Biomedical Signal Processing and Control 70 (2021) 102965

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Single upper limb functional movements decoding from motor imagery EEG signals using wavelet neural network

Xiaobo Zhou [a], Renling Zou [a,*], Xiayang Huang [b]

[a] University of Shanghai for Science and Technology, Shanghai, China
[b] The University of British Columbia, Vancouver, Canada

ARTICLE INFO

Keywords:
Wavelet neural network
Brain-computer interface
Motor imagery
Movement decoding
Wavelet packet decomposition

ABSTRACT

In this study, we propose a wavelet neural network (WNN) to improve the performance of movements decoding from motor imagery (MI) EEG signals. Fifteen healthy subjects performed six functional movements of single upper limb: forearm pronation/supination, hand open/close, and elbow flexion/extension. The wavelet packet decomposition (WPD) was used to decompose MI EEG into sub-bands, and the statistical features were extracted from the sub-bands to code the functional movements. Subsequently, the principal component analysis (PCA) was used to select the best feature vector. The factors of mother wavelet, multidimensional wavelet function, number of channels, and the time of imaging segment were optimized through experiments. As a result, the best accuracy of 86.27 ± 6.98% was achieved with the optimized coif1 mother wavelet with a six-level decomposition, Mexican Hat wavelet as a hyperparameter, the time of imaging segment set as 0.7 s, and 61 channels data. In addition, to further verify the efficiency of the WNN classifier, the comparative experiments were conducted with the support vector machine (SVM), linear discriminant analysis (LDA), k-nearest neighbor (KNN), and single hidden layer feedforward neural network (ANN). The results show that the accuracy of WNN is improved by about 15 ~ 40%, which proved the effectiveness of the WNN method. This study provides a new way to decode from motor imagery EEG in multi-movement classification.

1. Introduction

Brain-computer interface systems (BCIs) aim to establish communication between the computer and the human brain [1]. There are various methods of constructing BCIs, and motor imagery (MI) based BCIs provide a way to reflect human movement information through non-muscle channels [2,3]. The electroencephalography (EEG) signals can detect motion intention more quickly than electromyogram (EMG) signals do. It is widely applied in the fields of rehabilitation medicine, virtual reality and robot control, etc. [2,4,5]. Generally, BCIs commands are usually obtained from decoding neural and bioelectrical signals by using classification algorithms.

In the past decades, there has been significant progress in decoding research from MI EEG research. Wang et al. [2] used correlation of spatial patterns decoding right-hand and left-hand movements and obtained an average accuracy of 80%. Wang et al. [1] used the k-nearest neighbor (KNN) classifier to distinguish the right-hand and right-foot MI tasks. Pfurtscheller et al. [6] investigated the EEG of patients with complete spinal cord injury (SCI) during the imaging of right-hand, left-hand, and foot movements. The result was obtained by common spatial patterns (CSP) and linear discriminant analysis (LDA) classifier. Tabar et al. [7] used convolutional neural networks (CNN) and stacked autoencoders (SAE) decoding right-hand and left-hand movements. Schlogl et al. [8] used the support vector machine (SVM) decoding right-hand, left-hand, tongue, and foot movements and obtained an accuracy between 52% and 77%. Sreeja et al. [9] proposed a sparsity approach decoding right-hand, left-hand, tongue, and foot movements and obtained an average accuracy of 91.48%. The complex functional movements of the upper limb have been decoded to communicate better between the computer and the human brain [10–12]. Mohseni et al. [10] decoded five different upper limb movements, using a KNN classifier, an average classification accuracy of 94% is obtained. Ofner et al. [11] analyzed the encoding of single upper limb movements in the time-domain of low-frequency (< 3 Hz) MI EEG signals. They applied six-class movements including, elbow flexion, elbow extension, forearm pronation, forearm supination, hand open, and hand close. They used an LDA classifier and obtained an average accuracy from 15 healthy subjects of about 27%. Ofner et al. [12] investigated 10 patients of cervical

* Corresponding author.
E-mail address: zourenling@163.com (R. Zou).

https://doi.org/10.1016/j.bspc.2021.102965
Received 22 September 2020; Received in revised form 6 May 2021; Accepted 5 July 2021
Available online 15 July 2021

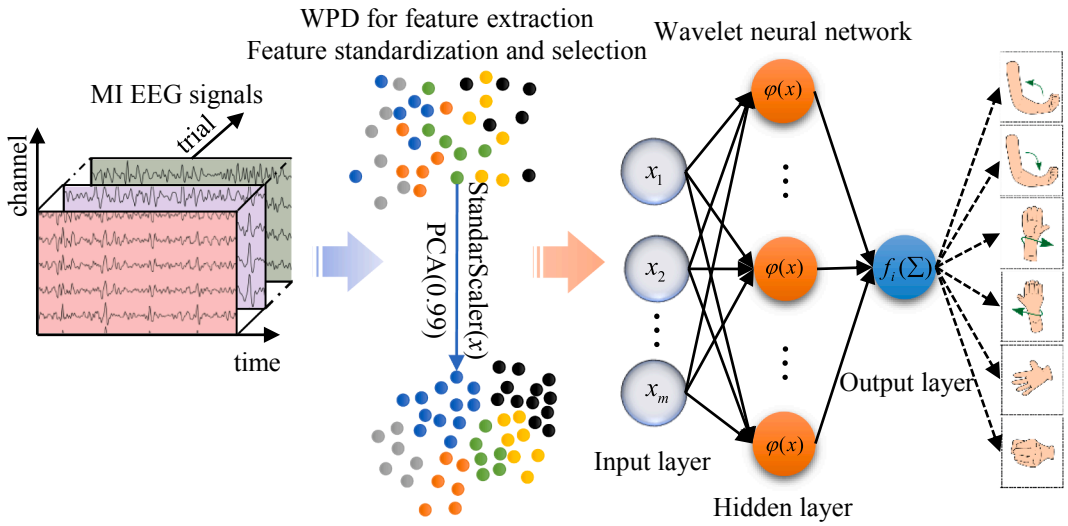

Fig. 1. The approach for the decoding single upper limb functional movements from MI EEG.

