applied sciences

MDPI

Article

A Novel Convolutional Neural Network Classification Approach of Motor-Imagery EEG Recording Based on Deep Learning

Amira Echtioui [1,*], Ayoub Mlaouah [2,3,4], Wassim Zouch [5], Mohamed Ghorbel [1], Chokri Mhiri [6,7] and Habib Hamam [2,8,9]

1 ATMS Lab, Advanced Technologies for Medicine and Signals, ENIS, Sfax University, Sfax 3038, Tunisia; mohamed.ghorbel@enetcom.usf.tn
2 Faculty of Engineering, Université De Moncton, Moncton, NB E1A3E9, Canada; ayoub.mlaouah@esprit.tn (A.M.); habib.hamam@umoncton.ca (H.H.)
3 Higher Institute of Computer Sciences and Mathematics of Monastir (ISIMM), Monastir 5000, Tunisia
4 Private Higher School of Engineering and Technology (ESPRIT), El Ghazela, Ariana 2083, Tunisia
5 Electrical and Computer Engineering Department, Faculty of Engineering, King Abdulaziz University (KAU), Jeddah 21589, Saudi Arabia; wzouch@kau.edu.sa
6 Department of Neurology, Habib Bourguiba University Hospital, Sfax 3029, Tunisia; mhiri.chokri@gmail.com
7 Neuroscience Laboratory "LR-12-SP-19", Faculty of Medicine, Sfax University, Sfax 3029, Tunisia
8 Spectrum of Knowledge Production and Skills Development, Sfax 3027, Tunisia
9 School of Electrical Engineering and Electronic Engineering, University of Johannesburg, Johannesburg 2006, South Africa
* Correspondence: amira.echtioui@stud.enis.tn or amira.echtioui@isims.usf.tn

Citation: Echtioui, A.; Mlaouah, A.; Zouch, W.; Ghorbel, M.; Mhiri, C.; Hamam, H. A Novel Convolutional Neural Network Classification Approach of Motor-Imagery EEG Recording Based on Deep Learning. *Appl. Sci.* **2021**, *11*, 9948. https://doi.org/10.3390/app11219948

Academic Editor: Vasudevan (Vengu) Lakshminarayanan

Received: 22 August 2021
Accepted: 18 October 2021
Published: 25 October 2021

Abstract: Recently, Electroencephalography (EEG) motor imagery (MI) signals have received increasing attention because it became possible to use these signals to encode a person's intention to perform an action. Researchers have used MI signals to help people with partial or total paralysis, control devices such as exoskeletons, wheelchairs, prostheses, and even independent driving. Therefore, classifying the motor imagery tasks of these signals is important for a Brain-Computer Interface (BCI) system. Classifying the MI tasks from EEG signals is difficult to offer a good decoder due to the dynamic nature of the signal, its low signal-to-noise ratio, complexity, and dependence on the sensor positions. In this paper, we investigate five multilayer methods for classifying MI tasks: proposed methods based on Artificial Neural Network, Convolutional Neural Network 1 (CNN1), CNN2, CNN1 with CNN2 merged, and the modified CNN1 with CNN2 merged. These proposed methods use different spatial and temporal characteristics extracted from raw EEG data. We demonstrate that our proposed CNN1-based method outperforms state-of-the-art machine/deep learning techniques for EEG classification by an accuracy value of 68.77% and use spatial and frequency characteristics on the BCI Competition IV-2a dataset, which includes nine subjects performing four MI tasks (left/right hand, feet, and tongue). The experimental results demonstrate the feasibility of this proposed method for the classification of MI-EEG signals and can be applied successfully to BCI systems where the amount of data is large due to daily recording.

Keywords: EEG; BCI; motor imagery; Common Spatial Pattern (CSP); Wavelet Packet Decomposition (WPD); deep learning; CNN; Long Short-Term Memory (LSTM); merged CNNs

1. Introduction

In recent years, the use of brain signals from EEG electroencephalography has been widely explored for various applications with a major focus on the field of biomedical engineering. A Brain-Computer Interface (BCI) system, also referred to as brain-machine interaction, bridges the gap between humans and computers by translating thoughts into commands, which can be used to communicate with external devices like exoskeletons, wheelchairs, prostheses or used for rehabilitation of sensory, motor, and cognitive disabilities [1]. Thought-based control of machines is a hot topic that is increasingly integrated into various applications.

BCIs were initially developed for people with disabilities, and in particular for those suffering from what is known as "locked-in syndrome" [2] which is a neurological pathology. The patient with this disease is usually quadriplegic and cannot move or speak. On the other hand, his consciousness and his cognitive and intellectual faculties are intact. People suffering from complete muscular paralysis can no longer autonomously communicate with the outside world, despite their full mental capacity. The development of BCIs would allow these people to regain their autonomy of motion. Non-invasive electroencephalography (EEG) of the scalp is an easy and inexpensive technique for recording brain activity.

Decoding the user's intent from EEG signals is one of the most difficult challenges in BCI. One of the main reasons is that EEG signals have non-stationary and complex properties [3]. For the MI task, it is particularly difficult to obtain high-quality data because it is not known what exactly the user has imagined. For this reason, recent advances in MI-based BCI approaches [4–6] enabled improving the decoding accuracy by using many feature extraction or classification methods based on advanced machine learning algorithms and deep learning (DL).

Recently DL algorithms have been intensively developed to make a profit from the progress of artificial intelligence technologies. Successful results have been obtained in the classification of MI tasks.

CNN is a class of artificial neural networks that presents one of the most commonly used DL architectures for image recognition tasks. Since CNN is specially designed to process input images, its architecture is made up of two main blocks.

The first block presents the particularity of this type of neural network since it functions as a feature extractor.

The second block is not an exclusive feature of CNN: it is, in fact, used at the end of all the neural networks used for classification.

CNN model has low network complexity and strong feature extraction capabilities, which can well solve the problem of difficult feature extraction from EEG signals. Therefore, it is feasible to classify the MI tasks of EEG signals based on the CNN model.

Our aims are oriented towards the high precision requirements of the classification of MI tasks of the EEG signal. Based on the fusion of CNNs architectures, we want to achieve an efficient classification of EEG signals from the Competition IV 2a dataset.

The CNN model with the Wavelet Packet Decomposition (WPD) and Common Spatial Pattern (CSP) algorithms were used for feature extraction and classification of the MI-EEG signals.

The elements of originality may be summarized as follows:

- The performance of classifiers is improved by pre-processing of the EEG signals. We proceed with removal of EOG channels and with applying a bandpass filter.
- We extracted frequential and spatial features by using WPD and CSP techniques respectively.
- We showed that the proposed method based on CNN1 model gives the highest value of accuracy compared to the state-of-the-art.

The rest of this paper is organized as follows. In Section 2, we briefly review related work. Section 3 presents our five proposed methods for the classification of MI tasks in EEG signals. In Section 4, we analyze the experimental results that verify the effectiveness of the proposed methods. Section 5 draws a conclusion and provides direction for future research.

2. Related Work

Recently, the DL and traditional algorithms were combined with other methods to extract meaningful information from EEG signals. For example, in [7], the authors proposed a method based on the combination of CNN with the gradient boosting (GB) algorithm.

These algorithms are widely used in the fields of EEG signals, but their performance and accuracy in processing EEG signals are not satisfactory. Many researchers started investigating the potential for using various DL models for EEG signals analysis [8]. DL models, in particular CNN, can extract more robust and discriminating features [9,10].

Other models, such as Long Short-Term Memory (LSTM) [11], the Recurrent Neural Network (RNN) [12], Deep Belief Network (DBN), and the Sparse AutoEncoder SAE [13], are useful in time series applications. Research shows that the DL technology has performed well in the field of EEG signal processing [14], which indicates that the automatically extracted features are better than the ones that are manually extracted.

In [15], the authors propose a method that combines the multi-branch 3D CNN and the 3D representation of the EEG. The strength of the application of 3D CNN is that the temporal and spatial characteristics of EEG signals can be extracted simultaneously, and the relationship between them can be fully utilized. In [16], the authors explore CNN in combination with Filter Bank Common Spatial Pattern (FBCSP), a spatial filtering algorithm, and their own approach to extract temporal characteristics. In [17], four different CNN models with increasing depths are used to learn the spatial and temporal characteristics, which are then merged and sent to either an autoencoder or to a multilayer perceptron for classification.

