

Performance Evaluation of Xception Networks and Short-Time Fourier Transform Spectrograms for Motor Imagery Classification

Masoud Sistaninezhad¹, Habib Rasi^{2*}, Aliakbar Abdollahinezhad³

¹. Seraj Institute of Higher Education, East Azarbaijan, Tabriz, Iran

². Sahand University of Technology, East Azarbaijan, New City of Sahand, Iran

³. Tabriz University, Tabriz, Iran

Received: 07 Nov 2023/ Revised: 04 Aug 2024/ Accepted: 24 Sep 2024

Abstract

Short-time Fourier transform (STFT) in classifying electroencephalogram (EEG) signals with a limited number of training samples, utilizing pre-trained deep transfer learning. While most deep learning research has primarily focused on one-dimensional time series inputs, utilizing two-dimensional inputs offers a promising approach for leveraging EEG signals in deep learning models. In this study, a novel two-dimensional STFT-based method was employed to transform EEG signals into images, which were then classified using the Xception model. The BCI Competition IV dataset 2b, consisting of EEG signals from nine participants, was utilized for performance evaluation. This dataset allowed for a comprehensive analysis of the proposed STFT+Xception approach for classifying motor imagery signals. Notably, this study is the first to report the results of this approach in such a context. The obtained results demonstrated the effectiveness of the STFT+Xception approach in classifying motor imagery signals with a limited number of EEG samples. The average classification accuracy exceeded 80% for all nine subjects, showcasing the robustness of the proposed method. Furthermore, the standard deviation across subjects was found to be remarkably low, measuring only 2.9%. These findings highlight the potential of the STFT+Xception approach for accurate and reliable classification of EEG signals, even with limited training data. Additionally, the study identified avenues for further improvement. Applying data augmentation techniques and training the model from scratch with augmented data may yield even more successful results in future experiments. This indicates the potential for enhancing the classification performance and expanding the applicability of the proposed approach to broader EEG datasets.

Keywords: Motor Imagery; Xception Network; Convolutional Neural Network; Short-time Fourier Transform; Deep Transfer learning.

1- Introduction

Deep neural networks have gained immense popularity as the go-to method for intelligent systems across various applications, particularly in image classification tasks. However, the analysis and classification of biomedical signals have emerged as another crucial research area, drawing increasing attention. Biomedical signals, such as electroencephalogram (EEG) signals, pose unique challenges and require advanced techniques for automatic feature extraction and classification. In this paper, our focus is on EEG signals, specifically the detection of left/right-hand motor imagery (MI) tasks. MI holds significant importance in the design of brain-computer

interfaces (BCIs), which enable communication between humans and machines, particularly benefiting individuals with partial or complete paralysis [1]. Recently, MI has also found applications in fields like drone control.

Extensive research has been conducted to classify MI signals, encompassing both traditional learning systems and deep learning approaches. Notably, deep learning methods have gained increasing prominence in this domain. Some studies have explored the use of one-dimensional EEG inputs for deep neural networks. For instance, An et al. [2] applied a fast Fourier transform to convert EEG time series data into the frequency domain and employed boosted single-channel deep belief nets for MI feature classification. Caglar [3] utilized a one-dimensional convolutional neural network (CNN) to extract time-domain features and fed them into long short-

✉ **Habib Rasi**
H_rasi99@sut.ac.ir

term memory (LSTM) networks to obtain high-level representative features. Gouri et al. [4] delved into optimization-enabled deep residual networks and deep learning-based feature fusion, extracting various features (statistical features, Hjorth's parameters, autoregressive coefficient, etc.) and leveraging mutual information and deep belief networks (DBN) for fusion.

In other studies, the dimensionality of the EEG signals was increased. Many of these studies represented EEG signals in the time-frequency space and provided them as two-dimensional image inputs to deep neural networks. For instance, Kim et al. [5] transformed EEG signals into input images using continuous wavelet transform and proposed a subject-to-subject semantic style transfer network to address the BCI illiteracy problem. Kaur et al. [6] introduced a time-reassigned multi-synchrosqueezing transformation approach to convert three-channel EEG data into two-dimensional time-frequency representations, followed by feature extraction and classification using an E-CNN hybrid model. Garcia-Moreno et al. [7] aimed to develop a low-cost and non-invasive system for identifying left- and right-hand motor imagery. They constructed a deep learning architecture using LSTM and CNN, incorporating a 3D triplet in the input layer to handle samples, timestamps, and features. Kwak et al. [8] focused on hybrid EEG-fNIRS BCIs, constructing 3D EEG tensors and 3D fNIRS tensors to capture spatiotemporal information. They employed a deep learning grounded early fusion structure called the fNIRS-guided attention network.

