Classification of Imagined Speech EEG Signals with DWT and SVM

ZHANG Lingwei, ZHOU Zhengdong*, XU Yunfei, JI Wentao,
WANG Jiawen, SONG Zefeng

(*State Key Laboratory of Mechanics and Control of Mechanical Structures*, *Nanjing University of Aeronautics and Astronautics*, *Nanjing* 210016, *China*)

Abstract: With the development of human–computer interaction technology, brain–computer interface (BCI) has been widely used in medical, entertainment, military, and other fields. Imagined speech is the latest paradigm of BCI and represents the mental process of imagining a word without making a sound or making clear facial movements. Imagined speech allows patients with physical disabilities to communicate with the outside world and use smart devices through imagination. Imagined speech can meet the needs of more complex manipulative tasks considering its more intuitive features. This study proposes a classification method of imagined speech Electroencephalogram (EEG) signals with discrete wavelet transform (DWT) and support vector machine (SVM). An open dataset that consists of 15 subjects imagining speaking six different words, namely, up, down, left, right, backward, and forward, is used. The objective is to improve the classification accuracy of imagined speech BCI system. The features of EEG signals are first extracted by DWT, and the imagined words are classified by SVM with the above features. Experimental results show that the proposed method achieves an average accuracy of 61.69%, which is better than those of existing methods for classifying imagined speech tasks.

Keywords: Brain-computer Interface (BCI), EEG, Imagined Speech, Discrete Wavelet Transform (DWT), Signal Processing, Support Vector Machine (SVM)

1 Introduction

Since the 1980s, brain–computer interface (BCI) has been a research topic in cognitive neuroscience. In its first international conference, BCI is defined as “a communication system that does not rely on normal output pathways composed of peripheral nerves and muscles” [1–2]. BCI realizes the connection between the human brain and external devices and the brain controls these devices by collecting and processing brain activity signals.

Common brain activity signals include electrocorticography (ECoG), functional near-infrared spectroscopy (fNIRS), functional magnetic resonance imaging (fMRI), and electroencephalography (EEG). EEG equipment is inexpensive, lightweight, and non-invasive and has a higher temporal resolution than other technologies. As a result, EEG technology is the most widely used BCI in many fields, such as neuroscience, psychology, cognitive science, and clinical diagnosis.

Imagined speech is the latest paradigm of BCI; it can extract the signals of brain activity when the user imagines making speech (but does not produce voice and facial organ activity) and then analyze and classify data to realize a certain number of output instructions.

Compared with other types of EEG signals, such as motor imagination, steady-state visual evoked potential (SSVEP), and P300, imagined speech is the most intuitive and is close to a form of verbal communication in daily life and thus can be used more naturally and comfortably.

The BCI system is composed of signal preprocessing, feature extraction, and feature classification; among which, the latter two are the crucial steps. Various machine learning approaches have been used to classify imagined speech EEG signals; these approaches include support vector machine (SVM) [3], linear discriminant analysis (LDA) [4], k-nearest neighbor [5], and random forests [6]. SVM and LDA are the most often utilized classification methods [7]. In [8], the authors used wavelet transform to extract the features of silently reading "yes" and "no" by LDA. In [9], SVM is used to classify the binary tasks of imagined speech of "yes" and "no."

In recent years, various signal decomposition techniques, such as eigenvalue decomposition [10], Fourier decomposition method [11], and discrete wavelet transform (DWT) [12], have been explored for extracting features. DWT is widely used for nonlinear and non-stationary signal decomposition and analysis and is very promising in analyzing EEG signals [13]. DWT has been used to classify epileptic EEG signals [14] and motor imagery BCI classification problems [15].

This study proposes a classification method of imagined speech EEG signals by using DWT and SVM to improve the classification accuracy of imagined speech. The diagram of the proposed methodology is shown in Fig. 1. An open dataset of imagined speech is collected, and a Butterworth filter is used to process the original signals. The features of the EEG signals are extracted by DWT decomposition. The SVM classifier is then used to classify imagined speech EEG signals with 10-fold cross-validation.

