# Fusion Convolutional Neural Network for Multi-Class Motor Imagery of EEG Signals Classification

**Amira ECHTIOUI [1,2], Wassim ZOUCH [3], Mohamed GHORBEL [1], Chokri MHIRI [4,5], Habib HAMAM [2,6,7]**

[1]ATMS Lab, Advanced Technologies for Medicine and Signals, ENIS, Sfax University, Sfax, Tunisia.
[2]Faculty of Engineering, Uni de Moncton, N-B, Canada.
[3]King Abdulaziz University (KAU), Jeddah, Saudi Arabia.
[4]Department of Neurology, Habib Bourguiba University Hospital, Sfax, Tunisia.
[5]Neuroscience Laboratory "LR-12-SP-19", Faculty of Medicine, Sfax University, Sfax, Tunisia.
[6]International Institute of Technology (IIT), Sfax, Tunisia.
[7]School of Elect. Eng. and Electronic Eng., Uni. of Johannesburg, South-Africa.

Email: amira.echtioui@stud.enis.tn

***Abstract*—Classification of EEG signals based on motor imagery is an important task in Brain-Computer Interface (BCI). Deep learning approaches have been successfully used in several recent applications to learn features and classify different types of data. However, the number of researches using these approaches in BCI applications is very limited. In this paper, we aim at using the fusion of Convolutional Neural Networks (CNN) methods to improve the classification performance of EEG motor imagery signals in the framework of e-health Internet of Things. We propose and compare two classification methods based on the fusion of two CNNs. Our results show that the fusion of the CNNs with the Long Short-Term Memory (LSTM) layers offers a better classification performance compared to other state-of-the-art methods. The classification performance achieved by our proposed method using the BCI competition IV 2a dataset in terms of accuracy value is 61.68%. This method can be successfully applied to BCI systems where the amount of data is large due to daily recording.**

***Keywords*—E-Health, EEG, BCI, Motor Imagery, CNN, LSTM**

## I. Introduction

The Brain-Computer Interface (BCI) is an alternative method of communication between a person and a computer that does not depend on the normal nerve pathways of the brain or muscles. The process usually begins by recording the person's brain activities and continues with signal processing to detect the person's intentions. The appropriate signal is then sent to an external device, such as a wheelchair, which is then controlled according to the detected signal.

Recently, electroencephalography (EEG) based machine learning was proposed for e-health Internet of Things by using the TensorFlow open-source platform of TensorFlow [1]. Our work is embedded in the same spirit of e-health and builds upon our previous works [2-3].

An important goal of BCI research is to improve certain functions for a healthy person via a new signaling pathway. BCI is currently being studied in a wide range of applications, including communication tools for patients with partial or total paralysis. Recent works showed that the electroencephalogram signals produced while the patient is mentally imagining different movements can be translated into different orders.

In this paper, we explore the merging of the two Convolutional Neural Network for Motor Imagery EEG (MI-EEG) signal classification. The originality of our proposed approach is that it leads to good results compared to state of the art methods in terms of accuracy.

The rest of the paper is organized as follows. Related works are exposed in Section II. At the end of this section, our main research contributions enumerated. Section III provides a description of the data set used and details the proposed methodology. In Section IV, the proposed framework is evaluated and compared to other competing methods, while Section V highlights the importance of this paper and offers some thoughts for the future.

## II. Related Works

Numerous studies have developed classification algorithms for EEG-based MI classification. Principal Component Analysis (PCA) [4] and Common Spatial Pattern (CSP) [5] are popular algorithms in MI searches for feature extraction, while Linear Discriminant Analysis (LDA) [6] and Support Vector Machine (SVM) [7] are frequently used in MI classification tasks.

