MEGNet: A MEG-Based Deep Learning Model for Cognitive and Motor Imagery Classification

Minerva Sarma*, Charles Bond*, Sanjeev Nara† and Haider Raza*
* School of Computer Science and Electronics Engineering, University of Essex, Colchester, United Kingdom.
† Department of Mathematics and Computer Science, Physics, Geography, Mathematics Institute,
Justus-Liebig-Universität Gießen, Gießen, Germany.

***Abstract*—Decoding complex patterns associated with task-specific activities embedded within magnetoencephalography (MEG) signals is pivotal for understanding brain functions and developing applications such as brain-computer interfacing. It is widely recognized that machine learning algorithms rely on feature extraction before undertaking decoding tasks. In this work, we introduce MEGNet, aiming to enhance the single-trial decoding framework of a compact deep neural network inspired by EEGNet, a model widely utilized in electroencephalography (EEG) studies. MEGNet accepts raw MEG signals, evoked responses and frequency spectrum as input. For validation, the MEG dataset containing motor and cognitive imagery tasks was used for classification. We performed pair-wise decoding of cognitive and motor tasks. Classification accuracy was evaluated using metric scores and benchmarked against ShallowConvNet and DeepConvNet. Our findings demonstrate that MEGNet can successfully decode between cognitive and mental imagery tasks. This MEGNet model surpasses existing feature extraction techniques, exhibiting consistent and stable mean accuracy of $64.76\% \pm 3\%$ across tasks and subjects. All codes are available at our GitHub repository: https://github.com/Charliebond125/MEGNet.git.**

***Index Terms*—Magnetoencephalography, Convolutional Neural Network, Deep Learning, Cognitive and Motor Imagery.**

I. Introduction

In recent years, the convergence of neuroscience and artificial intelligence has been a driving force behind the exploration of innovative techniques for understanding human brain activity. MEG harnesses high-resolution spatiotemporal data to decode real-time neural dynamics non-invasively [1], facilitating advanced deep learning methodologies. Spatiotemporal elements refer to the combination of spatial and temporal information in a dataset. In the context of MEG data, spatiotemporal elements represent the patterns of brain activity over both space and time. MEG data provides information about the timing and location of brain activity. The capacity to decode MEG signals on a trial-by-trial basis has enormous potential in a variety of fields, including cognitive neuroscience [1], brain-computer interfaces (BCIs) [2], and clinical applications [3]. The study of visual perception in neuroscience is intriguing, particularly due to the common neural foundation that has been identified between visual perception and mental imagery processes. It has been reported that both the perception of visual stimuli and the ability to generate mental images activate similar regions of the brain and cognitive processes [4]. However, despite this shared neural foundation, a crucial distinction exists between cognitive and motor imagery (MI). Cognitive imagery (CI) encapsulates the mental processes involved in visualizing scenarios, concepts, or objects, while MI pertains to the mental simulation of movement without actual physical execution. Exploration of CI and MI is significant because it provides insights into various aspects of human cognition and behaviour. Studies have shown that MI involves the generation, maintenance, manipulation, and temporal sequencing of motor images [5]. Additionally, cognitive and psychological measures have been found to impact the performance of MI brain-computer interfaces (MI-BCIs), with factors such as vividness of visual imagery, personality traits, and motivation playing a role [6]. CI and MI classification, within the context of this study, pertains to the discernment and categorization of distinct mental states and intentions through brain activity patterns. Remarkably, researchers have made a significant stride in categorically classifying visual and CI from MEG signal data [7], [8]. A large number of existing research studies have predominantly focused on decoding visual and motor intentions, often bypassing the complex interplay between cognitive processes and motor responses. This noticeable gap in research highlights the need for a comprehensive investigation into the neural correlates of cognitive and MI, and how these can be accurately classified using deep learning approaches.

Traditional methods for CI and MI classification have primarily relied on hand-engineered features extracted from neurophysiological signals. However, the intricate and dynamic nature of brain activity patterns poses challenges for conventional feature engineering approaches. Many approaches were presented to classify brain activities by segmenting MEG signal data into epochs and statistical features [9], [10]. Although the majority of research still relies on the use of hand-crafted features, many recent studies have explored the potential of deep learning approaches. Particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), offer a promising solution. These architectures excel in automatic feature learning from raw data, thereby bypassing the need for explicit feature extraction which refers to transforming raw data into a set of meaningful and representative features that can be used as input for a machine learning model. This process typically involves selecting or creating a subset of relevant features from the original data, which can be time-consuming and require domain expertise. The success of machine learning and deep learning algorithms in a variety of domains has inspired researchers to apply similar techniques to neuroimaging data processing including MEG and EEG signal classification.

