**Noname manuscript No.**
(will be inserted by the editor)

# Multi-Scale Learning for Multimodal Neurophysiological Signals: Gait Pattern Classification as an Example

**Feng Duan · Yizhi Lv · Zhe Sun · Junhua Li**

Received: date / Accepted: date

**Abstract** Neurophysiological signals are manifestations of the underlying brain activity, and they contain an abundance of neural information. The decoding and understanding of these signals is useful to develop robotic exoskeletons, benefitting device-aided motor rehabilitation. To this date, numerous efforts have been carried out to explore the relations between neurophysiological signals and locomotor capacity. Most of these studies focused on a single modality of neurophysiological signal and ignored its multiple modalities. In this study, the modalities from two kinds of biosensors were fused (electroencephalogram (EEG) and electromyogram (EMG)), and a novel deep learning model was proposed (multi-scale learning, MSL) to classify four walking patterns. The EEG and EMG data were collected during a walking experiment, where different walking conditions with and without exoskeleton-aided assistance were implemented (i.e. free-walking and exoskeleton-aided walking at zero, low, and high assistive forces). The performance achieved by the MSL model was compared to that of existing models, and the results show that multimodal MSL achieved the highest performance in terms of classification accuracy (89.33%). Moreover, the comparisons in our study show that an improved classification performance was obtained when a full 62-channel EEG

F. Duan
College of Artificial Intelligence, Nankai University, Jinnan, Tianjin, China
E-mail: duanf@nankai.edu.cn

Y. Lv
College of Artificial Intelligence, Nankai University, Jinnan, Tianjin, China
E-mail: lvyiq@foxmail.com

Z. Sun
Computational Engineering Applications Unit, Head Office for Information Systems and Cybersecurity, RIKEN, Wako-Shi, Japan
E-mail: zhe.sun.vk@riken.jp

J. Li
The School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, UK.
E-mail: junhua.li@essex.ac.uk

F. Duan and Y. Lv contributed equally to this work.
F.Duan, Z.Sun and J.Li are corresponding authors.

setting was used compared to using a subset of 20 channels located on the sensorimotor region. This work contributes to the improvement of neurophysiological signal decoding and promotes the development of rehabilitation technologies as well as exoskeleton-aided applications.

**Keywords** Convolutional Neural Network · Multi-Scale · EEG · EMG · Exoskeleton · Walking

## 1 Introduction

It is well known that the central nervous system initiates movement intention, and that this intention is then delivered to the corresponding limbs and muscles through the peripheral nervous system. The process is achieved by transmitting neurotransmitters between synapses, which generates bioelectricity in the nervous system. While this happens, several kinds of neurophysiological signals can be detected from around the human body, for instance, EEG [1] from the brain and EMG [2] from limbs and muscles. These signals are closely related to the human locomotor system, and investigating them is helpful to explore its underlying mechanisms. To this date, however, little is known about the relations between neurophysiological signals and body movement. These need to be further studied in order to help improve the motor abilities of people such as athletes or those with disabilities, and a better understanding will also pave a solid foundation for the development of rehabilitation devices such as robotic exoskeletons [3].

In the field of human-machine interface (HMI) [4, 5], EEG[6] and EMG signals are the most commonly used neurophysiological signals due to their portability and high temporal resolution. Steady-state visual evoked potentials [7], event-related potentials (ERP) [8, 9] and motor imagery [10] are the three typical paradigms in brain-computer interface (BCI) research. Many credible EEG processing algorithms have been investigated in recent years, such as common spatial pattern [11], principal component analysis (PCA) [12] and independent component analysis (ICA) algorithms [13]. On the other hand, due to its accessibility and its close relationship with muscular activities, EMG is often used in locomotion research. Feature extraction and selection are critical steps in EMG classification and regression. There are three types of features: time-domain features (integral EMG, wavelength, mean absolute value and root mean square, etc.) [14], spectral domain features [15] (mean power, mean frequency and max power spectrum, etc.) and time-frequency domain features (wavelet packet transform [16], short-time Fourier transform (STFT) [17], etc.). Additionally, the EEG signals can be decomposed into frequency bands [18], namely delta (1–3 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz) and gamma (30–100Hz). In a previous study, we found that the alpha, beta and theta bands were involved in consistently increasing EEG-EMG correlations [19].

However, most locomotion-related studies are based on single-modality, which has a limited performance due to having a lower classification accuracy and poor real-time capabilities. Using multimodal neurophysiological signals has recently become a popular research method [20], and our previous study found consistently increasing and decreasing EEG-EMG PSD correlations involving different brain regions [19]. Multimodal neurophysiological signals have also shown superiority in the HMI research field. Zhang et al. [21] developed a novel multimodal human-machine interface system using combinations of electrooculography (EOG), EEG and EMG to generate numerous control instructions. Ahmed Ben Said [22] presented a joint compression and classification approach of EEG and EMG signals using a deep learning approach for data compression and feature learning.

In the field of HMI, analyzing EEG and EMG signals in spectral or temporal scales is a typical practice, and some traditional methods, such as the Fourier transform [17] and feature extraction from the neurophysiological signals, are frequently used for signal processing and recognition. However, most existing methods analyze the neurophysiological signals using a single scale approach, which may miss out on some more hidden information. At present, several studies [23–25] have shown that analysing neurophysiological signals using a multi-scale approach can lead to obtaining richer and more distinguishable information, and it can also help to achieve a superior motion pattern classification.

Classical supervised machine learning algorithms are commonly used methods for motion pattern recognition and neurophysiological signal analysis. These include support vector machines (SVMs) [26], random forests (RFs)[27] and k-Nearest Neighbors (KNN) [28]. Wenyu Li et al [29] proposed a fusion method to explore and verify the feasibility of human-vehicle collaborative driving, and used KNN to classify EEG features and obtain the final control command. Deep learning models, such as convolution neural networks (CNNs) [30] and recurrent neural networks (RNNs) [31] are also promising in the machine learning research field. Compared with the classical machine learning methods, deep learning avoids the feature extraction and selection process [32], which has much higher learning capability. Deep learning algorithms are now widely used in the fields of computer vision[33, 34], natural language processing and time series prediction [35] among many others, as they usually provide a better performance than the classical algorithms. Considering the advantages that deep learning has to offer, it is unsurprising that they are also suitable for the analysis of EEG [36], EMG [37] and other bioelectric signals. Compared with classical machine learning methods, CNN is an end-to-end algorithm that can avoid the feature extraction step, and it therefore has a much higher learning ability. Goh et al. [38] used a deep neural network topology with shared weights to acquire the spatial and spectral representations for walking pattern recognition. Dai et al. [39] proposed a type of CNN with the addition of a variational autoencoder (VAE) to classify MI data and raised the classification performance for the BCI Competition IV dataset 2b to the current state of art. Zhai et al.'s [40] research showed that a CNN-based system consistently achieved a higher absolute performance compared to a SVM classifier with the NinaPro database of EMG.