Nowadays, Deep Learning (DL) methods have gained enormous success in speech [8], image [9], text [10], video processing [11] and other areas, which outperform state of the art methods. DL methods have attracted attention in many areas because of its superior performance. They can efficiently process non-stationary, non-linear data and learn the underlying features from signals. Some DL methods are used for EEG signal classification [12-13]. CNNs have been widely used in EEG classification based on MI-EEG because of their ability to learn characteristics from local receptive fields. Since the trained detector can be used to detect abstract features by repetition of convolutional layers, CNNs are suitable for complex EEG recognition tasks and have been used extensively by many researchers [14-17].

Pretreatment of raw EEG signals can improve the signal-to-noise ratio of the EEG and the accuracy of classification,

978-1-7281-8616-0/21/$31.00 ©2021 IEEE 

but it is not necessary [18]. CNNs are biologically inspired variants of multilayer perceptrons designed to use minimal pre-processing [18].

For example, the authors of [19-20] used CNN to directly classify raw EEG signals. In [21], the authors established a deeper layer of the neural network to decode imagination or perform tasks from raw EEG signals.

Our objective is to propose a new Deep Learning-based method to enhance the performance of the motor imagery classification. The main research contributions in this work may be summarized as follows:

- The performance of the proposed classifiers is improved by pre-processing the EEG. We proceed with a removal of EOG channels and with applying a bandpass filter.
- Use of Wavelet Packet Decomposition (WPD) and CSP techniques for the extraction of frequency and spatial features respectively. This allows obtaining a set of features that will be considered as input to our proposed models.
- Addition of LSTM layers in the CNN model. The accuracy value is then expected to increase.
- Merging a CNN model with a CNN that contains LSTM layers enabled using temporal features with frequency and spatial features extracted respectively by WPD and CSP.
- By merging the two CNN models, we succeeded to improve the accuracy rate of the classification.
- The comparative results show that the proposed fusion CNN-based method could offer the best performance in achieving the highest accuracy value compared to recent state of the art approaches.

## III. Methodology

In this section, we provide a description of the data set used in this research work. Then, we provide a detailed description of the proposed methodology to classify MI based on the merging of the CNNs models.

### A. Data set description

In this work, we used the BCI Competition IV 2a data set. This database has 22 recordings of EEG channels and 3 EOG channels of 9 healthy subjects recorded in two sessions where the sessions have 288 trials of four seconds each. EEG signals are sampled at 250 Hz and a bandwidth filtered between 0.5 and 100 Hz. The tasks consist of imagining the movements of the right hand, left hand, feet and tongue.

The timing of data acquisition is shown in Figure 1.

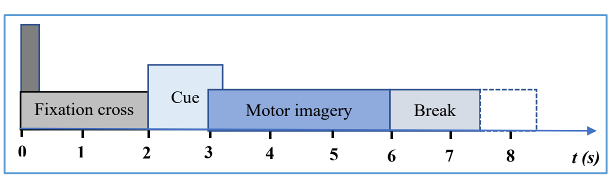


Fig. 1. The paradigm time sequence

### B. Proposed work

Our proposed methods start with a simple pre-processing involving the application of a bandpass filter from 7 to 30Hz, and the removal of the 3 EOG channels and we keep only 22 EEG channels. Then, to each EEG channel we apply the Wavelet Packet Decomposition (WPD) technique [22] followed by the Common Spatial Pattern (CSP) technique for the extraction of the frequency and spatial features respectively.

The extracted features are then considered as input for our two proposed methods based on the merging of CNNs. Figure 2 presents the general block diagram of our proposed methods.

#### 1) Wavelet Packet Decomposition

The WPD is an extension of Wavelet Decomposition (WD). It comprises several bases and different bases will result in different classification performance and cover the lack of fixed time-frequency decomposition in Discret Wavelet Transform (DWT) [23].

#### 2) Common Spatial Pattern

The CSP algorithm is a feature extraction method using spatial filters to maximize the discrimination of two classes. It has been widely used for feature extraction in EEG-based BCI systems for MI [24-25].