Craik and team [11] provided a detailed assessment of several deep learning algorithms used in EEG classification tasks, providing insights into potential modifications for MEG data interpretation. EEGNet, a small CNN particularly built for EEG-based BCIs, was introduced by [12], and captures important characteristics from EEG data effectively, making it an appealing choice for adaption into MEGNet for single-trial classification. In this study, we intend to suggest a fresh approach: converting EEGNet into MEGNet. We hypothesise that MEGNet's compact design, which was inspired by EEGNet's success in EEG-based BCIs, will efficiently capture significant spatiotemporal elements from MEG data, boosting single-trial classification accuracy. We used publicly accessible MEG datasets which include a wide range of cognitive activities and motor imaging paradigms. These tasks include hand imagery, feet imagery, subtraction imagery, and word generation imagery. The MEG signals underwent advanced artefact removal techniques and filtering processes. In addition, we will compare MEGNet's performance to that of other state-of-the-art existing models such as ShallowConvNet and DeepConvNet [13].

The rest of the paper is organized as follows: Materials and methods are presented, including dataset, data pre-processing and model training in Section 2, the performance analysis and results are described in Section 3, the study is discussed in Section 4, and the conclusions are summarized in Section 5.

II. Materials and Methods

A. Dataset

This study uses the magnetoencephalography (MEG) dataset, specifically designed for motor and cognitive imagery-based brain-computer interface (BCI) applications[10]. It consists of MEG signals recorded during four mental imagery tasks using a typical BCI imagery paradigm. The four tasks were:

1) Hand imagery- Participants were asked to imagine opening and closing their right hand.
2) Feet imagery- Participants were asked to imagine moving their toes up and down.
3) Subtraction imagery- Participants were asked to subtract 7 from a given number and imagine the result.
4) Word generation imagery- Participants were asked to imagine generating words starting with a given letter.

The dataset used in this study involved the recruitment of 20 healthy participants. Three subjects are not included due to noisy data present within the recordings The final dataset (N= 17, 14 Males (82.35%), 3 Females (17.64%), with a mean age of 28. The minimum age is 22, with the highest age being 40) Data was acquired using an Elekta Neuromag™ system [10], recorded with 306-channels (102 magnetometers and 204 planar gradiometers), including the use of MaxShield™. Each participant underwent two recording sessions on different days. Each session consists of two data runs due to session breaks. For better handling of the data, the authors have merged the sessions. The dataset includes 1,134 minutes of MEG recordings and a total of 6,800 imaging trials. By containing many MEG recordings and imaging trials, this dataset provides a valuable resource for investigating and developing brain-computer interface systems based on motor and cognitive imagery. The single trial classification was performed by [10] using a linear classifier i.e., a Support Vector Machine (SVM) classifier to estimate accuracies for the six binary tasks i.e. hand versus feet (H-F), hand versus word generation (H-W), hand versus subtraction (H-S), feet versus word generation (F-W), feet versus subtraction (F-S), and word generation versus subtraction (W-S). The SVM classifier was trained using the feature set of Session 1 data and evaluated on the feature set of Session 2 data. This study also compares the findings with the previous results presented by the authors of the dataset. The block diagram of the proposed system is illustrated in Fig. 1.

B. Data Pre-processing

MEG data was processed offline using the MNE Python [14] library. Bad channel detection, correction, jump artifacts, and head movements were corrected for, by the implementation of Signal Space Projection (SSS) and Maxwell Filtering. In the MNE Python Library, SSS and Maxwell Filtering are performed using one function to deal with environmental noise and artefacts. We implemented the inbuilt spatiotemporal Signal Space Separation method (tSSS) [15] which is activated by passing in a time value as an argument to the function. By incorporating the use of the Shannon-Nyquist Theorem, taking half the sampling frequency and dividing this by the duration to create evenly spaced temporal windows and passing this as an argument into the required function. Following this, the MEG data was then down-sampled to 500Hz. A Signal Space Projection method was used to remove and suppress the effects of eye blinks (EOG) and Electrocardiogram (ECG) artefacts from MEG Data. Notch filtering was performed using a power-line frequency of 50Hz, accounting for the 3^rd harmonic. Data was then bandpass filtered, using a double forward-backwards pass using second-order sectioning, and filtered between 1-40Hz. The data was baseline corrected in a window of -200ms to 0 ms, which was then epoched into segments using the time window of 2000ms at the window start, concluding at 6000ms for the window end, from the trial onset. It emerged that one of the subject's channels did not align between session 1 and session 2, as a result, this had to be discarded.