Taking the above aspects into account, this work proposes a novel deep learning model named multi-scale learning (MSL) used to classify four walking conditions (free-walking and exoskeleton-assisted walking at zero, low and high assistive forces) using multimodal neurophysiological signals (EEG and EMG). Multi-scale refers to temporal, spectral and spatial scales. Compared to other methods, the wavelet transform can keep spectral and temporal multi-scale features at the same time. The wavelet transform is extensively used in EEG [41], EMG and other neurophysiological signal analysis, particularly in research related to CNN or other deep learning models. In this research, the wavelet transform was used to process EEG and EMG signals instead of the Fourier transform to obtain temporal and spectral information at the same time. Generally, the works of this study can be summarized as follows:

– Advancing a deep learning model based on neurophysiological signals for multiple classes of walking pattern classification.
– Employing multimodal neurophysiological signals instead of single signal sources.
– Analyzing signals in temporal, spectral and spatial scales rather than in a single scale.

The rest of the paper is organized as follows: in Section II, the steps of the experiment and the data processing are described, and the architecture of MSL for multimodal neurophysiological signals is introduced. In Section III, the experiment results are presented, and in Section IV, a discussion of the research is provided. Finally, conclusions are discussed in Section V.

## 2 Methodology

### 2.1 Experiment

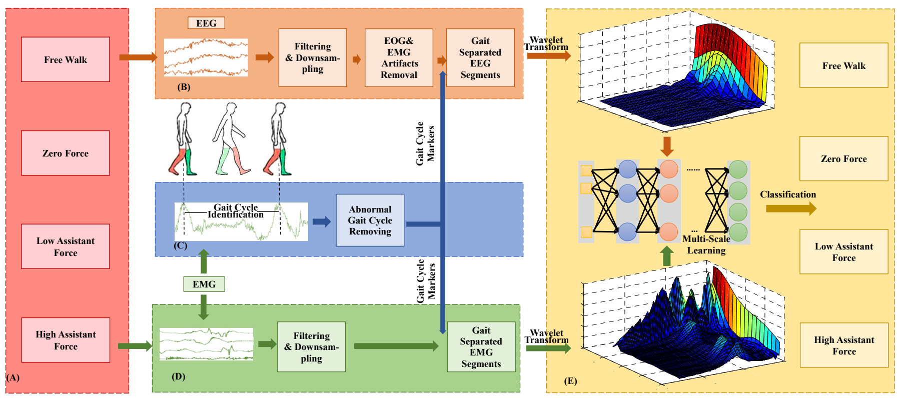


**Fig. 1** Illustration of the main steps in this study. (A) Exoskeleton power-assisted walking experiment in 4 walking conditions (FW, ZF, AFL, AFH), collecting EEG and EMG signals at the same time. (B) EEG signals were downsampled to 250 Hz and then bandpass-filtered (2-400Hz), followed by an ICA-based artifacts removal and segment separation. (C) Identification of gait cycles based on the EMG signals was carried out. The intervals between the periodic peaks corresponded to the gait cycles. Abnormal gait cycles were removed according to their magnitude pattern and length, and valid gait cycle markers used to separate EEG and EMG segments were figured out. (D) EMG signals were bandpass-filtered (2-400Hz) and then downsampled to 250 Hz. They were then split into gait-related segments. (E) EEG and EMG gait-related segments were decomposed through the wavelet transform to later on be classified by the MSL model.

Figure. 1 shows the workings of the main experiment of this study. During the experiment, subjects wore an exoskeleton on their right lower limbs and walked under four different walking conditions in a corridor of approximately 21 meters in length. The experiment was reviewed and approved by the Institutional Review

Board of the National University of Singapore. A total of 30 healthy males were recruited for this study. They possessed a normal or corrected visual acuity without a history of lower limb injury. After receiving an explanation of the procedure, all the subjects signed the consent form.

Subjects were required to perform under four exoskeleton-assisted walking conditions [42] [19], including:

- Free-walking (FW, subjects performed normal walking without an exoskeleton)
- Zero force (ZF, subjects were required to walk with an exoskeleton, but the exoskeleton provided no torque)
- Low assistive force (AFL, subjects were required to walk with an exoskeleton that provided a low assistive torque)
- High assistive force (AFH, subjects were required to walk with an exoskeleton that provided a high assistive torque)

An impedance controller was used to provide the torque assistance. 0.2 Nm/deg were implemented in the knee joint for the low assistive force condition, while 0.4 Nm/deg were implemented for the high assistive force condition.

In this study, EEG signals were recorded using an ANT ASA-Lab system (ANT BV, Netherlands), and 62 electrodes were used altogether (the location of electrodes is shown in Fig.2). Four EMG electrodes were stuck on the surfaces of four major muscles of the lower limb: the semitendinosus (SM), the gastrocnemius lateralis (GL), the tibialis anterior (TA) and the rectus femoris (RF). The sampling frequency of both EEG and EMG signals was of 1000Hz.

### 2.2 Preprocessing of multimodal Neurophysiological Signals

Firstly, gait cycles were identified based on the recorded EMG signals. More specifically, EMG signals were processed through the steps of detrending, centering and filtering (using a bandpass filter with frequency cutoffs of 2-400Hz). Subsequently, the signals were normalized using maximal voluntary contraction (MVC). Then, peak detection was applied to identify the gait cycles. Note that the EMG data used for gait cycle identification was not directly used for data analysis. The EEG data was analysed after carrying out the detrending and centering steps, and to reduce computing costs, the processed EMG signals were further downsampled to 250Hz before being used for classification purposes. A more detailed description of the procedure can be found in [43]). Abnormal gait cycles were removed, and the remaining gait cycles were stored with their corresponding markers indicating their starting and ending points.