#### 3) Merged CNNs models

We proposed two methods based on the fusion of the two models CNN1 and CNN2. The first CNNs fusion method (figure 3) includes:

- A CNN1 which contains 5 blocks, each block contains a convolution layer, Batch Normalization and Max Pooling, then a Flatten layer.
- A CNN2 which contains 4 blocks, each block contains a convolution layer and Max Pooling, followed then by an LSTM and a Flatten layers. LSTMs have introduced the concept of memory cells (units) with control gates. This helps to maintain the gradients as they are backpropagated during training and at the same time to maintain long- or short-term-time dependencies between inputs.
- Concatenation of the two CNNs.
- Merging of the two CNNs by Multi-Layers Perceptron.

The second CNNs fusion method (figure 4) involves:

- A CNN1 that contains 5 blocks, each block contains a convolution layer, Batch Normalization and Max Pooling, followed by a Flatten layer.
- A CNN2 which contains 4 blocks, each block contains a convolution layer, Batch Normalization and Max Pooling, and a final Flatten layer.
- Concatenation of the two CNNs.
- Merging of the two CNNs by Mono-Layers Perceptron.

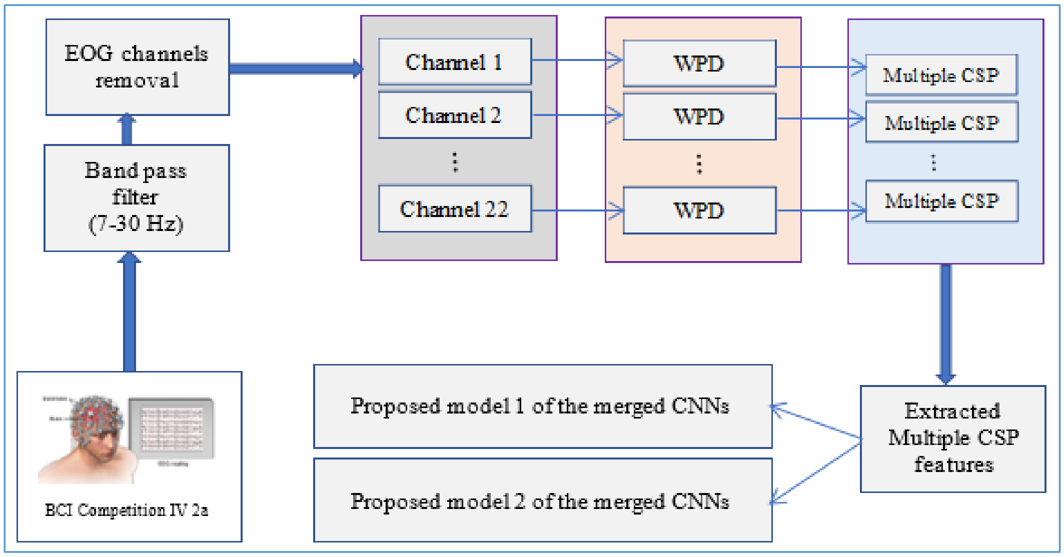


Fig. 2. Block diagram of our proposed methods.

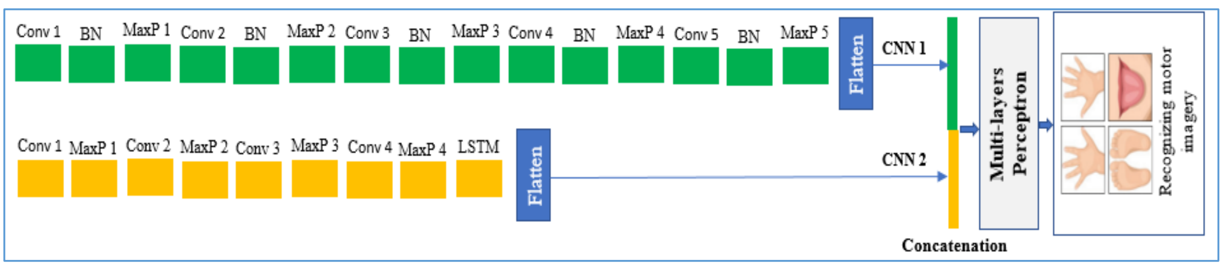


Fig. 3. Proposed model 1 of merged CNNs.

