Walking Imagery Evaluation in Brain Computer Interfaces via a Multi-view Multi-level Deep Polynomial Network

Baiying Lei, *Senior Member, IEEE*, Xiaolu Liu, Shuang Liang*, Wenlong Hang, Qiong Wang, Kup-Sze Choi, and Jing Qin

***Abstract*—Brain-computer interfaces (BCIs) based on motor imagery (MI) have been widely used to support the rehabilitation of motor functions of upper limbs rather than lower limbs. This is probably because it is more difficult to detect brain activities of lower limb MI. In order to reliably detect the brain activities of lower limbs to restore or improve the walking ability of the disabled, we propose a new paradigm of walking imagery (WI) in a virtual environment (VE) in order to elicit reliable brain activities and achieve a significant training effect. First, we extract and fuse both spatial and time-frequency features as a multi-view feature (MVF) to represent the patterns in the brain activity. Second, we design a multi-view multi-level deep polynomial network (MMDPN) to explore the complementarity among the features so as to improve the detection of walking from an idle state. Our extensive experimental results show that the VE-based paradigm significantly performs better than the traditional text-based paradigm. In addition, the VE-based paradigm can effectively help users to modulate brain activities and improve the quality of electroencephalography signals. We also observe that the MMDPN outperforms other deep learning methods in terms of classification performance.**

***Index Terms*—Walking imagery; Brain-computer interface; Virtual environment; Multi-view feature; Multi-view multi-level deep polynomial network**

I. INTRODUCTION

Brain-computer interfaces (BCIs) are capable of establishing a communication pathway between a user and the external world through brain activities [1-3], and electroencephalography (EEG) is often used to record oscillatory brain activities. Motor imagery (MI) is one of the most typical types of the BCI paradigm, and it is promising to apply MI-based BCIs in various areas (e.g., entertainment and convenient life for healthy people [4, 5]). More importantly, MI-based BCIs can help to restore motor functions of disabled patients [3], especially those with neuromuscular disorders. Numerous studies on the use of MI-based BCIs are carried out for the rehabilitation of upper limbs [6, 7]. For those patients who are lower-disabled, they also need to recover their walking ability, which is equally important to improve the capability of self-care of the rehabilitation population [8]. However, lower limb rehabilitations have only received limited attention due to a deeper location of the representation area in sensorimotor cortex, which makes it difficult to detect EEG signal reliably. Therefore, it is significant to obtain reliable brain activities associated with walking imagery (WI) tasks [9-11].

In order to improve the performance of MI-based BCIs, it is necessary to have a well-designed training approach to BCI control. It can help users to effectively modulate brain signals, which are the key elements to improve the robustness of BCIs. For example, some researchers consider a variety of feedback modalities (e.g., visual, haptic, or auditory [12, 13]), while others tend to increase the immersiveness of the visual feedback [14-17]. They leverage different technologies to create environments with higher degrees of fidelity, such as three-dimensional (3D) vision, virtual reality (VR), and augmented reality (AR) [14-17]. Although these feedback modalities can improve the BCI performance, they often lead to an inhibitory effect if the users incorrectly modulate their brain signals and fail to guarantee the reliability of the training calibration session [13, 17, 18].

In order to improve the reliability of the training calibration session, previous researchers devoted to designing suitable visual cues to improve the MI-based BCI performance [19, 20]. For example, when performing the upper limb MI training, the subjects are instructed to observe a static picture of a hand or a dynamic video of a hand grasping movement [19, 20]. Recently, static pictures of Chinese characters of the left/right hand writing are used as visual cues to perform MI tasks [21]. These visual cues are able to train the users more appropriately than the conventional approaches. The enhanced vividness of the visual cues can help users modulate their brain activities. However, the previous design of BCI training paradigm is seldom considered to strengthen the capacity of lower limb MI to record reliable brain activities. For this reason, it is important to find a suitable paradigm of WI with a reliable calibration session. Our previous study shows that virtual environments (VE) as visual cues are able to appropriately guide users to perform kinesthetic WI [22]. The subjects are requested to feel their feet naturally touch the ground and imagine walking in accordance with the movement rhythm of the virtual avatar. To reduce the interference of other limb movements on the brain activity, the established VE presents only the lower limbs of the walking avatar, and the results show that the VE-based paradigm can help users modulate their brain activities and increase the WI capacity.

This work was supported in part by National Basic Research Program of China, 973 Program (2015CB351706), National Natural Science Foundation of China (Nos. U1613219, 61871274), National Natural Science Foundation of Guangdong Province (2017A030313377), the Hong Kong Polytechnic University (PolyU 152040/16E). (*Corresponding author: Shuang Liang.*)

B. Lei, L. Xiao are with the National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China, 518060.

S. Liang is with School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications Nanjing, China, 210023 (e-mail: shuang.liang@njupt.edu.cn).

W. Hang is with the College of Computer Science and Technology, Nanjing Tech University, Nanjing, China, 211816.

Q. Wang is with Guangdong Provincial Key Laboratory of Computer Vision and Virtual Reality Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, 518055.

K-S Choi and J. Qin are with School of Nursing, Hong Kong Polytechnic University, Hung Hom, Hong Kong.

Despite the fact that the use of the VE-based paradigm obtains good training results during the imagination of walking, it is also essential to learn a robust feature representation to reveal the changes in brain activities in order to enhance BCI performance. Many previous feature extraction algorithms have been proposed to design BCI systems. For example, Wang *et al.* [23] introduced a wavelet-packet transform (WPT) algorithm to extract time-frequency features from EEG data. Saa *et al*. [24] employed power spectral density (PSD) to represent the frequency feature. Previous researchers [25-30] also adopted a spatial pattern (CSP) algorithm to analyze spatial features of imagined limb movements. However, these single feature-based methods fail to capture the complementarity of the features from multi-domains. It is argued that fused features are more promising for complex EEG data and can achieve better classification performance [31]. For this reason, a WPT-CSP algorithm was proposed and showed attractive performance [32, 33]. In this paper, we fuse CSP, PSD, and WPT features to form multi-view features (MVFs) in order to make full use of both spatial and time-frequency patterns.

