EEG Motor Imagery Classification using Fusion Convolutional Neural Network

Wassim Zouch[1][a] and Amira Echtioui[2][b]
[1]King Abdulaziz University (KAU), Jeddah, Saudi Arabia
[2]ATMS Lab, Advanced Technologies for Medicine and Signals, ENIS, Sfax University, Sfax, Tunisia

Keywords: Convolution Neural Network (CNN), Motor Imagery (MI) Classification, Electroencephalography (EEG).

Abstract: Brain-Computer Interfaces (BCIs) are systems that can help people with limited motor skills interact with their environment without the need for outside help. Therefore, the signal is representative of a motor area in the active brain system. It is used to recognize MI-EEG tasks via a deep learning techniques such as Convolutional Neural Network (CNN), which poses a potential problem in maintaining the integrity of frequency-time-space information and then the need for exploring the CNNs fusion. In this work, we propose a method based on the fusion of three CNN (3CNNs). Our proposed method achieves an interesting precision, recall, F1-score, and accuracy of 61.88%, 62.50%, 61.47%, 64.75% respectively when tested on the 9 subjects from the BCI Competition IV 2a dataset. The 3CNNs model achieved higher results compared to the state-of-the-art.

1 INTRODUCTION

Recently, EEG is widely used in research involving cognitive load (Qiao, et al., 2020), rehabilitation engineering (Sandheep et al. 2019) and disease detection (Usman et al., 2019) due to its relatively low financial cost (Lotte et al., 2018), its non-invasive nature, and its high temporal resolution.

MI-EEG (Pfurtscheller et al., 2001) is a popular field based on EEG, it allows to arouse a great interest on the part of researchers. MI-EEG databasets contain EEG recordings of imaginary body movements without any actual movement, to help people with disabilities control and control external devices (Royer et al., 2010).

Nowadays, researchers have started to study and apply various deep learning (DL) models for the analysis of the EEG signal (Muhammad et al., 2018).

DL models, especially CNN, have been successful for images

There is some research (Lee et al., 2017; Soleymani et al., 2018; Li et al., 2017; Zhang et al., 2017; Hariharan et al., 2015; Bhattacharjee et al., 2017; Ueki al., 2015) that has used intermediate characteristics of CNN layers to improve classification accuracy values.

CNN with a Stacked Automatic Encoder (SAE) has been proposed (Tabar et al., 2017). It provides better classification accuracy compared to traditional methods based on the BCI competition IV-2b dataset.

(Robinson et al., 2019) used a CNN model representation of multi-band and multi-channel EEG input to further improve classification accuracy.

(Zhao et al., 2019) proposed a new 3D representation of EEG signals, a multi-branch 3D CNN and the corresponding classification strategy. They got good performance.

In this research work, we proposed a new classification method based on the fusion of three CNNs to classify MI-EEGs.

The main research contributions to this work as follows:

- Pre-processing of the data: removal of three EOG channels and band pass filter;
- Feautres extraction by using Common Spatial Pattern (CSP) and Wavelet Packet Decomposition (WPD);

[a] https://orcid.org/0000-0003-1047-1968
[b] https://orcid.org/0000-0003-2041-1301

Zouch, W. and Echtioui, A.
EEG Motor Imagery Classification using Fusion Convolutional Neural Network.
DOI: 10.5220/0010975600003116
In *Proceedings of the 14th International Conference on Agents and Artificial Intelligence (ICAART 2022) - Volume 1*, pages 548-553
ISBN: 978-989-758-547-0; ISSN: 2184-433X

- The proposed method based on fusion of three CNNs allows for the classification of MI-EEG with an precision, recall, F1-score, and accuracy of 61.88%, 62.50%, 61.47%, 64.75% respectively
- The results show that the proposed method could give the best results compared to recent state of the art classification.