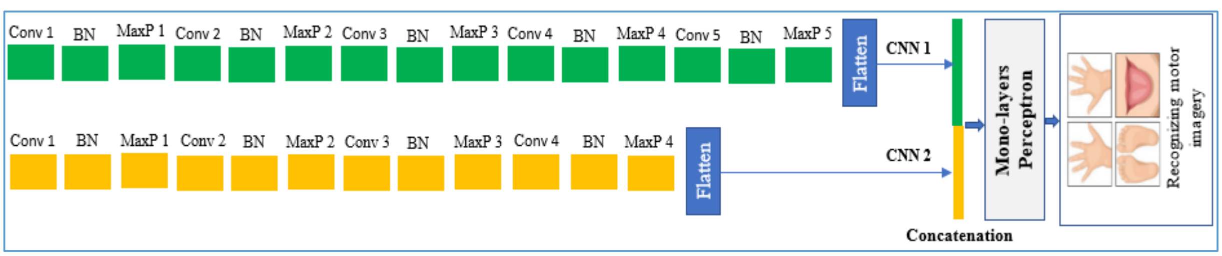


Fig. 4. Proposed model 2 of merged CNNs.

## IV. Results and Discussion

The performances of each proposed merged CNNs are evaluated by the accuracy, in percentage, that is defined as follows:

$$Accuracy = \left(\frac{TP+TN}{TP+TN+FP+FN}\right) \times 100 \quad (1)$$

where TP (True positive), TN (True negative), FP (False positive). and FN (False negative).

We fine-tuned each model by 500 epochs. The batch size is set to 32, the Adaptive Moment Estimation known as "Adam optimizer" is used to optimize the loss function. All the adopted models are trained by a cross-entropy loss function.

We divided the dataset, into two sets: one for training (80%) and one for validation (20%). The training set is used to train the model which is then tested by the testing set for the final evaluation.

We applied our proposed methods with different electrodes configuration: 3, 15, and 22 electrodes. And we tested them with the 4 activation functions Rectified Linear Units (ReLU), Exponential Linear Unit (ELU), Scaled Exponential Linear Unit (SELU) and Hyperbolic tangent (Tanh).

The tables from 1 to 3 give the accuracy values obtained by the application of two proposed methods based on the fusion of CNNs.

The electrodes used in this work:

- **3 electrodes:** C3, Cz, C4
- **15 electrodes:** FC3, FC1, FCz, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CPz, CP2, CP4
- **22 electrodes:** Fz, FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4, P1, Pz, P2, POz

We note that the first proposed method gave the highest accuracy value of 61.68%. This value is obtained with the ReLU activation function and the use of 22 electrodes.

The low accuracy value obtained in this work was 41.37%. We obtained this value after applying the first proposed method with 3 electrodes and the ReLU activation function.

The results we obtained in this work can be summarized as follows:

- The higher the number of electrodes, the higher the accuracy value, and vice versa.
- The addition of an LSTM layer has improved the classification results.
- The Multi-Layers Perceptron fusion method gave slightly higher accuracy values compared to the Mono-Layers Perceptron.

Returning to reference [26], we can say that our two proposed methods gave good accuracy values compared to the "Ensemble" method. In reference [26], the authors applied this machine learning method using the function of Matlab, namely, fitensemble. For the Ensemble method, the authors selected AdaBoostM2 as the classification algorithm, and the decision tree as the learner, with the ensemble learning cycle number set to 100.

The first method we proposed, based on merging the CNNs and adding an LSTM layer, gave the best accuracy value, which may be due to the simple pre-processing we applied to the dataset. In addition, the extraction of frequency features using WPD, spatial features using CSP, and temporal features improved the accuracy of the motor imagery classification.