To date, deep learning methods have witnessed a great success. For example, Lu *et al.* [34] used deep belief networks (DBNs) for MI classification. Tabar *et al.* [35] utilized stacked auto-encoder (SAE) to process EEG signals. Veres *et al.* [36] explored convolutional neural networks (CNNs) [37] to encode features. These algorithms usually require a substantial amount of training data to learn the features successfully. However, it is inconvenient and time-consuming to acquire a large amount of labeled EEG data, which leads to difficulty in training robust models. A deep polynomial network (DPN) is a new type of deep learning algorithm with fine feature representation for small datasets. It has the ability to compactly represent any functions on a finite sample dataset based on its structure. The derivation of DPN [38] is inspired by [39]. Specifically, a powerful layer-by-layer learning algorithm is employed in DPN, where the output of each node is a quadratic function of its inputs. The DPN algorithm attempts to build a network to represent the data in order to enhance the feature expression. Moreover, DPN runs in polynomial time, which is easy to control the depth and width of the network by only determining the number of layers and the number of nodes per layer in DPN encoding. Inspired by the promising performance by stacking DPNs [40], we build a multi-view multi-level DPN (MMDPN) framework to obtain both higher-level and integrated features to further boost the performance. To the best of our knowledge, there is no such deep learning study in BCIs for WI evaluation.

In summary, this paper explores the effectiveness of WI detection using a VE-based training paradigm and deep feature representation. We conduct experiments on self-collected EEG data and systematically evaluate the effect of visual cues on discriminating the WI state from the idle state. The experimental results verify that the VE-based training paradigm induces EEG patterns that are easier for single-trial WI detection. Furthermore, the proposed MMDPN framework is superior in the discrimination performance to other widely used deep learning frameworks. The rest of the paper is organized as follows. Section II elaborates the experimental design and methodology. Section III presents and discusses the experimental results. Finally, our conclusions are presented in Section IV.

II. METHODOLOGY

A. Participants

EEG signals are acquired from nine healthy subjects (aged 23-29 years old) and a patient with lower limb paralysis (aged 30 years old). All subjects are native speakers of Mandarin from China. None of them has prior knowledge about WI or any experience on BCIs training. Before the start of the experiment, they are informed the entire experimental procedure. During the experiment, each subject sits in a comfortable armchair about 50cm from a 24-inch computer screen placed in a quiet room. They are instructed to keep their bodies relaxed and avoid making any unnecessary movements and excessive blinking.

B. EEG signal collection

The EEG signals are recorded using a Biosemi ActiveTwo system, at a sampling rate of 512Hz. In the ActiveTwo system, the "ground" electrodes are replaced by two separate electrodes (CMS and DRL, as illustrated in Fig. 1). The CMS/DRL placement does not affect the signals measured between the other electrodes. For each subject, the entire preparation procedure of EEG acquisition takes about 20 minutes for applying the conductive gels to ensure the low impedance between the electrode and the participant's scalp. More details of this EEG acquisition system are available at the Biosemi website (http://www.biosemi.com).

The EEG electrode placements are labeled on the basis of the international 10-20 system. The 32 Ag/AgCl electrodes, including Fp1, AF3, F7, F3, FC1, FC5, T7, C3, CP1, CP5, P7, P3, Pz, PO3, O1, Oz, O2, PO4, P4, P8, CP6, CP2, C4, T8, FC6, FC2, F4, F8, AF4, Fp2, Fz, and Cz, are utilized for EEG signal collection. All the 32-channel EEG signals are adopted for data processing in our current study. The voltage offset at each active electrode is kept below 25 mV. The trigger signals, which can indicate the beginning of each trial in continuous EEG, are also synchronously recorded using the Biosemi ActiveTwo system via a parallel port.

C. Paradigms

The experimental paradigm is designed for walking detection with reference to the traditional BCI paradigm, which usually includes three stages of preparation, task performing (e.g., action or idle), and restoration. Visual cue is commonly used to instruct the subjects to perform the different classes of MI tasks, such as left/right hand movement or foot movement. The type of movement imagination can greatly affect the BCI performance. Some studies found that the subjects could better control the BCIs while performing the kinesthetic motor imagery. Therefore, we develop a VE-based BCI to guide subjects to consciously perform the imagination of lower limb movements. This paradigm is also used to provide relatively unified walking imagery action for different subjects.

During the experiments, subjects are instructed to perform kinesthetic WI via the visual guidance followed by the text-based and the VE-based paradigms, as shown in Fig. 2a. The stages of these two paradigms in a single-trial are illustrated in Fig. 2b and Fig. 2c, respectively. For the text-based paradigm: instructions appear on a computer screen in sequence as the guidance for five specific stages of WI training. The stages consist of 'idle', 'preparation', 'walking imagery', and 'restoration'. The subjects are asked to perform WI training with the help of this visual cue

platform.

For the VE-based paradigm, we adopt the same procedure as the text-based paradigm. However, we replace the texts with a 3D VE, where subjects perform WI following the movements of an avatar (see Fig. 2c). In order to reduce the interference introduced by the movements from other parts of the body, the computer screen only displays the lower limbs of the avatar. The avatar's movements provide subjects with the guidance on which experimental stage they shall perform. In the preparation stage, the avatar appears and stands in a virtual hallway. Then, the subjects imagine themselves walking along the hallway following the avatar's movement rhythm. The global location of the avatar in the VE can refer to the mini-map in the upper left corner of the screen, while there is a green progress bar indicating the time remained to complete the current stage of a single-trial.

