Classification of Motor Imagery EEG Using Wavelet Envelope Analysis and LSTM Networks

Jie Zhou, Ming Meng, Yunyuan Gao, Yuliang Ma, Qizhong Zhang

Institute of Intelligent Control and Robotics, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China
E-mail: 151060041@hdu.edu.cn

Abstract: Motor imagery (MI) based brain-computer interface (BCI) facilitates a medium to translate the human motion intentions using Motor imagery electroencephalogram (EEG) into control signals. A major challenge in BCI research is the identification of non-stationary brain electrical signals to categorize human motion intentions. We propose a novel method based on wavelet envelope analysis and long-term short-term memory (LSTM) classifier which consider the amplitude modulation characteristics and time series information of MI-EEG to classify EEG signals into multiple classes. First, the Hilbert transform (HT) and discrete wavelet transform (DWT) are combined to extract significant features which contains the underlying information of both amplitude modulation and frequency modulation of the EEG signals. Then, the wavelet envelope features are input into an LSTM classifier with input gates, forget gates, and output gates for classification. Finally, the experiment was conducted on the 2003 BCI competition data set III with 5-fold cross-validation, and experimental results show that the proposed method helps achieve higher classification accuracy.

Key Words: Brain computer interface, Discrete Wavelet Transform, Hilbert Transform, Long Short-Term Memory

1 INTRODUCTION

Brain Computer Interface (BCI) provides users a communication and control channel that do not depend on the brain's normal output channels of peripheral nerves and muscles [1]. The BCI system records brain activity, identifies the user's intention through the signal processing section and then controls the device by sending the appropriate signal to the external device. With BCI, those severely disabled people (such as stroke, spinal cord injury, traumatic brain injury, etc.) can hopefully re-establish their capabilities of environmental control.

Various techniques may be used to measure the brain activity but the BCI system based on electroencephalography (EEG) has high time resolution, relatively low cost and is more convenient for patients to use [2]. Motor Imagery (MI) is one of effective ways to obtain EEG signals. Though MI-based BCI system has made important developments, a low classification accuracy is still an important problem [3]. The EEG signal is essentially a very unstable time series signal, and the time series information helps to improve the Classification Performance of EEG Signals [4]. However, most classifiers do not consider this characteristic.

Due to the complex interconnections between billions of neurons, the recorded EEG signals are non-linear, non-stationary in nature. Therefore, it is very important for the BCI system to extract the features that can identify motor tasks effectively from complex MI-EEG. A large number of methods have been applied in EEG feature extraction, such as spectral analysis [5], Wavelet Transform (WT) [6], Hilbert transform (HHT) [7] and Short time Fourier transform (STFT) [8]. Among them, the wavelet transform (WT) algorithm can not only decompose the signal in multi-scale, but also have the characteristics of multi-resolution, which can effectively obtain the time and frequency information of the signal [9]. Xu et al. proposed a novel method of extracting EEG features based on discrete wavelet transform (DWT) and autoregressive (AR) model, in which the combination features of wavelet coefficient statistics and the sixth-order AR coefficients were used as input vectors for the classifier [10]. Wang et al. proposed a method based on Wavelet Transform and Independent Component Analysis to Extract the Time-Frequency and Spatial Features of EEG Signals [6]. On the other hand, the presence of amplitude modulation in bioelectrical processes is of fundamental nature, since it is a direct reflection of the control, synchronization, regulation and interaction in the nervous and other body systems [11]. Currently, researchers attempt to explore the ability of envelope analysis (EA) in the field of EEG signals processing [12]. Wu et al. used the sliding window ICA algorithm to detect the envelope and apply it to EEG classification [13]. Clerico et al. extract the mutual information of the envelopes of each frequency band in the EEG signal as a feature and achieves up to 20% gain when the proposed features were fused with spectral power and asymmetry index features [14].