C. Model Architecture and Configuration

1) MEGNet: The study utilized the MEGNet architecture, inspired by the compact EEGNet model [12]. The MEGNet architecture is illustrated in Fig 2. EEGNet's effectiveness has been demonstrated across four BCI paradigms, including P300 visual-evoked potentials, ERN, MRCP, and SMR [12]. MEGNet integrates a conventional 2D convolutional layer,

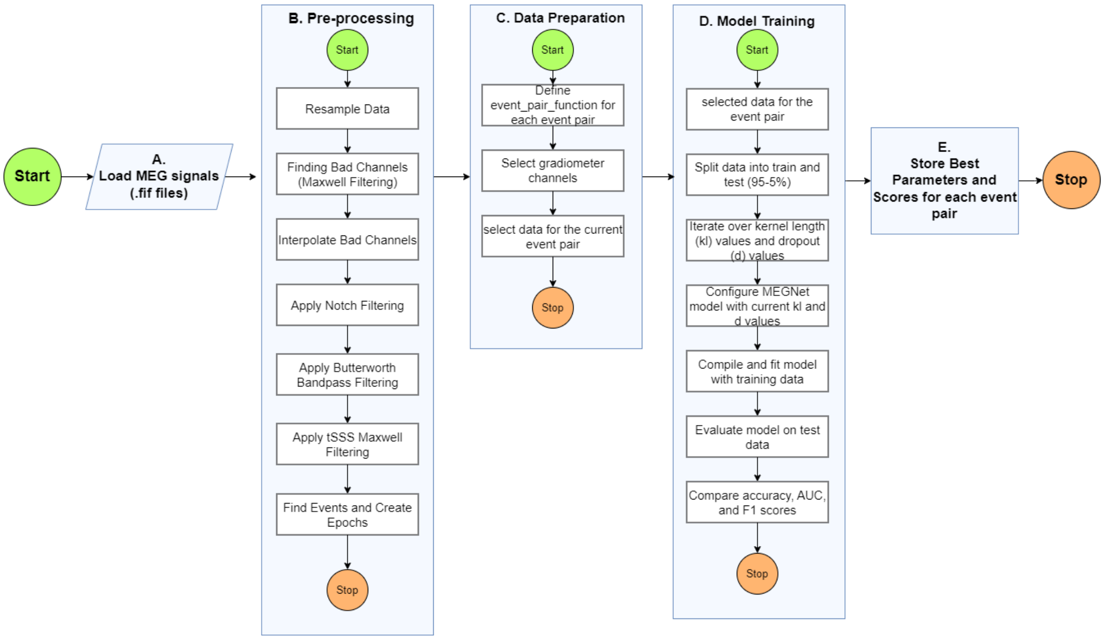

Fig. 1: MEG data analysis and deep learning classification implementation, with sample data from one participant as it goes through the classification pipeline. A. Loading of raw MEG signals of both sessions. B. Pre-processing pipeline of raw MEG signal data. C. Data Preparation of the pre-processed MEG data to meet the input requirements of the models. D. Model Training pipeline to find best parameters. E. Storing the best parameters along with accuracy score

depth-wise convolution, and a separable convolution— the latter combining depth-wise followed by point-wise convolution. This separation reduces parameters, minimizing overfitting risks. The model's adaptability allows for the extraction of spatial and temporal EEG features, a capability supported by [16], [17], and [18]. The tailored model for MEG data, detailed in Table I, was crucial for precise event classification from pre-processed MEG data. Hyper-parameters like kernel length and dropout rate were fine-tuned for optimal results, with the model compiled using categorical cross-entropy loss, the Adam optimizer, and an accuracy metric.

2) Shallow Convolutional Network and Deep Convolutional Network: The Shallow and Deep Convolutional Network serves as a foundational architecture for EEG and MEG event classification. Shallow Convolutional Networks (SCNN) and Deep Convolutional Networks (DCNN) have been used in EEG signal classification. SCNN has been proposed for motor imagery (MI) classification, achieving an accuracy of 68.77% on the BCI Competition IV-2a dataset [16]. On the other hand, DCNN has been used for alcoholism classification, achieving an average accuracy of 98% on the UCI-ML EEG dataset [19].

The models underwent a rigorous configuration and training process to achieve reliable performance. The meticulous configuration process involved the tuning of hyperparameters, including filter size, stride length, and activation functions, to optimize feature extraction. The model's compilation encompassed categorical cross-entropy loss, a stochastic gradient descent optimizer, and accuracy as the primary evaluation metric. The implementation of both Deep and Shallow ConvNets was accomplished through the utilization of the source code found in [12] with few modified parameters.

TABLE I: MEGNet modal Summary, where F1 = number of temporal filters, D = depth multiplier (number of spatial filters), F2 = number of pointwise filters, and N = number of classes, respectively.

