M. Lopez-Martin, A. Nevado, and B. Carro, "Detection of early stages of Alzheimer's disease based on MEG activity with a randomized convolutional neural network," Artif. Intell. Med., vol. 107, 2020, Art. no. 101924.