2 MATERIAL AND PROPOSED METHOD

2.1 Data Set Description

We used the BCI Competition IV 2a dataset (Leeb et al. 2008), featuring 22 scalp electrode positions. This dataset contains 9 subjects who are involved in the recordings that were made over two sessions. Each session contains 288 trials. The motive imagination task lasts 4 second. The imagined tasks are left/right hand, feet, and tongue.

2.2 Proposed Method

The proposed methodology (Figure 1) begins with the removal of the three EOG channels and the application of a band pass filter. Then, the application of the two techniques of features extraction WPD and CSP. Finally, the 3CNNs model proposed for the classification of MI tasks.

2.2.1 Pre-Processing

We applied a simple data pre-processing which consists in keeping only the 22 EEG channels and the application of a bandpass filter from 7 to 30 Hz.

2.2.2 Wavelet Packet Decomposition

WPD is extended from wavelet decomposition (WD). This technique includes multiple bases and different bases will result in different classification performance and cover the lack of fixed time-frequency decomposition in DWT (Xue et al., 2003).

2.2.3 Common Spatial Pattern

The CSP is efficient in constructing optimal spatial filters which discriminate 2 MI-EEG classes (Blankertz et al., 2008).

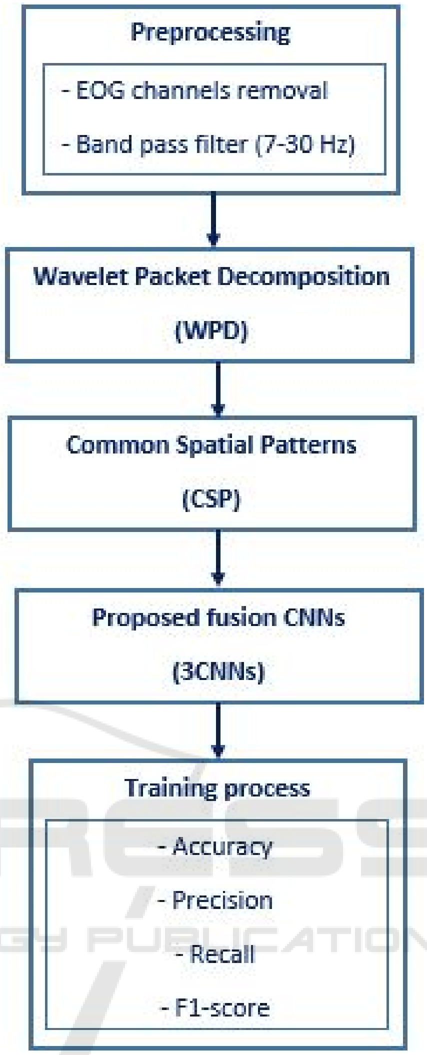

Figure 1: Flowchart of the proposed method.

2.2.4 Fusion of 3CNNS

Our fusion of CNNs contains three CNNs as shown in figure 2. Each CNN has 5 convolution blocks and Max Pooling, followed by a Flatten, then 4 dense layers. The concatenation of these 3 CNNs is followed by two dense layers. We have used the ReLu activation function in all convolutional layers and dense layers except in the last dense layer. The SoftMax activation function has been used for the last Dense layer.