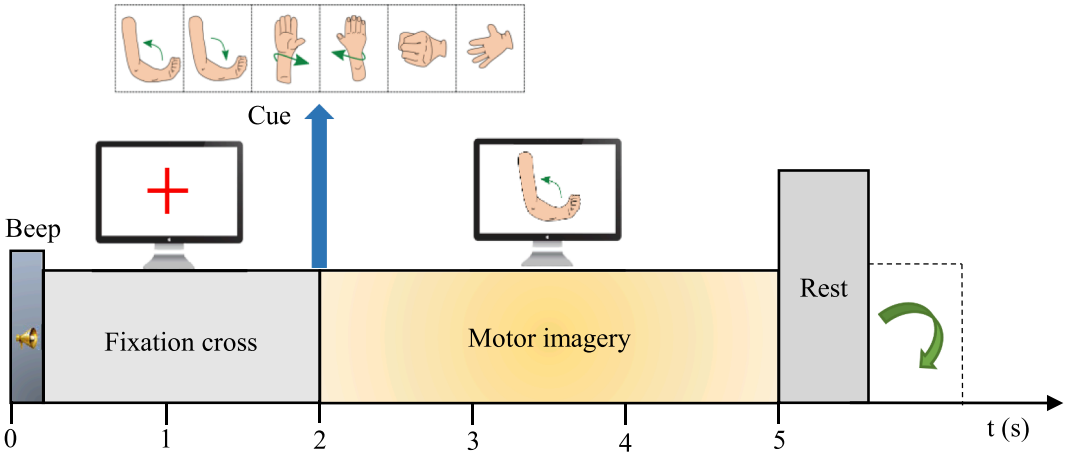

Fig. 2. Experimental paradigm.

SCI with hand open, forearm pronation, forearm supination, palmar grasp, and lateral grasp. Using an LDA classifier to decode five-class functional movements from low-frequency (< 3 Hz) MI EEG signals, an average classification accuracy rate of 45% is obtained.

The above decoding EEG process is in the usual way of SVM, CSP, LDA, KNN and CNN, etc. [1,3,5–7,12–17]. The wavelet neural network (WNN) is a new class of the neural network model proposed by Zhang et al. [18]. In this classifier, the multidimensional wavelet function is used as a hyperparameter and combines the advantages of feedforward neural network and wavelet analysis. So far, the WNN has been widely used in signals analysis, nonlinear time series analysis, and real-time prediction [19,20]. This method is very suitable for non-stationary and non-linear signals analysis. However, the WNN has not been used for MI EEG signal classification.

In this paper, we proposed a WNN model for single upper limb six-class functional movements decoding. In signal processing, we optimized the approach of feature vector construction. Then the parameters of WNN, the number of EEG channels, and the time of imaging segment may respectively affect the performance of decoding, and they are optimized in the method of experiment. We appraise the efficiency of the proposed WNN classifier by a comparison experiment with single hidden layer feedforward neural network (ANN), SVM, LDA, and KNN.

2. Methodology

The schematic representation of the WNN approach for decoding single upper limb functional movements from MI EEG signals is illustrated in Fig. 1. The steps followed in the proposed approach including signal preprocessing, data segmentation, feature extraction, feature standardization and selection, and WNN-based classification.

2.1. Dataset

The public EEG dataset from BNCI Horizon 2020 has been used in this study [10,21]. This dataset was collected from the 15 healthy subjects aged between 22 and 40 years with an average age of 27 years. In the paradigm, the six types of functional movement in the experiment include forearm pronation/supination, hand open/close, and elbow flexion/extension. This study only uses the first two continuous training sessions of the dataset of each subject who has completed six-class different upper limb functional movements which were performed 72 times respectively. The paradigm was trial-based, and cues were displayed on a computer screen in front of the subjects. The paradigm is shown in Fig. 2. At second 0, a cross appears on the computer screen and beeps. Afterward, at second 2, a cue was presented on the computer screen, indicating the required movement until the 5th second. At the end of the trial, subjects moved back to the starting position. The EEG

Table 1
Feature extraction and selection algorithm description.

Algorithm 1: Feature extraction and selection algorithm.
Input: EEG epochs.
Output: Feature vector.
1. Input EEG epochs, mother wavelet name, the max-level of decomposition.
2. Compute the mean, standard deviation, range, median, root mean square, skewness, kurtosis, and the sub-band ratio of WPD.
3. Standardization of features.
4. Compute covariance matrix of features.
5. Return the feature vector.

was sampled at 512 Hz, and it was filtered between 0.01 Hz and 200 Hz. Power line interference was suppressed with a notch filter at 50 Hz. The 61 channel signals were measured and reference was placed on the right mastoid, ground on AFz. Further details about this dataset can be found in [10].

2.2. Signal processing and feature extraction

We used EEGLAB [22] to remove the irrelevant channels from the raw data and to interpolate bad electrodes using channel spectra maps. Independent component analysis (ICA) was performed to identify components representing eye movements and muscle artifacts. The thresholds were set as brain components $< 5\%$. Then it was filtered between 8 and 30 Hz with notch-filter on [23]. Studies have shown that the period of motor imagery could not last more than 0.5 s [9,24]. Therefore, the filtered signals are divided into 0.5 s continuous imagery period segments to encode the corresponding movements, which is done by using MNE [25]. After segmentation, the dataset is randomly split into training and testing sets. The training sets are used for selecting the best features and training the classifier model, whereas the testing sets are used for evaluating the classification.

Feature extraction is crucial for the analysis and processing of EEG. The wavelet packet decomposition (WPD) used in this study provides a multi-level time–frequency decomposition of signals and has good time–frequency localization characteristics [13]. It overcomes the shortcomings of the Fourier transformation that has no localization in the time domain, and it is widely used in biomedical signal analysis. Extracting statistical features from each frequency band of WPD is the most common and an easy method to calculate the feature vector method. This article selects the mean, standard deviation, range, median, root mean square, skewness, kurtosis, and the sub-band ratio. Decomposing the signal in n layers will produce 2^n wavelet packet coefficients, and the algorithm principle expression of wavelet packet decomposition is [26,27]:

$$S_{\xi,\eta,2f} = \frac{1}{\sqrt{2}}\sum_k h_{k-2\eta} S_{\xi+1,k,f} \tag{1}$$

$$S_{\xi,\eta,2f+1} = \frac{1}{\sqrt{2}}\sum_k g_{k-2\eta} S_{\xi+1,k,f} \tag{2}$$

where $S_{\xi,\eta,2f}$ and $S_{\xi,\eta,2f+1}$ are the results of the next layer of wavelet decomposition; $S_{\xi+1,k,f}$ is the results of the upper layer decomposition; ξ is the scale coefficient; η is the position coefficient; f is the sampling frequency; g and h are quadrature conjugate low-pass and high-pass filters.

Standardization of datasets and using principal component analysis (PCA) to reduce dimensionality is a well-known requirement for a variety of machine learning algorithms [28,29]. PCA maps sample data to the principal component space of the feature matrix to preserve most of the data differences and has lower dimensions. The principal component vectors are formed orthogonal to each other, so they have the highest variance. The principal component P_k is formed from the combination of n-dimensional original data X, and its expression is as follows:

$$P_k = a_{k1}x_1 + a_{k2}x_2 + \cdots + a_n x_{kn}, \sum_i a_{ki}^2 = 1 \tag{3}$$

By calculating the covariance matrix and setting the threshold of the cumulative variance to 99.99%, sufficient data variance can be retained, and the dimensionality can be effectively reduced. Feature extraction and selection algorithm for MI EEG classification can be summarized in Table 1.