Classification is another important step in BCI signal analysis. At present, commonly used motor imagery EEG classifiers, such as support vector machine [15], Linear Discriminant Analysis (LDA) [16], k-nearest neighbor (KNN) [17] and logistic regression (LR) [17], do not take advantage of the time series information of EEG signal in the principle of discrimination. However, the EEG signal is a kind of time series signal, and the time series information

This work is supported by Zhejiang Provincial Natural Science Foundation of China under Grant LY18F030009 and LY17F030021 and National Natural Science Foundation of China under Grant 61372023 and 61671197.

978-1-5386-1243-9/18/$31.00 ©2018 IEEE

helps in enhancing the classification performance of EEG signals as these signals are highly non-stationary in nature [18]. In recent years, Recurrent Neural Network (RNN) is widely used in signal processing of time series and has become an effective model in the fields of speech recognition [19], robot motor recognition [20] and image text recognition [21]. However, the traditional RNN neural network has the problem of gradient explosion or gradient vanish during the training process, which causes the RNN neural network to be unable to use the information of the earlier time. That is to say, the more distant sequence input is, the smaller the influence of the correct change of weights is. So the result of training is often biased towards new information, that is, it doesn't have long memory function. In response to this problem, Hochreiter and Schmidhuber introduced memory units to the traditional RNN to store long time information, and proposed the Long Short Term Memory (LSTM) recurrent neural network [22].

In this paper, a new approach using wavelet envelope analysis and LSTM classifier is presented for MI-EEG classification. In order to capture the basic information of amplitude modulation and frequency modulation, a discrete wavelet transform is performed for EEG signals first. Then, in order to extract wavelet envelope, we select the specific wavelet sub-bands to carry out Hilbert Transform (HT). Finally, a LSTM classifier consisting of input layer, a LSTM hidden layer and output layer is used to recognize the MI-EEG signals.

This paper is organized as follows: In Section 2, the used methods including DWT-based EA and LSTM method are explained, and a description of the classification procedure is given. Section 3 presents the experimental research. The data source, experiment results and discussions are given in this Section. Section 4 concludes the paper.

2 METHODOLOGY

2.1 Discrete Wavelet Transform

Compared with Discrete Fourier transform (DFT) and STFT, the wavelet transform reconstructs the signal into a linear combination of wavelet bases weighted by wavelet coefficients, and has better time-frequency transform characteristics. Another characteristic of wavelet transform is to support high-frequency analysis with short-time scale and low-frequency analysis with long-term scale. Due to its multi-resolution characteristics, it is suitable to extract the features of non-stationary signals, such as the EEG signals.

WT analysis can be classified as two types: Continuous wavelet transform (CWT) and Discrete wavelet transform (DWT). According to the principle, the CWT is defined as follows [23]:

$$CWT(a,b)=\int_{-\infty}^{\infty}x(t)\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right)dt \qquad (1)$$

where ψ, a, b and $x(t)$ are referred as the wavelet function, scaling and shifting parameters and the signal to be processed, respectively. It can be seen that the one-dimensional signal becomes a two-dimensional signal after wavelet transformation.

Accordingly, Discrete Wavelet Transform (DWT) is:

$$DWT(j,k)=\int_{-\infty}^{\infty}x(t)\frac{1}{\sqrt{|2^{j}|}}\psi\left(\frac{t-2^{j}k}{2^{j}}\right)dt \qquad (2)$$

In practical application, the signal is decomposed into finite layers based on the Mallat algorithm [24]. The discrete wavelet transform (DWT) analyzed the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information [25]. In the procedure of multi-resolution decomposition of signal $x(t)$, each stage consisted of two digital filters and two down-samplers. The down sampled outputs of first high-pass and low-pass filters provided the detail D1 and the approximation A1, respectively. The first approximation A1, was further decomposed and this process was continued [26]. Fig. 1 illustrates the structure of wavelet decomposition.