When it comes to the EEG signals, the mean value of each channel was firstly subtracted from the amplitude of that channel to remove the mean value of the amplitude. Then, the EEG signals were downsampled (250Hz) and bandpass-filtered (0.5-45Hz). The EOG artifacts were reduced through an adaptive filtering method [44], and EMG artifacts were mitigated using our previous approach [45]. Then, the remaining artifacts of EOG and EMG were removed through ICA.

Additionally, the classification performances of two different EEG channel settings were compared: these settings included the whole brain, consisting of 62 channels, and the sensorimotor region, consisting of 20 channels. The 20-channel setting had been adopted in a previous study [38]. Both settings were used in a

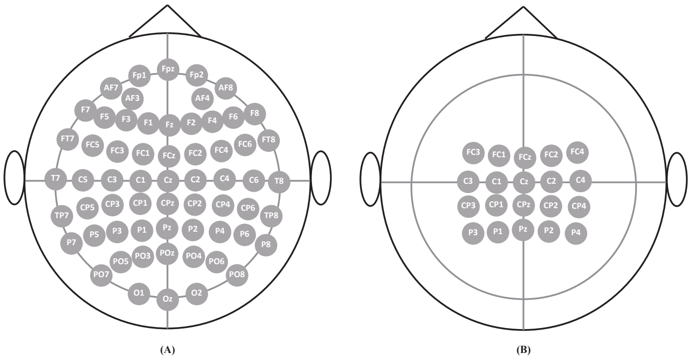


**Fig. 2** The channels included in the two settings. (A) 62 channels covering the whole brain. (B) 20 channels covering the sensorimotor region

single-modal case (EEG) and in a multimodal case (combining EEG and EMG). The channels included in these two settings are shown in Fig. 2.

The preprocessed data was checked, and the gait cycles were identified as being invalid when their magnitude pattern did not follow the standard pattern, or when their duration was extremely long or short. The data of subjects which had less than 30 valid gait cycles in any of the walking conditions was removed from the study. This resulted in data from 23 subjects, who were included in this study

## 2.3 Multi-Scale Learning

In this work, we propose a novel CNN-based deep learning model, named Multi-Scale Learning (MSL), which here was used to classify different walking conditions based on multimodal neurophysiological signals. MSL analysed EEG and EMG signals in the time and frequency domains simultaneously. Compared to the single scale analysis, MSL showed a higher learnability for multimodal neurophysiological signals. Figure. 3 presents the architecture of the MSL model.

### *2.3.1 Wavelet Transform*

The wavelet transform is a practical algorithm used in digital signal processing with many interesting applications. It is considered to be an extension of the Fourier transform, since it works on a multi-scale basis rather than at a single scale. The Morlet wavelet[46] was selected as the wavelet basis in this research.

$$\psi\left(t\right) = Ce^{t^2/2}cos\left(5t\right) \tag{1}$$

Preprocessed EEG and EMG signals would be decomposed into multi-scale tensors (temporal, spectral and spatial scales) through the wavelet transform. The optimal frame and spectral width was investigated in this study, as these determine the temporal and spectral resolution ratio of the decomposed signals. The parameters of the wavelet transform are summarized in TABLE I.

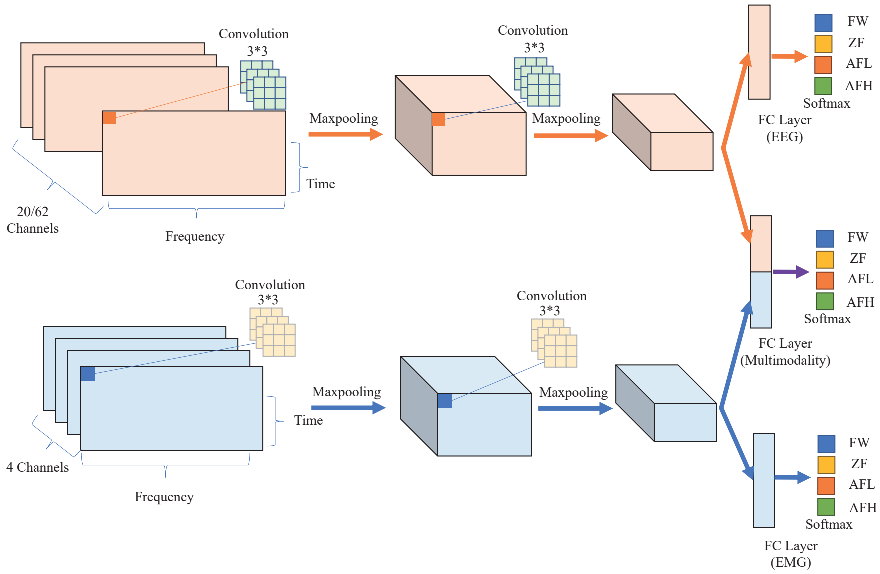


**Fig. 3** The architecture of the MSL. This deep learning model was trained to classify four different walking conditions using EEG and EMG signals. Signals are decomposed into time and frequency domains through the wavelet transform. For EEG and EMG signals, the model is comprised of two different convolution-pooling layers, fully connected layers (FC) and softmax layers. After the convolution and max-pooling layers, EEG and EMG signals are flattened into vectors and are combined into multimodal signals. ReLU activation functions and dropout technology are employed in each hidden layer.

**Table 1** Morlet Wavelet Transform Parameters

| | EEG | EMG |
|---|---|---|
| Frequency Band (Hz) | 1-45 | 2-125 |
| Spectral Width (Hz) | 0.5-2.0 | 0.5-2.0 |
| Frame Width (s) | 0.10-0.19 | 0.10-0.19 |

#### *2.3.2 Convolutional Neural Networks*

CNN is a novel neural network model with several convolution-pooling layer pairs and fully connected layers, designed for tensor analysis of data such as images. In this research, the multi-scale tensors, which had been decomposed from EEG and EMG signals, were used as inputs for the CNN model.