2.3. Wavelet neural network for classification

The fundamental idea of WNN is to adapt the wavelet basis to the training data, introduce variables into the neural network in the input layer, and transform the input variables into the expansion and translation of the mother wavelet in the hidden layer, and finally, estimate the approximate value of the target value in the output layer [18,30]. In this study, we implement multidimensional wavelet function $\varphi_j(x)$ as the hyperparameter in the hidden layer, then the activation function was used in the output layer. The hidden layers are just adding more neurons in between the input and output layers. Data in the input layer was labeled as X_m. The number of hidden layer neurons $N = 2m + \alpha, \alpha \in (0, 9)$ is determined by an empirical formula [19]. Neurons in the output layer depend on the number of classifications. The calculation of the output value Y is defined as [30]:

$$Y_{[i=1,2,3\ldots]} = g_\lambda(x, w) = w_{N+1} + \sum_{j=1}^{\lambda} w_j \cdot \varphi_j(x) + \sum_{k=1}^{m} w_k \cdot x_k \tag{4}$$

where $\varphi_j(x)$ is multidimensional wavelet, the network weights w contain the direct connections (w_k), weights (w_j), translation factors $(w_{\xi(ij)})$ and dilation factors $(w_{\zeta(ij)})$.

WNN is inappropriate to initialize the initial values to random values like the sigmoid neural network. The initial values of the direct connections and the weights are initialized in small random values between 0 and 1 [19]. However, the initialization of the translation factors and the dilation factors have a significant impact on the performance of the neural network. Various methods have been proposed for their initialization. Zhang et al. [30] have introduced a heuristic method as follows:

$$w_{(\xi)ij} = 0.5(N_i + M_i) \tag{5}$$

$$w_{(\zeta)ij} = 0.2(M_i - N_i) \tag{6}$$

where M_i and N_i are defined as the maximum and minimum of input X_m.

WNN algorithms originated from the feedforward neural network. The basic idea of backpropagation was used, which is to find the percentage of contribution of each weight to the error. The definition of E and updating is performed as [19,30]:

$$E = \sum \frac{(y_t - y_p)^2}{2} \tag{7}$$

$$w_k = w_k - l\frac{\partial E}{\partial w_k} \tag{8}$$

$$w_j = w_j - l\frac{\partial E}{\partial w_j} \tag{9}$$

$$w_{(\xi)ij} = w_{(\xi)ij} - l\frac{\partial E}{\partial w_{(\xi)ij}} \tag{10}$$

$$w_{(\xi)ij} = w_{(\xi)ij} - l\frac{\partial E}{\partial w_{(\xi)ij}} \tag{11}$$

where the y_t is the final output, the y_p is desired output, and the l is learning rate, respectively.

A critical decision related to the training of a WNN is when the

Table 2
Wavelet neural network algorithm description.

Algorithm 2: WNN algorithm.

Input: A set of EEG training data X_m, the number of hidden nodes and the number of classifications.

Output: Predicted label.

1. Randomly initialize weight vector direct connections (w_k), weights (w_j). The translation factors ($w_{\zeta(ij)}$) and dilation factors ($w_{\varsigma(ij)}$) initialize by formula (5) and (6).
2. Compute the multidimensional wavelets $\varphi_j(x)$.
3. Compute the hidden layer output using multidimensional wavelets $\varphi_j(x)$.
4. Compute the output using activation function $f(x)$.
5. Compute the error E by formula (7).
6. Update the parameters by formula (8), (9), (10) and (11).
7. If the parameters condition $< 10^{-8}$, stop and return predicted label. Else, return to step 3.

weight adjustment should end. The training is stopped when one of the following criteria is met the cost function reaches a fixed lower bound and the variations of the parameters reach a lower bound. In our experiment, the parameters of dropout were set to 10^{-8} [30,31]. WNN algorithm for EEG classification can be summarized in Table 2.

2.4. Evaluating the performance

The results are recorded and evaluated using 10-fold cross-validation. We used accuracy, kappa value, and precision to evaluate the classification performance of the proposed methods in the movement classification of a six-classification task [9,32]. The decoding timing cost is obtained by a PC with a hardware property of 1.6 GHz CPU with 8 GB RAM. Besides, we implemented the signal processing algorithm and WNN algorithm by Python3.6 [33]. Moreover, the paired t-test is used to statistically compare the performance parameters of different methods, and the level of significance was set as $P = 0.05$.

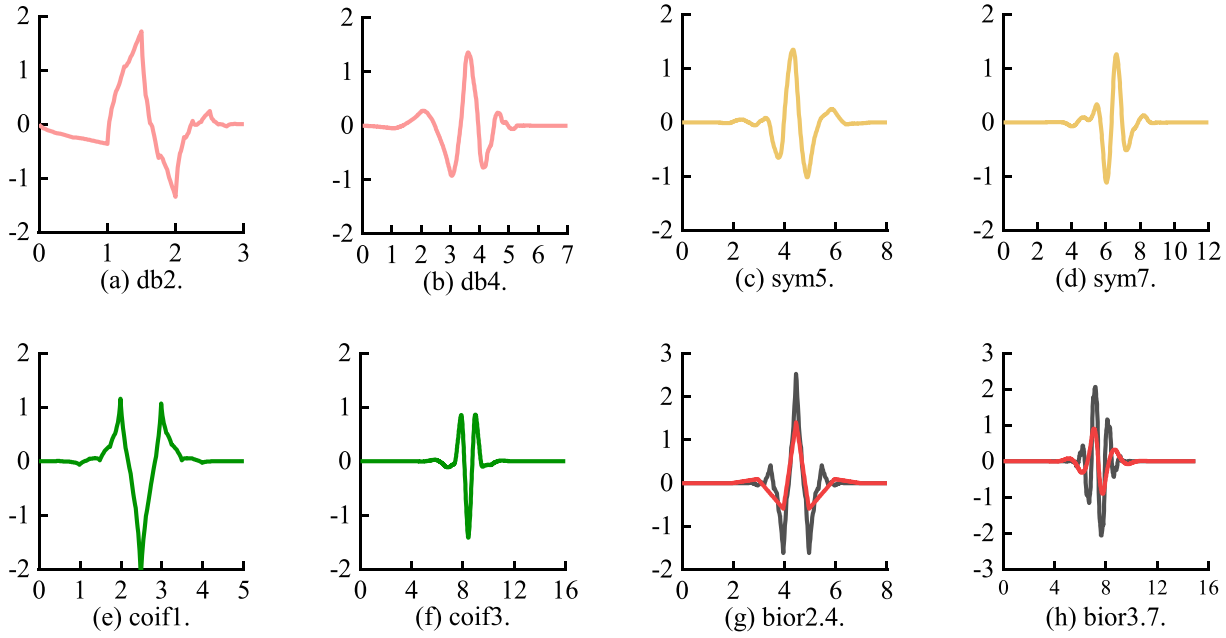

Fig. 3. Eight different mother wavelet functions.