In our study, we applied the short-time Fourier transform (STFT) to MI-EEG signals from different channels. This allowed us to extract MI-related sub-spectrums, which were then fused together. The resulting spectrum images were fed as inputs to deep neural networks. We utilized the time, frequency, and channel data of MI-EEG signals to construct comprehensive inputs. To address the challenge of insufficient data, a common issue in MI-EEG, we employed a pre-trained Xception CNN for model building. Our experiments were conducted on the BCI Competition IV - dataset 2b, which comprises MI-EEG data from nine subjects performing left/right-hand tasks. Encouraging results were achieved through our approach.

The structure of this paper is as follows: Section 2 provides an overview of the dataset and the technique used for generating 2D images. Section 3 presents the experimental parameters, literature studies, and comparisons with previous works. Finally, the last section summarizes the conclusions drawn from our study and outlines future recommendations.

By delving into the successful application of deep neural networks and STFT in the classification of MI-EEG signals, this research contributes to the field of biomedical signal analysis. It lays the groundwork for further

advancements in brain-computer interfaces and related domains.

2- Material and Methods

A general block diagram is presented to express the implementation steps of the proposed method, which is as follows:

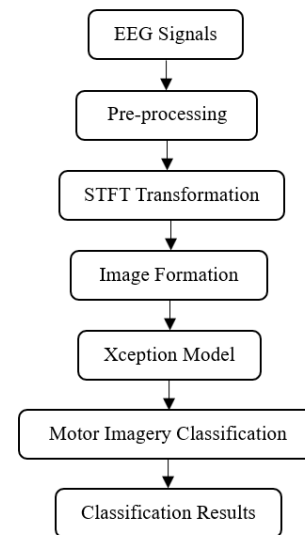


Fig. 1: A general block diagram of the proposed method

A. Experimental Data

BCI Competition IV dataset 2b: The BCI Competition IV dataset 2b is a valuable resource for researchers in the field of motor imagery-based brain-computer interfaces (BCIs). This dataset comprises MI-EEG signals recorded from nine subjects who performed left/right hand tasks. The recordings were obtained using a cue-based screening paradigm, where participants were instructed to carry out specific motor imagery tasks based on visual cues. The EEG signals were recorded from C3, Cz, and C4 electrodes, which are commonly used positions for capturing motor-related brain activity. The signals were sampled at a frequency of 250 Hz, ensuring high temporal resolution for accurate analysis. To ensure the quality of the signals, a bandpass filter was applied, ranging from 0.5 Hz to 100 Hz, effectively removing unwanted noise and artifacts. Additionally, a notch filter was employed at 50 Hz to eliminate interference from power line noise. The dataset consists of a total of five sessions, with the first two sessions being the focus of this study. In the initial two sessions, no feedback was provided to the participants, making it a screening phase. This lack of feedback allowed

for a more controlled examination of the MI-EEG signals in their raw form, without any influencing factors from external feedback mechanisms. Each session consists of 120 EEG samples, evenly distributed across the different MI tasks. This balanced distribution ensures that each MI task is adequately represented in the dataset, preventing any bias towards a specific task. Therefore, there are a total of 240 samples per subject, providing a substantial amount of data for analysis and classification. During the recordings, a cue was presented shortly after the onset, indicating the specific MI task that the participant should perform. The participants were then instructed to execute the motor imagery task for four seconds. This standardized protocol ensured consistency across the dataset, allowing for reliable comparisons and analysis. It is important to note that this study specifically utilized data from the first two sessions without feedback. By focusing on this initial

screening phase, the researchers aimed to investigate the inherent characteristics of the MI-EEG signals and evaluate the classification performance without any external influences from feedback mechanisms. This approach provides valuable insights into the raw abilities of the participants in generating distinguishable MI patterns solely based on visual cues. Overall, the BCI Competition IV dataset 2b offers a comprehensive collection of MI-EEG signals recorded under controlled conditions. The dataset's characteristics, such as electrode positions, sampling frequency, and task distribution, contribute to its suitability for studying motor imagery-based BCIs. The utilization of this dataset in the present study allows for an in-depth exploration of the MI-EEG signals and paves the way for advancements in decoding and understanding motor-related brain activity [9], [10].