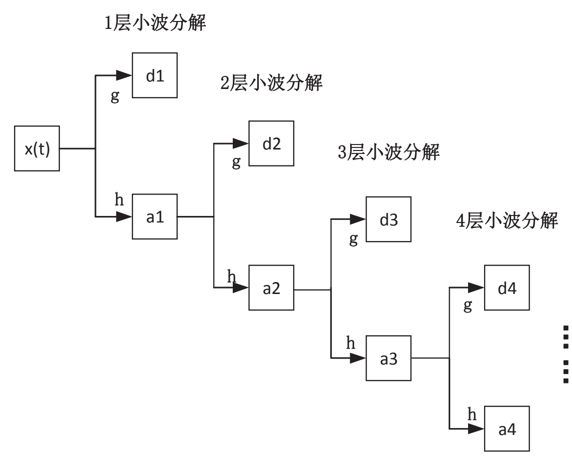

Fig. 1 DWT tree with 4 decomposition levels.

2.2 Hilbert Transform

In order to extract the envelope of each sub-band signal, we carry out the Hilbert transform to the signal of each sub-band. Hilbert transform is a classical method for the extraction of envelope signals. It can effectively extract the envelope of narrow-band carrier signals and is widely used in the envelope extraction of EEG signals. After the Hilbert transform, the amplitude of the signal is constant and the phase changes. The negative frequencies and the positive frequencies are respectively +90 and +90 phase shifts. The signal after Hilbert transform is:

$$\hat{x}(t)=x(t)*\frac{1}{\pi t}=\frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(\tau)}{t-\tau}d\tau \qquad (3)$$

So get the analytical signal:

$$g(t)=x(t)+j\hat{x}(t) \qquad (4)$$

And the amplitude of $g(t)$ is the envelope of the original signal:

$$A(t)=\sqrt{x^2(t)+\hat{x}^2(t)} \quad (5)$$

2.3 Long Short Term Memory Network

In the traditional neural network, the gradient signal multiplies the weight matrix between the neurons of recurrent hidden layer many times (as many as the number of time steps) in the back-propagation process. Therefore, the weight of the transformation matrix is important for the gradient signal.

If the weights in the matrix are small, this can lead to what is called vanishing gradient. Conversely, if the weights in the matrix are large, this can lead to what is called exploding gradients. These issues are the main motivation behind the LSTM model which introduces a new structure called a memory cell. As you can see from the Fig. 2, LSTM specially designed memory unit to preserve historical information, while the update and utilization of historical information are controlled by 3 gates: input gate, forget gate, output gate [27-28].

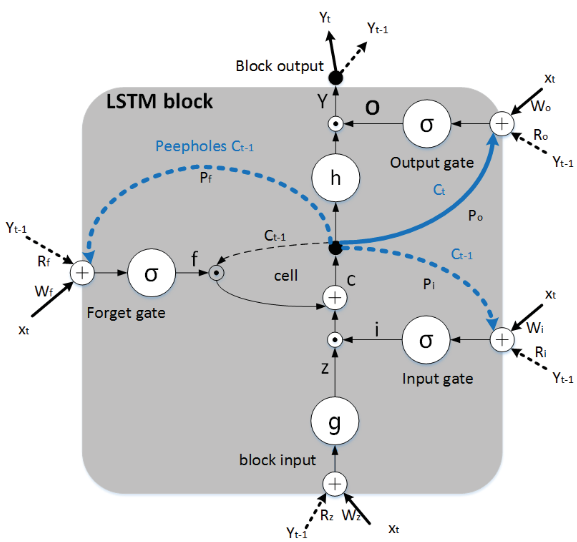

Fig. 2 Illustration of an LSTM memory cell.