The features of the mu (8–13 Hz) and beta (13–30 Hz) bands of EEG have been explored by some researchers using Stacked AutoEncoder (SAE) and CNN for MI-EEG classification [18,19]. In [20], the authors convert the EEG signal into images by using the Short-Term Fourier Transform (STFT).

Several researchers propose DL models [21–27] to extract intermediate features and successfully merge models with different architectures. In [28], the authors propose a deep belief network method using a Restricted Boltzmann Machine (RBM) for the classification of MI.

The training phase of in-DL models is difficult on small data sets because they can have millions of parameters, which usually requires huge training data. There are not many public EEG data sets available, and those that are available are limited in size. This limits the scope of applying deep networks in this area. However, techniques, such as transfer learning, are of nature to present new avenues of using deep networks that are first pre-trained on large datasets and then fine-tuned for smaller datasets. These techniques show better performance and reduce the training time of the deep models [29].

Numerous variants of the CNN models were used for the image classification with good accuracy. One of them is the merging of several CNNs for feature aggregation.

Many researchers merge multiple CNN models and features [21–29] to extract intermediate features and merge models with architectures and have had some success. Although the researchers applied the CNN model and other DL models on the MI EEG data to obtain good accuracy values, they could not achieve major improvements over technical machine learning [13,30]. This is due to the fact that the EEG signal is characterized by a low SNR, low spatial characteristics, and is difficult to interpret due to its non-stationarity.

In this work, we propose five methods for the classification of MI-EEG signals. These methods are based on Artificial Neural Network (ANN), Convolutional Neural Network 1 (CNN1), CNN2, CNN1 with CNN2 merged, and the modified CNN1 with CNN2 merged. These methods start with a pre-processing step where the EOG channels have been removed and a bandpass filter (7–30 Hz) has been applied to the EEG data. Then we extract frequency characteristics using Wavelet Packet Decomposition (WPD) and spatial characteristics using Common Spatial Pattern (CSP). The obtained characteristics are transmitted to the proposed classifiers: ANN, CNN1, CNN2, merged CNNs, and modified merged CNNs. Each CNN has a different depth. Our CNN-based method reports improved performance for EEG MI data using its frequency and spatial characteristics. We also show that merging the two CNN models and adding LSTM layers in the CNN model does not always give better results.

3. Material and Methods

3.1. Data Set Description

The well-known database "Competition IV 2a" [31] is employed to train then to test our method. It enables the comparison of our results with those from state-of-the-art methods.

The dataset includes recordings of EEG signals for 9 subjects when they were comfortably sitting and performing MI tasks. Data are taken from 22 EEG (named Fz, FC3, FC1, FCz, FC2, FC4, C5, C3, C1, Cz, C2, C4, C6, CP3, CP1, CPz, CP2, CP4, P1, Pz, P2, POz) and 3 EOG channels. The electrodes are arranged according to the standard 10 to 20 system across the scalp of the person.

The subjects accomplish four MI tasks. These tasks represent the Left hand, Tongue, Feet, and Right Hand. Also, data are split up into short runs. It is worth noting that each run carries 48 trials of each activity of the MI.

The collection of the data consists of two sessions in two days, where the session comprises six runs with a short break between them. Thus, for each MI activity, a total of 288 trials is collected.

Each session is started by recording approximately 5 min of EEG data to estimate the influence of the 3 EOG channels. This recording is divided into 3 blocks: (1) 2 min with eyes open (looking at a fixing cross on the screen), (2) 1 min with eyes closed, and (3) 1 min with eye movements (Figure 1). Due to technical issues, the EOG block is shorter for Subject 4 and only contains the eye movement condition. The timing of data acquisition is shown in Figure 2. At the start of a test (t = 0 s), a fixing cross appears on the black screen. In addition, a short acoustic warning tone is emitted. After 2 s (t = 2 s), a cue in the form of an arrow pointing left, right, down, or up (corresponding to one of the 4 classes) appears and remains on screen for 1.25 s. This prompted subjects to perform the desired MI task until the fixation cross disappeared from the screen at t = 6 s. Finally, a short break with a black screen is used.

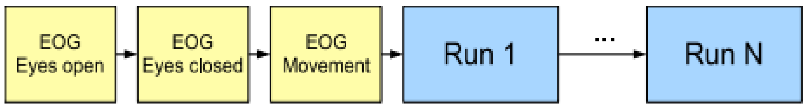

Figure 1. Timing scheme of one session.

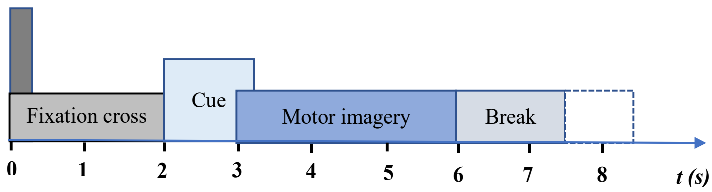

Figure 2. The paradigm time sequence.

Furthermore, the data were sampled at a frequency of 250 Hz and were filtered by a band pass-filter between 0.5 Hz to 100 Hz. In order to remove line noise, another 50 Hz notch filter was utilized.

3.2. Proposed Work

In this section, we will detail our proposed method for the classification of MI-EEG signals. Our proposed method (Figure 3) begins with the application of a 7 to 30 Hz bandpass filter on the BCI Competition IV 2a dataset. Then, we eliminate the three EOG channels and keep only the 22 EEG channels. Subsequently, we apply to each EEG channel, the WPD method by the extraction of frequency characteristics followed by the CSP algorithm for the extraction of spatial characteristics.

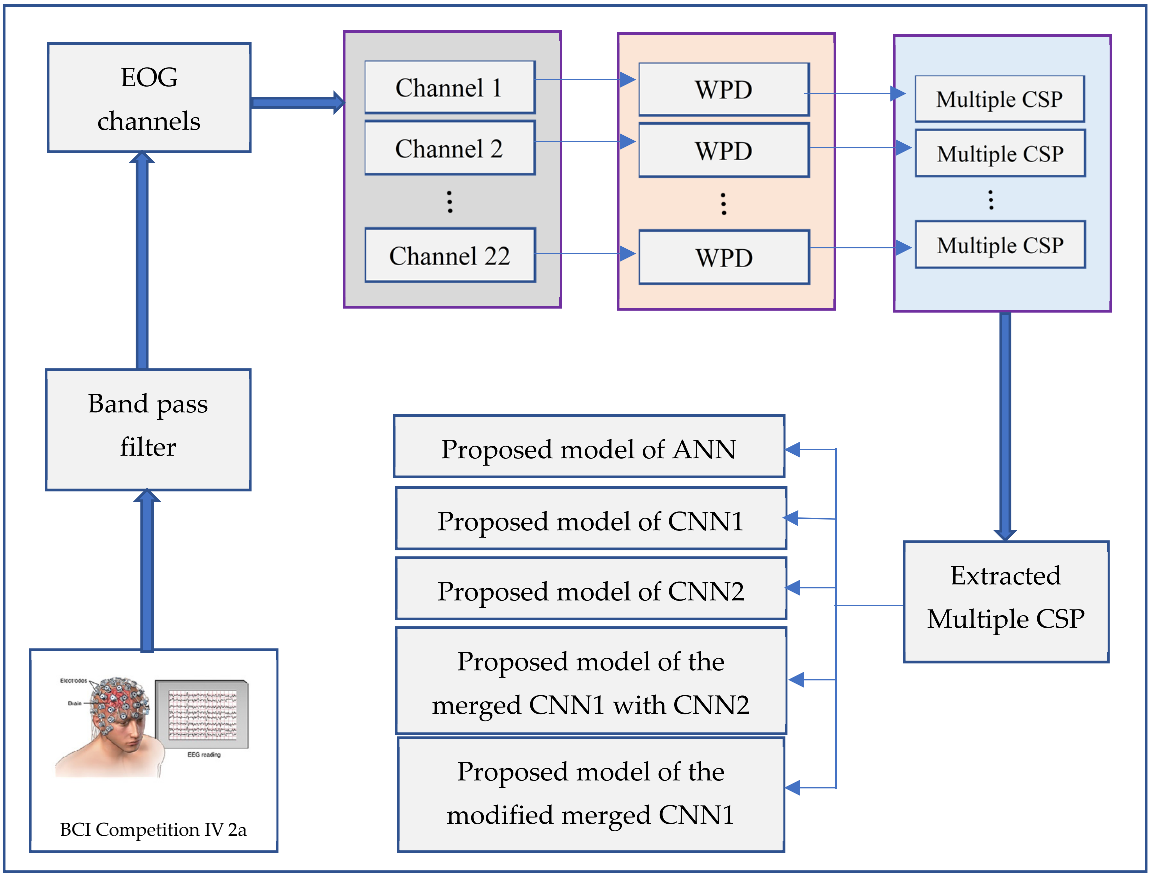

Figure 3. Block diagram of the proposed methodology.