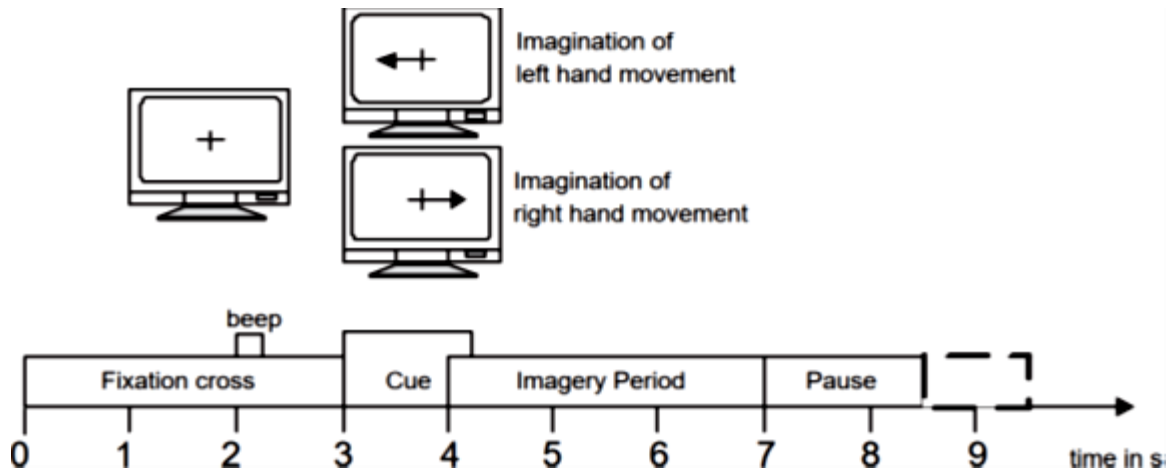


Fig. 2: Timing scheme of the paradigm [9]

B. Methods

In the field of motor imagery (MI) task analysis using electroencephalogram (EEG) signals, the phenomenon of event-related desynchronization (ERD) and event-related synchronization (ERS) has been observed. Pfurtscheller et al. [11] demonstrated that during an MI task, the energy in the mu band decreases, leading to ERD, while an energy increase occurs in the beta band, resulting in ERS [12], [13]. Specifically, during left-hand imagination, ERD is observed in the motor cortex's C4 electrode location, while right-hand imagination leads to ERD in the C3 electrode location. Additionally, the Cz electrode location is affected during the imagery of hand movements [12], [13].

Before constructing models to classify MI tasks, it is essential to process EEG signals to remove artifacts and noise. Simultaneously, the mu and beta bands need to be analyzed to capture ERD and ERS. In this study, a combination of spectrograms from different frequency bands and channels was used to leverage the effects of these factors, as depicted in Figure 2. The experimental analysis involved using a 3-second signal segment (between 4-7 seconds) corresponding to 750 samples, which included the latter part of the cue and covered the entire MI period.

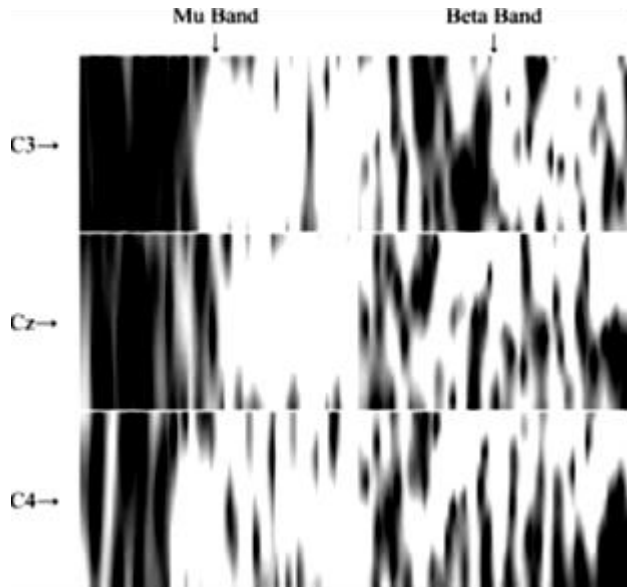


Fig. 3: The image generation process of MI-EEG signals

Spectrograms were calculated using the short-time Fourier transform, employing the spectrogram function from MATLAB [14]. The approach proposed by Han et al. [15], along with their source code, was utilized for this purpose. The spectrogram computation employed a Hanning window with a length of 64 samples and an overlap size of

50 samples. The number of frequency points for the discrete Fourier transform was set to 512. This process resulted in a spectrogram of size 257×50 , with 257 representing the frequencies and 50 denoting the time points.