D. The process of collecting EEG data

The procedures of the single-trial and the whole experiments are illustrated in Fig. 2c. The aforementioned two paradigms follow the same procedures. These procedures are introduced to all subjects, and then they sign the informed consent document prior to their participation.

The single-trial EEG signals are composed of brain activities recorded from four stages as shown in Fig. 2. At the beginning of each trial, a word ('*idle*') representing idle state is shown on the screen for 5 seconds, during which subjects keep themselves in a relaxed state. This stage is followed by a preparation period of 2 seconds to remind them to be ready to start movement imaginations. Then subjects feel their feet naturally touch the ground and imagine walking. This is referred to the stage of WI. They are asked to avoid moving their own bodies or blinking eye in the above three stages. At the end of lower limb MI stage, subjects take a rest to adjust themselves in a restoration period of 5 seconds. We use trigger signals to label the beginning of each single-trial. In the whole processes, the screen first displays a word ('*initialization*') to indicate the start of an experimental run. It starts with an initialization period of 5 seconds, which gives the subjects time to adjust their experimental state. The subjects are allowed to take a rest to relieve fatigue between any two runs. Two sessions of EEG signals corresponding to two different days for each subject are collected. One session consists of 10 experimental runs, and each contains 10 trials. The experiments are performed five times for each paradigm (i.e., 2D text-based paradigm or 3D VE-based paradigm), respectively, and the order is random. Each trial includes two control groups (idle state or WI state), that is, there are totally100 trials of idle and 100 trials of WI for each paradigm.

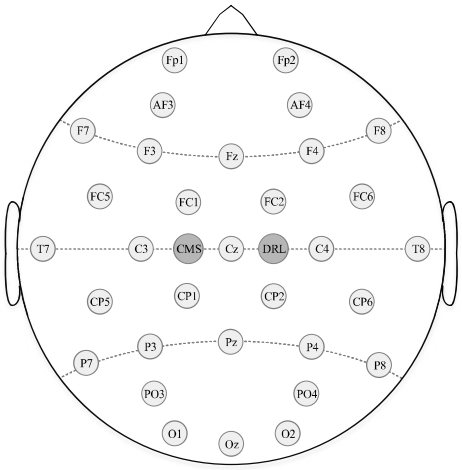

Fig. 1. EEG electrode positions used in Biosemi ActiveTwo system, and CMS/DRL electrodes marked in shades of dark green.

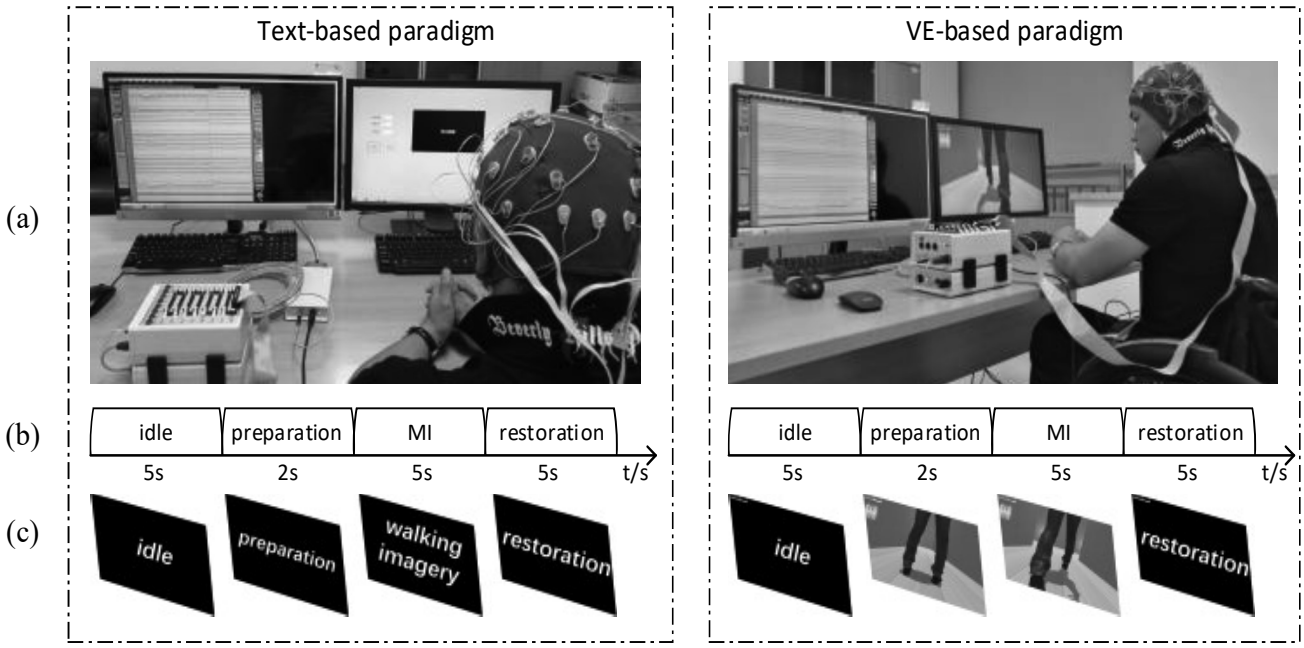

Fig.2. Experimental setup for EEG data collection based on two different training paradigms (Text-based and VE-based). The training paradigms are used to guide subjects to perform WI tasks. In each scenario, EEG signals are acquired using Biosemi ActiveTwo system, which are installed in the left computer. Meanwhile, the right computer is used to provide the visual cues for each trial and send corresponding trigger signals. Each single-trial contains four stages corresponding to idle (5s), preparation (2s), walking imagery (5s), and restoration (5s). An experimental run begins with 5-seconds initialization stage, being followed by performing 10 single-trials. Subjects can have a rest to relieve fatigue between any two runs. Experiments are carried out a total of 5 times for each paradigm, respectively, with random order.