Layer (Type)	Output Shape	Param #
InputLayer	(None, Channels, Samples, F1)	0
Conv2D	(None, Channels, Samples, F1)	64
BatchNormalization	(None, Channels, Samples, F1)	32
DepthwiseConv2D	(None, 1, Samples, F1 * D)	3264
BatchNormalization	(None, 1, Samples, F1 * D)	64
Activation	(None, 1, Samples, F1 * D)	0
AveragePooling2D	(None, 1, DownSampled, F1 * D)	0
Dropout	(None, 1, DownSampled, F1 * D)	0
SeparableConv2D	(None, 1, DownSampled, F2)	512
BatchNormalization	(None, 1, DownSampled, F2)	64
Activation	(None, 1, DownSampled, F2)	0
AveragePooling2D	(None, 1, DownSampled // N, F2)	0
Dropout	(None, 1, DownSampled // N, F2)	0
Flatten	(None, F2)	0
Dense	(None, Classes)	1986
Activation	(None, Classes)	0

D. Model Compilation and Training

The pre-processed magnetoencephalography (MEG) data underwent further refinement to meet the input requirements of the MEGNet, Deep and Shallow ConvNets. To ensure that the model focuses on relevant spatiotemporal information, reducing computational complexity but still capturing the

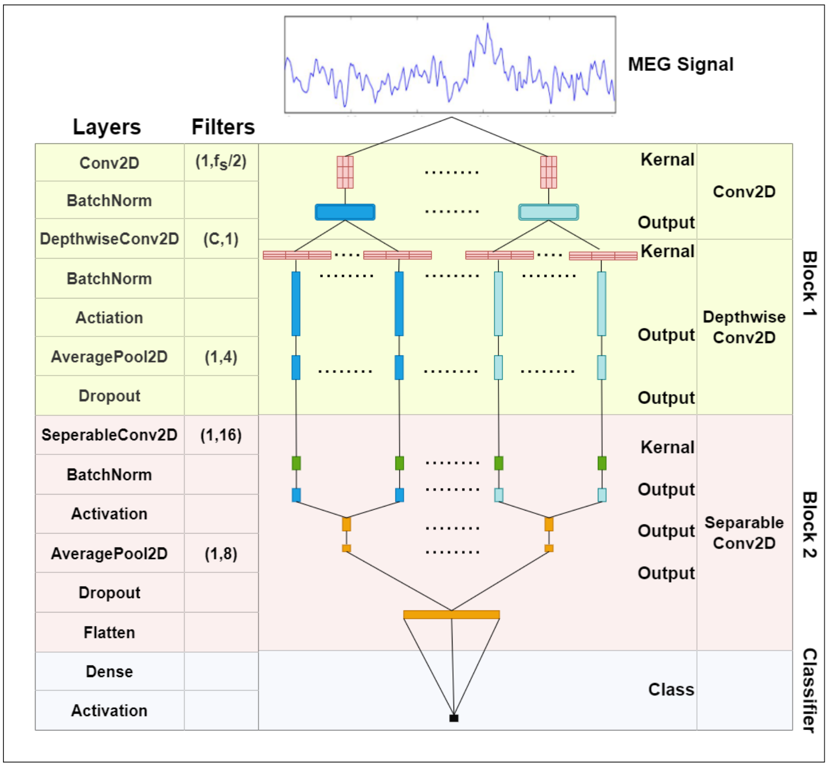

Fig. 2: The visualization of the MEGNet architecture, taken in its entirety, is a comprehensive representation of the convolutional kernel connectivity between inputs and outputs. The network initiates with a temporal convolution which is the first part of Block 1, to acquire knowledge of frequency filters. It then proceeds to utilize a depthwise convolution, located in the second part of Block 1, which is connected to each feature map individually, to learn frequency-specific spatial filters. The separable convolution (Block 2), is a combination of a depthwise convolution that learns a temporal summary for each feature map individually, and a pointwise convolution that learns how to optimally mix the feature maps together.

distinct neural activity, gradiometer channels were specifically chosen to capture the changes in the magnetic field gradient.

In consideration of the research objective, distinct event pairs were defined as the foundation for binary classification tasks. Subsequently, the dataset corresponding to each event-pair was extracted from the pre-processed datasets. Min-max scaling was utilized for data normalization, which mapped the values to the range of -1 to 1. This normalization process was implemented to prevent variations in data magnitudes from impacting the model's performance. The data was also reshaped to conform to the input format of the models, which includes dimensions denoting trials, channels, and time samples. Moreover, an additional dimension representing the number of MEG electrodes (channels) was included to ensure compatibility with the MEGNet, Deep and Shallow ConvNets architecture.

For the intra-subject learning, the epochs from both sessions were combined to create a comprehensive dataset for each subject. This merging process aimed to maximize data utilization as the amount of available data is limited in our case; we have 200 trials in each session and capture a broader range of variability. For each subject, the combined epochs were initially split in the ratio of approximately 95% training and 5% testing subsets. Within the training subset, a further subdivision was made to create a validation set.

To address the potential issue of class imbalance when creating the splits, a stratification sampling[20] technique is used to ensure that each class or category is represented proportionally in the sample. It involves dividing the dataset into subgroups based on the classes and then sampling from each stratum in a way that maintains the original class distribution, thereby reducing bias in the model assessment. This approach facilitates a more accurate representation of the model's performance.