Table 1. Accuracy values obtained by the proposed models of the merged CNNs with 3 electrodes.

| | Proposed model 1 of the merged CNNs | | | | Proposed model 2 of the merged CNNs | | | |
|---|---|---|---|---|---|---|---|---|
| | **ReLU** | **ELU** | **SELU** | **Tanh** | **ReLU** | **ELU** | **SELU** | **Tanh** |
| **Subject 1** | 43.10% | 60.34% | 55.17% | 50.00% | 53.45% | 53.45% | 53.45% | 53.45% |
| **Subject 2** | 55.17% | 41.38% | 36.21% | 43.10% | 43.10% | 46.55% | 48.28% | 51.72% |
| **Subject 3** | 53.45% | 68.97% | 67.24% | 62.07% | 65.52% | 60.34% | 60.34% | 58.62% |
| **Subject 4** | 34.48% | 37.93% | 37.93% | 32.76% | 32.76% | 32.76% | 34.48% | 36.21% |
| **Subject 5** | 43.10% | 37.93% | 44.83% | 44.83% | 24.14% | 39.66% | 29.31% | 37.93% |
| **Subject 6** | 31.03% | 36.21% | 34.48% | 34.48% | 31.03% | 24.14% | 31.03% | 32.76% |
| **Subject 7** | 39.66% | 48.28% | 41.38% | 48.28% | 51.72% | 44.83% | 53.45% | 51.72% |
| **Subject 8** | 41.38% | 39.66% | 48.28% | 50.00% | 44.83% | 44.83% | 51.72% | 39.66% |
| **Subject 9** | 31.03% | 39.66% | 41.38% | 44.83% | 36.21% | 34.48% | 43.10% | 48.28% |
| **Average** | **41.37%** | **45.59%** | **45.21%** | **45.59%** | **42.52%** | **42.33%** | **45.01%** | **45.59%** |

Table 2. Accuracy values obtained by the proposed models of the merged CNNs with 15 electrodes.

| | Proposed model 1 of the merged CNNs | | | | Proposed model 2 of the merged CNNs | | | |
|---|---|---|---|---|---|---|---|---|
| | **ReLU** | **ELU** | **SELU** | **Tanh** | **ReLU** | **ELU** | **SELU** | **Tanh** |
| **Subject 1** | 62.07% | 63.79% | 67.24% | 65.52% | 74.14% | 68.97% | 70.69% | 70.69% |
| **Subject 2** | 70.69% | 65.52% | 60.34% | 55.17% | 65.52% | 65.52% | 62.07% | 60.34% |
| **Subject 3** | 81.03% | 82.76% | 82.76% | 84.48% | 81.03% | 81.03% | 84.48% | 81.03% |
| **Subject 4** | 48.28% | 46.55% | 53.45% | 53.45% | 43.10% | 39.66% | 44.83% | 44.83% |
| **Subject 5** | 37.93% | 50.00% | 44.83% | 43.10% | 46.55% | 39.66% | 43.10% | 37.93% |
| **Subject 6** | 39.66% | 34.48% | 32.76% | 36.21% | 41.38% | 37.93% | 44.83% | 43.10% |
| **Subject 7** | 63.79% | 63.79% | 60.34% | 60.34% | 63.79% | 56.90% | 58.62% | 55.17% |
| **Subject 8** | 72.41% | 70.69% | 63.79% | 62.07% | 67.24% | 75.86% | 74.14% | 70.69% |
| **Subject 9** | 50.00% | 39.66% | 43.10% | 48.28% | 46.55% | 48.28% | 46.55% | 41.38% |
| **Average** | **58.42%** | **57.47%** | **56.51%** | **56.51%** | **58.81%** | **57.09%** | **58.81%** | **56.12%** |

Table 3. Accuracy values obtained by the proposed models of the merged CNNs with 22 electrodes.