For the version of LSTM used in this paper, LSTM is implemented by the following composite function:

$$z_t = g(W_z x_t + R_z y_{t-1} + b_z) \quad (6)$$

$$i_t = \sigma(W_i x_t + R_i y_{t-1} + P_i \odot c_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma(W_f x_t + R_f y_{t-1} + P_f \odot c_{t-1} + b_f) \quad (8)$$

$$c_t = z_t \odot i_t + c_{t-1} \odot f_t \quad (9)$$

$$o_t = \sigma(W_o x_t + R_o y_{t-1} + P_o \odot c^t + b_o) \quad (10)$$

$$y_t = h(c_t) \odot o_t \quad (11)$$

where σ, g and h are pointwise nonlinear activation functions. The logistic sigmoid ($\sigma(x)=(1/1+e^{-x})$) is used as the gate activation function and the hyperbolic tangent ($g(x)=h(x)=\tanh(x)$) is usually used as the block input and output activation function. Pointwise multiplication of two vectors is denoted by $\odot$ [28].

2.4 Description of the Classification Procedure

In this paper, discrete wavelet transform and Hilbert transform are used to extract the wavelet envelope features of EEG signals, and the LSTM recurrent neural network was used as a classifier to recognize the MI-EEG signals. The block diagram of the algorithm is shown in Fig. 3.

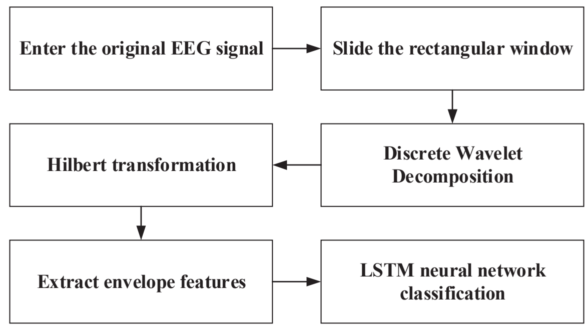

Fig. 3 Block diagram of the proposed method.

A single sub-band feature sequence contains too little frequency domain information. In this paper, a combination of multiple sub-band feature sequences is used to provide different frequency-domain information. However, since the time series of each sub-band envelope is inconsistent, multiple sub-band time series cannot be input to the LSTM classifier at the same time and vectorize the combination of multiple sub-band time series. Thus, a sliding rectangular window is used to obtain a fixed-length time series, and the statistical feature combination of the multiple sub-bands at each window is computed as the feature of each window. The procedure of feature extraction for the motor imagery task was as follows:

(1) Slide over the original EEG using a rectangular window of overlapping windows. Each window contains a single EEG segment.

(2) The DWT is used to decompose each EEG segment inside the sliding window into different frequency bands.

(3) The HT is conducted on the decomposed sub-bands of DWT, then sub-band envelopes are obtained.

(4) Statistical features are calculated to characterize the envelope spectrum of EEG signals. The following statistical features were used to characterize the time-frequency distribution of the envelope spectrum:

1) Mean of the envelope spectrum in each sub-band.

2) Energy of the envelope spectrum in each sub-band.

3) Standard deviation of the envelope spectrum in each sub-band.

4) The max value of the envelope spectrum in each sub-band.

(5) The statistical features of different frequency bands under the same time window are combined into a feature vector, and the feature vectors are combined into a feature time series according to a time window series.

The classification process can be divided into two parts: the first part is to optimize the classifier based on training data set. The wavelet envelope feature of the training set is input into the LSTM classifier, and the whole network is trained by supervised learning. In the course of training, a full back propagation algorithm (Full BPTT) algorithm is used to transfer the error from the last layer to the next layer to minimize the error and update the network parameter set. The second part is the analysis of test data set. Input the wavelet envelope feature of the test set into the trained LSTM classifier and evaluate the performance of the LSTM classifier based on the classification results.

3 EXPERIMENTAL RESEARCH

3.1 Experimental Data

The data set III of BCI Competition 2003 provided by Dr.Gert Pfurtscheller from Graz University of Technology were used to evaluate the classification performance of our proposed method [29]. The EEG signals in this data set were measured using channels C3, Cz, and C4 for one normal female subject who was required to control a feedback bar by the imagination of left or right hand movements. The data set includes 140 training trails and 140 test trails, and the specific experimental paradigm is shown in the Fig. 4 below. Each trial lasted 9s. At t= 2s, an acoustic stimulus indicates the beginning of the trial, and a cross “+” was displayed for 1s; At t= 3s,a left or right arrow was displayed as a cue, and the subject was, at the same time, asked to do motor imagery along the direction of the cue; Then there was a period of rest, and the subjects relaxed and prepared for the next set of experiments.