The extracted features from CSP are considered as input for our five offered models: ANN, CNN1, CNN2, Merged CNNs, and Modified merged CNNs.

3.2.1. Wavelet Packet Decomposition

Coifman et al. introduced the wavelet packet and proposed the concept of orthogonal wavelet packet based on orthogonal wavelets. The Wavelet Packet Transformation allows decomposing of the low-frequency part of the signal and the high-frequency part of the signal in a more detailed way. Moreover, this decomposition has neither redundancy nor omission, in order to allow a better time-frequency localization analysis capability than the wavelet transform for vibration signals containing medium and high-frequency information. Towards improving the accuracy of EEG recognition. It is an idea that can be understood as the decomposition of space.

The bandwidth of the EEG signal was chosen from the frequency band of 8 to 31 Hz. Figure 4 shows the frequency decomposition at each step that is obtained from filter banks.

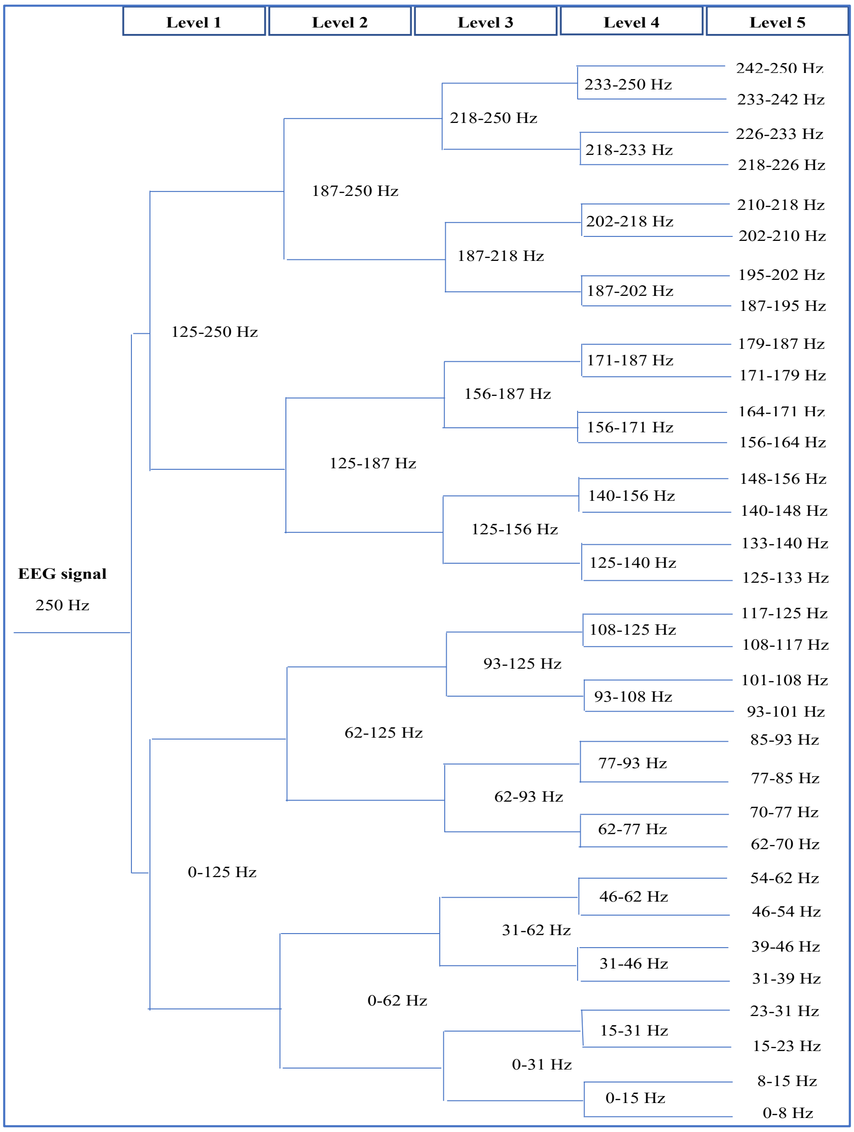

Figure 4. Five level EEG signal decomposition using WPD.

3.2.2. Common Spatial Pattern

CSP algorithm is a feature extraction method. It uses the theory of the diagonalization of the covariance matrix in a two-class signal. The main idea of CSP is to find the optimal projection matrix, which maximizes the variance for one class while minimizing the variance for the other class. Consequently, this algorithm achieves a maximum, unlike the two classes, for example, the right hand and left hand.

3.2.3. Proposed ANN Model

Our ANN model is composed of a first layer, a last layer, and between them, several intermediate layers. The first layer contains three elements: Dense layer, activation function, and Dropout layer. Instead of Sigmoid, the activation may be through a Rectified Linear Unit (ReLU), Tanh, an Exponential linear unit (ELU), or Scaled Exponential Linear Unit (SELU). The intermediate layers are composed of seven blocks. Each block starts with a Batch Normalization layer followed by the dense layer, activation function, and dropout layer. The last layer is dense with the SoftMax activation function.

Batch normalization is a technique for improving the speed, performance, and stability of artificial neural networks.

The most common activation function for ANNs is the ReLu: It is less computationally expensive than some other common activation functions like Tanh and Sigmoid because the mathematical operation is simpler and the activation is sparser. Since the function outputs 0 when $x \leq 0$, there is a considerable chance that a given unit does not activate at all.

$(x) = \max\,(0, x)$. Its outputs x when x is positive and outputs 0 otherwise.

We tested the four activation functions (Figure 5): ReLu, ELU, SELU, and Tanh.

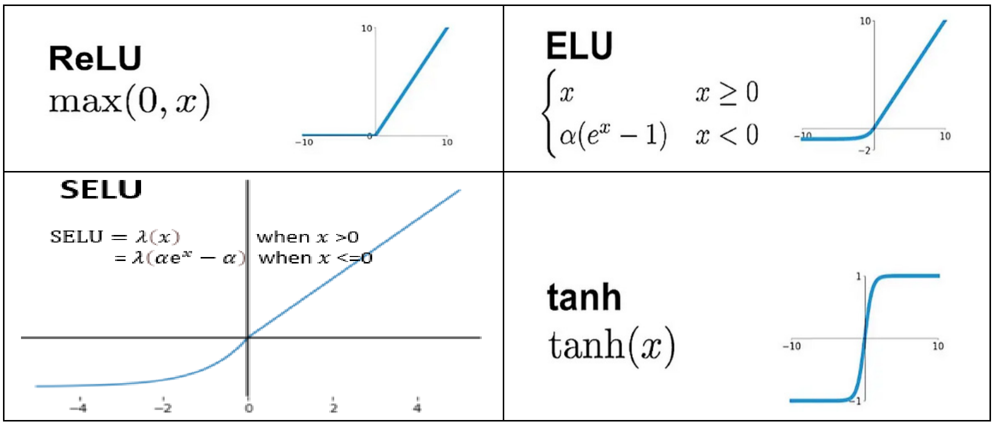

Figure 5. Mathematical Expressions of ReLU, ELU, SELU, and Tanh activation functions.