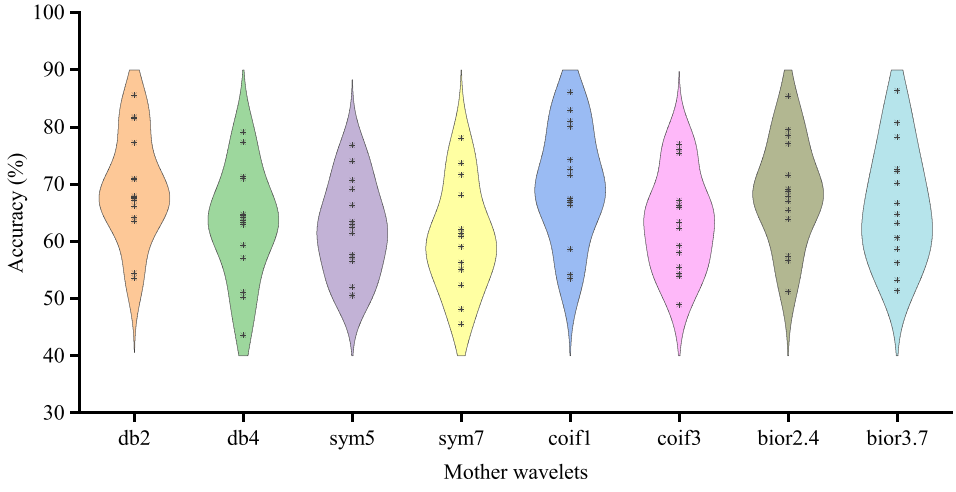

Fig. 4. The average accuracy based on eight different mother wavelets for six-level decompositions.

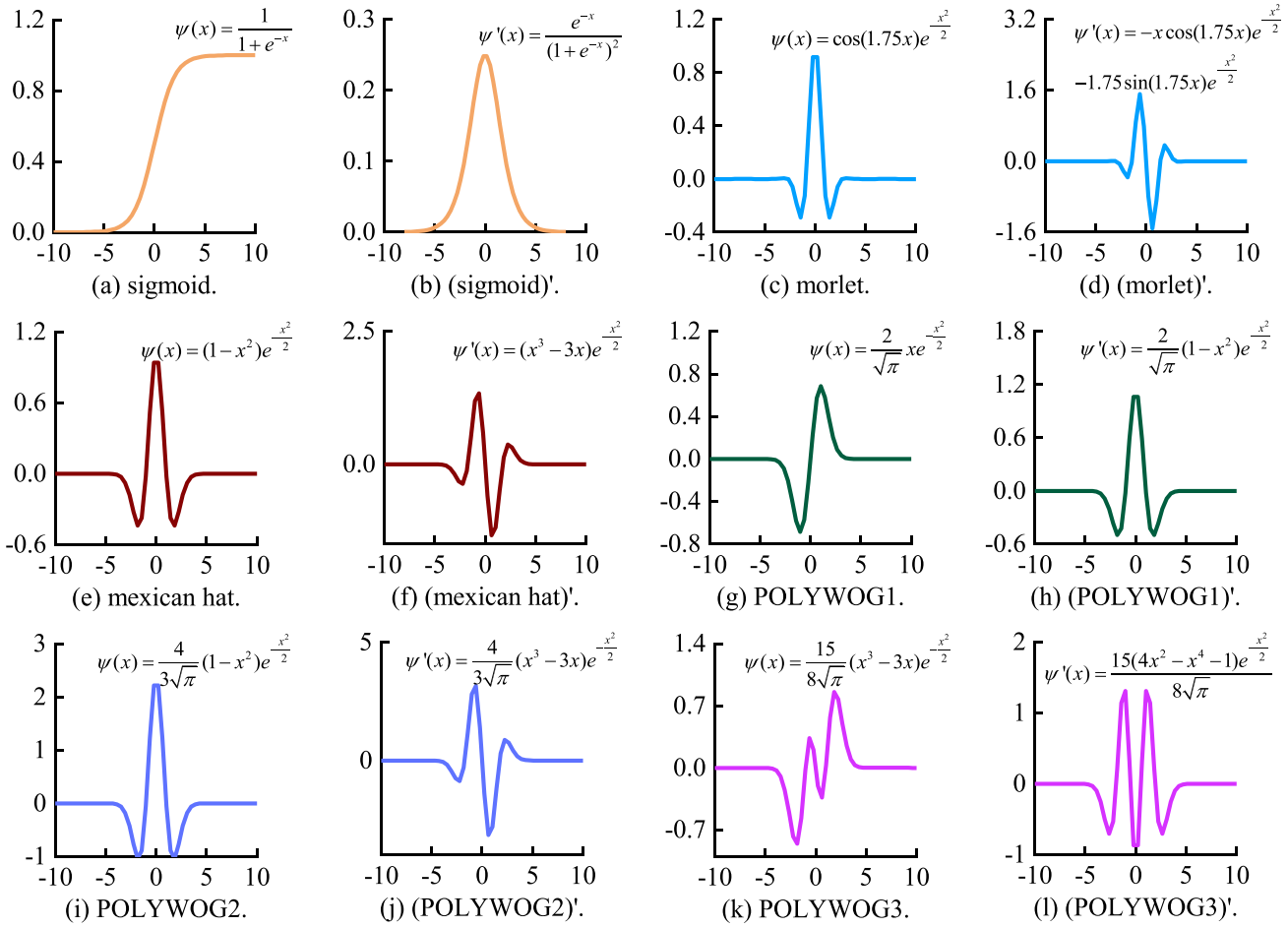

Fig. 5. Activation functions and their derivatives.

3. Results and discussion

3.1. Choice of mother wavelet

The WPD was used to decomposed MI EEG signals at the feature extraction stage and extract statistical features from each frequency band. However, a different mother wavelet function has a different time–frequency characteristic. In order to explore the influence of the time–frequency characteristics of different mother wavelets on the results, we experimented with eight distinct mother wavelet functions, namely db2, db4, sym5, sym7, coif1, coif3, bior2.4, and bior3.7, respectively [12,34]. ANN is used as the baseline classifier. It may be noted that db2 and db4 belong to Daubechies wavelets; sym5 and sym7 belong to symlet wavelets; coif1 and coif3 belong to coiflet wavelets; bior2.4 and bior3.7 belong to biorthogonal wavelets. All the above eight types of mother wavelet functions have biorthogonality and compact

Table 3
The average classification accuracy obtained by10-flod cross-validation (mean ± std).