The EEGNet model used the filter size of the first convolutional block as half of the sampling frequency rate. In this study, a grid search implementation was employed with various combinations of kernel length and dropout. By attaining the maximum training epochs or through the implementation of the early stopping strategy, the optimal weights of the network were recorded. The utilization of model checkpoints was to preserve the superior performing model as determined by the validation accuracy, while concurrently achieving an equilibrium between bias and variance. This feat was accomplished through a comprehensive evaluation process through assiduous experimentation and scrupulous analysis.

Several experiments were conducted to compare the neural networks from the EEGNet family (Shallow ConvNet, Deep ConvNet and E/MEGNet) [21][8][12],[22]. For experimental subjects who partook in various experiments on different days, the information was managed as though distinct subjects had taken part instead, which is referred to as an independent days configuration. We aimed to rank the neural networks; therefore, to evaluate the model's capacity to classify various event pairs, the accuracy score is calculated from the model's ability to correctly classify binary pair-wise configuration.

E. Significance investigation of data

The significance of the data variance was investigated. Two subjects were randomly selected to employ the t-test and compute the p-value to scrutinize the divergence between them. The resulting t-statistic value of -0.023 and the associated p-value of 0.982 provide insight into our assessment. The t-statistic measures the level of differentiation between groups, and in this study, a value near zero indicates limited distinction. On the other hand, the p-value represents the probability of observing such a discrepancy by chance. In this case, the high p-value suggests that the observed variance among subjects falls within the range of random fluctuations, indicating a lack of statistically significant differentiation.

III. Results

In pursuit of understanding and harnessing the capabilities of MEGNet, we embarked on a comprehensive comparison of three distinct networks MEGNet (derived from EEGNet), ShallowConvNet, and DeepConvNet. To comprehend the strengths and limitations of these networks in the context of event prediction using MEG data. Our investigation delved into the nuanced differences between EEGNet (MEGNet), ShallowConvNet, and DeepConvNet. These networks, each with its unique configuration, were put to the test to gauge their predictive prowess. The architecture variations spanned from MEGNet's specialized depth-wise and separable convolutions to the simpler yet potent design of ShallowConvNet and the more complex layers of DeepConvNet. To unravel the networks' true potential, we cautiously evaluated their performance across a set of event pairs. These pairs, encompassing diverse cognitive tasks, provided a robust and varied ground for assessment. The selection of event pairs and their corresponding functions added a layer of specificity to the evaluation, ensuring a comprehensive exploration of the networks' capabilities.

The experiment involved careful tuning of hyperparameters to optimize the model's performance. Through meticulous experimentation with different values of kernel length and dropout rate, we discerned the configurations that led to the highest accuracy score. This iterative process underscored the importance of hyperparameter tuning in fine-tuning the model's predictive capabilities. As we turned our attention to individual event pairs, we observed nuanced variations in the model's performance. All the neural network models showcased varying degrees of accuracy scores across different event pairs. These results provided a deeper understanding of the brain's responses to different cognitive tasks, shedding light on the intricacies that define our cognitive experiences.

TABLE II: The classification accuracy (%) for Hand vs Feet (H-F)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	58	70	50	60
Sub 3	74	70	80	70
Sub 4	50	70	60	80
Sub 6	47	60	50	70
Sub 7	51	50	70	60
Sub 9	51	50	80	50
Sub 11	47	50	50	50
Sub 12	56	60	80	70
Sub 13	49	70	80	50
Sub 14	50	50	70	50
Sub 15	80	70	90	50
Sub 16	57	60	70	60
Sub 17	55	60	40	50
Sub 18	53	70	70	60
Sub 19	54	90	50	80
Sub 20	64	50	60	50
Mean	56	62.5	65.625	60
Std	9.37	11.25	14.59	10.95

TABLE III: The classification accuracy (%) for Hand vs Subtraction (H-S)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	53	60	80	50
Sub 3	95	80	90	80
Sub 4	50	70	40	70
Sub 6	50	60	80	70
Sub 7	62	70	60	50
Sub 9	83	50	60	50
Sub 11	91	80	90	60
Sub 12	51	50	80	50
Sub 13	86	60	80	50
Sub 14	50	50	70	60
Sub 15	56	50	70	50
Sub 16	57	70	80	50
Sub 17	71	50	60	60
Sub 18	88	80	90	60
Sub 19	56	80	70	60
Sub 20	90	60	100	50
Mean	68.06	63.75	75	57.5
Std	17.57	12.0	15.05	9.30

Tables 2-5 present the accuracy score of the SVM classifier reported by [10] along with the performance of three distinct neural network architectures: MEGNet, ShallowConvNet, and DeepConvNet, across six pair-wise binary classification tasks involving motor and cognitive imagery-based brain-computer interfaces. Our main focus lies in understanding how each model performs in terms of mean accuracy and the variability of these accuracies across the tasks. When it comes to distinguishing between hand and feet movements given in Table II, we observed that the MEGNet model exhibited a superior mean accuracy of 62.%, surpassing the traditional SVM classifier's mean accuracy of 56%. This indicates the potential of deep learning in decoding intricate cognitive tasks. ShallowConvNet hovered around 65.62%, and Deep-