Convolutional layers can extract local features on temporal and spectral scales simultaneously, which is important for signal classification and motion recognition. The convolution can be expressed as follows:

$$x_{i,j}^{l} = \sum_{a=0}^{n-1} \sum_{b=0}^{m-1} w_{a,b}^{l} \times y_{a+i,b+j}^{l-1} \tag{2}$$

The convolutional kernel $w^l$ has size $n \times m$ and is applied across sub-regions of output layer $l-1$, $y^{l-1}$ of size $N \times M$, producing output layer $l$, $x^l$ of size $(N-n+1) \times (M-m+1)$. $w_{a,b}^{l}$ denotes the element of $w^l$ in the $a^{th}$ row and $b^{th}$

**Table 2** Model Parameters

| Parameters | 20 Channels EEG | 62 Channels EEG | EMG |
|---|---|---|---|
| Number of Convolution Filters | 40, 80 | 124, 124 | 8, 16 |
| SGD Learning Rate | 0.01 | 0.01 | 0.01 |
| Momentum Coefficient | 0.9 | 0.9 | 0.9 |
| Epoch | 200 | 200 | 200 |
| Batch Size | 32 | 32 | 32 |
| Parameters | 20 Channels EEG & EMG | 62 Channels EEG & EMG | |
| Number of Convolution Filters | 40, 80 & 8, 16 | 124, 124 & 8, 16 | |
| SGD Learning Rate | 0.01 | 0.01 | 0.01 |
| Momentum Coefficient | 0.9 | 0.9 | 0.9 |
| Epoch | 200 | 200 | 200 |
| Batch Size | 32 | 32 | 32 |

column, and $x^l_{i,j}$ denotes the element of the output layer $x^l$ in the $i^{th}$ row and $j^{th}$ column.

Compared to the sigmoid or tanh activation functions, the Rectified Linear Unit (ReLU) function [47] can solve the vanishing gradient and explosion gradient problems. In this study, the ReLU function was used between convolutional layers and fully connected (FC) layers. ReLU is defined as follows:

$$\sigma(x) = x^+ = max(0, x) \tag{3}$$

The softmax activation function is used at the output layer to turn the output array $z$ into predicted class probabilities:

$$softmax(z_i) = \frac{e^{z_i}}{\sum_k e^{z_k}} \tag{4}$$

where $z_i$is an element in array $z$.

Cross-entropy was selected as the cost function:

$$L_i = -log(f(z_i)) \tag{5}$$

As a classical and efficient optimization technique, the stochastic gradient descent (SGD) algorithm was used to optimize the cost function. The Nesterov accelerated gradient method is an improved form of the classical momentum algorithm which can accelerate convergency efficiently. The expression of SGD with Nesterov momentum is shown below:

$$v_w^{(i)} = \beta v_w^{(i-1)} + (1-\beta)\frac{\partial L}{\partial(W + \beta v_w^{(i-1)})} \tag{6}$$

$$W \leftarrow W - \eta v_w^{(i)} \tag{7}$$

where $v_w^{(i)}$ is velocity; $\beta$ is the momentum parameter, which controls how quickly the velocity can change and how much the local gradient influences long term movement; $W$ represents the weights and bias of the neural network, and finally, $\eta$ is the learning rate.

The computing platform used in this study was an Alienware(R) R17 with i9-8950H, 32GB ram, GTX1080, and the operating system used was Ubuntu 18.04 with Pytoch 1.40. All parameters of the model are listed in TABLE II. All classification results reported in this paper were obtained using five-fold cross-validation.

In summary, the structure of the MSL model can be described as a sequence: EEG&EMG signals - Morlet wavelet transform - Convolution - ReLU - Convolution - ReLU - FC - Dropout - ReLU - FC - Dropout - ReLU - FC - Softmax - Class Prediction.

*2.3.3 Multimodal Neurophysiological Signal Fusion*

In this study, a comparison of the classification performances between two EEG channel settings was also carried out: these two settings included the whole brain, containing 62 channels, and the sensorimotor region, containing 20 channels (the channels included in these two settings are shown in Fig.2). The 20-channel setting had been adopted in a previous study [38]. The EMG signals came from the tibialis anterior, gastrocnemius lateralis, rectus femoris and semitendinosus muscles. Since the EEG and EMG signals have a different number of channels, and also have different output sizes after the wavelet transform (according to TABLE I), these two signals should be combined. Specifically, two kinds of convolutional layer pairs were used to process EEG and EMG signals separately and then flatten and combine them. This process is illustrated in the right-hand part of Fig. 3.

## 3 Results

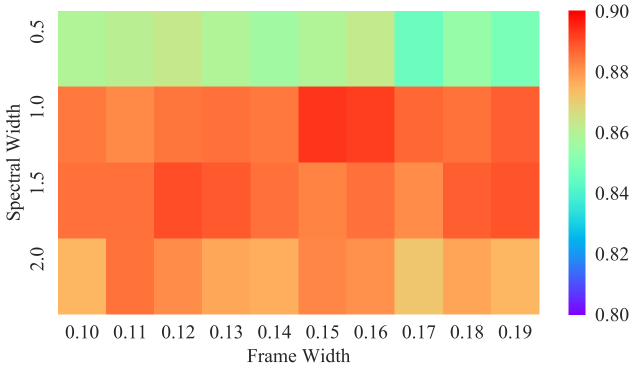


**Fig. 4** 5-fold cross-validation accuracy (%) map of MSL using different frame and spectral widths for 62 EEG channels with a multimodal neurophysiological signal source.

To investigate the optimal coefficients of the wavelet transform, the frame and spectral widths were tested on a range from 0.1 to 0.19 (with a 0.01 interval) for the frame width, and from 0.5 to 2 (with a 0.5 interval) for the spectral width. The classification accuracy of the different parameters based on 62-channel and 20-channel multimodal signals is shown in Fig. 4 and Fig. 5. The parameters which corresponded to the highest classification accuracy of the two kinds of multimodal signals were selected as the optimal coefficients. According to the two-way ANOVA results, the effect of the frame width on the classification accuracy was

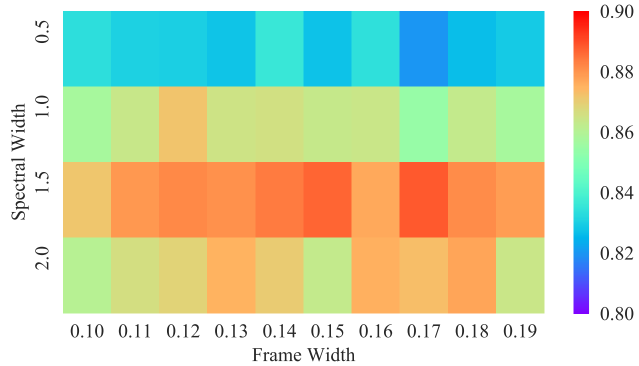


**Fig. 5** 5-fold cross-validation accuracy (%) map of MSL using different frame and spectral widths for 20 EEG channels with a multimodal neurophysiological signal source.