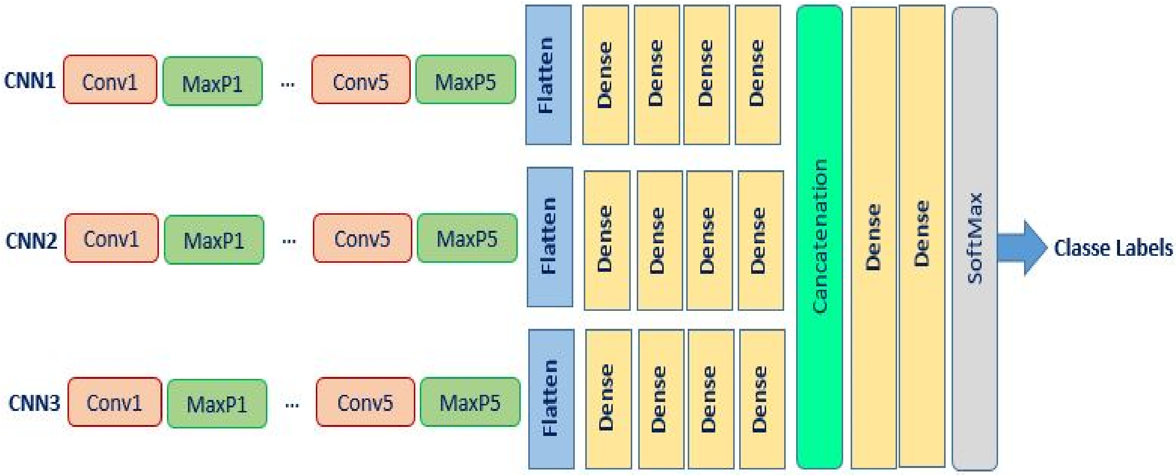

Figure 2: Flowchart of the proposed fusion of 3CNNs.