2 Materials and Methods

2.1 Dataset

The dataset used in this work was recorded in the Faculty of Engineering at the National University of Entre Ríos (UNER) by Pressel Curette et al. [16]. The subjects consist of 15 Argentine volunteers (seven women and eight men). EEG signals were recorded while the subjects performed overt and imagined speech tasks corresponding to the production of Spanish words and vowels. Only signals corresponding to imagined word production were analyzed. Thus, the data consist of trials in which participants imagined speaking six Spanish words including "arriba," "abajo," "derecha," "izquierda," "adelante," and "atrás" (corresponding to English words up, down, left, right, backward, and forward). The experimental protocol for the imagined word tasks required participants to imagine speaking one of the prompted words at three times during the 4-second trial period. Prior to the trial period, stimuli were presented visually and audibly,

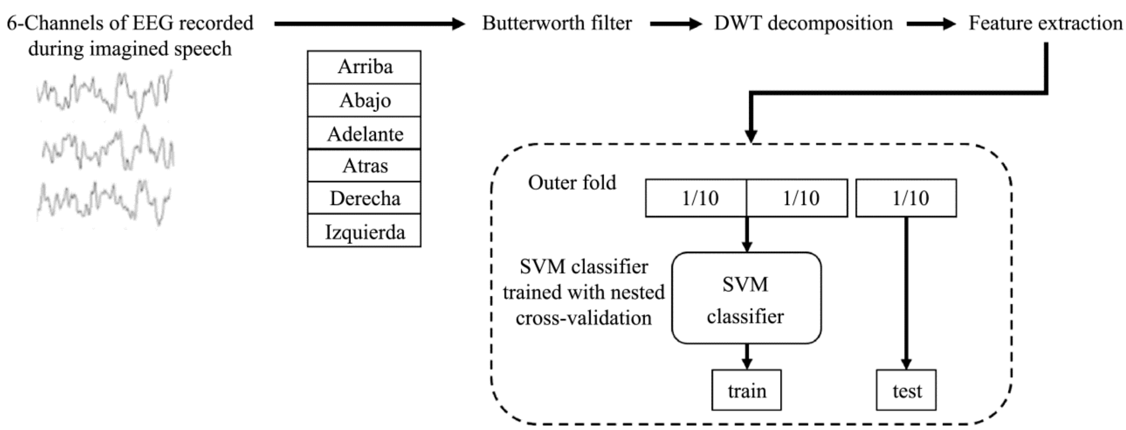

Fig.1 Depiction of the Proposed Methodology

showing each subject the word for 2 seconds. The EEG signals were recorded using an 18-channel Grass analogue amplifier and sampled at 1024 Hz. The electrodes were positioned according to the 10–20 international system at positions F3, F4, C3, C4, P3, and P4 (Fig. 2).

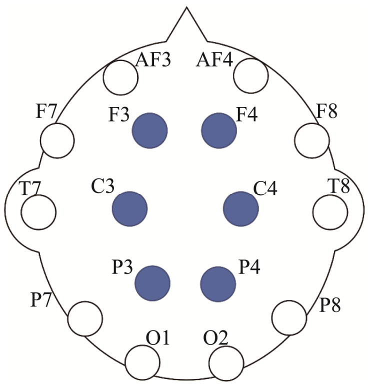

Fig.2 EEG Electrodes Used to Acquire Data

2.2 Data Preprocessing

EEG signal is a very weak bioelectrical signal that is susceptible to the interference of blinking, swallowing, and other muscle activities as well as power supply and sound. Therefore, the EEG signal should be preprocessed before feature extraction and classification to improve the signal quality and classification accuracy.

EEG data usually contain power frequency noise and artifact information and thus need to be filtered before feature extraction and classification. The alpha and beta rhythms in the language regions of the brain are significantly oscillated when the subjects are imagining speech. Therefore, band pass filtering should be performed on EEG signals to obtain α and β rhythm signals during the processing of imagined speech EEG signals. In this work, Butterworth band pass filter (0–40 Hz) was used to filter the EEG signal to remove noise and artifacts.

2.3 DWT Feature Extraction

Wavelet theory is a time-frequency domain signal analysis theory, and its basic idea is to use a family of functions to represent or approximate a signal; this family of functions is constituted by the translation and expansion of different scales of a basic wavelet (also known as mother wavelet). Wavelet transform is suitable for processing non-stationary signals, such as EEG signals, because of its characteristics of multi-resolution analysis and adaptive signal processing. In practice, the signal is usually processed discretely, which is termed as DWT, to facilitate calculation and reduce computation. The DWT principle is shown in Fig. 3, where S represents the original input signal, and A and D signals are generated by two complementary filters. Signal A represents the low-frequency component of the signal, which is the approximations of the original signal. Signal D represents the high-frequency component of the signal, which is the detail of the original signal.