E. MMDPN framework

Fig. 3 shows the flowchart of the proposed method. The framework contains four procedures: feature preprocessing, feature extraction, feature encoding, and prediction. After the recorded raw data are preprocessed and the EEG data in reference-free form are obtained. Three types of features, including CSP, PSD, and WPT, are extracted. Each type of the feature is normalized and contains two rhythms based on the EEG characteristics of WI: *mu* frequency bands (8∼13Hz) and *beta* frequency bands (13∼30 Hz). Firstly, these features are fused linearly as MVFs and are trained by a technique named linear discriminant analysis (LDA). Secondly, each feature is encoded by a first-level DPN (DPN-1). These features are encoded again by a second-level DPN (DPN-2). Finally, MMDPN features are obtained to train the LDA classifier.

1) Preprocessing

In the preprocessing stage, we use the trigger signals to segment continuous raw EEG recordings into idle epochs and WI epochs for each paradigm. The recorded EEG signals are converted into reference-free forms with common average references.

2) Feature extraction

The mu frequency bands and $beta$ frequency bands are relevant to walking imagination [10]. The feature extraction is performed on these two bands, respectively. For CSP features, all EEG segments are filtered with a fifth order Butterworth band-pass filter in the range of mu frequency band and $beta$ frequency band [7, 41, 42]. CSP is used to extract EEG features for each frequency band. For PSD features, the spectral entropy of PSD within mu and $beta$ frequency bands are estimated, respectively. For WPT features, we choose two nodes of the wavelet packet: $U(5,1)$ (8∼16Hz) and $U(4,1)$ (16∼32Hz). The former corresponds to the mu frequency bands, while the latter corresponds to the $beta$ frequency bands. The entropy of each node is calculated as the WPT feature.

3) Feature Encoding

In the stage of feature encoding, DPNs [38] are used. The DPN is a newly proposed deep architecture, which is capable to compactly learn complex functions on a finite training set. The DPN is built in a layer-by-layer fashion, in which the output of each node is a quadratic function of its input. A DPN integrates features between different samples and different dimensions through the network. By presenting the output with hierarchical features, the DPN greatly improves the feature representation ability. Moreover, the DPN runs in polynomial time, and is easy to train without relying on complex heuristics. In the following, we briefly describe the construction process of the DPN.

Given a set of training samples $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), ..., (\mathbf{x}_m, y_m)\}$, a coefficient vector $(w_1, w_2, ..., w_m)$ and m polynomials $\mathbf{P}_1, \mathbf{P}_2, ..., \mathbf{P}_m$, according to Lemma in [38], we can obtain the equation $\sum_{i=1}^{m} w_i \mathbf{P}_i(\mathbf{x}_j) = y_j$, in which $(y_1, y_2, ..., y_m)$ can be any set of values.

For the first layer of DPN, given a set of values $\{(\langle \mathbf{w}, [\mathbf{1}\ \mathbf{x}_1]\rangle, \langle \mathbf{w}, [\mathbf{1}\ \mathbf{x}_2]\rangle, ..., \langle \mathbf{w}, [\mathbf{1}\ \mathbf{x}_m]\rangle) : \mathbf{w} \in \mathbb{R}^{d+1}\}$ attained by degree-1 polynomials functions on training samples, we can find a set of vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_{d+1}$ to make $\{(\langle \mathbf{w}_j, [\mathbf{1}\ \mathbf{x}_1]\rangle, \langle \mathbf{w}_j, [\mathbf{1}\ \mathbf{x}_2]\rangle, ..., \langle \mathbf{w}_j, [\mathbf{1}\ \mathbf{x}_m]\rangle)\}_{j=1}^{d+1}$ linearly independent with a basis-construction method. Specifically, we use a matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_{d+1}]$ to map $[\mathbf{1}\ \mathbf{X}]$ into the constructed basis. The $m \times (d+1)$ matrix F^1, in which $F_j^1 = \langle \mathbf{W}_j, [\mathbf{1}\ \mathbf{X}] \rangle, j = 1, 2, \ldots d+1$ is the output of j-th node of the first layer, spans all possible values attained by degree-1 polynomials functions on training samples.

For the second layer of the DPN, let $F^2 = \left[(F_1^1 \circ F_1^1) \cdots (F_1^1 \circ F_{|F|}^1) \cdots (F_{|F|}^1 \circ F_1^1) \cdots (F_{|F|}^1 \circ F_{|F|}^1) \right]$, where $F_1^1 \circ F_1^1$ refer to their Hadamard product. We form a new matrix $\left[F \tilde{F}^2 \right]$ that spans all values attained by the degree-2 polynomial functions on training samples. Then, we can find linearly independent columns of $\left[F F^2 \right]$ to be the basis for $\left[F \tilde{F}^2 \right]$. Thus, each column of F^2 corresponds to a node of the second layer of DPN.

We then redefine F as the matrix $\left[F F^2 \right]$, and the above processes can be extended to construct layer $3, 4, \ldots, z$. Let $\tilde{F}^z = \left[\left(F_1^{z-1} \circ F_1^1 \right) \cdots \left(F_1^{z-1} \circ F_{|F_1|}^1 \right) \cdots \left(F_{|F^{z-1}|}^{z-1} \circ F_1^1 \right) \cdots \left(F_{|F^{z-1}|}^{z-1} \circ F_{|F_1|}^1 \right) \right]$, and we can finally construct the z-th layer of the DPN. Fig. 3a shows the basic architecture of the DPN with four layers.