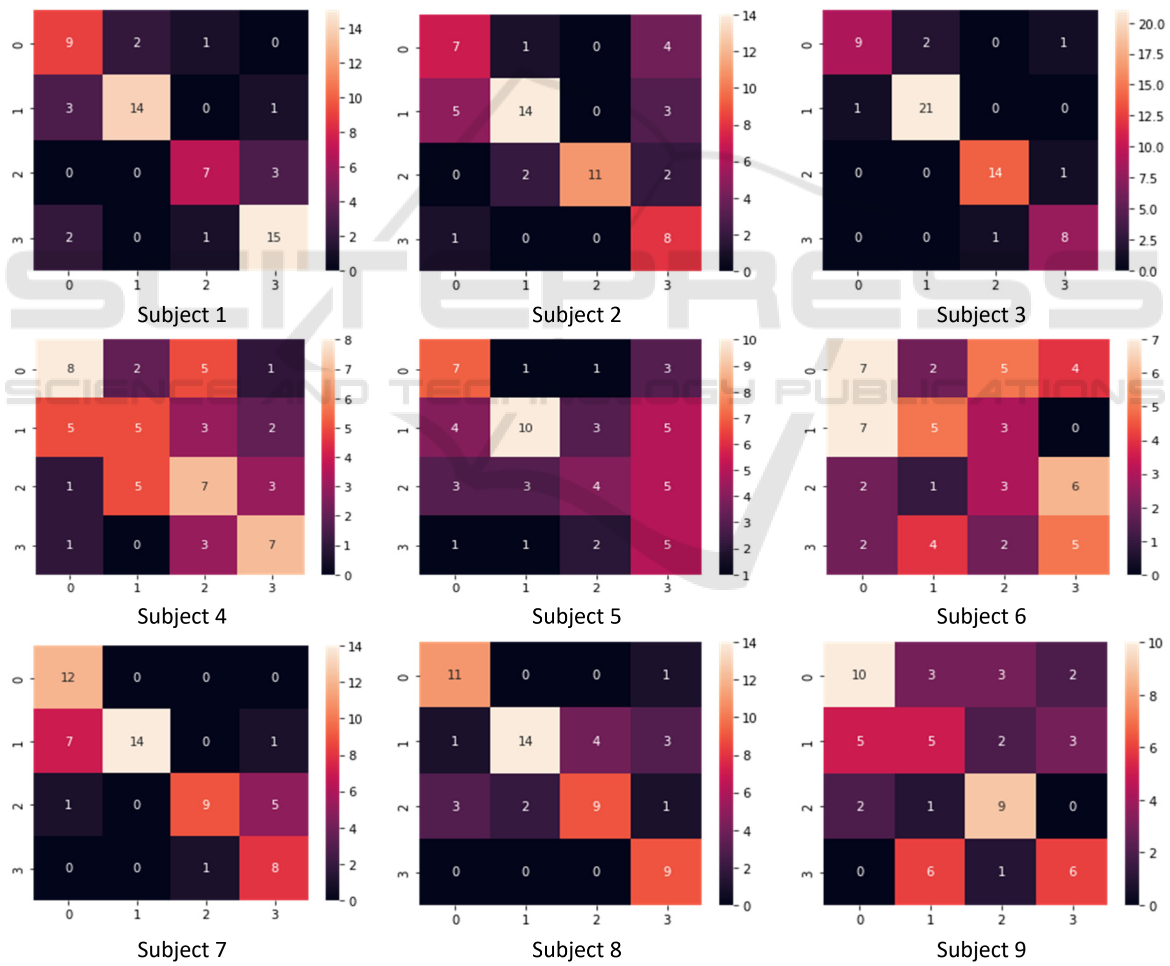

Figure 3: Confusion matrices of classification accuracy for the proposed methods.

3 RESULTS AND DISCUSSION

3.1 Metrics Evaluation

The four metrics used for the evaluation are:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$precision = \frac{TP}{Tp + FP} \quad (2)$$

$$recall = \frac{TP}{FN + TP} \quad (3)$$

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (4)$$

Where: TP: True Positive; TN: True Negative; FP: False Positive and FN: False Negative.

3.2 Results and Discussion

We provide in Figure 3, the confusion matrix for the proposed method based on the fusion of 3 CNNS. Diagonal elements indicate that the number of points for which the predicted label is equal to the true label. Moreover, we can also notice that the non-diagonal elements are those which are badly labeled by the classifier.

The performance measures obtained for each subject are shown in Tables 1 to 9.

Table 1: Classification Report for the proposed method for subject 1 (%).

	Precision	Recall	f1-score
Left hand	64	75	69
Right hand	88	78	82
Feet	78	70	74
Tongue	79	83	81
Average	**77.25**	**76.5**	**76.5**

Table 2: Classification Report for the proposed method for subject 2 (%).

	Precision	Recall	f1-score
Left hand	54	58	56
Right hand	82	64	72
Feet	100	73	85
Tongue	47	89	62
Average	**70.75**	**71**	**68.75**

Table 3: Classification Report for the proposed method for subject 3 (%).

	Precision	Recall	f1-score
Left hand	90	75	82
Right hand	91	95	93
Feet	93	93	93
Tongue	80	89	84
Average	**88.5**	**88**	**88**

Table 4: Classification Report for the proposed method for subject 4 (%).

	Precision	Recall	f1-score
Left hand	53	50	52
Right hand	42	33	37
Feet	39	44	41
Tongue	54	64	58
Average	**47**	**47.75**	**47**

Table 5: Classification Report for the proposed method for subject 5 (%).

	Precision	Recall	f1-score
Left hand	47	58	52
Right hand	67	45	54
Feet	40	27	32
Tongue	28	56	37
Average	**45.5**	**46.5**	**43.75**

Table 6: Classification Report for the proposed method for subject 6 (%).

	Precision	Recall	f1-score
Left hand	39	39	39
Right hand	42	33	37
Feet	23	25	24
Tongue	33	38	36
Average	**34.25**	**33.75**	**34**

Table 7: Classification Report for the proposed method for subject 7 (%).

	Precision	Recall	f1-score
Left hand	60	100	75
Right hand	100	64	78
Feet	90	60	72
Tongue	57	89	70
Average	**76.75**	**78.25**	**73.75**

Table 8: Classification Report for the proposed method for subject 8 (%).

	Precision	Recall	f1-score
Left hand	73	92	81
Right hand	88	64	74
Feet	69	60	64
Tongue	64	100	78
Average	**73.5**	**79**	**74.25**

Table 9: Classification Report for the proposed method for subject 9 (%).

	Precision	Recall	f1-score
Left hand	59	56	57
Right hand	33	33	33
Feet	60	75	67
Tongue	55	46	50
Average	**51.75**	**52.5**	**51.75**

From tables 1 to 9, we can notice that subject 3 gives the best values of precision, recall and F1-score. The latter reached 88.5%, 88%, and 88% of precision, recall and F1-score respectively.