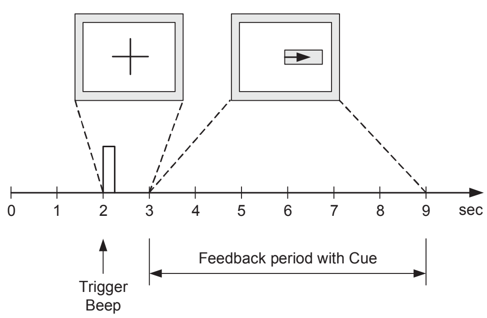

Fig. 4 Time scheme of the 2003 experimental paradigm.

3.2 Feature Extraction

In this study, DWT is employed to decompose each EEG segment inside the sliding window for channels C3, C4 and Cz. The size of rectangle window is 1s and half of the adjacent window overlap. Besides, only the data between t = 3.5s and 7.5 s were used, so a time series with a length of 7 is obtained through a sliding window. Each EEG segment is decomposed into 5 levels by Daubechies 4-tap wavelet. Table 1 represents the frequency distribution of the wavelet envelope coefficients of the EEG signals at 128 Hz. In this paper, four sub-bands are selected for Hilbert transform. Fig. 5 shows the sub-band wavelet coefficients we selected for a segment of EEG.

Table 1: Frequency band of EEG signal using fifth level decomposition.

Decomposition levels	Sub-bands	Frequency range (Hz)
1	D1	32-64
2	D2	16-32
3	D3	8-16
4	D4	4-8
5	D5	2-4
5	A5	0-2

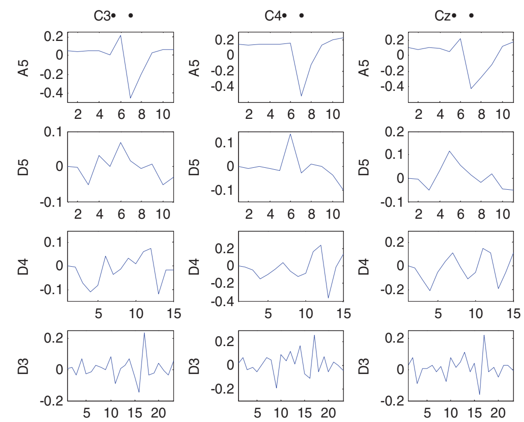

Fig.5 The wavelet decomposition of the C3, C4 and Cz.

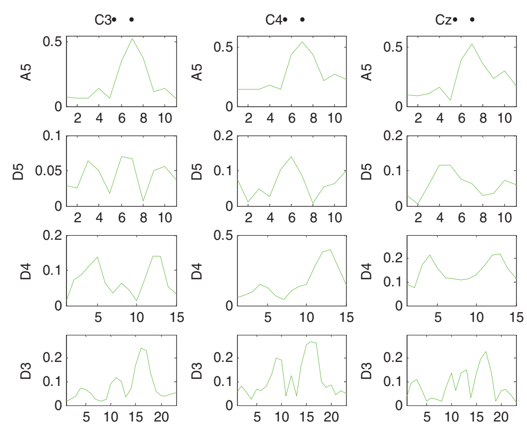

Fig. 6 The wavelets envelope of the decomposed sub-bands for a segment of EEG.

HT is performed on the DWT decomposition sub-band to obtain the sub-band envelope, as shown in Fig. 6. Combining all the wavelet envelope statistical features extracted from the selected sub-bands of C3, C4 and Cz, a 1×48 dimensional feature vector is obtained as a feature of each EEG time series. Finally, the formed 7×48 dimensional feature time series is used as the input of LSTM neural network.