4) Multi-view multi-level deep polynomial network

For a single DPN based algorithm, it lacks diversity, relevance, and discrimination. It fails to fuse and capture the linear combination of multiple feature representations. To address this shortcoming, we can propose a multi-view and multi-level DPN. The MMDPN algorithm explores the complementary information of all the features and enhances the feature representation via the multi-level DPN. The MMDPN for multiple feature representation is shown in Fig. 2b.

In our work, we focus on efficient integration of the features of CSP, PSD, and WPT. In the first level, the features of CSP, PSD, and WPT are encoded by the DPN (DPN-1) to obtain a high-level feature representation. In the second level, all three types of higher-level features are cascaded in the DPN encoding (DPN-2). The ultimate learned features include both intrinsic attribute of each modality and the relativities among all modalities. Therefore, the MMDPN efficiently exploits the complementarity among multiple features, which exhibits a highly non-linear correlation between multi-view features.

Our method combines both common and complementary information of the EEG data to boost the WI performance. In addition, the hierarchical representation of each level can reduce the noise and extract useful information. If only one layer of the network in the basic DPN encoding is used, it is equivalent to the dimensionality reduction of principle component analysis method. To obtain an effective encoded feature matrix for WI classification, we construct a multi-level DPN network with different numbers of layers with different width. The encoded features are obtained by the direct connection of the learned hierarchical features from different layers. As a result, the proposed MMDPN further enhances the WI representation.

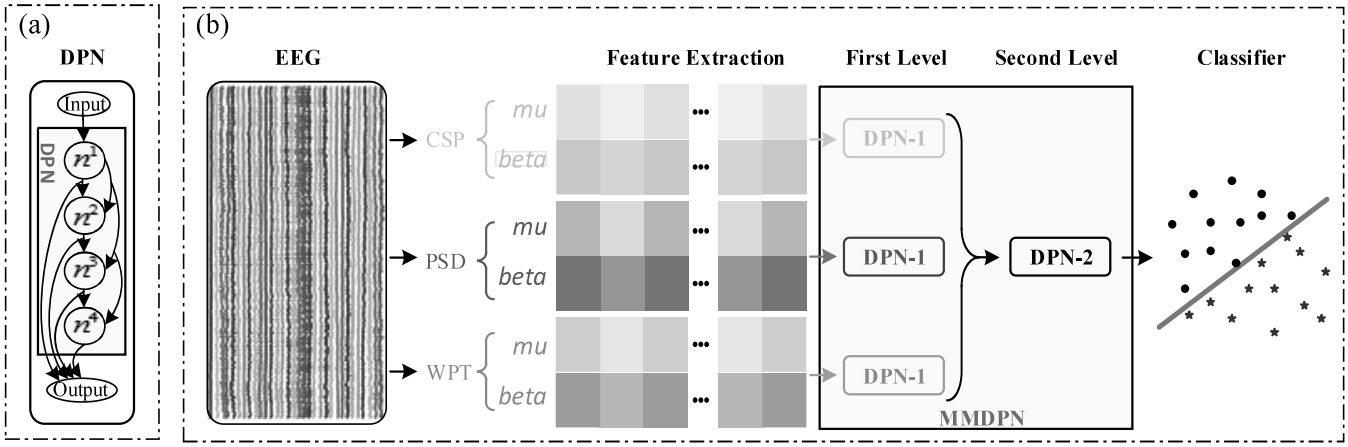

Fig.3. Illustration of flowchart of proposed method using three features (e.g., CSP, PSD, WPT). (a) Schematic diagram of DPN architecture with four layers. (b) Raw EEG signals with *mu* and *beta* frequency bands after preprocessing to obtain three features. Each feature is encoded by DPN (a.k.a., DPN-1), and fused and decoded again by DPN (a.k.a., DPN-2). The final output is used to train linear discriminant analysis classifier.

III. Experimental Results

A. Experimental settings

In this section, we show the experimental results with our proposed method. Our experimental results are based on binary classification: idle state vs. WI state. The EEG data of two control groups (idle state and WI state) are used for training the model and evaluating the classification performance. In our experiments, there are 100 samples of idle class and 100 samples of WI class for each subject in each paradigm. 50 samples of idle class and 50 samples of WI class are randomly selected for training and the remaining ones are used for testing. The experimental processing ignores problem of non-stationary EEG signals. To further verify the effectiveness of our DPN method, it is compared with the recent popular deep learning techniques (e.g., DBN, SAE and CNN) with the same multi-view multi-level framework (a.k.a., MMDBN, MMSAE, and MMCNN). The above processing is carried out for 10 times and the average results are reported for each subject in each paradigm.

To validate the performance of different paradigms on WI tasks, assuming that the WI state corresponds to the positive class and the idle state to the negative class. When the WI is classified as the WI, it is regarded as a true positive (TP). When the WI is classified as the idle state, it is regarded as a false positive (FP). The metrics are as below: The true positive rate (TPR), the false positive rate (FPR), Accuracy (ACC) = (TP+TN)/(TP+TN+FP+FN), Sensitivity (SEN) = TP/(TP+FN); Specificity (SPEC) = TN/(TN+FP), Youden's index (Youden) = SEN+SPEC-1, F1 score (F1) = 2×PPV×SEN/ (PPV+SEN) (PPV = TP/(TP+FP), NPV = TN/(TN+FN)), AUC is the area under receiver operating characteristics curve (ROC). All the signal analysis programs are performed using MATLAB 2017a (MathWorks Inc.).