Precision, recall, and F1-score values for subjects 1, 2, 7, and 8 vary between 68.75% and 78.25%.

Subjects 4, 5, and 6 have the precision, recall, and F1-score values too low compared to the values obtained by subjects 1, 2, 3, 7, 8, and 9.

According to Table 10, our proposed method based on the fusion of 3CNNs gives a value of precision, Recall, F1-Score and accuracy of 62.80%, 63.69%, 61.97%, 62.45% respectively.

Table 10: Classification Report for the proposed method (%).

	Precision	Recall	F1-score	Accuracy
Subject 1	77.25	76.50	76.50	77.59
Subject 2	70.75	71.00	68.75	68.97
Subject 3	88.50	88.00	88.00	89.66
Subject 4	47.00	47.75	47.00	46.55
Subject 5	45.50	46.50	43.75	44.83
Subject 6	34.25	33.75	34.00	34.48
Subject 7	76.75	78.25	73.75	74.14
Subject 8	73.50	79.00	74.25	74.14
Subject 9	51.75	52.50	51.75	51.72
Average	**62.80**	**63.69**	**61.97**	**62.45**

Table 11 presents a comparison between the proposed method and some state-of-the-art methods, in terms of classification accuracy. The methods proposed by (Nguyen et al., 2017) are evaluated based on the BCI Competition VI 2a dataset.

The proposed CNN offered a good improvement in accuracy value compared to the methods presented in table 11.

For the Ensemble method (Nguyen et al., 2017), the authors proposed "Adaptive Boosting for Multiclass Classification 'AdaBoostM2' as a classification approach, the decision tree as a learner. The number of epochs for the Ensemble method is fixed at 100. This model is able to identify MI tasks with an accuracy value of 58.22%.

Alternatively, the Euclidean distance metric is used in the implementation of the K-Nearest Neighbor (KNN) classifier (Nguyen et al., 2017). This algorithm can give a good classification if the number of characteristics is large enough. But the accuracy of KNN can be severely degraded by the presence of noisy or irrelevant characteristics, which influences the accuracy value (58.88%).

Table 11: Classification accuracy.

	Accuracy
Proposed method	62.45%
Ensemble (Nguyen et al., 2017)	58.22%
KNN [(Nguyen et al., 2017)	58.80%

We notice that our proposed method based on the fusion of 3CNNs gives the best accuracy values are equal to 62.45%.

These results prove that the proposed method based on CNNs fusion leads to better performance by exhibiting the highest accuracy value compared to the state of the art.

4 CONCLUSIONS

In this work, we have proposed a new method of classification of MI tasks based on the merger of the three CNNs. The results obtained by merging three CNN models prove that these models can extract different types of features representing EEG data at different abstract levels. In future Work we are planning to test the proposed technique for real time EEG classification.

REFERENCES

Bhattacharjee, P., and Das, S. (2017). Two-stream convolutional network with multi-level feature fusion for categorization of human action from videos. *In Pattern Recognition and Machine Intelligence (Lecture Notes in Computer Science)*, vol. 10597. Cham, Switzerland: Springer, 2017.

Blankertz, B., Dornhege, G., Krauledat, M., Müller, K.-R., and Curio, G. (2007). The non-invasive Berlin brain-computer interface: fast acquisition of effective performance in untrained subjects. *Neuroimage* 37, 539–550.

Hariharan, B., Arbeláez, P., Girshick, R., and Malik, J. (2015). Hypercolumns for object segmentation and fine-grained localization. *In Proc. IEEE Int. Conf. Comput. Vis*., Jun. 2015, pp. 447–456.

Lee, J., and Nam, J. (2017). Multi-level and multi-scale feature aggregation using pretrained convolutional neural networks for music auto-tagging, *IEEE Signal Process. Lett*., vol. 24, no. 8, pp. 1208–1212, Aug. 2017, doi:10.1109/LSP.2017.2713830.