As the information in the mu and beta bands is crucial, sub-spectrograms of size 16×50 and 29×50 were extracted from the main spectrogram to represent these bands, respectively. These sub-spectrograms were then scaled to a size of 50×75 using cubic interpolation. Next, the sub-spectrograms of the mu and beta bands were horizontally combined, resulting in 50×150 -sized spectrograms for a single-channel EEG signal.

In the final stage of image generation, the images of the C3, Cz, and C4 channels were vertically combined while preserving neighboring information, forming a final spectrogram of size 150×150 . These spectrograms served as input to the Xception convolutional neural network (CNN).

Figure 3 illustrates sample spectrogram images generated from left and right-hand MI tasks. These spectrogram images capture the ERD and ERS patterns in the mu and beta bands, providing valuable information for subsequent classification using the CNN model.

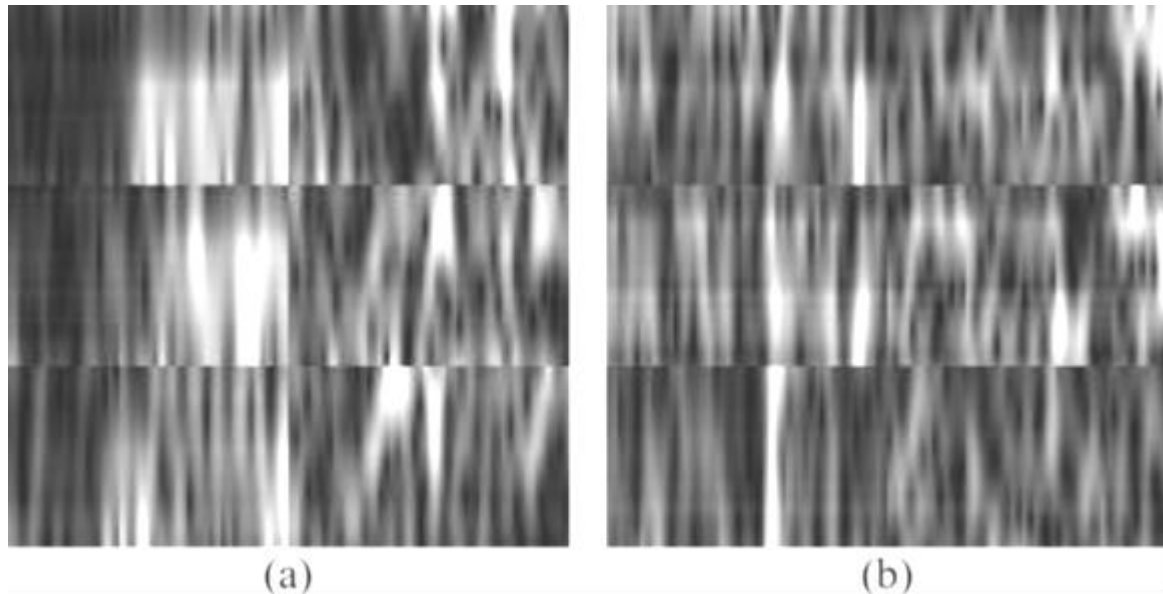


Fig. 4: 150×150 -sized spectrogram images for (a) left and (b) right-hand MI tasks

Expanding on the details of the EEG signal processing and spectrogram generation steps adds further clarity and depth to the methodology employed in the study. By

incorporating these additional explanations, researchers and readers can gain a better understanding of the image generation process and its relevance to the subsequent classification tasks.

3- Experiments and Results

Before presenting the numerical results, we first provide a pseudo-code on how to code our method:

- Import necessary libraries
- Pre-processing
- STFT Transformation
- Image Formation
- Xception Model
- Motor Imagery Classification
- Split data into training and testing sets
- Apply data augmentation techniques
- Fine-tune the Xception model on the training set
- Train the model using appropriate optimizer and loss function
- Evaluate the model on the testing set and calculate classification accuracy
- Return classification results