TABLE IV: The classification accuracy (%) for Hand vs Word (H-W)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	53	70	60	50
Sub 3	94	60	80	60
Sub 4	50	70	80	70
Sub 6	54	90	60	80
Sub 7	69	70	60	50
Sub 9	67	60	80	70
Sub 11	86	50	80	60
Sub 12	56	60	100	50
Sub 13	90	50	100	50
Sub 14	62	50	50	60
Sub 15	65	80	80	50
Sub 16	62	70	70	70
Sub 17	57	60	60	50
Sub 18	88	50	60	50
Sub 19	61	60	60	60
Sub 20	91	80	90	50
Mean	69.06	64.37	73.125	58.12
Std	15.35	12.09	15.37	9.81

TABLE V: The classification accuracy (%) for Feet vs Word (F-W)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	61	60	60	70
Sub 3	70	80	90	70
Sub 4	50	70	90	50
Sub 6	58	80	80	90
Sub 7	87	100	80	70
Sub 9	50	60	90	50
Sub 11	83	50	70	50
Sub 12	68	50	70	50
Sub 13	91	50	70	60
Sub 14	54	70	50	50
Sub 15	69	70	80	50
Sub 16	62	60	70	70
Sub 17	55	80	80	80
Sub 18	45	50	60	60
Sub 19	57	60	80	70
Sub 20	87	80	90	50
Mean	65.43	66.87	75.62	61.87
Std	14.66	14.47	12.09	12.76

TABLE VI: The classification accuracy (%) for Feet vs Subtraction (F-S)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	51	80	80	70
Sub 3	69	50	90	60
Sub 4	50	70	60	70
Sub 6	57	80	80	50
Sub 7	66	80	60	50
Sub 9	75	80	80	80
Sub 11	85	50	70	50
Sub 12	59	60	60	70
Sub 13	87	60	60	50
Sub 14	55	50	80	50
Sub 15	75	50	100	50
Sub 16	70	50	80	60
Sub 17	70	70	70	100
Sub 18	64	50	70	50
Sub 19	63	80	60	80
Sub 20	91	70	90	80
Mean	67.93	65.62	74.37	63.75
Std	12.42	12.63	12.63	15.43

TABLE VII: The classification accuracy (%) for Subtraction vs Word (S-W)

	Models			
	SVM (FB1)	MEGNet	ShallowConvNet	DeepConvNet
Freq. band	8-12 Hz	1-40 Hz	1-40 Hz	1-40 Hz
Sub 1	51	70	70	50
Sub 3	74	70	80	50
Sub 4	49	60	70	50
Sub 6	53	80	80	80
Sub 7	58	70	50	70
Sub 9	70	60	60	50
Sub 11	90	80	70	50
Sub 12	63	70	50	50
Sub 13	57	70	50	60
Sub 14	65	50	60	30
Sub 15	59	60	60	50
Sub 16	72	80	80	70
Sub 17	47	80	70	60
Sub 18	60	50	50	80
Sub 19	62	80	50	60
Sub 20	77	50	50	50
Mean	62.93	67.5	62.5	56.87
Std	11.43	11.25	11.83	13.02

ConvNet stood at about 60%. This translates to MEGNet and ShallowConvNet performing fairly well, whereas DeepConvNet lagged slightly behind. What‘s interesting is that while ShallowConvNet scored higher on average, MEGNet showed more consistent accuracy results across different scenarios. In distinguishing hand movements from mental subtractions given in Table III, MEGNet exhibited an average accuracy of 63.75%, outperforming the SVM classifier with a mean accuracy of 56%, ShallowConvNet settled at around 75%, and DeepConvNet emerged as the leader with roughly 57.5%. This time, ShallowConvNet took the lead with its higher average accuracy and a moderate level of stability. ShallowConvNet remained steady, but MEGNet demonstrated consistency in its performance. In the case of foot movements versus mental subtractions given in Table VI, MEGNet maintained an average accuracy of approximately 65.62%, ShallowConvNet stayed around 74.34%, and DeepConvNet reached an average of about 63.75%. MEGNet displayed consistent accuracy results and ShallowConvNet held its ground, while DeepConvNet showed a bit more variation in its performance. While distinguishing hand movements from imagined words given in Table IV, MEGNet achieved an average accuracy of around 64.37%, ShallowConvNet peaked at roughly 73.12%, and DeepConvNet settled at around 58.12%. ShallowConvNet claimed the highest average accuracy, yet MEGNet and DeepConvNet seemed to trade some accuracy for consistency. For foot movements from imagined words given in Table V, MEGNet showcased an average accuracy of 66.87%, ShallowConvNet rose to around 75.62%, and DeepConvNet lagged a bit at 61.87%. ShallowConvNet shone brightly in terms of both accuracy and stability, while MEGNet maintained a competitive edge. Lastly, in the case of mental subtractions versus imagined words given in Table VII, MEGNet stayed around 67.5%, ShallowConvNet hovered at 62.5%, and DeepConvNet scored around 56.87%. MEGNet demonstrated stable performance, ShallowConvNet retained its accuracy with moderate fluctuations, and DeepConvNet showcased a wider range of results. When comparing the current research results with the previous work conducted by the dataset author [10], where a Support Vector Machine (SVM) classifier was employed, notable differences and advancements become apparent. In the previous study, the SVM classifier achieved accuracy levels that ranged from approximately 50% to 95% across various cognitive event pairs. Interestingly, the current deep learning models, including MEGNet, ShallowConvNet, and DeepConvNet, exhibited competitive or even improved accuracy in many cases, indicating the potential for neural networks to outperform traditional feature extraction methods and machine learning classifiers in capturing intricate patterns within MEG data.