Li, E., Xia, J., Du, P., Lin, C., and Samat, A. (2017). Integrating multilayer features of convolutional neural networks for remote sensing scene classification, *IEEE Trans. Geosci. Remote Sens*., vol. 55, no. 10, pp. 5653–5665, Oct. 2017.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F. (2018). A review of classification algorithms for EEG-based brain computer interfaces: a 10 year update[J]. *J Neural Eng* 15(3):031005

Muhammad, G., Masud, M., Amin, S. U., Alrobaea, R., Alhamid, M. F. (2018). Automatic seizure detection in a mobile multimedia framework, *IEEE Access*, vol. 6, pp. 45372–45383

Nguyen, T., Hettiarachchi, I., Khosravi, A., Salaken, S.M., Bhatti, As., Nahavandi, S. (2017). Multiclass_EEG_Data_Classification_using_Fuzzy_Systems, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, Naples, Italy, pp. 1-6.

Pfurtscheller, G., Neuper, C. (2001). Motor imagery and direct brain computer communication[J]. *Proc IEEE* 89(7):1123–1134

Qiao, W., Bi, X. (2020). Ternary-task convolutional bidirectional neural turing machine for assessment of EEG-based cognitive workload[J]. *Biomedl Signal Process Control* 57:101745

R. Leeb, C., Brunner, G., Müller-Putz, G., Schlögl, A., and Pfurtscheller, G. (2008). BCI competition 2008-Graz data set A and B. *Inst. Knowl. Discovery*, Lab. Brain-Comput. Interfaces, Graz Univ. Technol., Graz, Austria, Tech. Rep. 1–6, 2008, pp. 136–142.

Robinson, N., Lee, S. W., & Guan, C. (2019). EEG Representation in Deep Convolutional Neural Networks for Classifcation of Motor Imagery. *In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE*. https://doi.org/10.1109/SMC.2019.8914184

Royer, A. S., Doud, A. J., Rose, M. L., He, B. (2010). EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies[J]. *IEEE Trans Neural Syst Rehabil Eng* 18(6):581–589

Sandheep. P., Vineeth, S., Poulose, M., Subha, D. P. (2019). Performance analysis of deep learning CNN in classification of depression EEG signals, *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India*, 2019: 1339–1344.

Soleymani, S., Dabouei, A., Kazemi, H., Dawson, J., and Nasrabadi, N. M. (2018). Multi-level feature abstraction from convolutional neural networks for multimodal biometric identification. [Online]. Available: https://arxiv.org/abs/1807.01332

Tabar, Y. R. & Halici, U. (2017). A novel deep learning approach for classifcation of eeg motor imagery signals. *J. Neural Eng.* 14(1), 016003. https://doi.org/10.1088/1741-2560/14/1/016003

Ueki, K., and Kobayashi, T. (2015). Multi-layer feature extractions for image classification-Knowledge from deep CNNs. *In Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, London, U.K., Sep. 2015, pp. 9–12.

Usman, S. M., Khalid, S., Akhtar, R., Bortolotto, Z., Bashir, Z., Qiu, H. (2019). Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: review of available methodologies[J]. *Seizure* 71:258–269

Xue, J. Z., Zhang, H., Zheng, C. X. (2003). Wavelet packet transform for feature extraction of EEG during mental tasks, *In Proceedings of the Second International Conference on Machine Learning and Cybernetics, Xi'an*.

Zhang, P., Wang, D., Lu, H., Wang, H., and Ruan, X. (2017). Amulet: Aggregating multi-level convolutional features for salient object detection. *In Proc. IEEE Int. Conf. Comput. Vis*., Jun. 2017, pp. 202–211.

Zhao, X. et al. (2019). A multi-branch 3D convolutional neural network for EEG-based motor imagery classifcation*. IEEE Trans. Neural Syst. Rehabil. Eng*. 27(10), 2164–2177. https://doi.org/10.1109/TNSRE.2019.2938295