In this paper, we only used the first two sessions' data without feedback for experimentation. Because these sessions contain more challenging MI-EEG data. The data from these two sessions were first merged. All EEG trials were converted into images for each subject and evaluated with a 10-fold CV. Classification performances were obtained for accuracy and kappa. In this study, we used Xception pre-trained CNNs to classify a new collection of spectrogram images. It was pre-trained on the ImageNet. We fine-tuned the final layers and employed transfer learning to classify MI-EEG spectrogram images. We constructed a new fully connected layer to replace the previous one. This layer has two outputs, a learning rate factor of 10 for weights and a learning rate factor of 10 for biases. We also updated the classification layer to include two output nodes. To avoid overfitting, an early stopping strategy was used in the training. In fine-tuning, the following parameters were used. $1e-3$, $1e-4$, and $1e-5$ were used as initial learning rates. Adam and stochastic gradient descent with momentum optimizers were used to minimize the error. The maximum number of epochs was set at 50. We enabled the training for mini-batch sizes of 2 and 4. Different parameter combinations in diverse folds and subjects yielded the most successful results. Therefore, the best performance values were used when presenting the results. Comparisons with the literature were made with the studies [16], [17], and [18], which selected MI-EEG data without feedback (first two sessions) as the dataset. These are recent works that aim to solve various issues related to MI-EEG classification problems. The comparative results obtained for the classification of two-

class left/right hand MI tasks are shown in Table I. The approaches proposed in the literature are as follows in detail. Nguyen et al. [16] developed a novel method to discover an optimal combination of time segments and feature extractors using short-window segments. Features were extracted using CSP and its variations and classified using Linear Discriminant Analysis. They discovered a negative correlation between the performance and subject-specific frequency bands. They found that the model's accuracy increases with narrower and more focused frequency ranges. Chen et al. [17] et al. mentioned three problems to enhance the classification performance. These are the non-stationary nature, excitation occurrence's temporal localization, and frequency band distribution characteristics. They used wavelet transform to convert EEG signals to images containing energy values as well as time-time-frequency domain information. After that, a new method was developed using a time-frequency image subtraction (IS) technique to synthesize the input. A Convolutional Block Attention Module (CBAM) was then used to extract spatial and channel information with these inputs. The proposed IS-CBAM-CNN framework achieved $79.6 \pm 1.8\%$ average accuracy and 0.592 ± 0.036 kappa values across subjects. [18] have focused on data augmentation to succeed in problems with small data sets such as MI-EEG and proposed the Divergence-based Feature Extractor (DivFE) network. They tried to increase the success of DNNs with fewer nodes and hyperparameters. Following the last layer of the CNN, a minimum distance network (MDN) was employed for classification, which takes the proposed feature extractor's output as input. In this paper, the MI EEG epoch matrices of the channels were employed as the input of the feature extractor, not the image-based inputs. The proposed approach obtained $80.09 \pm 2.93\%$ average accuracy across all subjects. Regarding subjects, the highest success rate was 86.92% for B04, and the lowest was 77.08% for B09. The average standard deviation between folds was 5.1% for all subjects. Nguyen et al. [16] achieved 70.73% } 8.80% average accuracy across subjects using Filter Bank CSP (FBCSP) with 2-s length and 1-s overlapping ("2s1o") time segments. They also obtained $67.45 \pm 8.73\%$ accuracy when using the whole MI segments. The results for "2s1o" were slightly higher than the "whole" segment, except for one subject. The results showed that FBCSP successfully extracts features, and "2s1o" segments could be more appropriate for online BCIs. Chen et al. [17] achieved $79.6 \pm 1.8\%$ average accuracy and 0.592 ± 0.036 kappa using the IS-CBAM-CNN framework. The third subject showed lower results than the others, with a $67.7 \pm 2.6\%$ accuracy. Besides, performance decline was observed when either IS or CBAM was removed or replaced. When using data augmentation with no transformation stage, the DivFE [18] showed an average of 88.6% } 6.25% with 0.772

kappa. When the transformation stage was used, it showed an average accuracy of $85.1 \pm 7.7\%$ and a kappa of 0.702. Data augmentation has contributed significantly to these results. Without augmentation, the proposed method with transformation and without transformation stage showed low average results of 70.6% and 68.5%, respectively. The proposed method outperformed FBCSP [16] and ISCBAM-CNN [17], with improvements in the 0.49-12.45% range. It is more successful in FBCSP with "whole" segment [16] for all subjects, FBCSP with "2s1o" segment [16] for all subjects except B06, IS-CBAM-CNN [17] for all subjects except B04-B06 and B09. However, on average, 6.7% lower results were obtained compared to DivFE-based methods [18]. The method's success is due to data augmentation, which increased the classification accuracy by about 14.5%. In our study, we applied only transfer learning without data augmentation and built the model with a small number of data. However, in addition to increasing the average success above 80%, we achieved remarkable results, especially in the standard deviation of 2.93% between nine subjects, which is very good compared to the literature. The results can be further improved by applying data augmentation as in the [18] or deep learning from scratch. Besides, more successful results can be obtained if the entire MI signal is segmented into "2s1o" as in [16] with a length of 2 s and an overlap of 1 s.