The proposed ANN model consists of 117,060 parameters. We have used the ADAM optimizer for weight updates and categorical cross-entropy loss. Then, we trained the network by using a batch size of 16 and 1000 training epochs. Figure 6 illustrates the proposed ANN architecture used to classify MI tasks.

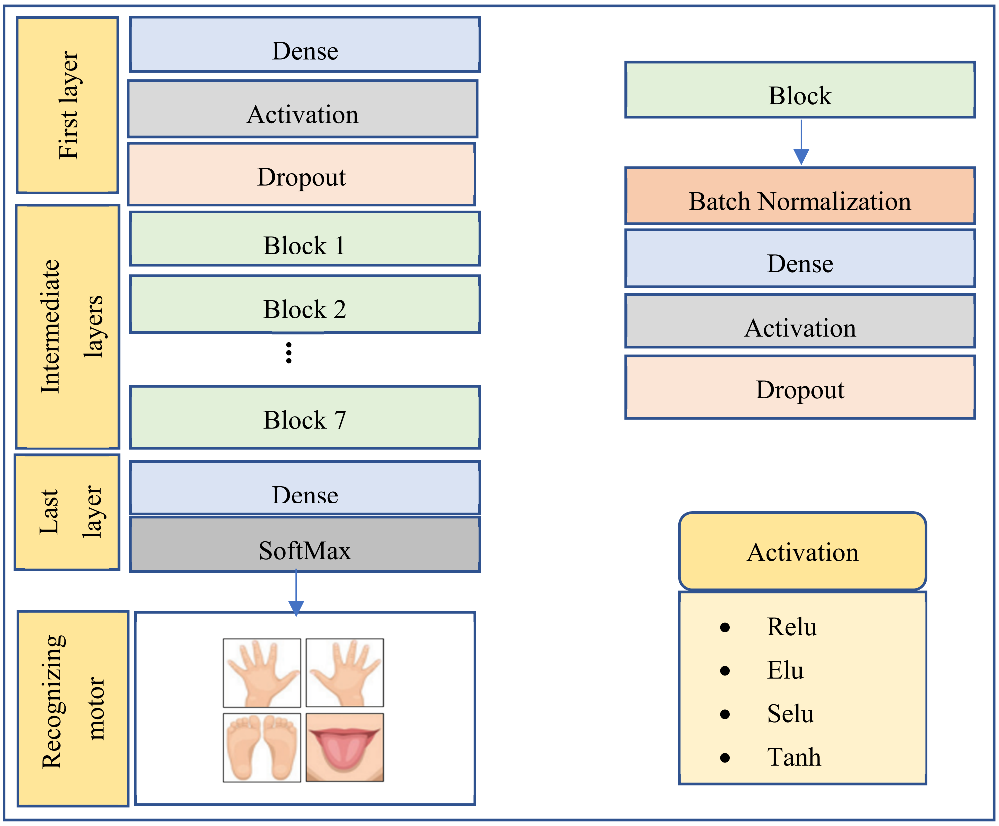

Figure 6. The proposed ANN model.

3.2.4. Proposed CNN Models

We have proposed two CNN models; the first one (Figure 7a) starts with 6 blocks. Each block has a convolutional layer, an activation function, and a Max Pooling step. Then we find the Flatten layer. This layer consists in converting the data into a one-dimensional table which will be considered as input in the next layer. We flatten the output of the convolution layers to create a single long feature vector. And it is connected to the final classification model, called the fully connected layer. In other words, we put all the discrete data on a single line and make connections with the 6 dense layers.

The convolutional layer is the key component of CNN and always constitutes at least their first layer. Its goal is to identify the presence of a set of features in the images received as input. The Max pooling operation consists in reducing the size of the data, while preserving their important characteristics. The Flatten layer makes it possible to flatten the tensor and reduce its dimension. We have used four activation functions: ReLU, SELU, ELU, and Tanh.

Our second CNN model (Figure 7b) has 4 blocks (like the first model) followed by an LSTM layer. The LSTM is a type of RNN.

The idea behind this choice of neural network architecture is to divide the signal between what is important in the short term, through the hidden state (analogous to the output of a simple RNN cell), and what is important in the long term, through the cell state, which will be explained below. Thus, the global functioning of an LSTM can be summarized in 3 steps:

1. Detecting relevant information from the past, taken from the cell state through the forget gate.
2. Select, from the current input, those that will be relevant in the long term via the input gate. These will be added to the cell state, which acts as a long memory.
3. Draw from the new cell state the important short-term information to generate the next hidden state through the output gate.

The LSTM layer is followed by a Flatten layer and 3 Dense layers.

CNNs are commonly used to solve problems with spatial data. LSTMs are best suited for analyzing temporal and sequential data. However, it is useful to think of our CNN 2 architecture as defining two sub-models: the CNN model for feature extraction and the LSTM model for interpreting features across time steps.

The proposed CNN1 model consists of 23,767,300 parameters and 696,324 parameters for the proposed CNN2 model. We have used for the two CNN, the ADAM optimizer which for weight updates and categorical cross-entropy loss. Then, we trained the network by using a batch size of 16 and 1000 training epochs.

The ADAM optimizer is a gradient descent method used for the minimization of an objective function which is written as a sum of differentiable functions. Categorical cross-entropy is a loss function that is used in multi-class classification tasks.

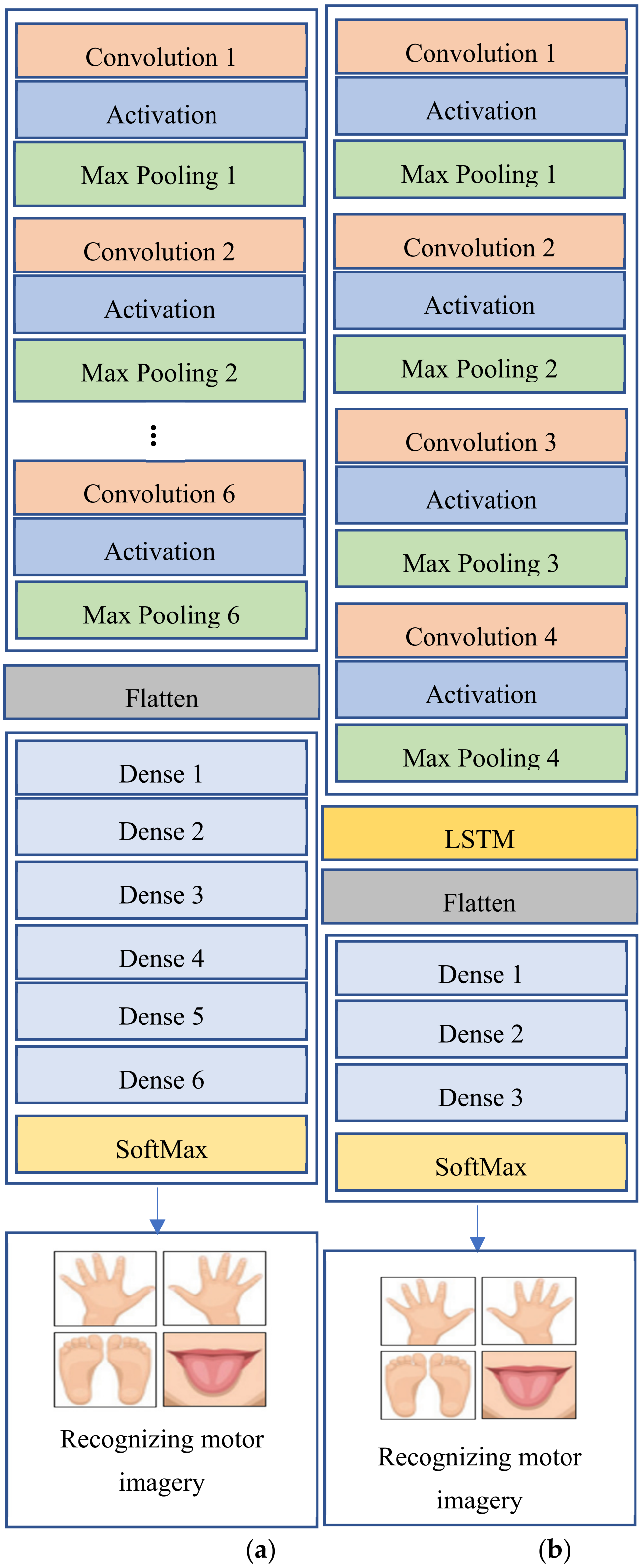

Figure 7. The proposed (**a**) CNN1 and (**b**) CNN2 model.