not significant ($p > 0.05$), while the spectral width had a significant effect on the classification accuracy ($p < 10^{-6}$). As shown in Fig. 4 and Fig. 5, for the 62-channel setting, the best parameters correspond to a frame width of 0.15 and a spectral width of 1, while for the 20-channel setting, these correspond to 0.17 and 1.5, respectively. Therefore, a frame width of = 0.17 and a spectral width of = 1.5 were used for the 20 EEG channel multimodal signals, while a frame width of =0.15 and a spectral width of = 1 were used for the 62 EEG channels multimodal signals.

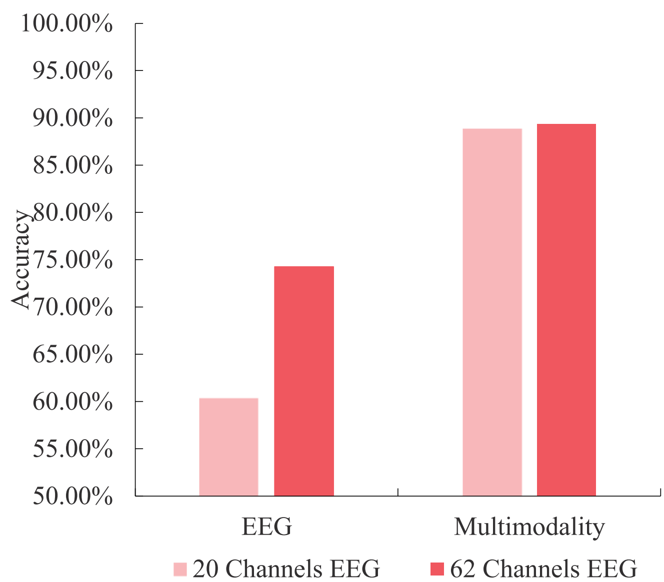


**Fig. 6** The average classification accuracy of the 20-channel EEG and 62-channel EEG settings for a single EEG signal source and for a multimodal signal source.

The classification performance of the 20 channels that cover the sensorimotor region was compared with that of all 62 channels covering the whole brain region, with all the signals being based on the optimal coefficients of the wavelet transform found prior. The results are shown in Fig. 6. The average classification accuracy of the 20 EEG channels was 60.34±10.48% for single EEG signals, and 88.82±7.13% for multimodal signals, while the accuracy of the 62 EEG channels was 74.26±8.62% for single EEG signals and 89.33±6.07% for multimodal signals. The results show that the classification accuracy of the 62-channel EEG setting is always better than the 20-channel EEG setting, regardless of whether the EMG signal is used or not. Therefore, in this study, the EEG channel setting of 62 channels was used.

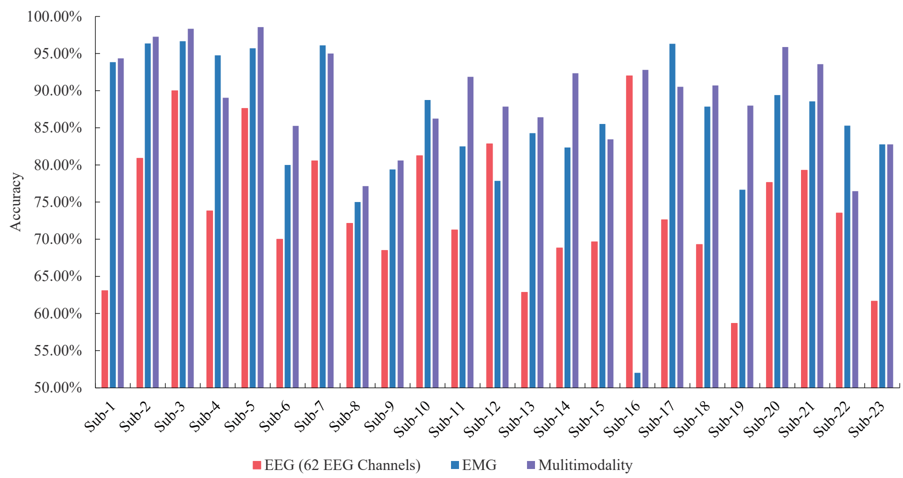


**Fig. 7** The average 5-fold cross-validation classification accuracies of three kinds of signal sources for the 23 subjects.

The classification accuracies of MSL for all subjects based on EEG, EMG and multimodal neurophysiological signals (using the 62-channel EEG setting) are shown in Fig. 7. The accuracies of the three kinds of signals for every subject are shown in Fig. 8. The average classification accuracy was 85.87±9.70% when only using EMG, 74.26±8.62% when only using EEG and 89.33±6.07% when using multimodal neurophysiological signals. According to one-way analysis of variance (ANOVA), the effect of using the whole signals was significant for $F_{(2,66)} = 21.99$, $p < 10^{-6}$. The classification accuracy values of all the subjects between the three signal groups were compared with a one tail t-test. Suppose the population means of the three signal groups were $\mu_1$(EEG), $\mu_2$(EMG), and $\mu_3$(multimodal). The p-values of the three hypotheses were $p < 10^{-3}$ for $\mu_1 < \mu_2$, $p < 10^{-6}$ for $\mu_1 < \mu_3$ and $p < 0.05$ for $\mu_2 < \mu_3$. The input signal is a key factor for MSL, and multimodal neurophysiological signals were found to be more discriminative than a single neurophysiological signal. Figure. 9 shows the confusion matrix of MSL for multimodal neurophysiological signals. It can be observed that the diagonal elements of the matrix show a higher accuracy than the other elements, which is a