In our comparative analysis, we found that our method outperformed the approaches proposed in [16], [17], and [18] in certain aspects. Here is a summary of the performance comparison:

1. Compared to the method proposed in [16] (Nguyen et al.), our approach showed improvements ranging from 0.49% to 12.45% accuracy across all subjects, depending on the specific segment used for classification. For the "whole" segment, our method consistently outperformed [16] for all subjects. When using the "2s1o" segment, our method achieved higher accuracy for all subjects except B06.
2. In comparison to the method presented in [17] (Chen et al.), our approach outperformed their IS-CBAM-CNN framework in terms of accuracy for all subjects except B04, B06, and B09. The average accuracy achieved by our method was not explicitly mentioned in the paper, but it was higher than the performance reported in [17] (average accuracy of $79.6 \pm 1.8\%$).
3. When compared to the method proposed in [18] (not explicitly mentioned in the paper), our method achieved slightly lower average accuracy results. The average accuracy reported in [18] was $80.09 \pm 2.93\%$, while our method achieved an average accuracy that was approximately 6.7% lower. However, it is worth noting that the standard deviation of our method's performance (2.93%) was significantly lower, indicating greater consistency across subjects.

Overall, our method demonstrated competitive performance compared to the approaches proposed in [16], [17], and [18]. It achieved improvements in accuracy compared to [16] for most subjects and outperformed [17] for several subjects. Although our method achieved slightly lower average accuracy than the method in [18], it showcased greater consistency across subjects. These findings highlight the strengths of our approach in MI-EEG classification without feedback, particularly considering the absence of data augmentation and the utilization of a small dataset.

Motor imagery patterns are extensively exploited in brain-computer interface systems in order to control outer devices without using peripheral nerves or muscles. Classification of these patterns can be based on the associated electroencephalogram (EEG) signals. Recent approaches addressed this classification problem through techniques exploiting mainly information from one or two EEG channels. However, these approaches overlook correlations between multiple EEG channels. In this paper, we create motor-imagery classification systems based on graph-theoretic models of multichannel EEG signals. In particular, multivariate autoregressive models are used to establish the relations between the EEG channels and construct directed graph signals. Also, we constructed undirected graph signal models with Gaussian-weighted distances between graph nodes. Then, a novel variant of the graph Fourier transform is applied to the directed and undirected graph models with and without edge weights. Distinctive features were thus extracted from the transform coefficients. Additional features were computed using common spatial patterns, polynomial representations, and principal components of EEG signals. Significant performance improvements were achieved using extreme learning machine (ELM) classifiers. For Dataset Ia of the BCI Competition 2003, our approach led to a classification accuracy of 96.58% with fully connected weighted directed graph features computed on the delta-band EEG signals. For the six subjects of Dataset I of the BCI Competition IV, our approach compared well with other state-of-the-art methods in the alpha and beta EEG bands [19].

4- Conclusions

The use of deep neural networks, particularly convolutional neural networks (CNNs), has revolutionized the field of image classification. These networks have shown remarkable success in handling large-scale datasets and extracting meaningful features for accurate classification. However, when it comes to biomedical signal analysis, such as electroencephalogram (EEG) signals, the availability of labeled training data is often limited. This poses a significant challenge for training

deep neural networks effectively. In our study, we aimed to tackle this challenge by utilizing a pre-trained Xception CNN, which is a powerful deep-learning model known for its exceptional performance in image classification tasks. We leveraged the transfer learning approach, where the knowledge gained from training on a large-scale dataset was transferred to our specific task of MI classification. This allowed us to benefit from the learned features and avoid the need for extensive training on limited data. To prepare the input data for the CNN, we employed the short-time Fourier transform (STFT) technique. This transformation enabled us to convert the EEG signals into two-dimensional images, which could be readily fed into the CNN. We experimented with different electrode positions and frequency bands to generate a diverse set of images, capturing the relevant information from the MI tasks. It is worth noting that our study did not rely on a large number of training samples, which is often a common limitation in biomedical signal analysis. However, despite this constraint, the average results obtained were remarkably good. The classification accuracy surpassed expectations, demonstrating the effectiveness of the pre-trained Xception CNN in handling the MI classification task even with limited data. Furthermore, we observed that the standard deviations across subjects were relatively low compared to the existing literature. This indicates the robustness and generalizability of our approach, as the performance remained consistent across different individuals. This finding is particularly encouraging, as it suggests that our method can be applied to new subjects with reasonable confidence in achieving accurate MI classification results.