3.3 Classification Based on a LSTM Based Recurrent Network

The number of network layers and LSTM units in the LSTM recurrent neural network structure have a great influence on the performance of the obtained features. LSTM networks with a small number of LSTM cells may not be enough to simulate complex functions. On the other hand, neural networks with excessive LSTM cells may lead to over-fitting the training set and lose its ability to promote, which is the main desired characteristic of the neural network [30]. At present, there is no general method to select the appropriate number of hidden layers. Generally, the optimal neural network architecture is found by trial and error. According to this, this paper discusses the influence of the number of hidden layer units with one hidden layer on the experimental results firstly. The experiment results are shown in Table 2. As can be seen from the table, for a MI-EEG signal, a network with one hidden layer consisting of 25 LSTM cells results in higher classification accuracy. In addition to one hidden layered network architecture, we further experimented with larger networks with two or three hidden layers, but we did not observe significant improvements in training and test results compared to one hidden layer. The number of output was 1 and samples with target outputs imagining left-hand movement, imagining right-hand movement were given the binary target values of 0, 1, respectively.

Table 2: Classification accuracy with different number of neurons in one hidden layer

The neurons number of hidden layer	**Classification accuracy**
23	88.57%
24	89.29%
25	91.43%
26	90.08%

Besides, the classification results of the LSTM recurrent neural network are also affected by the size of the training set and test set. Cross-validation is used to evaluate the predictive performance of the model, especially the performance of the training model for new data. It can reduce over-fitting to some extent, and can also obtain as much effective information as possible from limited data. Therefore, this paper uses 5-fold cross validation to calculate the classification accuracy.

3.4 Comparison with Other Methods

The feature extraction method of EEG signals based on DWT and HT is used to extract the 7×48 dimensional tensor features and send them to LSTM classifier. The results are as follows: the classification accuracy is 91.43%.

The results of the algorithm classification in this paper are compared with the BCI2003 competition results. The highest classification accuracy of the top three in the BCI competition 2003 were 89.29%, 84.29% and 82.86%, respectively [6]. By comparison, the classification accuracy obtained by this method was also improved compared with the best results in the competition.

In order to prove the effectiveness of the LSTM classifier proposed, the proposed algorithm is compared with the traditional RNN network, Gated Recurrent Unit (GRU, the variant of LSTM), KNN and Logistic Regression classification. The classification results of the five classifiers are shown in Fig. 7 below. The RNN and GRU neural network models have the same single hidden layer structure as the LSTM classifier and the KNN classifier has a K value of 5 and the penalty of LR classifier is L2.

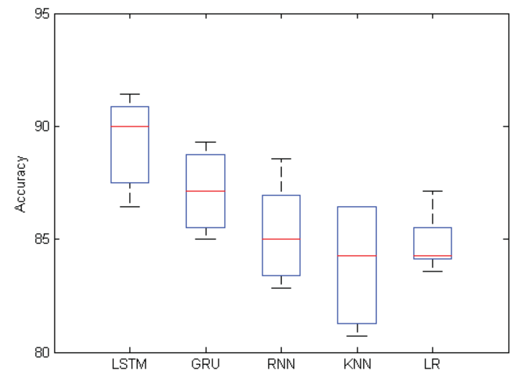

Fig. 7 Classification accuracy of different classifier.

The upper and lower lines extending vertically from the box (whiskers) in Fig. 7 represent the maximum value and the minimum value, respectively. The blue rectangle indicates the range between the 25th and 75th percentiles of the data, and the red horizontal line indicates the median of the data. As can be seen from the Fig. 7, the bit lines and maximum values of the RNN, LSTM, and GRU classifiers are higher than the KNN and LR classifiers. It means that the time series information has a positive influence on the classification of MI-EEG. In addition, the classification accuracy of LSTM unit is superior to the traditional RNN and GRU unit in RNN neural network, and it has a better network generalization ability.