B. Results

Table I shows the evaluation results for the text-based paradigm and the VE-based paradigm using CSP, PSD, WPT, and MVF. The results are reported in terms of the averaged values across all subjects with ten evaluations minus or plus their standard deviations. Table II shows the t-test statistical analysis between two paradigms with the same method.

For all the features, we find that the VE-based paradigm outperforms the text-based paradigm. Table II shows that there are significant differences between the two paradigms. The results prove that the VE-based paradigm can help subjects better modulate their brain activities, which significantly improve the performance of WI detection. For the two paradigms, MVF obtains the best performance among all the four types of features. It proves that MVF integrates spatial features and time-frequency features to fully leverage information of WI features and achieves the best performance.

Fig. 4 shows the ROC curves of two paradigms with three independent features and MVF. The curves of the VE-based paradigm are close to the upper left corner of the figure than the text-based paradigm, which proves the superiority of the VE-based paradigm over the text-based paradigm. It is also obvious that ACC of the VE-based paradigm is higher than that of the text-based paradigm in the same subject. The MVF's curves are also close to the upper left corner of the figure among all features. In addition, the multi-view features are better than the single view feature from which we can explore the feature complementarities.

Table III shows the performance comparison of the VE-based paradigm using MVF, MDPN, and MMDPN. Fig. 5 illustrates the results summarized in Table III. There are a significant difference between MVF and MDPN ($p \leq 0.001$) and a significant difference between MVF and MMDPN ($p \leq 0.0008$) in Table IV. From the mean results, MVF is lower than MDPN or MMDPN, which fully demonstrates the attractive feature encoding method via DPN. From the significance analysis in Table IV, we can see that SPEC cannot satisfy the significant difference at $alpha = 0.05$ level. However, all other metrics satisfy the significance requirement. Also, ACC, Youden, F1, and AUC meet the significant difference at $alpha = 0.01$ level. Hence, it proves that the MDPN can leverage the complementarity and correlation between the single features and MMDPN can better decode WI feature than MDPN.

Table I. Performance comparison between Text-based paradigm and VE-based paradigm using various features and metrics (%).

Paradigm	Feature	ACC	SEN	SPEC	Youden	F1	AUC
Text	CSP	67.27±08.37	67.72±08.28	66.82±08.71	34.54±16.74	67.35±08.24	73.06±10.36
	PSD	61.09±04.92	62.36±07.47	59.82±05.06	22.18±09.84	61.34±05.59	64.71±05.50
	WPT	65.46±05.72	65.82±06.86	65.10±05.24	30.92±11.45	65.40±06.01	70.56±07.01
	MVF	**69.45±08.83**	**70.40±09.67**	**68.50±08.26**	**38.90±17.66**	**69.58±09.01**	**74.90±09.93**
VE	CSP	81.19±07.86	82.54±08.34	79.84±07.94	62.38±15.72	81.37±07.83	88.21±06.47
	PSD	71.96±05.09	79.60±09.41	64.32±04.12	43.92±10.18	73.72±05.67	74.49±04.55
	WPT	74.96±06.52	79.40±08.99	70.52±05.20	49.92±13.05	75.88±06.66	80.92±06.77
	MVF	**81.81±07.50**	**83.42±08.73**	**80.20±06.90**	**63.62±14.99**	**81.94±07.61**	**88.30±05.75**

Table V presents the mean and standard deviation of performance comparison for MMDBN, MMSAE, MMCNN and MMDPN of the VE-based paradigm using 6 metrics. Fig. 6 (VE) corresponds to Table V. Table VI shows the t-test results among MMDPN and MMDBN, MMSAE, as well as MMCNN in the VE-based paradigm. Table V shows that the healthy subject S7 has the best performance of WI detection. The evaluation values are higher than most of the healthy subjects for the disable subject S10 with lower limbs paraplegia. It also proves that the disabled subject still can perform WI tasks and outperform the normal subjects. This observation suggests the potential for the disabled in other aspects of rehabilitation.

For subject S8 using MMSAE as the feature decoding method, all the WI instances are judged as idle. Both TP and FP are 0, PPV and F1 is NaN. This shows that MMSAE is unstable for this data. In Table V, it can be found that the MMDPN has the highest values for all five evaluation metrics except SPEC. Fig. 6 shows the bar diagram for performance comparison for MMDBN, MMSAE, MMCNN, and MMDPN in the text-based paradigm. In terms of average value of ten subjects, MMDPN obtains the best performance of all the multi-view multi-level frameworks.

Table II T-test between text-based paradigm and VE-based paradigm.

Feature	ACC	SEN	SPEC	Youden	F1	AUC
CSP	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
DSP	0.0002	<0.0001	0.1107	0.0002	<0.0001	0.0006
WPT	0.0003	<0.0001	0.022	0.0003	<0.0001	0.0002
MVF	0.0002	0.0002	0.0005	0.0002	0.0002	0.0002

Table III Performance comparison of VE-based paradigm based on MVF, MDPN and MMDPN method (%).