Subject	Six different hyperparameter functions					
	Sigmoid	Morlet	Mexican Hat	PLOYWOG1	PLOYWOG2	PLOYWOG3
S01	71.85 ± 2.90	89.28 ± 6.03	**91.01 ± 1.13**	90.51 ± 1.75	89.62 ± 1.59	83.32 ± 1.24
S02	67.32 ± 2.45	83.07 ± 1.86	**86.77 ± 0.75**	84.82 ± 1.07	84.41 ± 0.52	76.34 ± 1.62
S03	67.20 ± 2.94	70.53 ± 1.34	72.56 ± 0.81	72.74 ± 1.76	**72.63 ± 1.78**	61.66 ± 2.54
S04	68.86 ± 2.15	83.14 ± 0.90	85.45 ± 1.00	85.16 ± 0.96	**85.85 ± 1.24**	76.91 ± 1.55
S05	64.04 ± 2.38	65.64 ± 2.87	**72.70 ± 2.87**	71.54 ± 0.72	72.47 ± 1.32	61.23 ± 2.75
S06	72.67 ± 2.46	70.74 ± 0.93	73.81 ± 0.54	73.90 ± 1.73	**74.29 ± 1.42**	49.12 ± 2.65
S07	65.89 ± 3.09	84.21 ± 1.08	86.09 ± 1.05	86.51 ± 1.15	**87.55 ± 1.38**	67.96 ± 2.31
S08	76.67 ± 1.89	90.15 ± 1.44	90.65 ± 0.45	**91.95 ± 0.60**	91.19 ± 1.23	77.25 ± 2.46
S09	72.67 ± 2.52	78.07 ± 1.41	**86.48 ± 2.11**	80.30 ± 1.27	81.68 ± 0.75	61.32 ± 7.10
S10	77.03 ± 1.29	94.97 ± 0.43	**96.48 ± 0.41**	95.68 ± 1.01	94.61 ± 0.73	93.36 ± 0.87
S11	66.66 ± 4.36	81.96 ± 1.42	85.09 ± 1.15	85.57 ± 1.21	**85.82 ± 1.58**	61.71 ± 2.24
S12	67.78 ± 3.93	85.16 ± 0.89	**88.15 ± 0.99**	86.92 ± 1.64	87.17 ± 1.18	61.05 ± 4.47
S13	76.26 ± 3.04	90.74 ± 0.76	91.91 ± 0.64	91.33 ± 0.82	**92.01 ± 0.81**	84.96 ± 2.44
S14	66.64 ± 2.94	80.11 ± 2.03	**86.48 ± 1.93**	81.91 ± 1.33	82.57 ± 0.92	61.32 ± 2.37
S15	67.97 ± 1.55	81.57 ± 1.14	**85.94 ± 1.27**	83.37 ± 0.79	83.00 ± 1.96	77.88 ± 0.81
Mean	69.97 ± 4.23	81.96 ± 8.14	**85.24 ± 7.01**	84.15 ± 7.16	84.32 ± 6.83	70.36 ± 12.05

The numbers in bold indicate the highest performance metrics.

Table 4
Classification performance based on various classifiers.

Classifier	10-flod cross-validation (mean ± std)		
	Accuracy (%)	Kappa	Precision
LDA	44.84 ± 4.84	0.34 ± 0.06	0.45 ± 0.05
WNN	**85.24 ± 7.01**	**0.82 ± 0.08**	**0.85 ± 0.07**
SVM	50.37 ± 5.98	0.40 ± 0.07	0.51 ± 0.06
KNN	67.36 ± 6.73	0.61 ± 0.08	0.68 ± 0.07
ANN	69.97 ± 4.23	0.64 ± 0.05	0.70 ± 0.04

The numbers in bold indicate the highest performance metrics.

support, and they are commonly used in biomedical signal processing [34]. The waveforms of each mother wavelet function are shown in Fig. 3.

The obtained average accuracy of 15 subject's classification results is based on eight different mother wavelet functions shown in Fig. 4. The experimental results have shown that the coif1 mother wavelet used in WPD with six-level decomposed to extraction feature providing better classification results than other mother wavelets. A paired *t*-test showed that there is a significant difference in performance among the different mother wavelet functions (p-value < 0.05). Moreover, although there is no significant difference between coif1 and db2 mother wavelet, the coif1 mother wavelet has the maximum average classification accuracy rate. Therefore, coif1 mother wavelet function was selected as the best mother wavelet function for MI EEG signals feature extraction in our study.

3.2. Hyperparameter optimization

The activation functions were used in applicable deep learning architectures across all studies. Moreover, for the hidden layer, most of the studies employing the hidden layer for the deep learning architectures used sigmoid or rectified linear unit as the layer's activation function [35]. The activation function has a substantial impact on the neural network as well as the different mother wavelet used in WPD has a significant impact on classification performance. The mother wavelet activation function of WNN can be any of the multidimensional wavelet functions. But in literature, three of them are commonly adopted, Morlet, Mexican Hat, and Gaussian derivative [36,37]. In this study, we selected Morlet, Mexican Hat, and POLYWOG family wavelets for experiments to select the best multidimensional wavelet activation function in MI EEG classification. The hyperparameter functions and their derivatives are shown in Fig. 5.

To further appraise the efficiency, the report on comparative performance analysis of the proposed method with six different hyperparameter functions is shown in Table 3. It was observed that using the Mexican Hat wavelet function, the obtained maximum mean reached 85.24 ± 7.01%. Compared to the other activation functions, Mexican Hat has a significant difference in classification performance with sigmoid, Morlet, and POLYWOG3 (p-value < 0.05). Since EEG signals are non-liner and non-stationary in nature [38], many sensorimotor functions are non-strictly-periodic EEG signals related to the production of rhythmic movements [39]. These characteristics have caused the deep neural network difficulty in selecting the network model under fit the MI EEG in the research of decoding MI EEG and the structure of the neural network model. Time-frequency characteristics and multi-resolution capabilities make wavelet functions become a powerful tool for processing EEG signals. WNN enhances the fitting ability of the network model to MI-EEG by introducing a multidimensional wavelet as a hyperparameter. Therefore, Mexican Hat wavelet function was selected as the best hyperparameter function for WNN classifier in our study.

3.3. Performance of various classifiers

To further appraise the efficiency of the proposed WNN classifier, we compared WNN with the four classifiers: SVM, LDA, KNN, and ANN. In order to apply the above four classifiers to our six-class classification experiment usefully, the one-vs-one technique was used. The SVM, LDA and KNN were implemented using Scikit-learn [40]. The parameter of SVM classifier *liner* was set as kernel and penalty parameter $C = 10$. For KNN classifier, the number of nearest neighbors was chosen as 50. The parameter of LDA classifier shrinkage was set as *auto* and solver as *lsqr*.