| | **Proposed model 1 of the merged CNNs** | | | | **Proposed model 2 of the merged CNNs** | | | |
|---|---|---|---|---|---|---|---|---|
| | **ReLU** | **ELU** | **SELU** | **Tanh** | **ReLU** | **ELU** | **SELU** | **Tanh** |
| **Subject 1** | 67.24% | 72.41% | 74.14 | 74.14% | 77.59% | 74.14% | 70.69% | 72.41% |
| **Subject 2** | 67.24% | 70.69% | 67.24 | 63.79% | 70.69% | 65.52% | 63.79% | 65.52% |
| **Subject 3** | 79.31% | 82.76% | 84.48 | 77.59% | 79.31% | 82.76% | 82.76% | 84.48% |
| **Subject 4** | 55.17% | 51.72% | 41.38 | 48.28% | 43.10% | 51.72% | 46.55% | 43.10% |
| **Subject 5** | 46.55% | 37.93% | 46.55 | 48.28% | 41.38% | 48.28% | 43.10% | 43.10% |
| **Subject 6** | 39.66% | 31.03% | 29.31 | 37.93% | 41.38% | 31.03% | 37.93% | 41.38% |
| **Subject 7** | 81.03% | 65.52% | 63.79 | 60.34% | 62.07% | 65.52% | 72.41% | 65.52% |
| **Subject 8** | 70.69% | 63.79% | 77.59 | 72.41% | 79.31% | 70.69% | 70.69% | 67.24% |
| **Subject 9** | 48.28% | 44.83% | 44.83 | 56.90% | 44.83% | 55.17% | 39.66% | 51.72% |
| **Average** | **61.68%** | **57.85%** | **58.81** | **59.96%** | **59.96%** | **60.53%** | **58.62%** | **59.38%** |

## V. CONCLUSION

In this paper, we have proposed and compared two methods for classifying motor imagery tasks based on the merging of CNN models. With a simple pre-processing and an extraction of frequency, time, and spatial characteristics, our results show that the fusion of the CNNs with the LSTM layer offers a better classification performance compared to other state-of-the-art methods with an accuracy value of 61.68%.

In our future work, we are thinking about merging other DL methods. We want to refine the CNN or other models and the merging methods to further improve the accuracy of classification. We would like to discover such robust features that would allow experts in the field to use the proposed methods in advanced BCI systems.


## ACKNOWLEDGMENT

This work was carried out under the MOBIDOC scheme, funded by the European Union (EU) through the EMORI program and managed by the National Agency for the Promotion of Scientific Research (ANPR).


## REFERENCES


[1] I.A. Pap, S. Oniga, A Alexan, "Machine Learning EEG Data Analysis For eHealth IoT System", IEEE-AQTR, 2020. DOI: 10.1109/AQTR 49680.2020.9129966

[2] A. Akrout, A. Echtioui, R. Khemakhem, M. Ghorbel, "Artificial and Convolutional Neural Network of EEG-Based Motor imagery classification: A Comparative Study", 20th International conference on Sciences and Techniques of Automatic control & computer engineering (STA), pp. 46-50, 2020.

DOI: 10.1109/STA50679.2020.9329317

[3] A. Echtioui, W. Zouch, M. Ghorbel, C. Mhiri, H. Hamam, "Multi-class Motor Imagery EEG Classification Using Convolution Neural Network", In Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 1: SDMIS, ISBN 978-989-758-484-8, pp. 591-595, 2021. DOI: 10.5220/0010425905910595

[4] K. Yan, R. Sukthankar, "PCA-SIFT: a more distinctive representation for local image descriptors," IEEE Computer Society Conference on Computer Vision & Pattern Recognition, 2004.

[5] M. Grosse-Wentrup, M. Buss, "Multi-class Common Spatial Patterns and Information Theoretic Feature Extraction," IEEE transactions on bio-medical engineering, 55(8): pp. 1991-2000, 2008.