3.2.5. Proposed Merged CNNs Model

We merged the two models CNN1 and CNN2, as shown in Figure 8. We build this fusion by using monolayers perceptron. The resulting multi-layer CNNs features from the concatenation layer are fed to the monolayers perceptron.

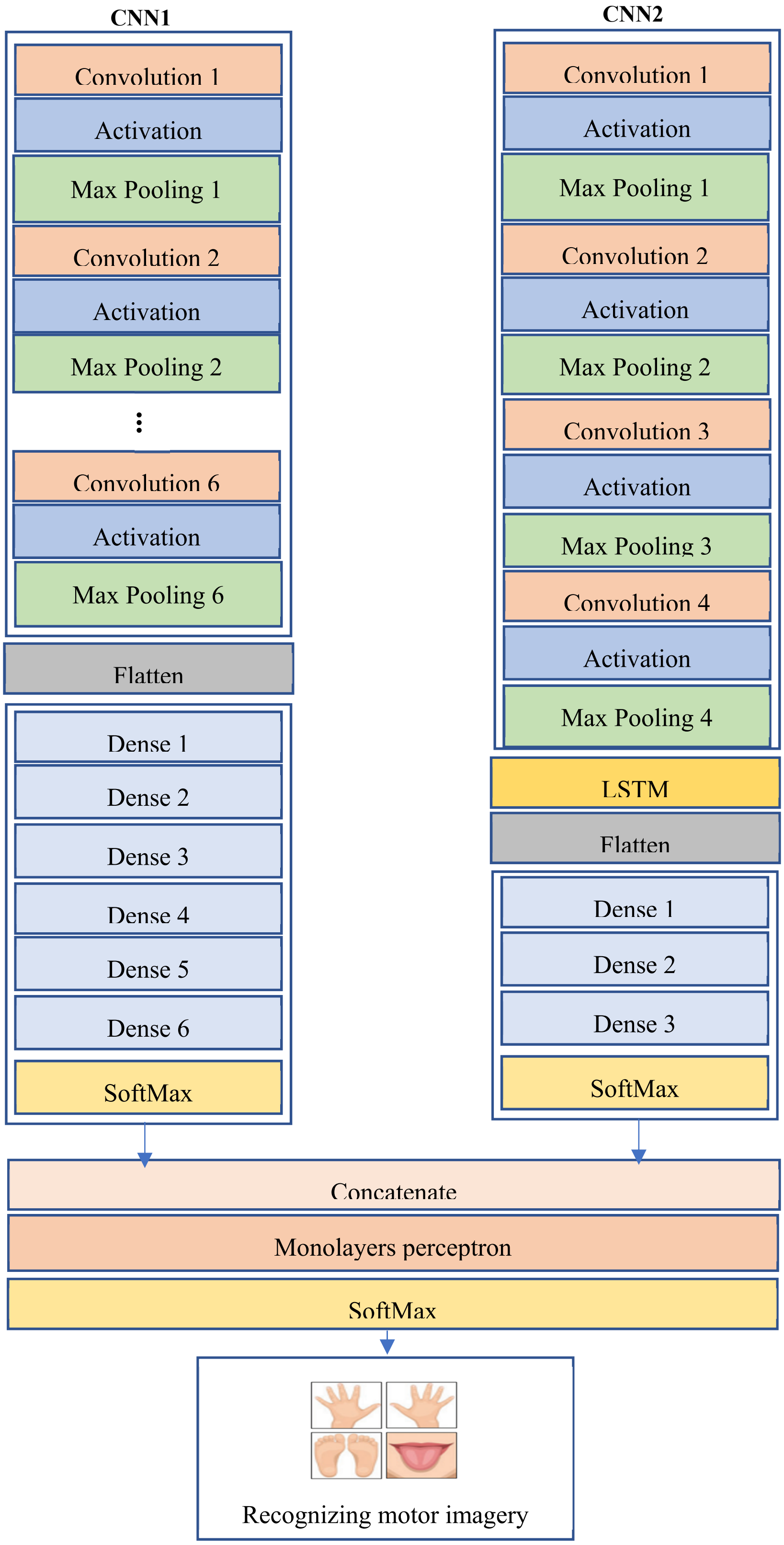

Figure 8. The proposed merged CNNs model.

The monolayers perceptron consists of one hidden layer, each having 8 nodes. The monolayers perceptron method is then trained on the combined feature vector, and the output is sent to the SoftMax layer to get the probability score for the MI classes.

3.2.6. Proposed Modified Merged CNNs Model

We have modified our proposed method of fusing CNN1 and CNN2 while reducing the number of blocks and dense layers, and also eliminating the Max Pooling step. In addition, we combine CNN1 and CNN2 modified (Figure 9) by removing the final SoftMax classification layer from each of them and concatenating the features using perceptron monolayers.

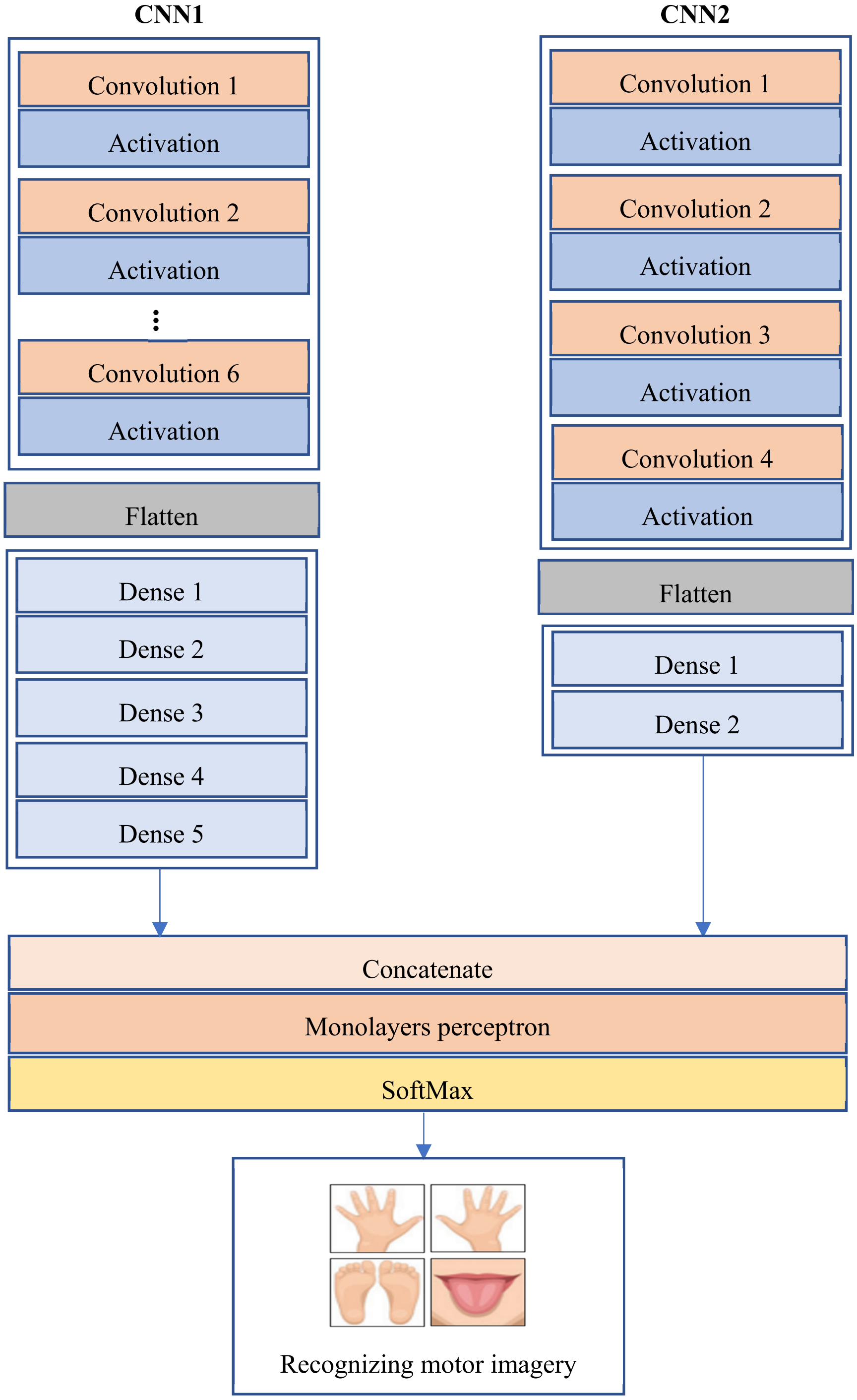

Figure 9. The proposed modified merged CNNs model.