The classification effectiveness of various classifiers is appraised in accuracy, kappa value and precision parameters. The results obtained using the LDA, WNN, SVM, KNN and ANN classifiers are shown in Table 4. The best results of parameters were obtained by WNN as accuracy of 85.24 ± 7.01%, kappa of 0.82 ± 0.08 and precision of 0.85 ± 0.07. Compared with LDA, SVM, ANN and KNN, the accuracy of WNN is improved by about 15 ~ 40%, which proves the effectiveness of the method.

3.4. Performance in different time of imaging segment

The research has shown that the signal motor imaging time is completed within 0.5 s [9,24], thus this paper defaults to set the time of imaging segment of previous experiments to 0.5 s. However, motor imagery duration has a significant impact on event-related desynchronization and event-related synchronization, which will affect the accuracy of motor imagery classification [41,42].

Therefore, we set the time of imaging segment step as 0.1 s between

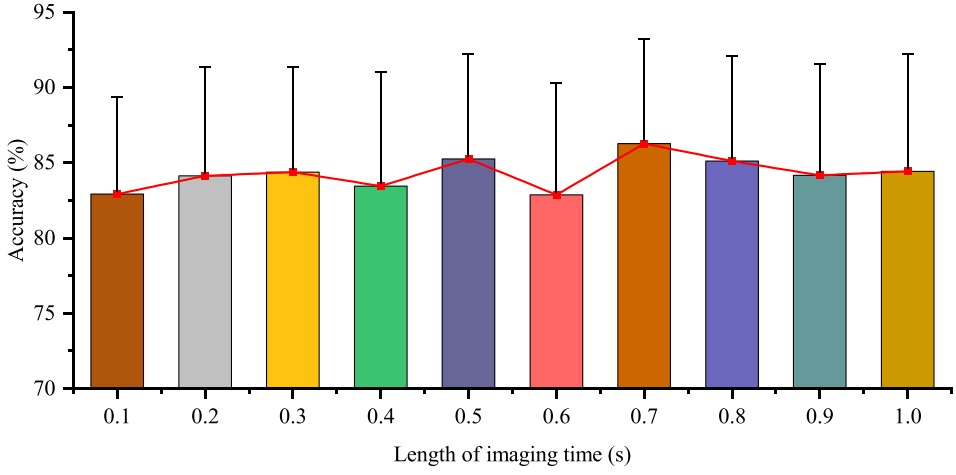

Fig. 6. The average accuracy based on different time of imaging segment.

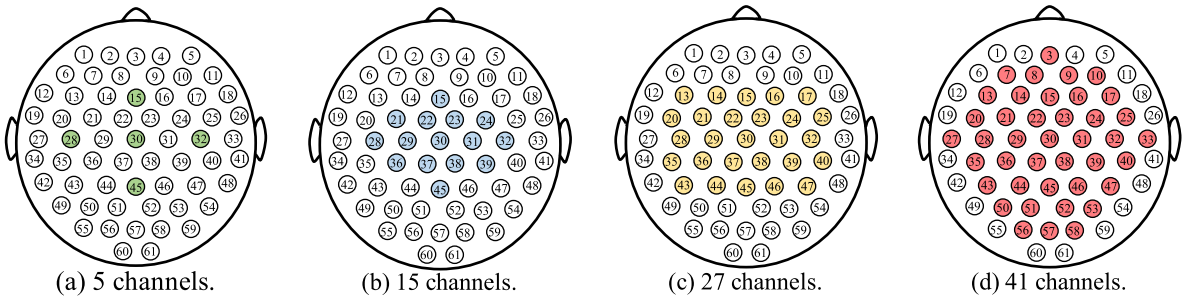

Fig. 7. Four different channel selection patterns.

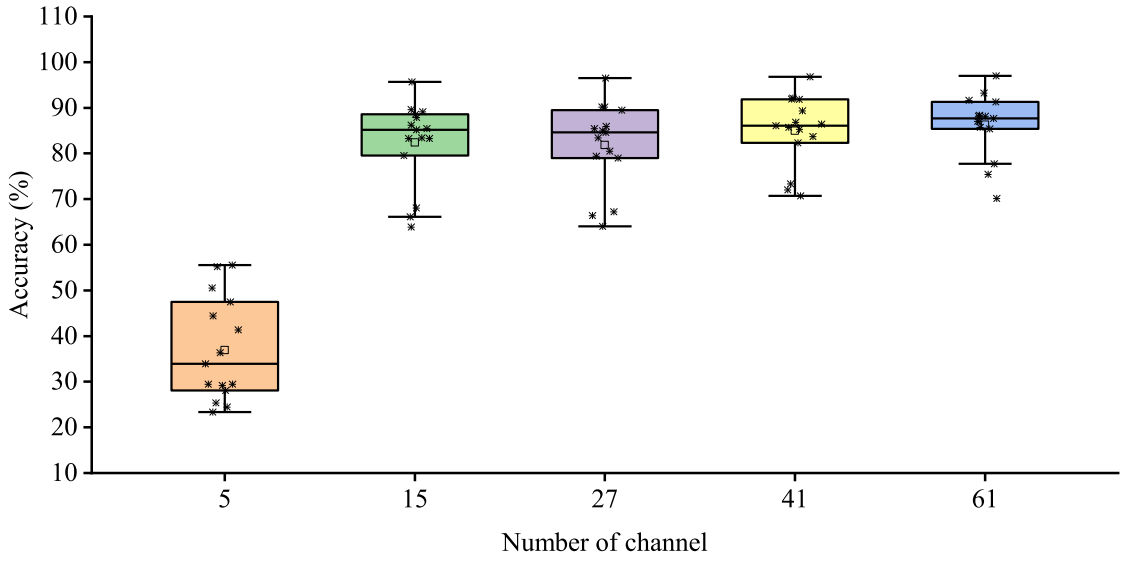

Fig. 8. The average accuracy based on five different channels selected patterns.

0.1 s and 1.0 s for the experiments to determine the optimal the time of imaging segment. The experimental results are shown in Fig. 6. It is interesting to note that when the time of imaging segment set to 0.1 s, the average accuracy rate will be significantly reduced (p-value < 0.05). When the time of imaging segment was set to 0.1 s, the number of signals significantly decreases, which may be the reason for poor signal quality. When the time of imaging segment was set to 1.0 s, the average accuracy rate will be significantly reduced (p-value < 0.05) as well. So, we selected 0.7 s as the time of imaging segment for decoding MI EEG and a better accuracy was obtained.

3.5. Performance of channel selection

The previous studies have shown that the motor cortex of the human brain gets activated during MI tasks [10,43]. Moreover, numerous channels may represent redundant information, high costs, user difficulties, and timely procedure, affecting the classifier's classification effect [9,12,43]. A simple approach to improve the classification efficiency is to choose the channels of EEG. Ofner et al. [10] studied the activation area of the brain when the right arm performs six-class functional movements. The results indicate that mainly premotor areas, primary motor cortex, somatosensory cortex, and posterior parietal cortex convey movement information. Therefore, according to the activation area of the brain, this paper carried out a comparative experiment of choosing channels. The channel selection scheme is shown in Fig. 7.