[6] E.I. Altman, G. Marco, F. Varetto, "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks," Journal of Banking & Finance, 18(3): pp. 505-529, 1994.

[7] C. Schuldt, I. Laptev, B. Caputo, " Recognizing human actions: a local SVM approach," International Conference on Pattern Recognition, 2004.

[8] A. Graves, A. Mohamed, G. Hinton, "Speech recognition with deep recurrent neural networks," ICASSP, 38(2003): pp. 6645-6649, 2013.

[9] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the Acm, vol. 60, no. 2, pp. 1106-1114, 2012.

[10] K.M. Hermann, T. Kočiský, E. Grefenstette, et al., "Teaching machines to read and comprehend," pp. 1693-1701, 2015.

[11] A. Karpathy, G. Toderici, S. Shetty, et al., "Large-Scale Video Classification with Convolutional Neural Networks," Computer Vision and Pattern Recognition. IEEE, pp. 1725-1732, 2014.

[12] P. Bashivan, I. Rish, M. Yeasin, N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," arXiv[Preprint]. arXiv:1511.06448, 2015.

[13] I.A. Corley, Y. Huang, "Deep EEG super-resolution: upsampling EEG spatial resolution with Generative Adversarial Networks," in 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI) (Las Vegas, NV), pp. 100–103, 2018.

[14] S.U. Amin, M. Alsulaiman, G. Muhammad, M.A. Mekhtiche, M.S. Hossain, Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion. Fut. Gener. Comput. Syst. 101, pp. 542–554, 2019.

[15] M.A. Jaoude, J. Jing, H. Sun, C.S. Jacobs, K.R. Pellerin, M.B.Westover, et al., "Detection of mesial temporal lobe epileptiform discharges on intracranial electrodes using deep learning", Clin. Neurophysiol. 131, 133–141, 2020.

[16] Y. Hou, L. Zhou, S. Jia, X. Lun, "A novel approach of decoding EEG four-class motor imagery tasks via scout ESI and CNN", J. Neural Eng, 2019.

[17] J. Zhang, R. Yao, W. Ge, J. Gao, "Orthogonal convolutional neural networks for automatic sleep stage classification based on single-channel EEG", Comput. Methods Prog. Biomed, 2020.

[18] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proc. IEEE 86, pp. 2278–2324, 1998.

[19] Z. Tang, C. Li, S. Sun, "Single-trial EEG classification of motor imagery using deep convolutional neural networks", Optik 130, pp. 11–18, 2017.

[20] H. Dose, J.S., Møller, H.K. Iversen, S. Puthusserypady, "An end-toend deep learning approach to MI-EEG signal classification for BCIs", Expert Syst. Appl. 114, pp. 532–542, 2018.

[21] R.T. Schirrmeister, J.T. Springenberg, L.D.J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, et al., "Deep learning with convolutional neural networks for EEG decoding and visualization", Hum. Brain Mapp. 38, pp. 5391–5420, 2017.

[22] J.Z. Xue, H. Zhang, C.X. Zheng, "Wavelet packet transform for feature extraction of EEG during mental tasks", Presented at Proceedings of

the Second International Conference on Machine Learning and Cybernetics, Xi'an, 2003.

[23] D. Giri, U.R. Acharya, R.J. Martis, S.V. Sree, T.C. Lim, T. Ahamed, J.S. Suri, Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. Knowl. Based Syst. 37, pp. 274–282, 2013.

[24] Y. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki , "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface", J. Neurosci. Methods 255 (1) pp. 85–91, 2015.

[25] J.S. Kirar, R. Agrawal, "Optimal Spatio-spectral variable size subbands filter for motor imagery brain computer Interface", Procedia Comput. Sci. 84, pp. 14–21, 2016.

[26] N. Thanh, H. Imali, K. Amin, G.B. Lee, P.L Chee, N. Saeid, "Classification of Multi-Class BCI Data by Common Spatial Pattern and Fuzzy System", IEEE access, 2018.