The proposed merged CNNs model consists of 24,463,660 parameters and 24,691,972 parameters for the proposed modified merged CNNs model. We have used for the two CNN, the ADAM optimizer for weight updates and categorical cross-entropy loss. Then, we trained the network by using a batch size of 16 and 1000 training epochs.

4. Results and Discussion

In this section, we analyzed our proposed methods to determine the optimal classifier. According to the instruction of the dataset, a classifier was trained and tested for each subject.

Training a simple DL model takes numerous hours and even days on laptops (typically on CPUs) which leads to an impression that DL requires big systems to run and execute. GPUs and TPUs, on the other hand, can train these models in a matter of minutes or seconds. But not everyone can afford a GPU because they are expensive. That is where Google Colab comes into play. Google Colab is a free cloud service hosted by Google to encourage Machine Learning and Artificial Intelligence research. It is a powerful platform for learning and quickly developing DL models in Python (3.7 version)

As mentioned in Section 3.1, each subject has 288 trials; we used 80% of trials for training and 20% of trials for testing.

4.1. Performance Analysis for Subject 3

The performance measures obtained for subject 3 are shown in Table 1.

Table 1. Classification Report for the proposed methods based on (a) ANN, (b) CNN1, (c) CNN2, (d) Merged CNN1 with CNN2, (e) Modified merged CNN1 with CNN2 model (Subject 3).

		Precision	Recall	F1 Score
ANN	Left hand	81%	87%	84%
	Right hand	73%	80%	76%
	Feet	94%	94%	94%
	Tongue	80%	71%	75%
CNN1	Left hand	71%	83%	77%
	Right hand	83%	91%	87%
	Feet	100%	80%	89%
	Tongue	75%	67%	71%
CNN2	Left hand	50%	58%	54%
	Right hand	84%	73%	78%
	Feet	81%	87%	84%
	Tongue	78%	78%	78%
Merged CNN1 with CNN2	Left hand	90%	75%	82%
	Right hand	84%	95%	89%
	Feet	92%	73%	81%
	Tongue	64%	78%	70%
Modified Merged CNN1 with CNN2	Left hand	82%	75%	78%
	Right hand	91%	91%	91%
	Feet	82%	93%	87%
	Tongue	75%	67%	71%

Table 1 shows the precision, recall, and F1 score of the proposed methods based on ANN, CNN1, CNN2, Merged CNN1 with CNN2, Modified merged CNN1 with CNN2 model for Subject 3. We notice that the five models gave very similar results.

We provide the confusion matrix for the proposed methods in Figure 10. The diagonal elements demonstrate the number of points for which the predicted label is equal to the true label, while off-diagonal elements are those that are mislabeled by the classifier. The higher the diagonal values of the confusion matrix, the better, showing many correct predictions.

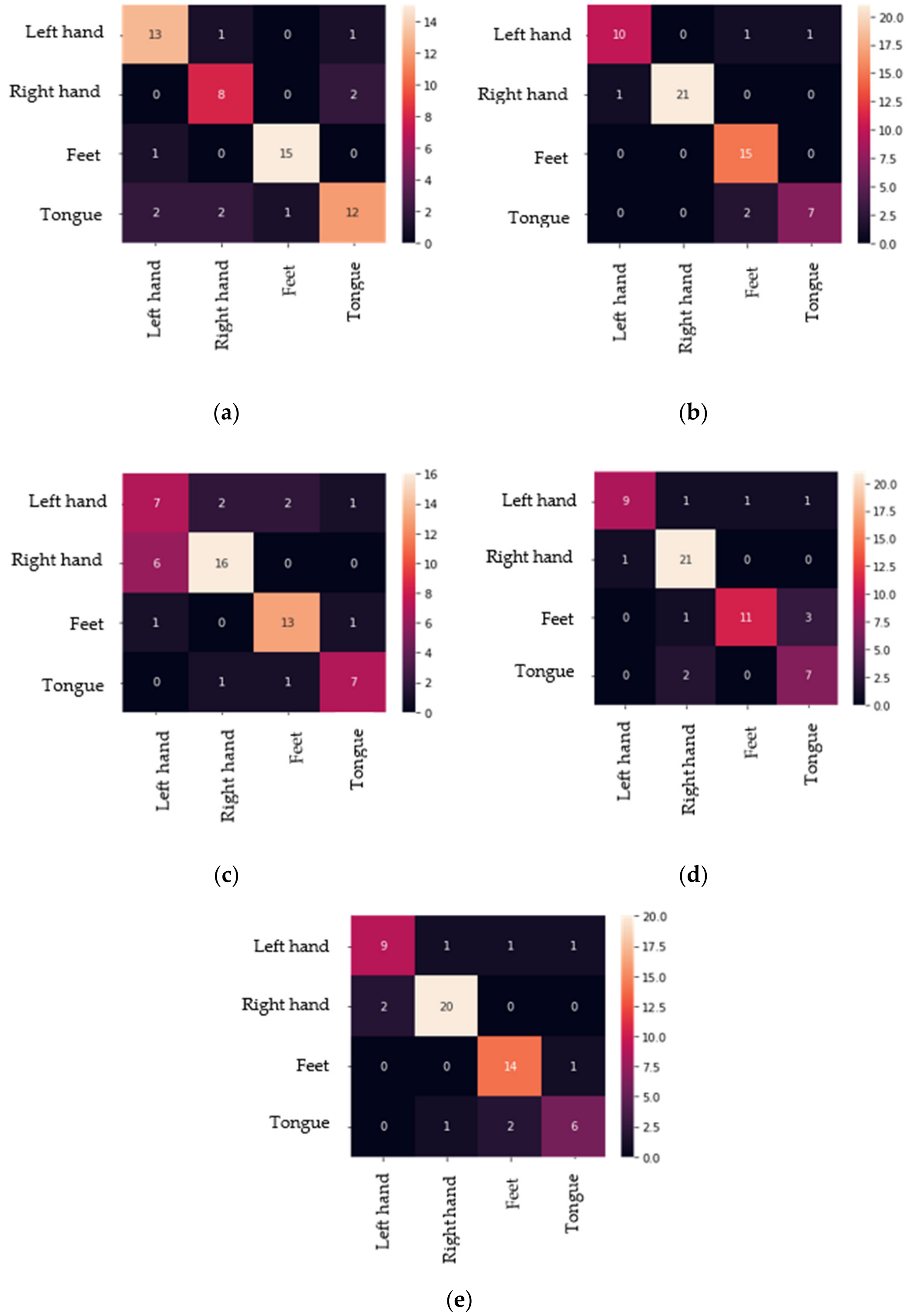

Figure 10. Confusion matrices of classification accuracy for the proposed methods based on (**a**) ANN, (**b**) CNN1, (**c**) CNN2, (**d**) Merged CNN1 with CNN2, (**e**) Modified merged CNN1 with CNN2 model (Subject 3).

4.2. Performance Analysis for All Subjects

Tables 2–6 give the accuracy values obtained by the application of our proposed methods.

Table 2. Classification accuracy of the proposed ANN methods (%).