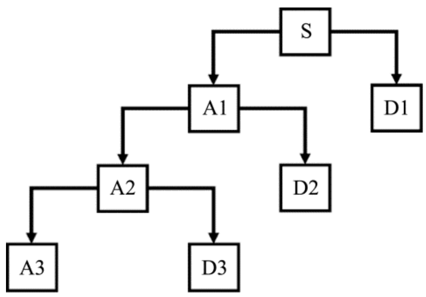

Fig.3 DWT Principle

In DWT, the choice of mother wavelet is important. Daubechies (dbN, N represents the wavelet order) extremal phase wavelet refers to a particular family of wavelets. The N in the function dbN is the same as the vanishing moment order of the wavelet function. In general, the larger the vanishing moment is, the flatter the corresponding filter will be; however, the calculation cost increases, and the real-time performance of signal processing worsens. In most EEG studies, Daubechies2 (db2) or Daubechies4 (db4) is used as the mother wavelet to guarantee real-time signal processing. In the present study, db2 was used as the mother wavelet with three decomposition levels; three detail values (D1, D2, D3) and one approximation (A3) were used to extract features (Fig. 4). The low-frequency signal A1 and the high-frequency signal D1 were obtained by the first decomposition. Subsequent decompositions were carried out according to the set decomposition scale to obtain A2, A3, D2, and D3 signals. The standard deviation (SD) and root mean square (RMS) of A3, D1, D2, and D3 signals were calculated as the EEG signal features. Six EEG acquisition channels (F3, F4, C3, C4, P3, and P4) were involved, and each channel held eight features. Thus, the total number of features is 48.

2.4 SVM Classifier

SVM is based on statistical learning theory and structural risk minimization criterion; the former avoids the limitation of sample size of a classification model, and the latter avoids model problems in model training. Therefore, SVM has strong promotion and discrimination ability, and its ultimate goal is to find an optimal classification hyperplane according to input data samples [17]. The nonlinear classification problem in the input space can be transformed into a linear classification problem in a dimensional feature space by nonlinear transformation. Linear SVMs can be learned in a high-dimensional feature space. EEG data cannot be linearly discriminated, so nonlinear SVMs are adopted.

The key to design a linear inseparable SVM classifier is to determine the kernel function, which is used to avoid the transformation of higher dimension. The inner product of the vector of higher dimension is equivalent directly to the parameter of lower dimension by substituting the kernel function. Non-separable mode is transformed into linearly separable problem by nonlinear mapping to higher dimension space. The formula is as follows:

$$K(x,z) = \langle \phi(x), \phi(z) \rangle \tag{1}$$

where $\langle , \rangle$ is the inner product, and K (x, z) is the kernel function.

By comparing the similarity among the samples, the Gaussian kernel gathers similar samples in a space characterized by similarity, and the samples are linearly separable. Radial basis function (RBF), as the most commonly used type of the Gaussian kernel, has small deviation in the treatment of most problems. The RBF kernel function can be formulated as follows:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|}{2\sigma^2}\right) \tag{2}$$

where σ is the kernel parameter.