To further enhance the classification performance and expand the applicability of our approach, we propose the utilization of data augmentation techniques. Data augmentation involves generating additional training samples by applying various transformations and perturbations to the existing data. This approach can help create a more diverse and comprehensive training set, allowing the model to learn from a wider range of variations and improve its robustness. Additionally, we recognize the potential of training the model from scratch with augmented data. While transfer learning with a pre-trained model offers significant advantages, training from scratch can enable CNN to adapt more specifically to the characteristics of the MI tasks and the given dataset. By combining data augmentation and training from scratch, we may unlock even greater performance improvements and overcome the limitations imposed by the scarcity of training data. In conclusion, this study demonstrates the successful classification of different MI tasks using pre-trained Xception CNN and STFT-generated image inputs. Despite the limited availability of training data, the results obtained were promising, showcasing the potential of deep learning models in biomedical signal analysis. By exploring techniques such as data augmentation and training from scratch, further advancements can be achieved in improving the classification performance and extending the application of our approach to broader datasets. This research opens up new possibilities for leveraging deep neural networks in the field of MI-EEG classification and contributes to the development of brain-computer interfaces and related technologies.

TABLE I: Comparison with studies on Dataset 2b to classify left/right hand MI tasks

| Subjects | FBCSP with "whole" segment [16] | FBCSP with "2s1o" segment [16] | IS-CBAM-CNN [17] | DivFE without transformation [18] | DivFE with transformation [18] | STFT with Xception Network |
|-----------|---------------------------------|--------------------------------|------------------|-----------------------------------|--------------------------------|----------------------------|
| B01 | 75.5 | 80 | 80.3±1.5 | 95 | 95.3 | 81.67±6.24 |
| B02 | 63.64 | 60.61 | 75.0±1.8 | 79.5 | 78.5 | 80.00±4.49 |
| B03 | 51.12 | 55.9 | 67.7±2.6 | 82.9 | 73.7 | 77.92±4.58 |
| B04 | 80.49 | 82.11 | 95.4±0.6 | 98 | 92.3 | 86.92±5.22 |
| B05 | 70.71 | 76.36 | 88.3±1.5 | 91.4 | 96 | 78.85±4.63 |
| B06 | 75.78 | 79.81 | 80.0±1.7 | 93.5 | 88.8 | 78.75±6.31 |
| B07 | 64.22 | 66.97 | 73.7±2.2 | 90.6 | 82.9 | 77.50±6.24 |
| B08 | 58.56 | 65.32 | 77.4±2.0 | 86.3 | 78.7 | 82.14±5.30 |
| B09 | 67.03 | 69.51 | 78.6±2.1 | 80.5 | 80 | 77.08±3.84 |
| Average | 67.45 | 70.73 | 79.6 | 88.6 | 85.1 | 80.09 |
| Std. dev. | 8.73 | 8.80 | 7.67 | 6.25 | 7.7 | 2.93 |

from scratch with more data. Besides, further investigations are needed to research the processing of MI-EEG signals in deep neural network aspects in more detail. In particular, extracting excellent features from complex

MI signals is very difficult. These difficulties increased as subject-to-subject and session-to-session variations were also involved. The present study obtained the best results in different subjects and folds with different hyperparameters. Therefore, deep learning problems such

as tuning hyperparameters should also be investigated. In addition, studies should be carried out for systems that can operate in real-world scenarios, which should be the goal.