	Relu	Elu	Selu	Tanh
Subjects 1	72.41	60.34	58.62	63.79
Subjects 2	60.34	58.62	53.45	58.62
Subjects 3	75.86	77.59	84.48	82.76
Subjects 4	29.31	31.03	31.03	31.03
Subjects 5	51.72	53.45	55.17	58.62
Subjects 6	48.28	46.55	44.83	27.59
Subjects 7	74.14	70.69	62.07	65.52
Subjects 8	60.34	67.24	62.07	60.34
Subjects 9	44.83	41.38	46.55	44.83
Mean value	**57.47**	**56.32**	**55.36**	**54.78**

Table 3. Classification accuracy of the proposed CNN1 methods (%).

	Relu	Elu	Selu	Tanh
Subjects 1	70.69	79.31	81.03	68.97
Subjects 2	58.62	67.24	56.90	68.97
Subjects 3	89.66	89.66	81.03	86.21
Subjects 4	50.00	44.83	53.45	53.45
Subjects 5	58.62	53.45	55.17	43.10
Subjects 6	37.93	29.31	31.03	36.21
Subjects 7	62.07	77.59	70.69	62.07
Subjects 8	65.52	74.14	74.14	75.86
Subjects 9	39.66	46.55	53.45	48.28
Mean value	**59.19**	**62.45**	**61.87**	**68.77**

Table 4. Classification accuracy of the proposed CNN2 methods (%).

	Relu	Elu	Selu	Tanh
Subjects 1	65.52	67.24	0.63.79	67.24
Subjects 2	41.38	46.55	0.56.90	50.00
Subjects 3	74.14	68.97	0.72.41	77.59
Subjects 4	32.76	39.66	39.66	43.10
Subjects 5	41.38	37.93	44.83	43.10
Subjects 6	36.21	29.31	32.76	34.48
Subjects 7	75.86	72.41	63.79	75.86
Subjects 8	51.72	50.00	48.28	50.00
Subjects 9	50.00	50.00	55.17	48.28
Mean value	**52.10**	**55.55**	**53.08**	**54.40**

Table 5. Classification accuracy of the proposed merged CNNs methods (%).

	Relu	Elu	Selu	Tanh
Subjects 1	56.90	65.52	34.48	63.79
Subjects 2	43.10	43.10	41.38	50.00
Subjects 3	36.21	60.34	63.79	53.45
Subjects 4	36.21	36.21	27.59	29.31
Subjects 5	32.76	43.10	37.93	29.31
Subjects 6	20.69	20.69	20.69	37.93
Subjects 7	29.31	58.62	67.24	70.69
Subjects 8	15.52	36.21	55.17	55.17
Subjects 9	36.21	50.00	50.00	41.38
Mean value	**34.10**	**45.97**	**44.25**	**47.89**

Table 6. Classification accuracy of the proposed modified merged CNNs methods (%).

	Relu	Elu	Selu	Tanh
Subjects 1	75.86	77.59	81.03	77.59
Subjects 2	55.17	63.79	65.52	63.79
Subjects 3	87.93	86.21	87.93	84.48
Subjects 4	46.55	55.17	50.00	43.10
Subjects 5	50.00	46.55	48.28	41.38
Subjects 6	36.21	22.41	31.03	34.48
Subjects 7	63.79	65.52	75.86	70.69
Subjects 8	75.86	74.14	77.59	74.14
Subjects 9	63.79	65.52	58.62	72.41
Mean value	**61.68**	**61.87**	**64.00**	**62.45**

From these tables, we can conclude that the proposed CNN1 method achieves the best average classification accuracy of 68.77% with the Tanh function.

On the other hand, the comparison between the results obtained by the proposed models CNN1 and CNN2 shows that the addition of the LSTM layers in the CNN model does not improve this model. In addition, merging the two CNNs without adding LSTM layers gives acceptable results.

According to Tables 5 and 6, removing the Max Pooling steps, LSTM layers, and the final SoftMax classification layer from CNN1 and CNN2, and concatenating the features using a perceptron monolayer, improve the results of motor imagery tasks classification. The modified merged CNNs give an improvement in accuracy but the accuracy achieved by CNN1 was the best overall despite being simple.

We can also conclude that based on the comparison of the subject-specific precision obtained by our proposed models, each model gave better results for a different subject.

A comparison of the classification accuracies of the proposed method based on CNN1 and other state-of-the-art methods is presented in Table 7.

As shown in Table 7, our proposed CNN1 method is better than all other machine learning methods with an average classification accuracy of 68.77%, while the NB method is the worst with an average classification accuracy of 58.20%. On the other hand, the average classification accuracy of the FBCSP is 67.21%, which is indicative of the result of the BCI competition.

FBCSP [32] applied to the BCI competition achieved acceptable results with the handcrafted characteristics for each subject. However, it is impossible to find such an optimal craft characteristic for each new subject; therefore, all methods were evaluated under the same conditions in our experiments.

Table 7. Comparison of the classification accuracy (%).

	FBCSP [32]	KNN [32]	LDA [33]	NB [34]	Ensemble [34]	SVM [34]	FLS [34]	Proposed CNN1
Subject 1	73.50	71.60	64.58	63.90	59.40	68.80	71.90	68.97
Subject 2	59.40	51.10	58.08	49.30	46.50	48.30	53.10	68.97
Subject 3	61.90	52.80	61.11	75.70	72.60	76.00	76.40	86.21
Subject 4	71.50	91.00	67.34	63.20	59.00	61.80	66.70	53.45
Subject 5	61.40	65.20	66.55	39.90	39.90	41.30	39.20	43.10
Subject 6	70.10	61.30	59.72	43.80	35.10	41.00	42.40	36.21
Subject 7	69.60	71.10	69.79	70.50	65.60	77.80	73.30	62.07
Subject 8	62.00	67.10	67.10	80.20	70.80	81.60	80.20	75.86
Subject 9	75.50	68.60	60.06	78.50	75.00	79.20	81.60	48.28
Average	**67.21**	**66.64**	**63.81**	**62.80**	**58.20**	**64.00**	**65.00**	**68.77**

In [33], the authors used the Latent Dirichlet Allocation (LDA) to classify Motor Imagery. They achieved an accuracy value of 63.81%.

In reference [34], the authors compared their proposed PSO-based Fuzzy Logic System (FLS) with many competing approaches, including Naïve Bayes (NB), AdaBoostM2 ensemble (Ensemble), and Support Vector Machines (SVM). They applied these machine learning methods using the functions of Matlab, namely, respectively; fitcnb, fitensemble, and fitcecoc. For the SVM method, they used the one-vs-all encoding scheme with a binary SVM learner using the Gaussian kernel. And 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.

Our proposed method based on CNN1 gave the best accuracy value that may be due to the simple preprocessing we applied to the dataset. In addition, the extraction of frequency characteristics using WPD and spatial characteristics using CSP, improved the rate of classification accuracy of MI.

First, our experimental results confirmed the feasibility of the approach based on the CNN1 in the EEG domain, then the effectiveness of the proposed method compared to the methods of the state of the art. We see that the resulting networks successfully learn important characteristics of MI-EEG signals, thus improving overall performance. However, there are still several difficult issues to resolve:

- In this work, merging the two CNNs and adding the LSTM layers did not improve the classification result. These two points can work in this direction either by adding other temporal characteristics, for example, or by applying other fusion methods other than concatenation and the merging by the perceptron monolayer.
- According to Table 7, the methods of the state of the art sometimes surpass our proposed method based on the CNN1 models for a few subjects. It is expected that various architectural extensions, such as adding more convolutional layers, can be applied to our proposed method based on CNN1; however, the effects of these extensions for the EEG domain are not clear. Therefore, we will attempt to extend the architecture of the merged CNNs to determine if the architectural change can affect the overall performance of the system.

As discussed in this section, the use of methods based on the fusion of CNNs can be improved in several directions. We will discuss the aforementioned issues to improve the performance of MI-EEG classification.

5. Conclusions

In this paper, we proposed five methods for the classification of four-class motor imagery EEG signals using DL. The results obtained by the application of the proposed CNN1 model gave better results compared to our proposed and state-of-the-art methods.