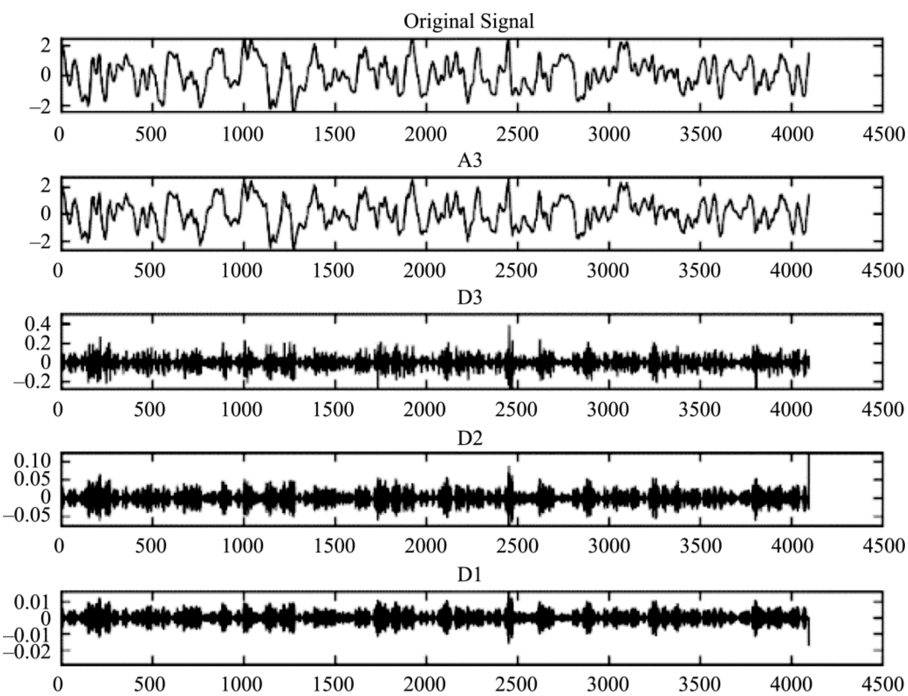

Fig.4 DWT Decomposition of an EEG Signal

RBF should determine fewer parameters than other kernel functions and has a lower degree of complexity. The RBF kernel function has a wide convergence domain and can be applied to samples with arbitrary distribution. RBF is an ideal mapping kernel function unless special kernel function is selected because of the special distribution of samples. Therefore, the RBF kernel function was used in this work.

3 Results

The classification performance of the proposed method for each subject with their own EEG signals was evaluated with 10-fold cross-validation. The experimental results show that the classification accuracy can reach 79.78%, and the mean classification accuracy is 61.69%. The accuracy of the classification with the proposed method is shown in Fig.5.

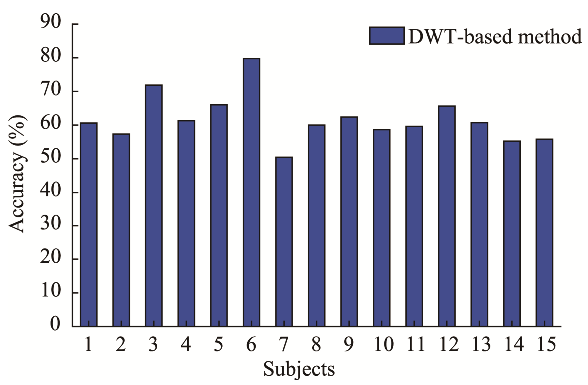

Fig.5 Accuracy of Each Subject

The classification performance of the proposed method with the mixed EEG signals of all subjects was evaluated. The experimental results show that the highest accuracy of classification with the proposed method only achieves 26.75%. The classification accuracy of the proposed method with mixed EEG signals of all subjects is lower than that with each subject's EEG signals but is higher than the classification accuracy reported in [16], which used the same EEG signal dataset. The comparison of the classification accuracy is shown in Fig.6.

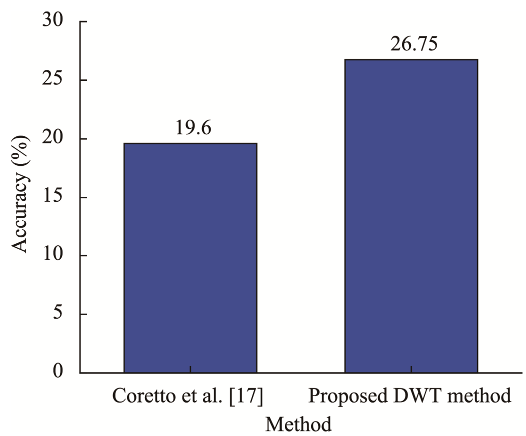

Fig.6 Accuracy of the Proposed Method with All Subjects' EEG Signals Mixed Together

Table 1 shows the comparison of the mean accuracy of the proposed method and those in other studies that used the same dataset. For the same dataset, the accuracy of the proposed method is the highest (61.9%). The classification accuracy of the proposed method with the mixed EEG signals of all the subjects is 26.75%, which is lower than that with individual's signal.

The method proposed by [16] achieved 19.6% accuracy only because the features extracted by relative wavelet energy (RWE) cannot well represent the differences between different imagined speech signals. However, the information presented by different features of EEG signals has significant differences, which lead to the reduction of the classification accuracy.