IV. Discussion

Previous studies [23][10] have reported significant classifier performance at the individual level, which was reliant on dramatic feature extraction and used multivariate pattern analysis to decode MEG signals. Conversely [7] used a linear discriminant analysis (LDA) classifier with 5-fold cross-validation to classify and evaluate the brain response to visual stimuli. In their research [24] used EEGNet to categorize objects, specifically faces, tools, animals, and scenes captured from MEG data with very high levels of accuracy in both binary and multi-class classification settings. The prevailing challenge however lies in the decoding of motor tasks from cognitive imagery using the high spatial and temporal precision of MEG, leading towards identifiable areas responsible for each related task; an objective that is addressed in our current study.

The consistency of our results with the previous literature [10] proved the capabilities of deep learning models to perform vast feature extraction and decoding of motor tasks from cognitive imagery in MEG signals without the use of additional algorithms such as the Common Spatial Pattern (CSP) [10] or Independent Component Analysis (ICA) [25]. The study reveals that the classification performance of the EEGNet-inspired MEGNet model and ShallowConvNet were similar across the subjects achieving high accuracy levels, especially in tasks that involve distinguishing between motor and cognitive imagery. This propensity for accuracy could be attributed to its shallow structure that enables efficient extraction of key features. However, ShallowConvNet also exhibits more variability, which suggests that its performance might be sensitive to specific task characteristics. ShallowConvNet efficiently captures prominent features from MEG signals but may struggle with capturing nuanced patterns present in more complex tasks. The MEGNet architecture allows for more comprehensive pattern extraction, capturing both spatial and temporal characteristics of MEG signals, making it a suitable choice for various MEG tasks. In contrast, DeepConvNet displayed poorer performance showcasing also the highest variability of all the models, highlighting its sensitivity to task complexity and architecture parameters. DeepConvNet, as its name suggests, adopts a more complex architecture with a greater number of layers. This suggests that DeepConvNet requires more data to be trained effectively than MEGNet. These observations emphasize the importance of a balanced consideration between model performance and stability in the context of brain-computer interface applications.

V. Conclusions

In the present study, we conducted a comprehensive analysis of three deep neural network models, namely MEGNet, ShallowConvNet, and DeepConvNet on MEG signal classification. The comparison was conducted on a dataset containing motor and cognitive imagery tasks. We performed pair-wise decoding of cognitive and motor tasks. Through an extensive evaluation of accuracy and stability, we provide valuable insights into the strengths and limitations of each model. Our findings underscore the significance of the trade-off between accuracy and stability in model selection for these applications. The MEGNet model shows consistent performance and generalizes well across paradigms compared to ConvNet. It also performs comparably to the reference algorithms even with limited training data across all tested paradigms. ShallowConvNet excels in terms of accuracy, although its performance varies with task intricacies. DeepConvNet, while occasionally competitive, exhibits more significant variability. In our study, MEGNet exhibited an average accuracy of 64.37% (± 3%), positioning it slightly below the performance of other established models like ShallowConvNet. Despite this, it is essential to underscore the unique effectiveness and advantages offered by MEGNet. The model's capabilities may extend beyond a singular accuracy metric, showcasing strengths in specific scenarios or applications. Moreover, it is imperative to acknowledge the limitations inherent in MEGNet. Recognizing these limitations provides valuable insights and perspectives for future research and enhancement. By presenting a nuanced understanding of both the strengths and limitations, we aim to foster a comprehensive evaluation of MEGNet's applicability in diverse contexts. This study serves as a stepping stone for further exploration and refinement of MEGNet's potential in neuroimaging applications.

Acknowledgement

HR was supported by the Economic and Social Research Council (ESRC) funded Business and Local Government Data Research Centre under Grant ES/S007156/1. SN acknowledges the support from "The Adaptive Mind" funded by the Excellence Program of the Hessian Ministry of Higher Education, Science, Research and Art. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

REFERENCES

[1] J. Gross, "Magnetoencephalography in cognitive neuroscience: a primer," Neuron, vol. 104, no. 2, p. 189–204, 2019.