Subject	Method	ACC	SEN	SPEC	Youden	F1	AUC
S1	MVF	76.90±04.53	77.40±05.34	76.40±06.85	53.80±09.07	77.02±04.49	85.62±04.50
	MDPN	83.10±03.54	85.00±06.62	**81.20±04.92**	66.20±07.08	83.34±03.83	89.69±04.20
	MMDPN	**83.60±04.30**	**86.80±05.75**	80.40±06.92	**67.20±08.60**	**84.11±04.11**	**91.00±02.95**
S2	MVF	75.40±06.54	72.40±12.10	78.40±07.99	50.80±13.07	74.28±07.84	81.96±06.50
	MDPN	80.90±05.88	79.00±10.47	82.80±05.01	61.80±11.75	80.26±06.91	86.93±06.35
	MMDPN	**83.90±03.11**	**84.80±07.38**	**83.00±03.68**	**67.80±06.21**	**83.91±03.82**	**89.67±03.37**
S3	MVF	83.80±02.53	85.80±06.00	81.80±03.19	67.60±05.06	84.04±02.90	90.59±02.74
	MDPN	87.60±03.44	90.60±06.60	84.60±03.78	75.20±06.88	87.88±03.72	94.45±02.78
	MMDPN	**88.40±02.95**	**92.20±05.53**	**84.60±04.01**	**76.80±05.90**	**88.78±03.12**	**94.87±01.99**
S4	MVF	80.50±03.03	83.60±05.87	77.40±04.81	61.00±06.06	81.03±03.20	87.62±04.22
	MDPN	88.00±02.40	88.40±05.48	87.60±03.75	76.00±04.81	87.99±02.63	94.10±02.02
	MMDPN	**90.40±01.35**	**92.00±03.40**	**88.80±03.91**	**80.80±02.70**	**90.55±01.29**	**96.41±01.18**
S5	MVF	88.90±04.01	94.40±05.06	83.40±04.12	77.80±08.02	89.46±03.90	94.12±03.46
	MDPN	92.70±02.71	95.20±02.53	90.20±04.37	85.40±05.42	92.90±02.56	**97.16±01.74**
	MMDPN	**93.40±02.72**	**96.20±03.19**	**90.60±05.89**	**86.80±05.43**	**93.62±02.53**	96.92±01.99
S6	MVF	74.00±03.27	75.00±08.55	73.00±04.03	48.00±06.53	74.06±04.36	82.97±03.56
	MDPN	82.50±02.72	84.00±04.22	81.00±05.83	65.00±05.44	82.77±02.49	91.35±02.56
	MMDPN	**84.10±03.11**	**84.00±04.81**	**84.20±06.76**	**68.20±06.21**	**84.10±02.84**	**92.56±02.81**
S7	MVF	97.20±01.40	98.80±01.69	95.60±02.27	94.40±02.80	97.25±01.37	99.44±00.94
	MDPN	99.30±00.82	99.40±00.97	99.20±01.03	98.60±01.65	99.30±00.82	99.97±00.05
	MMDPN	**99.50±00.71**	**99.40±00.97**	**99.60±00.84**	**99.00±01.41**	**99.50±00.71**	**99.97±00.06**
S8	MVF	85.50±02.99	89.20±05.90	81.80±04.85	71.00±05.98	85.97±03.05	90.13±02.47
	MDPN	87.40±02.63	91.20±04.44	83.60±02.80	74.80±05.27	87.83±02.69	90.20±02.92
	MMDPN	**89.20±02.04**	**91.20±04.34**	**87.20±04.73**	**78.40±04.09**	**89.40±02.02**	**93.32±01.95**
S9	MVF	73.30±06.09	75.80±07.39	70.80±08.18	46.60±12.19	73.92±05.91	81.02±07.27
	MDPN	82.40±02.76	81.20±04.83	**83.60±02.95**	64.80±05.51	82.14±03.17	90.32±02.63
	MMDPN	**83.20±03.71**	**83.60±06.45**	82.80±03.43	**66.40±07.41**	**83.18±04.11**	**90.61±02.18**
S10	MVF	82.60±03.89	81.80±05.85	83.40±04.90	65.20±07.79	82.42±04.12	89.56±03.96
	MDPN	88.40±02.67	84.00±05.66	92.80±03.79	76.80±05.35	87.80±03.06	96.09±01.42
	MMDPN	**90.10±02.38**	**85.40±04.72**	**94.80±02.35**	**80.20±04.76**	**89.56±02.72**	**97.04±01.02**

Table V. Performance comparison for VE-based paradigm based on MMDBN, MMSAE, MMCNN, and MMDPN methods (%).