Table 1 Comparison of the Classification Accuracy

Study	Method	Accuracy	Signal
Coretto et al. [16]	DWT+RF	19.6%	Individual
García-Salinas et al. [18]	Parallel Factor Analysis	59.70%	Individual
Cooney et al. [19]	CNN	35.68%	Individual
Lee et al. [20]	Siamese Neural Network	31.40%	Individual
Proposed Method	DWT+SVM	26.75%	Mixed
Proposed Method	DWT+SVM	61.69%	Individual

4 Discussion

The proposed method with DWT and SWM for the classification of imagined speech EEG signals was evaluated with the open dataset of Imagine Speech with desirable results. The experimental classification accuracies suggest that the imagined speech features extracted by DWT can be discriminated well by SVM with RBF kernel function.

The proposed method has satisfactory performance, with mean classification accuracy of 61.69% and the highest accuracy of 79.78% for the classification of EEG signals. The method outperforms other techniques in previous studies. However, the classification accuracy varies greatly, and the difference between the highest value and the lowest value reaches 28.0%. Furthermore, the classification accuracy of the proposed method with the mixed EEG signals of all subjects is 26.75% only (Fig.6), which is lower than that with each subject's signals.

Therefore, imagined speech EEG signals significantly differ among different subjects, and higher classification accuracy can be achieved by training each subject's imagined speech EEG signal separately.

5 Conclusion

Imagined speech recognition is the most intuitive and convenient type of BCI for patients with severe speech disorders and has attracted increasing research attention.

This study proposes the classification method of imagined speech EEG signals with DWT and SVM. The proposed method was evaluated with the open dataset of Imagined Speech and successfully discriminated six imagined words representing different orientations. The method achieved higher classification accuracy for imagined speech EEG signals compared with previously reported techniques.

This study also compared the classification accuracy for each subject's EEG signals and that for EEG signals mixed together. The imagined speech EEG signal varied significantly and is more effective to discriminate each person's EEG signal individually than EEG signals mixed together.

Acknowledgement

This work was supported in part by the Fundamental Research Funds for the Central Universities (xcxjh20210104).

References

[1] T. M. Vaughan, "Guest editorial brain-computer interface technology: a review of the second international meeting," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 2, pp. 94-109, June 2003, https://doi.org/10.1109/TNSRE.2003.814799.

[2] Putze, F., & Schultz, T. (2014). Adaptive cognitive technical systems. Journal of neuroscience methods, 234, 108-115. https://doi.org/10.1016/j.jneumeth.2014.06.029.

[3] C. Cooney, R. Folli and D. Coyle, "Mel Frequency Cepstral Coefficients Enhance Imagined Speech Decoding Accuracy from EEG," 2018 29th Irish Signals and Systems Conference (ISSC), 2018, pp. 1-7.

[4] Y. Song and F. Sepulveda, "Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state," 2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings, 2014, pp. 568-571, https://doi.org/10.1109/BioCAS.2014.6981789.

[5] Hashim N., Ali A., Mohd-Isa WN. (2018) Word-Based Classification of Imagined Speech Using EEG. In: Alfred R., Iida H., Ag. Ibrahim A., Lim Y. (eds) Computational Science and Technology. ICCST 2017. Lecture Notes in Electrical Engineering, vol 488. Springer, Singapore. https://doi.org/10.1007/978-981-10-8276-4_19.

[6] Gonzalez-Castaneda, E. F., Torres-Garcia, A. A., Reyes-Garcia, C. A., & Villasenor-Pineda, L. (2016). Sonification and textification: proposing methods for classifying unspoken words from EEG signals. Biomedical Signal Processing & Control, 37(AUG.), 82-91. https://doi.org/10.1016/j.bspc.2016.10.012.

[7] Cooney, C., Korik, A., Raffaella, F., & Coyle, D. (2019). Classification of imagined spoken word-pairs using

convolutional neural networks. In G. R. Muller-Putz, J. C. Ditz, & S. C. Wriessnegger (Eds.), Proceedings of the 8th Graz Brain Computer Interface Conference 2019: Bridging Science and Application (Vol. 2019, pp. 338-343). (Proceedings of the 8th Graz Brain-Computer Interface Conference 2019). Verlag der Technischen Universitat Graz. https://doi.org/10.3217/978-3-85125-682-6-62.