[2] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An meg-based brain–computer interface (bci)," *Neuroimage*, vol. 36, no. 3, pp. 581–593, 2007.

[3] S. Braeutigam, "Magnetoencephalography: fundamentals and established and emerging clinical applications in radiology," *International Scholarly Research Notices*, vol. 2013, 2013.

[4] N. Dijkstra, S. E. Bosch, and M. A. van Gerven, "Shared neural mechanisms of visual perception and imagery," *Trends in cognitive sciences*, vol. 23, no. 5, pp. 423–434, 2019.

[5] S. N. Kraeutner, S. N. Eppler, A. Stratas, and S. G. Boe, "Generate, maintain, manipulate? exploring the multidimensional nature of motor imagery," *Psychology of Sport and Exercise*, vol. 48, p. 101673, 2020.

[6] N. Leeuwis, A. Paas, and M. Alimardani, "Vividness of visual imagery and personality impact motor-imagery brain computer interfaces," *Frontiers in Human Neuroscience*, vol. 15, p. 634748, 2021.

[7] S. Nara, H. Raza, M. Carreiras, and N. Molinaro, "Decoding numeracy and literacy in the human brain: insights from meg and mvpa," *Scientific Reports*, vol. 13, no. 1, p. 10979, 2023.

[8] I. A. Abdellaoui, J. G. Fernandez, C. Sahinli, and S. Mehrkanoon, "Deep brain state classification of meg data," *arXiv preprint arXiv:2007.00897*, 2020.

[9] T. N. Alotaiby, S. R. Alrshoud, S. A. Alshebeili, and M. H. Alhumaid, "Automatic epileptic tendency screening using statistical features of meg data and svm," in *2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA)*. IEEE, 2019, pp. 1–3.

[10] D. Rathee, H. Raza, S. Roy, and G. Prasad, "A magnetoencephalography dataset for motor and cognitive imagery-based brain-computer interface," *Scientific Data*, vol. 8, no. 1, p. 120, 2021.

[11] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (eeg) classification tasks: a review," *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.

[12] 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, 2018.

[13] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[14] A. Gramfort, M. Luessi, E. B. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. Hämäläinen, "Mne software for processing MEG and EEG data," vol. 86, pp. 446–460, 02 2014.

[15] S. Taulu and J. Simola, "Spatiotemporal signal space separation method for rejecting nearby interference in meg measurements," *Physics in Medicine and Biology*, vol. 51, p. 1759–1768, 04 2006.

[16] Y. Han, B. Wang, J. Luo, L. Li, and X. Li, "A classification method for EEG motor imagery signals based on parallel convolutional neural network," *Biomedical Signal Processing and Control*, vol. 71, p. 103190, 2022.

[17] H. Raza, A. Chowdhury, S. Bhattacharyya, and S. Samothrakis, "Single-trial EEG classification with EEGNet and neural structured learning for improving bci performance," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.

[18] M. Riyad, M. Khalil, and A. Adib, "Mi-EEGNET: A novel convolutional neural network for motor imagery classification," *Journal of Neuroscience Methods*, vol. 353, p. 109037, 2021.

[19] T. Wen, Y. Du, T. Pan, C. Huang, Z. Zhang *et al.*, "A deep learning-based classification method for different frequency EEG data," *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.

[20] T. Amr, *Hands-On Machine Learning with scikit-learn and Scientific Python Toolkits: A practical guide to implementing supervised and unsupervised machine learning algorithms in Python*. Packt Publishing Ltd, 2020.

[21] R. T. Schirrmeister, J. T. Springenberg, D. Josef, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, pp. 5391–5420, 08 2017.

[22] D. Rathee, H. Cecotti, and G. Prasad, "Single-trial effective brain connectivity patterns enhance discriminability of mental imagery tasks," *Journal of Neural Engineering*, vol. 14, p. 056005, 08 2017.

[23] I. Simanova, M. Van Gerven, R. Oostenveld, and P. Hagoort, "Identifying object categories from event-related EEG: toward decoding of conceptual representations," *PloS one*, vol. 5, no. 12, p. e14465, 2010.

[24] R. Shi, Y. Zhao, Z. Cao, C. Liu, Y. Kang, and J. Zhang, "Categorizing objects from meg signals using EEGNet," *Cognitive Neurodynamics*, pp. 1–13, 2021.

[25] A. H. Treacher, P. Garg, E. Davenport, R. Godwin, A. Proskovec, L. G. Bezerra, G. Murugesan, B. Wagner, C. T. Whitlow, J. D. Stitzel *et al.*, "Megnet: automatic ica-based artifact removal for meg using spatiotemporal convolutional neural networks," *NeuroImage*, vol. 241, p. 118402, 2021.