In order to further verify the effectiveness of the proposed MI-EEG recognition method, we made a comparative study, mainly including the recently proposed DWT feature extraction method based on the same data set, the SVM classifier and their best classification accuracy, as shown in Table 3.

Table 3: Studies with Regard to the MI Classification Using DWT-based Feature Extraction and SVM Classifier.

Methods	**Feature extraction**	**Classifier**	**Classification accuracy**
[31]	DWT-RWE	SVM	85.70%
[26]	DWT	FSVM	87.86%
[6]	DWT-ICA	GA-SVM	90.71%
This paper	DWT-EA	LSTM	91.43%

From Table 3, we can see that the proposed method is better than the others and it can effectively improve the classification accuracy of MI-EEG. Therefore, we can come to the conclusion that the amplitude modulation characteristics and time series information is a good way to improve the classification performance in MI task.

4 CONCLUSIONS

According to the presence of amplitude modulation in bioelectrical processes and time series information of MI-EEG signals, this paper proposed a method using Wavelet Envelope Analysis and LSTM network for motor imagery EEG classification. The method decomposes the MI-EEG signals using DWT and then extracts the envelope wavelet of decomposition sub-bands by HT. In order to compose multiple sub-band signals into feature time series, sliding rectangular windows are used to generate the time series and the statistical features of multiple sub-bands are combined as the eigenvectors of each time window. Then, the wavelet envelop feature time series are fed into the LSTM classifier. Due to the proposed method considers the time-frequency and envelope information of motor imagery EEG signals, as well as its time series information, it can obtain high classification accuracy in motor imagine classification experiment with LTSM classifier.

REFERENCES

[1] J.R. Wolpaw, N. Birbaumer et al. Brain-computer interface technology: a review of the first international meeting. IEEE Transactions on Rehabilitation Engineering, Vol.8, No.2, 164-173, 2000.

[2] L. F. Nicolas-Alonso, J. Gomez-Gil, Brain Computer Interfaces, a Review, Sensors, Vol.12, No.2, 1211-1279, 2012.

[3] S. Marchesotti, M. Bassolino, A. Serino, H. Bleuler, O. Blanke, Quantifying the role of motor imagery in brain-machine interfaces, Scientific reports, 1-6, 2016.

[4] S. Pattnaik, M. Dash, S. K. Sabut, DWT-based feature extraction and classification for motor imaginary EEG signals, International Conference on Systems in Medicine and Biology (ICSMB), 186-201, 2016.

[5] B. Yang , A. Zhang, Power spectral entropy analysis of EEG signal based-on BCI, Proceedings of the 32nd Chinese Control Conference, 4513-4516, 2013.

[6] Y. Wang, X. Li, H. Li, C. Shao, L. Ying, S. Wu, Feature extraction of motor imagery electroencephalography based on time-frequency-space domains, Journal of Biomedical Engineering, Vol.31, No.5, 955-961, 2014.

[7] X. Ruan, K. Xue, M. Li, Feature extraction of SSVEP-based brain-computer interface with ICA and HHT method, Proceeding of the 11th World Congress on Intelligent Control and Automation, 2418-2423, 2014.

[8] Y. R. Tabar, U. Halici, A novel deep learning approach for classification of EEG motor imagery signals, Journal of Neural Engineering, Vol.14, No.1, 016003, 2017.

[9] T. T. Day, Applying a Locally Linear Embedding Algorithm for Feature Extraction and Visualization of MI-EEG, Journal of Sensors, Vol.2016, No.2, 1-9, 2016.

[10] B. G. Xu, A. G. Song, S. M. Fei, Feature extraction and classification of EEG in online brain-computer interface, Acta Electronica Sinica, Vol.39, No.5, 1025-1030, 2011.

[11] A. T. Bondar, A. I. Fedotchev, Concerning the amplitude modulation of the human EEG, Human Physiology, Vol.26, No.4, 393-399, 2000.

[12] M. Li, W. Chen, T. Zhang, Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble, Biomedical Signal Processing & Control, Vol.31, No.1, 357-365, 2017.