The results of these experiments are shown in Fig. 8. Sreeja et al. [9] selected 18 channels near the C3 and C4 obtained better performance. However, the difference is that the subjects performed six functional movements of a single upper limb in this study. In previous experiments, the coif1 mother wavelet with a six-level decomposition, Mexican Hat wavelet as a hyperparameter, the time of imaging segment set as 0.7 s, and select 61 channels data achieved the best accuracy of 86.27 $\pm$ 6.98%. When only used 5 channels (FCz, C3, Cz, C4 and CPz) the obtained the accuracy of was 36.95 $\pm$ 11.31%. When the EEG data of 41 channels were used the accuracy reached 84.98 $\pm$ 7.69%. Experiment results show that the 33% channel (61 reduced to 41) decreasing caused only about 2% accuracy decreases, but computing time reduced 20 s. The data of these EEG channels contains the main motor information. The appropriate number of channels can improve classification efficiency.

4. Conclusion

In this work, we performed a WNN classifier model by decoding single upper limb six-class functional movements. We optimized the factors that may affect the performance of decoding and compared them by experiment. The experimental results demonstrate that the accuracy of WNN improved by about 15 $\sim$ 40% compared to other classifiers. In summary, the contributions of this paper are that:

- We implemented a WNN model by decoding single upper limb six-class functional movements from MI EEG signals.
- WPD was used to extract the feature information of MI EEG. The different mother wavelet has a significant effect on the result. Experiment results show that coif1 is most suitable for WPD to

extract feature values. There is a significant difference between coif1 and db4, sym5, sym7 coif3, bior2.4, bior3.7 mother wavelets.

- The appropriate segment time is a critical factor of decoding MI EEG. Experiment results show that the time of imaging segment set as 0.1 s or 1.0 s, the accuracy significantly decreases. Choosing suitable segment time can not only reduce the computational complexity but also ensure accuracy.
- The appropriate number of channels can improve the classification efficiency. Experiment results show that the 33% channel decreasing caused only about 2% accuracy decreases and computing time reduced 20 s.

CRediT authorship contribution statement

Xiaobo Zhou: Investigation, Conceptualization, Methodology, Software, Data curation, Writing - original draft, Visualization. **Renling Zou:** Investigation, Supervision, Writing - review & editing, Funding acquisition, Project administration. **Xiayang Huang:** Investigation, Supervision, Writing - review & editing, Funding acquisition, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the National Natural Science Foundation of China [Grant No. 61803265], China's National Key R&D Programmes [Grant No. 2018YFB1307200], and the Medical-industrial cross-project of the University of Shanghai for Science and Technology [Grant No. 1019308505] for funding this study.