The proposed fusion models show an acceptable classification rate of the EEG MI signals; therefore, it would be very interesting to apply the multi-layer CNN fusion models on other EEG datasets. We wish to study other feature fusion methods in order to improve the performance of our proposed method.

Author Contributions: Conceptualization, methodology, simulation and writing original draft, A.E.; methodology, A.M. Supervision, validation and writing—review & editing, W.Z.; Visualization, supervision, M.G. Validation, C.M. Project administration, H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research 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), grant number 739.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is public and can be found at the following links: BCI Competition IV dataset 2a http://www.bbci.de/competition/iv/, accessed on 30 January 2020.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tariq, M.; Trivailo, P.M.; Simic, M. EEG-Based BCI Control Schemes for Lower-Limb Assistive-Robots. *Front. Hum. Neurosci.* **2018**, *12*, 312. [CrossRef]
2. Jacome, D.E.; Morilla-Pastor, D. Unreactive EEG: Pattern in Locked-In Syndrome. *Clin. Electroencephalogr.* **1990**, *21*, 31–36. [CrossRef] [PubMed]
3. Hossain, M.S.; Muhammad, G. Environment Classification for Urban Big Data Using Deep Learning. *IEEE Commun. Mag.* **2018**, *56*, 44–50. [CrossRef]
4. Khan, M.A.; Das, R.; Iversen, H.K.; Puthusserypady, S. Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application. *Comput. Biol. Med.* **2020**, *123*, 103843. [CrossRef] [PubMed]
5. Echtioui, A.; Zouch, W.; Ghorbel, M.; Mhiri, C.; Hamam, H. A novel method for classification of EEG motor imagery signals. In Proceedings of the 17th International Wireless Communications & Mobile Computing Conference—IWCMC 2021, Harbin, China, 28 June–2 July 2021; pp. 1648–1653.
6. Echtioui, A.; Zouch, W.; Ghorbel, M.; Mhiri, C.; Hamam, H. Fusion Convolutional Neural Network for Multi-Class Motor Imagery of EEG Signals Classification. In Proceedings of the 17th International Wireless Communications & Mobile Computing Conference—IWCMC 2021, Harbin, China, 28 June–2 July 2021; pp. 1642–1647.
7. Xu, F.; Rong, F.; Miao, Y.; Sun, Y.; Dong, G.; Li, H.; Li, J.; Wang, Y.; Leng, J. Representation Learning for Motor Imagery Recognition with Deep Neural Network. *Electronics* **2021**, *10*, 112. [CrossRef]
8. Mehmood, R.M.; Du, R.; Lee, H.J. Optimal Feature Selection and Deep Learning Ensembles Method for Emotion Recognition From Human Brain EEG Sensors. *IEEE Access* **2017**, *5*, 14797–14806. [CrossRef]
9. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computeraided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef]
10. Echtioui, A.; Zouch, W.; Ghorbel, M.; Mhiri, C.; Hamam, H. 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, Vienna, Austria, 4–6 February 2021; pp. 591–595, ISBN 978-989-758-484-8.
11. Nicola, M.; Acharya, U.R.; Filippo, M. Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals. *Comput. Biol. Med.* **2019**, *106*, 71–81.
12. Leilei, S.; Bo, J.; Haoyu, Y.; Jianing, T.; Chuanren, L.; Hui, X. Unsupervised EEG feature extraction based on echo state network. *Inf. Sci.* **2019**, *475*, 1–17.
13. Tabar, Y.R.; Halici, U. A novel deep learning approach for classification of EEG motor imagery signals. *J. Neural. Eng.* **2016**, *14*, 016003. [CrossRef]
14. Chambon, S.; Galtier, M.N.; Arnal, P.J.; Wainrib, G.; Gramfort, A. A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 758–769. [CrossRef]
15. Zhao, X.; Zhang, H.; Zhu, G.; You, F.; Kuang, S.; Sun, L. A Multi-Branch 3D Convolutional Neural Network for EEG-Based Motor Imagery Classification. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2019**, *27*, 2164–2177. [CrossRef] [PubMed]
16. Sakhavi, S.; Guan, C.; Yan, S. Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 5619–5629. [CrossRef]
17. Amin, S.U.; Alsulaiman, M.; Muhammad, G.; Mekhtiche, M.A.; Hossain, M.S. Deep Learning for EEG motor imagery classification based on multi-layer CNNs feature fusion. *Future Gener. Comput. Syst.* **2019**, *101*, 542–554. [CrossRef]

18. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
19. Bengio, Y.; Lamblin, P.; Popovici, D.; Larochelle, H. Greedy layer-wise training of deep networks. *Adv. Neural Inform. Process. Syst.* **2006**, *153*, 19.
20. Yang, H.; Sakhavi, S.; Ang, K.K.; Guan, C. On the use of convolutional neural networks and augmented CSP features for mul-ti-class motor imagery of EEG signals classification. In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, Milan, Italy, 25–29 August 2015; pp. 2620–2623.
21. Lee, J.; Nam, J. Multi-level and multi-scale feature aggregation using pretrained convolutional neural networks for music auto-tagging. *IEEE Signal Process. Lett.* **2017**, *24*, 1208–1212. [CrossRef]
22. Li, E.; Xia, J.; Du, P.; Lin, C.; Samat, A. Integrating Multilayer Features of Convolutional Neural Networks for Remote Sensing Scene Classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5653–5665. [CrossRef]
23. Soleymani, S.; Dabouei, A.; Kazemi, H.; Dawson, J.; Nasrabadi, N.M. Multi-level feature abstraction from convolutional neural networks for multimodal biometric identification. *arXiv* **2018**, arXiv:1807.01332.
24. Zhang, P.; Wang, D.; Lu, H.; Wang, H.; Ruan, X. Amulet: Aggregating Multi-level Convolutional Features for Salient Object Detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 202–211. [CrossRef]
25. Bhattacharjee, P.; Das, S. Two-Stream Convolutional Network with MultiLevel Feature Fusion for Categorization of Human Action from Videos Pattern Recognition and Machine Intelligence. In Proceedings of the International Conference on Pattern Recognition and Machine Intelligence, Kolkata, India, 5–8 December 2017; Springer: Cham, Switzerland, 2017; Volume 10597.
26. Hariharan, B.; Arbelaez, P.; Girshick, R.; Malik, J. Hyper-columns for object segmentation and fine-grained localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 447–456.
27. Ueki, K.; Kobayashi, T. Multi-layer feature extractions for image classification—Knowledge from deep CNNs. In Proceedings of the International Conference on Systems, Signals and Image Processing (IWSSIP), London, UK, 10–12 September 2015; pp. 9–12.
28. Lu, N.; Li, T.; Ren, X.; Miao, H. A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *25*, 566–576. [CrossRef]
29. Hao, Y.; Yang, J.; Chen, M.; Hossain, M.S.; Alhamid, M.F. Emotion-Aware Video QoE Assessment Via Transfer Learning. *IEEE MultiMedia* **2018**, *26*, 31–40. [CrossRef]
30. Schirrmeister, R.T.; Springenberg, J.T.; Fiederer, L.D.J.; Glasstetter, M.; Eggensperger, K.; Tangermann, M.; Hutter, F.; Burgard, W.; Ball, T. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **2017**, *38*, 5391–5420. [CrossRef] [PubMed]
31. BCI Competition IV. Available online: http://www.bbci.de/competition/iv/ (accessed on 30 January 2020).
32. Kwon-Woo, H.; Jin-Woo, J. Motor Imagery EEG Classification Using Capsule Networksy. *Sensors* **2019**, *19*, 2854.
33. Ikhtiyor, M.; Taegkeun, W. Efficient Classification of Motor Imagery Electroencephalography Signals Using Deep Learning Methods. *Sensors* **2019**, *19*, 1736.
34. Thanh, N.; Imali, H.; Amin, K.; Lee, G.-B.; Chee, P.L.; Saeid, N. Classification of Multi-Class BCI Data by Common Spatial Pattern and Fuzzy System. *IEEE Access* **2018**, *6*, 27873–27884.