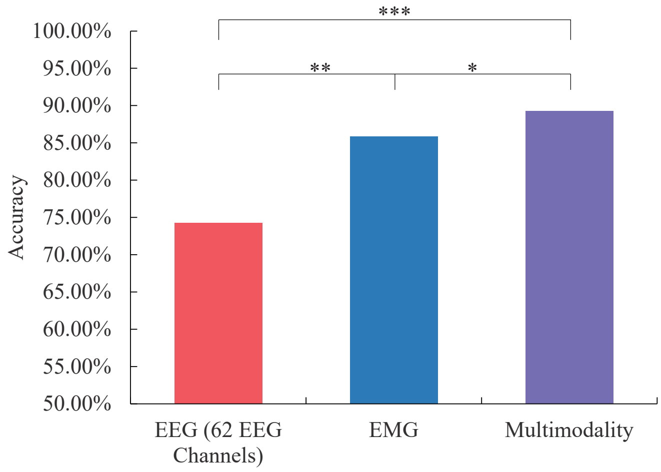


**Fig. 8** The average classification accuracy of three kinds of signal sources. One-tailed paired t-tests of the three signal sources are shown. (* indicates $p < 0.05$, ** indicates $p < 10^{-3}$, *** indicates $p < 10^{-6}$)

clear demonstration of the effectiveness of the proposed algorithm. Compared with the other assistive walking conditions, AFH and AFL are difficult to discriminate between each other, while FW shows a significantly high accuracy. The loss-epoch curve of MSL is shown in Fig. 10

In previous studies, researchers proposed several algorithms to classify these four gait patterns based on the same dataset. The previous methods had lower classification accuracies, such as 73.80% for SVM, 75.90% for random forest-F-score (RF-FS), 76.20% for SVM-PCA, 76.30% for SVM-FS and 77.80% for Spatio–Spectral Representation Learning (SSRL) [38] (shown in Fig. 11). MSL, on the other hand, achieved a much higher classification performance with a 89.33% average accuracy.

In addition, before the paired t-test, all accuracy results from the different signal sources were proven to correspond to the normal distribution through the Kolmogorov–Smirnov test.

## 4 Discussion

In this present study, we proposed a novel deep learning model named Multi-Scale Learning for multimodal neurophysiological signals and used it to classify four exoskeleton assistive gait patterns. We tested the MSL based using EEG, EMG and multimodal neurophysiological signals. Compared with single EEG and single EMG signal sources, the multimodal neurophysiological signal source showed a higher capacity to separate between the different walking conditions. Furthermore, to determine the effect of different EEG settings on the classification accuracy, 20-channel EEG signals covering the sensorimotor region were compared to 62-channel EEG signals covering the whole brain. The results showed that

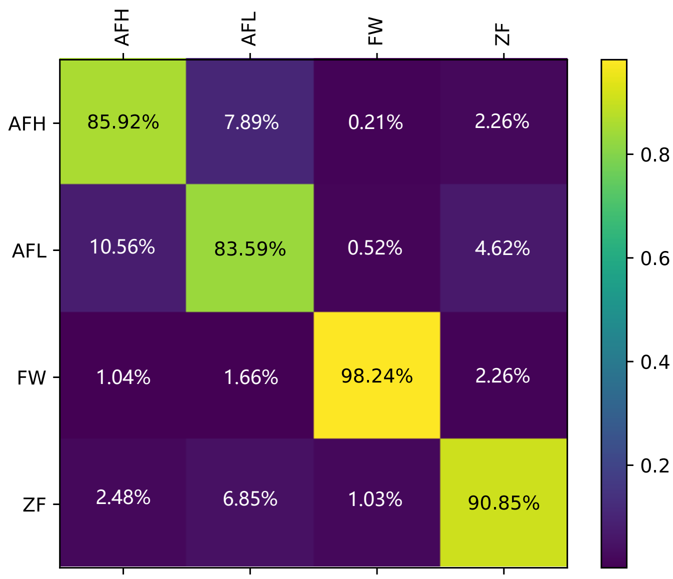


**Fig. 9** The confusion matrix of MSL for multimodal neurophysiological signals.

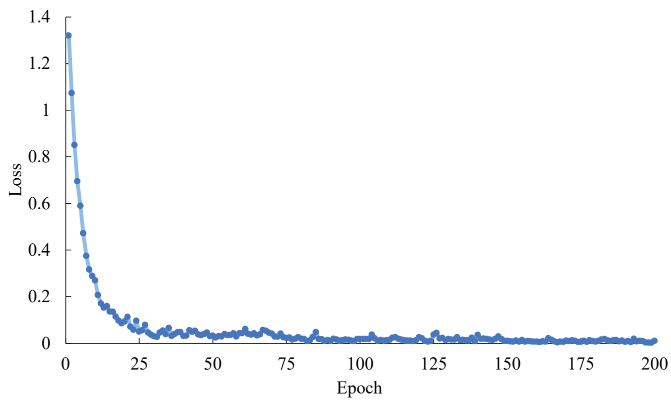


**Fig. 10** The loss-epoch curve of MSL.

62-channel EEG signals can achieve a higher classification accuracy for either single EEG signal sources and for the multimodal neurophysiological signal source. Furthermore, note that the method presented in this work achieved the best classification performance among all the other methods previously used in other studies based on this database. Collectively, these results demonstrate that multi-scale learning for multimodal neurophysiological signals displays its superiority for gait pattern classification.

According to a previous study ([38])

, SSRL only used the EEG signal from the sensorimotor region. The differences in pattern recognition abilities between different brain regions indicate that

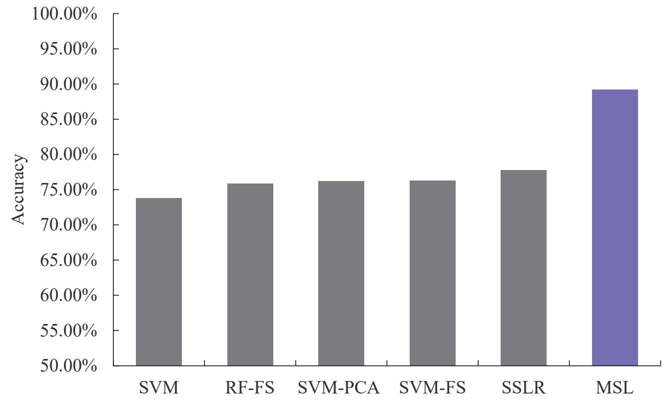


**Fig. 11** The average classification accuracy of MSL for multimodal neurophysiological signals compared to that achieved by former algorithms.

although the sensorimotor region is the most relevant area when it comes to motion, the other regions also contain some important information, and the global brain signal as a whole leads to a higher separability between the types of motion. However, it is worth noticing that the use of fewer channels would bring benefits in practical applications due to its ease and simplicity. Thus, a 20-channel EEG setting may be more efficient in clinical trials. In 7, it can be seen that using multimodal signals did not always result in the highest classification accuracy out of the three kinds of signal sources. This might be due to the larger signal mismatch between EEG and EMG. In a future study, we will take this into consideration when designing a classification model, or will adopt adaptive approaches to mitigate this issue.