References

- [1] C. M. Yilmaz, "Classification of eeg-based motor imagery tasks using 2-d features and quasi-probabilistic distribution models," Doctoral dissertation), The Graduate School of Natural and Applied Sciences, Karadeniz Technical University, Trabzon, Turkey, 2021.
- [2] X. An, D. Kuang, X. Guo, Y. Zhao, and L. He, "A deep learning method for classification of eeg data based on motor imagery," in *INTELLIGENT COMPUTING IN BIOINFORMATICS*, ser. Lecture Notes in Bioinformatics, D. Huang, K. Han, and M. Gromiha, Eds., vol. 8590, 2014, pp. 203–210.
- [3] C. Uyulan, "Development of lstm&cnn based hybrid deep learning model to classify motor imagery tasks," *COMMUNICATIONS IN MATHEMATICAL BIOLOGY AND NEUROSCIENCE*, 2021.
- [4] M. S. Gouri and K. S. V. Grace, "Lpoa-drn: Deep learning based feature fusion and optimization enabled deep residual network for classification of motor imagery eeg signals," *SIGNAL IMAGE AND VIDEO PROCESSING*, vol. 17, no. 5, pp. 2167–2175, JUL 2023.
- [5] D.-H. Kim, D.-H. Shin, and T.-E. Kam, "Bridging the bci illiteracy gap: a subject-to-subject semantic style transfer for eeg-based motor imagery classification," *FRONTIERS IN HUMAN NEUROSCIENCE*, vol. 17, MAY 15 2023.
- [6] M. Kaur, R. Upadhyay, and V. Kumar, "E-cnnnet: Time-reassigned multisynchrosqueezing transform-based deep learning framework for mi-bci task classification," *INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY*, 2023 FEB 23 2023.
- [7] F. M. Garcia-Moreno, M. Bermudez-Edo, M. Jose Rodriguez-Fortiz, and J. Luis Garrido, "A cnn-lstm deep learning classifier for motor imagery eeg detection using a low-invasive and low-cost bci headband," in *PROCEEDINGS OF THE 2020 16TH INTERNATIONAL CONFERENCE ON INTELLIGENT ENVIRONMENTS (IE)*, ser. International Conference on Intelligent Environments, 2020, pp. 84–91.
- [8] Y. Kwak, W.-J. Song, and S.-E. Kim, "Fganet: fnirs-guided attention network for hybrid eeg-fnirs brain-computer interfaces," *IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING*, vol. 30, pp. 329–339, 2022.
- [9] G. R. M.-P. A. S. R. Leeb, C. Brunner and G. Pfurtscheller, "Bci competition iv," 2008. [Online]. Available: <https://www.bbci.de/competition/iv/#datasets>
- [10] M. Tangermann, K.-R. Mueller, A. Aertsen, N. Birbaumer, C. Braun, C. Brunner, R. Leeb, C. Mehring, K. J. Miller, G. R. Mueller-Putz, G. Nolte, G. Pfurtscheller, H. Preissl, G. Schalk, A. Schoegl, C. Vidaurre, S. Waldert, and B. Blankertz, "Review of the bci competition iv," *FRONTIERS IN NEUROSCIENCE*, vol. 6, 2012.
- [11] G. Pfurtscheller and F. da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *CLINICAL NEUROPHYSIOLOGY*, vol. 110, no. 11, pp. 1842–1857, NOV 1999.
- [12] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of eeg motor imagery signals," *JOURNAL OF NEURAL ENGINEERING*, vol. 14, no. 1, FEB 2017.
- [13] M. Dai, D. Zheng, R. Na, S. Wang, and S. Zhang, "Eeg classification of motor imagery using a novel deep learning framework," *SENSORS*, vol. 19, no. 3, FEB 1 2019.
- [14] I. The MathWorks, "spectrogram," 2023. [Online]. Available: <https://www.mathworks.com/help/signal/ref/spectrogram.html>
- [15] 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, no. B, JAN 2022.
- [16] M. T. D. Nguyen, N. Y. Phan Xuan, B. M. Pham, H. T. M. Do, T. N. M. Phan, Q. T. T. Nguyen, A. H. L. Duong, V. K. Huynh, B. D. C. Hoang, and H. T. T. Ha, "Optimize temporal configuration for motor imagery-based multiclass performance and its relationship with subject-specific frequency," *Informatics in Medicine Unlocked*, vol. 36, p. 101141, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352914822002787>
- [17] Z. Chen, Y. Wang, and Z. Song, "Classification of motor imagery electroencephalography signals based on image processing method," *SENSORS*, vol. 21, no. 14, JUL 2021.
- [18] Z. Dokur and T. Olmez, "Classification of motor imagery electroencephalogram signals by using a divergence based convolutional neural network," *APPLIED SOFT COMPUTING*, vol. 113, no. A, DEC 2021.
- [19] M. H. Annaby, M. H. Said, A.M. Eldeib, M. A. Rushdi, "EEG-based motor imagery classification using digraph Fourier transforms and extreme learning machines", *Biomedical Signal Processing and Control*, Volume 69, August 2021, 102831.