[13] X. Guo, X. Wu, X. Gong, L. Zhang, Envelope detection based on online ICA algorithm and its application to motor imagery classification, 6th International IEEE/EMBS Conference on Neural Engineering (NER), 1058-1061, 2013.

[14] A. Clerico, R. Gupta, T. H. Falk, Mutual information between inter-hemispheric EEG spectro-temporal patterns: A new feature for automated affect recognition, 7th International IEEE/EMBS Conference on Neural Engineering (NER), 914-917, 2015.

[15] Y. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki, Optimizing spatial patterns with sparse filter bands for motor-imagery based brain-computer interface, Journal of Neuroscience Methods, Vol.255, No.1, 85-91, 2015.

[16] İ. Yılmaz, S. Demir, T. Taşdizen, M. Çetin, Semi-supervised adaptation of motor imagery based BCI systems, 23nd Signal Processing and Communications Applications Conference (SIU), 1841-1844, 2015.

[17] M. Z. Ilyas, P. Saad, M. I. Ahmad, A. R. I. Ghani, Classification of EEG signals for brain-computer interface applications: Performance comparison, International Conference on Robotics, Automation and Sciences (ICORAS), 1-4, 2016.

[18] S.Pattnaik, M. Dash, SK. Sabut, DWT-based feature extraction and classification for motor imaginary EEG signals, International Conference on Systems in Medicine and Biology (ICSMB), 186-201, 2016.

[19] R. Zazo, A. Lozanodiez, J. Gonzalezdominguez, D. T. Toledano, J. Gonzalezrodriguez, Language Identification in Short Utterances Using Long Short-Term Memory (LSTM) Recurrent Neural Networks, Plos One, Vol.11, No.1, e0146917, 2016.

[20] D. N. T. How, C. K. Loo, K. S. M. Sahari, Behavior recognition for humanoid robots using long short-term memory, International journal of advanced robotic systems, Vol.13, No.6, 1-14, 2016.

[21] C. Zhai, Z. Chen, J. Li, B. Xu, Chinese Image Text Recognition with BLSTM-CTC: A Segmentation-Free Method, 7th Chinese Conference on Pattern Recognition (CCPR), Vol.663, No.2, 525-536, 2016.

[22] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation, Vol.9, No.8, 1735-1780, 1997.

[23] H. Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, Expert Systems with Applications, Vol.36, No.2, 2027-2036, 2009.

[24] S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.11, No.7, 674-693, 1989.

[25] A. Subasi, EEG signal classification using wavelet feature extraction and a mixture of expert model, Expert Systems with Applications, Vol.32, No.4, 1084-1093, 2007.

[26] Q. Xu, H. Zhou, Y. Wang, J. Huang, Fuzzy support vector machine for classification of EEG signals using wavelet-based features, Medical Engineering and Physics, Vol.31, No.7, 858-865, 2009.

[27] F.A. Gers, J. Schmidhuber, F. Cummins, Learning to Forget: Continual Prediction with LSTM, Neural Computation, Vol.12, Bo.10, 2451-2471, 2000.

[28] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, J. Schmidhuber, LSTM: A Search Space Odyssey, IEEE Transactions on Neural Networks and Learning Systems, Vol.28, No.10, 2222-2232, 2017.

[29] B. Blankertz, K. R. Muller, G. Curio, T. M. Vaughan, G. Schalk, J. R. Wolpaw et al, The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials, IEEE Transactions on Biomedical Engineering, Vol.51, No.6, 1044-1051, 2004.

[30] N. F. Güler, E. D. Übeyli, I. Güler, Recurrent neural networks employing Lyapunov exponents for EEG signals classification, Expert systems with applications, Vol.29, No.3, 506-514, 2005.

[31] H. B. Zhao, H. Wang, C. S. Li, Brain-computer interface design based on relative wavelet energy, Journal of Northeastern University: Natural Science, Vol.30, No.8, 1103-1106, 2009.