The classification accuracy results for different signal sources suggests that a multimodal neurophysiological signal source has a better classification performance than single signal sources, as it contains more motion-related information. It should be noted that the EMG signal is more separable than the EEG signal, which is indicative that EMG may have a higher correlation with body movement than EEG. This inference is also in accordance with common sense.

It is worth noting that in the confusion matrix, the FW condition shows a significantly higher accuracy in contrast to the AFH and AFL conditions. This suggests that neurophysiological signal patterns differ greatly depending on whether an exoskeleton is being worn or not, but that this difference becomes relatively insignificant for varying degrees of power assistance.

The major strength of this study was the novel algorithm which combined EEG and EMG as a multimodal signal source to improve the classification accuracy, and used the wavelet transform and CNN to process signals in temporal and spectral dimensions concurrently. Nevertheless, our study had several limitations. Firstly, only male subjects took part in the experiment; therefore, we should employ both more male and female subjects in future studies. Secondly, the relationship between brain regions and classification accuracy was not explored completely, and should be further researched.

In summary, the results prove that MSL used for multimodal neurophysiological signals has a higher classifiability compared to previously used methods. In future gait-related or other HMI research, using multimodal neurophysiological signals may be a better choice than using a single signal source. Meanwhile, the structure of MSL can provide reference and guidance for further studies.

## 5 Conclusion

Based on this study, it was demonstrated that the multimodality of the combination of EEG and EMG signals provided a superior performance compared to using a single modality of EEG or EMG in the classification of walking patterns. In order to effectively extract the walking-related features of the multimodal signals, a novel deep learning model was proposed, named Multi-Scale Learning (MSL). According to the results evaluated in the 4-class walking pattern classification, the MSL with multimodal signals achieved the highest classification accuracy (89.33%). The proposed model can be transferred to analogous classification scenarios and could benefit the development of a closed-loop motor rehabilitation system. In the future, we plan to adapt the proposed model to perform regression tasks to obtain continuous estimations of variables such as joint angle or moment, and plan to employ data augmentation methods to improve the model's performance.

## References

1. W. J. Ren and M. Han. Classification of eeg signals using hybrid feature extraction and ensemble extreme learning machine. *Neural Process. Lett.*, 50(2):1281–1301, 2019.
2. X. G. Xi, C. Yang, J. H. Shi, Z. Z. Luo, and Y. B. Zhao. Surface electromyography-based daily activity recognition using wavelet coherence coefficient and support vector machine. *Neural Processing Letters*, 50(3):2265–2280, 2019.
3. A. J. Young and D. P. Ferris. State of the art and future directions for lower limb robotic exoskeletons. *Ieee Transactions on Neural Systems and Rehabilitation Engineering*, 25(2):171–182, 2017.
4. Susanna Yu Gordleeva, Sergey A. Lobov, Nikita A. Grigorev, Andrey O. Savosenkov, Maxim O. Shamshin, Maxim V. Lukoyanov, Maxim A. Khoruzhko, and Victor B. Kazantsev. Real-time eeg-emg human-machine interface-based control system for a lower-limb exoskeleton. *IEEE Access*, 8:84070–84081, 2020.
5. Jiaxin Ma, Yu Zhang, Andrzej Cichocki, and Fumitoshi Matsuno. A novel EOG/EEG hybrid human–machine interface adopting eye movements and ERPs: Application to robot control. *IEEE Transactions on Biomedical Engineering*, 62(3):876–889, 2014.
6. Md Rashed-Al-Mahfuz, Mohammad Ali Moni, Shahadat Uddin, Salem A. Alyami, Matthew A. Summers, and Valsamma Eapen. A deep convolutional neural network method to detect seizures and characteristic frequencies using epileptic electroencephalogram (eeg) data. *IEEE J. Transl. Eng. Health Med.*, 9:2000112–2000112, 2021.
7. Setare Amiri, Reza Fazel-Rezai, and Vahid Asadpour. A review of hybrid brain-computer interface systems. *Advances in Human-Computer Interaction*, 2013, 2013.
8. Louise Giraudet, Jean-Paul Imbert, Marie Berenger, Sebastien Tremblay, and Mickael Causse. The neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and EEG/ERP measures. *Behavioural brain research*, 294:246–253, 2015.
9. Leandro da Silva-Sauer, Luis Valero-Aguayo, Alejandro de la Torre-Luque, Ricardo Ron-Angevin, and Sergio Varona-Moya. Concentration on performance with p300-based BCI systems: A matter of interface features. *Applied ergonomics*, 52:325–332, 2016.
10. Hohyun Cho, Minkyu Ahn, Sangtae Ahn, Moonyoung Kwon, and Sung Chan Jun. EEG datasets for motor imagery brain–computer interface. *GigaScience*, 6(7):gix034, 2017.

11. Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4):441–446, 2000.
12. Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics intelligent laboratory systems*, 2(1-3):37–52, 1987.
13. Pierre Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.
14. Angkoon Phinyomark, Pornchai Phukpattaranont, and Chusak Limsakul. Feature reduction and selection for EMG signal classification. *J Expert systems with applications*, 39(8):7420–7431, 2012.
15. Angkoon Phinyomark, Franck Quaine, Sylvie Charbonnier, Christine Serviere, Franck Tarpin-Bernard, and Yann Laurillau. EMG feature evaluation for improving myoelectric pattern recognition robustness. *Expert Systems with applications*, 40(12):4832–4840, 2013.
16. H. Wang, L. Zhang, and L. X. Yao. Application of genetic algorithm based support vector machine in selection of new eeg rhythms for drowsiness detection. *Expert Syst. Appl.*, 171:7, 2021.
17. Daniel Griffin and Jae Lim. Signal estimation from modified short-time Fourier Transform. *IEEE Transactions on Acoustics, Speech, Signal Processing*, 32(2):236–243, 1984.
18. J. J. Newson and T. C. Thiagarajan. Eeg frequency bands in psychiatric disorders: A review of resting state studies. *Front. Hum. Neurosci.*, 12:24, 2019.
19. Junhua Li, Georgios N Dimitrakopoulos, Pavithra Thangavel, Gong Chen, Yu Sun, Zhao Guo, Haoyong Yu, Nitish Thakor, and Anastasios Bezerianos. What are spectral and spatial distributions of EEG-EMG correlations in overground walking? an exploratory study. *IEEE Access*, 7:143935–143946, 2019.
20. A. S. Alharthi, S. U. Yunas, and K. B. Ozanyan. Deep learning for monitoring of human gait: A review. *IEEE Sens. J.*, 19(21):9575–9591, 2019.
21. Jinhua Zhang, Baozeng Wang, Cheng Zhang, Yanqing Xiao, and Michael Yu Wang. An EEG/EMG/EOG-based multimodal human-machine interface to real-time control of a soft robot hand. *Frontiers in neurorobotics*, 13:7, 2019.
22. A. Ben Said, A. Mohamed, T. Elfouly, K. Harras, and Z. J. Wang. Multimodal deep learning approach for joint EEG-EMG data compression and classification. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 2017.
23. Tian Wang, Anastasios Bezerianos, Andrzej Cichocki, and Junhua Li. Multi-kernel capsule network for schizophrenia identification. *IEEE Transactions on Cybernetics*, page DOI: 10.1109/TCYB.2020.3035282, 2020.
24. Pengbo Zhang, Xue Wang, Weihang Zhang, and Junfeng Chen. Learning spatial-spectral-temporal EEG features with recurrent 3d convolutional neural networks for cross-task mental workload assessment. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 27(1):31–42, 2019.
25. Martin Lamos, Radek Marecek, Tomas Slavicek, Michal Mikl, Ivan Rektor, and Jiri Jan. Spatial-temporal-spectral eeg patterns of BOLD functional network connectivity dynamics. *J. Neural Eng.*, 15(3), 2018.
26. Christopher J.C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.
27. Leo Breiman. Random forests. *Machine Learning archive*, 45(1):5–32, 2001.
28. Thomas Cover and P. E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1):21–27, 1967.
29. Wenyu Li, Feng. Duan, Shili Sheng, C. G. Xu, Rensong Liu, Zhiwen Zhang, and Xue Jiang. A human-vehicle collaborative simulated driving system based on hybrid brain-computer interfaces and computer vision. *IEEE Transactions on Cognitive and Developmental Systems*, 10(3):810–822, 2018.
30. H. F. Hu, Z. K. Liao, and X. Xiao. Action recognition using multiple pooling strategies of cnn features. *Neural Process. Lett.*, 50(1):379–396, 2019.
31. Zachary C Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.
32. Rui Zhang, Feiping Nie, Xuelong Li, and Xian Wei. Feature selection with multi-view data: A survey. *Information Fusion*, 50:158–167, 2019.
33. Guoqing Wang, Hu Han, Shiguang Shan, and Xilin Chen. Cross-domain face presentation attack detection via multi-domain disentangled representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

34. Enze Xie, Yuhang Zang, Shuai Shao, Gang Yu, Cong Yao, and Guangyao Li. Scene text detection with supervised pyramid context network. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9038–9045, 2019.
35. Z. Y. Han, J. Zhao, H. Leung, A. Ma, and W. Wang. A review of deep learning models for time series prediction. *IEEE Sens. J.*, 21(6):7833–7848, 2021.
36. Alexander Craik, Yongtian He, and Jose L Contreras-Vidal. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering*, 16(3):031001, 2019.
37. Panagiotis Tsinganos, Bruno Cornelis, Jan Cornelis, Bart Jansen, and Athanassios Skodras. Deep learning in EMG-based gesture recognition. In *Proceedings of the 5th International Conference on Physiological Computing Systems*, pages 107–114, 2018.
38. Sim Kuan Goh, Hussein A. Abbass, Kay Chen Tan, Abdullah Al-Mamun, Nitish Thakor, Anastasios Bezerianos, and Junhua Li. Spatio-spectral representation learning for electroencephalographic gait-pattern classification. *IEEE Trans. Neural Syst. Rehabil. Eng.*, 26(9):1858–1867, 2018.
39. Mengxi Dai, Dezhi Zheng, Rui Na, Shuai Wang, and Shuailei Zhang. EEG classification of motor imagery using a novel deep learning framework. *Sensors*, 19(3):551, 2019.
40. Xiaolong Zhai, Beth Jelfs, Rosa HM Chan, and Chung Tin. Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network. *Frontiers in neuroscience*, 11:379, 2017.
41. Neep Hazarika, Jean Zhu Chen, Ah Chung Tsoi, and Alex Sergejew. Classification of eeg signals using the wavelet transform. *Signal processing*, 59(1):61–72, 1997.
42. Junhua Li, Nitish Thakor, and Anastasios Bezerianos. Unilateral exoskeleton imposes significantly different hemispherical effect in parietooccipital region, but not in other regions. *Sci. Rep.*, 8, 2018.
43. Junhua Li, Gong Chen, Pavithra Thangavel, Haoyong Yu, Nitish Thakor, Anastasios Bezerianos, Yu Sun, and Ieee. A robotic knee exoskeleton for walking assistance and connectivity topology exploration in eeg signal. In *6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, Proceedings of the IEEE RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics, pages 1068–1073, 2016.
44. P. He, G. Wilson, and C. Russell. Removal of ocular artifacts from electro-encephalogram by adaptive filtering. *Med. Biol. Eng. Comput.*, 42(3):407–412, 2004.
45. Junhua Li, Yu Chen, Fumihiko Taya, Julian Lim, Kianfoong Wong, Yu Sun, and Anastasios Bezerianos. A unified canonical correlation analysis-based framework for removing gradient artifact in concurrent EEG/fMRI recording and motion artifact in walking recording from EEG signal. *MED. BIOL. ENG. COMPUT.*, 55(9):1669–1681, 2017.
46. Christopher Torrence and Gilbert P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78, 1998.
47. Alex Krizhevsky, Ilya Sutskever, and E. Geoffrey Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, pages 84–90, 2017.