References

[1] S. Chaudhary, S. Taran, V. Bajaj, S. Siuly, A flexible analytic wavelet transform based approach for motor-imagery tasks classification in BCI applications, Comput Methods Programs Biomed 187 (2020), 105325, https://doi.org/10.1016/j.cmpb.2020.105325.
[2] T. Wang, J. Deng, B. He, Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns, Clin Neurophysiol 115 (2004) 2744–2753, https://doi.org/10.1016/j.clinph.2004.06.022.
[3] R. Djemal, A.G. Bazyed, K. Belwafi, S. Gannouni, W. Kaaniche, Three-Class EEG-Based Motor Imagery Classification Using Phase-Space Reconstruction Technique, Brain Sci 6 (2016) 1–19, https://doi.org/10.3390/brainsci6030036.
[4] J. Leon, J.J. Escobar, A. Ortiz, J. Ortega, J. Gonzalez, P. Martin-Smith, J.Q. Gan, M. Damas, Deep learning for EEG-based Motor Imagery classification: Accuracy-cost trade-off, PLoS One 15 (2020), e0234178, https://doi.org/10.1371/journal.pone.0234178.
[5] S.U. Amin, M. Alsulaiman, G. Muhammad, M.A. Bencherif, M.S. Hossain, Multilevel Weighted Feature Fusion Using Convolutional Neural Networks for EEG Motor Imagery Classification, Ieee Access 7 (2019) 18940–18950, https://doi.org/10.1109/Access.2019.2895688.
[6] G. Pfurtscheller, P. Linortner, R. Winkler, G. Korisek, G. Muller-Putz, Discrimination of motor imagery-induced EEG patterns in patients with complete spinal cord injury, Comput Intell Neurosci (2009), 104180, https://doi.org/10.1155/2009/104180.
[7] Y.R. Tabar, U. Halici, A novel deep learning approach for classification of EEG motor imagery signals, J Neural Eng 14 (2017), 016003, https://doi.org/10.1088/1741-2560/14/1/016003.
[8] A. Schlogl, F. Lee, H. Bischof, G. Pfurtscheller, Characterization of four-class motor imagery EEG data for the BCI-competition 2005, J Neural Eng 2 (2005) L14–22, https://doi.org/10.1088/1741-2560/2/4/L02.
[9] S.R. Sreeja, D. Samanta, Classification of multiclass motor imagery EEG signal using sparsity approach, Neurocomputing 368 (2019) 133–145, https://doi.org/10.1016/j.neucom.2019.08.037.
[10] M. Mohseni, V. Shalchyan, M. Jochumsen, I.K. Niazi, Upper limb complex movements decoding from pre-movement EEG signals using wavelet common spatial patterns, Comput Methods Programs Biomed 183 (2020), 105076, https://doi.org/10.1016/j.cmpb.2019.105076.
[11] P. Ofner, A. Schwarz, J. Pereira, G.R. Muller-Putz, Upper limb movements can be decoded from the time-domain of low-frequency EEG, PLoS One 12 (2017), e0182578, https://doi.org/10.1371/journal.pone.0182578.
[12] P. Ofner, A. Schwarz, J. Pereira, D. Wyss, R. Wildburger, G.R. Muller-Putz, Attempted Arm and Hand Movements can be Decoded from Low-Frequency EEG from Persons with Spinal Cord Injury, Sci Rep 9 (2019) 7134, https://doi.org/10.1038/s41598-019-43594-9.
[13] Y. Zhang, Y. Zhang, J. Wang, X. Zheng, Comparison of classification methods on EEG signals based on wavelet packet decomposition, Neural Computing and Applications 26 (2014) 1217–1225, https://doi.org/10.1007/s00521-014-1786-7.
[14] M. Yazici, M. Ulutas, M. Okuyan, A Comprehensive sLORETA Study on the Contribution of Cortical Somatomotor Regions to Motor Imagery, Brain Sci 9 (2019) 14, https://doi.org/10.3390/brainsci9120372.
[15] Z.C. Tang, C. Li, S.Q. Sun, Single-trial EEG classification of motor imagery using deep convolutional neural networks, Optik 130 (2017) 11–18, https://doi.org/10.1016/j.ijleo.2016.10.117.
[16] H. Dose, J.S. Moller, H.K. Iversen, S. Puthusserypady, An end-to-end deep learning approach to MI-EEG signal classification for BCIs, Expert Systems with Applications 114 (2018) 532–542, https://doi.org/10.1016/j.eswa.2018.08.031.
[17] 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, Future Gener Comp Sy 101 (2019) 542–554, https://doi.org/10.1016/j.future.2019.06.027.
[18] Q. Zhang, Using wavelet network in nonparametric estimation, IEEE Trans Neural Netw 8 (1997) 227–236, https://doi.org/10.1109/72.557660.
[19] F. Duan, L.L. Dai, W.N. Chang, Z.Q. Chen, C. Zhu, W. Li, sEMG-Based Identification of Hand Motion Commands Using Wavelet Neural Network Combined With Discrete Wavelet Transform, Ieee T Ind Electron 63 (2016) 1923–1934, https://doi.org/10.1109/Tie.2015.2497212.
[20] H.Y. Wen, G.F. Yang, W.T. Wu, Research on the Real-Time Prediction Model of the Traffic Flow Based on Wavelet Neural Network, Applied Mechanics and Materials 241–244 (2012) 2088–2094, https://doi.org/10.4028/www.scientific.net/AMM.241-244.2088.
[21] BNCI Horizon 2020, Upper limb movement decoding from EEG (001-2017). http://www.bnci-horizon-2020.eu/database/data-sets, 2021 (accessed 6 May, 2021).
[22] A. Delorme, S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, J Neurosci Methods 134 (2004) 9–21, https://doi.org/10.1016/j.jneumeth.2003.10.009.
[23] N. Lu, T. Li, X. Ren, H. Miao, A Deep Learning Scheme for Motor Imagery Classification based on Restricted Boltzmann Machines, IEEE Trans Neural Syst Rehabil Eng 25 (2017) 566–576, https://doi.org/10.1109/TNSRE.2016.2601240.
[24] C. Neuper, R. Scherer, S. Wriessnegger, G. Pfurtscheller, Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain-computer interface, Clin Neurophysiol 120 (2009) 239–247, https://doi.org/10.1016/j.clinph.2008.11.015.
[25] A. Gramfort, M. Luessi, E. Larson, D.A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, M. Hamalainen, MEG and EEG data analysis with MNE-Python, Front Neurosci 7 (2013) 267, https://doi.org/10.3389/fnins.2013.00267.
[26] W. Ting, Y. Guo-zheng, Y. Bang-hua, S. Hong, EEG feature extraction based on wavelet packet decomposition for brain computer interface, Measurement 41 (2008) 618–625, https://doi.org/10.1016/j.measurement.2007.07.007.
[27] Y. Zhang, B. Liu, X. Ji, D. Huang, Classification of EEG Signals Based on Autoregressive Model and Wavelet Packet Decomposition, Neural Processing Letters 45 (2016) 365–378, https://doi.org/10.1007/s11063-016-9530-1.
[28] A. Subasi, M. Ismail Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, Expert Systems with Applications 37 (2010) 8659–8666, https://doi.org/10.1016/j.eswa.2010.06.065.
[29] J. Kevric, A. Subasi, Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system, Biomedical Signal Processing and Control 31 (2017) 398–406, https://doi.org/10.1016/j.bspc.2016.09.007.
[30] A.K. Alexandridis, A.D. Zapranis, Wavelet neural networks: a practical guide, Neural Netw 42 (2013) 1–27, https://doi.org/10.1016/j.neunet.2013.01.008.
[31] B. Yang, X. Zhu, Y. Liu, H. Liu, A single-channel EEG based automatic sleep stage classification method leveraging deep one-dimensional convolutional neural network and hidden Markov model, Biomedical Signal Processing and Control 68 (2021), 102581, https://doi.org/10.1016/j.bspc.2021.102581.
[32] D. Wang, D. Miao, G. Blohm, Multi-class motor imagery EEG decoding for brain-computer interfaces, Front Neurosci 6 (2012) 151, https://doi.org/10.3389/fnins.2012.00151.
[33] https://www.python.org, 2021 (accessed 6 May, 2021).
[34] G. Baldazzi, G. Solinas, J. Del Valle, M. Barbaro, S. Micera, L. Raffo, D. Pani, Systematic analysis of wavelet denoising methods for neural signal processing, J Neural Eng (2020), 066016, https://doi.org/10.1088/1741-2552/abc741.
[35] A. Craik, Y. He, J.L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review, J Neural Eng 16 (2019), 031001, https://doi.org/10.1088/1741-2552/ab0ab5.
[36] W. Puchalsky, G.T. Ribeiro, C.P. da Veiga, R.Z. Freire, L.D. Coelho, Agribusiness time series forecasting using Wavelet neural networks and metaheuristic optimization: An analysis of the soybean sack price and perishable products demand, Int J Prod Econ 203 (2018) 174–189, https://doi.org/10.1016/j.ijpe.2018.06.010.
[37] M. Akraminia, M.J. Mahjoob, M. Tatari, Active noise control using adaptive POLYnominal Gaussian WinOwed wavelet networks, Journal of Vibration and Control 21 (2014) 3020–3033, https://doi.org/10.1177/1077546313520025.
[38] A. Khosla, P. Khandnor, T. Chand, A comparative analysis of signal processing and classification methods for different applications based on EEG signals, Biocybern Biomed Eng 40 (2020) 649–690, https://doi.org/10.1016/j.bbe.2020.02.002.

[39] B. Chemin, G. Huang, D. Mulders, A. Mouraux, EEG time-warping to study non-strictly-periodic EEG signals related to the production of rhythmic movements, J Neurosci Methods 308 (2018) 106–115, https://doi.org/10.1016/j.jneumeth.2018.07.016.
[40] https://scikit-learn.org/stable, 2021 (accessed 6 May, 2021).
[41] C.S. Nam, Y. Jeon, Y.J. Kim, I. Lee, K. Park, Movement imagery-related lateralization of event-related (de)synchronization (ERD/ERS): motor-imagery duration effects, Clin Neurophysiol 122 (2011) 567–577, https://doi.org/10.1016/j.clinph.2010.08.002.
[42] Y. Jeon, C.S. Nam, Y.J. Kim, M.C. Whang, Event-related (De)synchronization (ERD/ERS) during motor imagery tasks: Implications for brain-computer interfaces, Int J Ind Ergonom 41 (2011) 428–436, https://doi.org/10.1016/j.ergon.2011.03.005.
[43] K. Zhu, S. Wang, D.Z. Zheng, M.X. Dai, Study on the effect of different electrode channel combinations of motor imagery EEG signals on classification accuracy, J Eng-Joe 2019 (2019) 8641–8645, https://doi.org/10.1049/joe.2018.9073.