[8] Alborz Rezazadeh Sereshkeh, Rozhin Yousefi, Andrew T Wong, Frank Rudzicz & Tom Chau (2019) Development of a ternary hybrid fNIRS-EEG brain–computer interface based on imagined speech, Brain-Computer Interfaces, 6:4, 128-140, https://doi.org/10.1080/2326263X.2019.1698928

[9] Sereshkeh AR, Trott R, Bricout A, Chau T. Online EEG Classification of Covert Speech for Brain-Computer Interfacing. Int J Neural Syst. 2017 Dec;27(8):1750033. doi: 10.1142/S0129065717500332. Epub 2017 Jun 13. PMID: 28830308. https://doi.org/10.1142/s0129065717500332

[10] Sriraam, N., Raghu, S. Classification of Focal and Non-focal Epileptic Seizures Using Multi-Features and SVM Classifier. J Med Syst 41, 160 (2017). https://doi.org/10.1007/s10916-017-0800-x

[11] Singh Pushpendra. 2018Novel Fourier quadrature transforms and analytic signal representations for nonlinear and non-stationary time-series analysis R. Soc. open sci.5181131181131. https://doi.org/10.1098/rsos.181131

[12] Zeng, W., Li, M., Yuan, C. et al. Identification of epileptic seizures in EEG signals using time-scale decomposition (ITD), discrete wavelet transform (DWT), phase space reconstruction (PSR) and neural networks. Artif Intell Rev 53, 3059–3088 (2020). https://doi.org/10.1007/s10462-019-09755-y.

[13] Shahbakhti, Mohammad & Beiramvand, Matin & Nazari, Mojtaba & Broniec-Wójcik, Anna & Augustyniak, Piotr & Rodrigues, Ana & Wierzchoń, Michał & Marozas, Vaidotas. (2021). VME-DWT: An Efficient Algorithm for Detection and Elimination of Eye Blink from Short Segments of Single EEG Channel. IEEE Transactions on Neural Systems and Rehabilitation Engineering. http://dx.doi.org/10.1109/TNSRE.2021.3054733.

[14] S. Raghu, Natarajan Sriraam, Yasin Temel, Shyam Vasudeva Rao, Alangar Satyaranjandas Hegde, Pieter L. Kubben. (2019). Performance evaluation of dwt based sigmoid entropy in time and frequency domains for automated detection of epileptic seizures using SVM classifier. Computers in Biology and Medicine, 110(C), 127-143, https://doi.org/10.1016/j.compbiomed.2019.05.016.

[15] Rajashekhar, U. & Neelappa, D. & Rajesh, L. (2021). Electroencephalogram (EEG) signal classification for brain–computer interface using discrete wavelet transform (DWT). International Journal of Intelligent Unmanned Systems. ahead-of-print. http://dx.doi.org/10.1108/IJIUS-09-2020-0057.

[16] Germán A. Pressel Coretto, Iván E. Gareis, and H. Leonardo Rufiner "Open access database of EEG signals recorded during imagined speech", Proc. SPIE 10160, 12th International Symposium on Medical Information Processing and Analysis, 1016002 (26 January 2017); https://doi.org/10.1117/12.2255697.

[17] Zhou, J., Huang, S., Wang, M. et al. Performance evaluation of hybrid GA–SVM and GWO–SVM models to predict earthquake-induced liquefaction potential of soil: a multi-dataset investigation. Engineering with Computers (2021). https://doi.org/10.1007/s00366-021-01418-3.

[18] Jesús S. García-Salinas, Luis Villaseñor-Pineda, Carlos A. Reyes-García, Alejandro A. Torres-García. (2018, October). Tensor decomposition for imagined speech discrimination in EEG. In Mexican International Conference on Artificial Intelligence (pp. 239-249). Springer, Cham. https://doi.org/10.1016/j.bspc.2019.01.006.

[19] C. Cooney, R. Folli and D. Coyle, "Optimizing Layers Improves CNN Generalization and Transfer Learning for Imagined Speech Decoding from EEG," 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), 2019, pp. 1311-1316, https://doi.org/10.1109/SMC.2019.8914246.

[20] Lee, Dong-Yeon & Lee, Minji & Lee, Seong-Whan. (2020). Classification of Imagined Speech Using Siamese Neural Network. 2979-2984. http://dx.doi.org/10.1109/SMC42975.2020.9282982.

Author Biographies

ZHANG Lingwei is currently a M.Sc. candidate at Nanjing University of Aeronautics and Astronautics. His main research interest includes brain-computer interface.

E-mail: nuaazlw@nuaa.edu.cn

ZHOU Zhengdong is currently an associate professor in Nanjing University of Aeronautics and Astronautics. His main research interests include BCI, deep learning, virtual reality, etc.

E-mail: zzd_sc@nuaa.edu.cn