Subject	Method	ACC	SEN	SPEC	Youden	F1	AUC
S1	MMDBN	77.30±10.08	75.20±28.41	79.40±17.46	54.60±20.16	73.66±20.07	90.59±05.88
	MMSAE	82.00±04.32	78.20±10.13	**85.80±06.14**	64.00±08.64	81.03±05.64	89.51±03.60
	MMCNN	81.60±05.15	83.20±05.59	80.00±06.46	63.20±10.29	81.90±05.05	87.90±03.15
	MMDPN	**83.60±04.30**	**86.80±05.75**	80.40±06.92	**67.20±08.60**	**84.11±04.11**	**91.00±02.95**
S2	MMDBN	75.40±09.12	74.00±17.07	76.80±13.64	50.80±18.24	74.37±10.74	82.98±11.88
	MMSAE	74.40±08.54	**88.60±09.29**	60.20±20.43	48.80±17.08	77.83±05.54	80.87±17.66
	MMCNN	78.60±05.15	75.40±07.95	81.80±06.36	57.20±10.29	77.78±05.87	86.11±05.60
	MMDPN	**83.90±03.11**	84.80±07.38	**83.00±03.68**	**67.80±06.21**	**83.91±03.82**	**89.67±03.37**
S3	MMDBN	79.80±04.85	77.40±12.62	82.20±09.73	59.60±09.70	78.92±06.35	87.85±04.44
	MMSAE	83.80±04.76	**94.40±06.02**	73.20±11.93	67.60±09.51	85.47±03.54	91.31±06.80
	MMCNN	88.30±03.06	90.20±04.05	**86.40±03.98**	76.60±06.11	88.51±02.99	93.20±02.52
	MMDPN	**88.40±02.95**	92.20±05.53	84.60±04.01	**76.80±05.90**	**88.78±03.12**	**94.87±01.99**
S4	MMDBN	83.40±06.67	85.80±12.59	81.00±13.00	66.80±13.34	83.53±07.27	89.04±06.10
	MMSAE	82.70±07.85	**96.40±03.37**	69.00±17.34	65.40±15.69	85.19±05.78	84.67±15.63
	MMCNN	87.00±03.13	88.40±05.87	85.60±05.80	74.00±06.25	87.15±03.13	93.52±02.29
	MMDPN	**90.40±01.35**	92.00±03.40	**88.80±03.91**	**80.80±02.70**	**90.55±01.29**	**96.41±01.18**
S5	MMDBN	92.00±03.16	94.00±03.40	90.00±06.32	84.00±06.32	92.21±02.92	**97.84±01.51**
	MMSAE	90.90±04.20	91.00±03.43	**90.80±06.48**	81.80±08.40	90.97±03.96	95.78±03.56
	MMCNN	91.70±02.50	93.00±03.16	90.40±02.07	83.40±04.99	91.79±02.52	97.14±01.34
	MMDPN	**93.40±02.72**	**96.20±03.19**	90.60±05.89	**86.80±05.43**	**93.62±02.53**	96.92±01.99
S6	MMDBN	79.60±08.24	**84.40±07.59**	74.80±12.04	59.20±16.47	80.66±07.39	87.93±09.02
	MMSAE	83.50±06.06	79.60±10.41	**87.40±09.62**	67.00±12.12	82.66±06.70	90.34±06.10
	MMCNN	83.70±01.95	81.60±03.75	85.80±03.46	67.40±03.89	83.33±02.08	**92.59±02.11**
	MMDPN	**84.10±03.11**	84.00±04.81	84.20±06.76	**68.20±06.21**	**84.10±02.84**	92.56±02.81
S7	MMDBN	96.10±02.28	93.60±05.56	98.60±03.13	92.20±04.57	95.94±02.47	98.36±01.92
	MMSAE	96.30±03.23	95.40±04.62	97.20±07.00	92.60±06.47	96.31±03.01	98.13±02.99
	MMCNN	98.80±01.32	99.00±01.41	98.60±01.65	97.60±02.63	98.80±01.31	99.47±00.84
	MMDPN	**99.50±00.71**	**99.40±00.97**	**99.60±00.84**	**99.00±01.41**	**99.50±00.71**	**99.97±00.06**
S8	MMDBN	86.20±05.75	89.60±06.24	82.80±10.38	72.40±11.50	86.75±05.27	88.80±06.27
	MMSAE	84.70±12.84	79.00±29.23	**90.40±06.98**	69.40±25.68	NaN	85.31±25.71
	MMCNN	88.20±03.29	87.60±05.72	88.80±04.83	76.40±06.59	88.09±03.50	**93.90±03.09**
	MMDPN	**89.20±02.04**	**91.20±04.34**	87.20±04.73	**78.40±04.09**	**89.40±02.02**	93.32±01.95
S9	MMDBN	78.40±09.19	76.00±07.72	80.80±13.60	56.80±18.38	78.07±08.31	81.17±12.23
	MMSAE	78.00±06.45	**89.20±07.73**	66.80±19.33	56.00±12.89	80.51±03.49	83.82±05.03
	MMCNN	82.50±02.80	76.60±02.84	**88.40±05.32**	65.00±05.60	81.43±02.64	89.60±02.61
	MMDPN	**83.20±03.71**	83.60±06.45	82.80±03.43	**66.40±07.41**	**83.18±04.11**	**90.61±02.18**
S10	MMDBN	82.50±04.67	82.40±09.28	82.60±11.16	65.00±09.35	82.43±04.75	90.91±05.07
	MMSAE	84.30±09.90	**88.40±08.68**	80.20±16.90	68.60±19.80	85.23±08.64	89.87±09.05
	MMCNN	88.90±01.60	86.40±02.95	91.40±03.27	77.80±03.19	88.61±01.62	95.06±01.38
	MMDPN	**90.10±02.38**	85.40±04.72	**94.80±02.35**	**80.20±04.76**	**89.56±02.72**	**97.04±01.02**

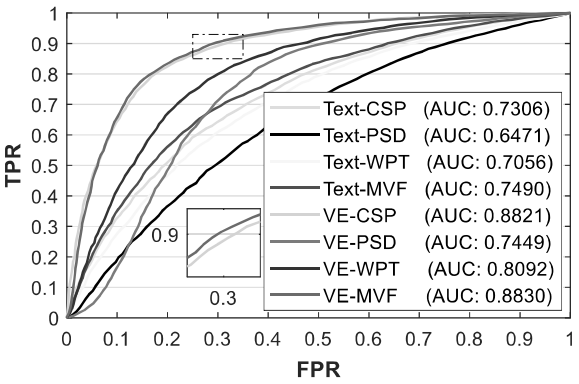

Fig.4. ROC curves of two paradigms with three single features and MVF (positive class: WI state; negative class: idle state).

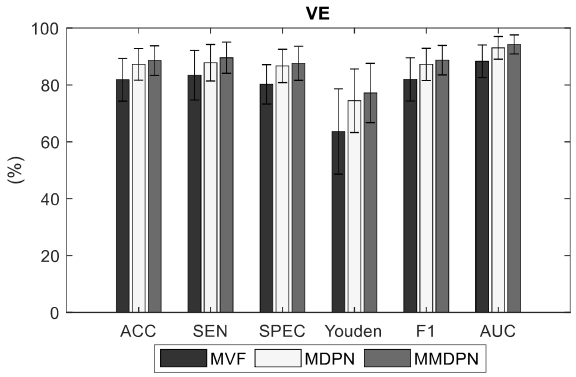

Fig. 5 Performance comparison among MVF, MDPN and MMDPN in VE-based paradigm using various metrics.

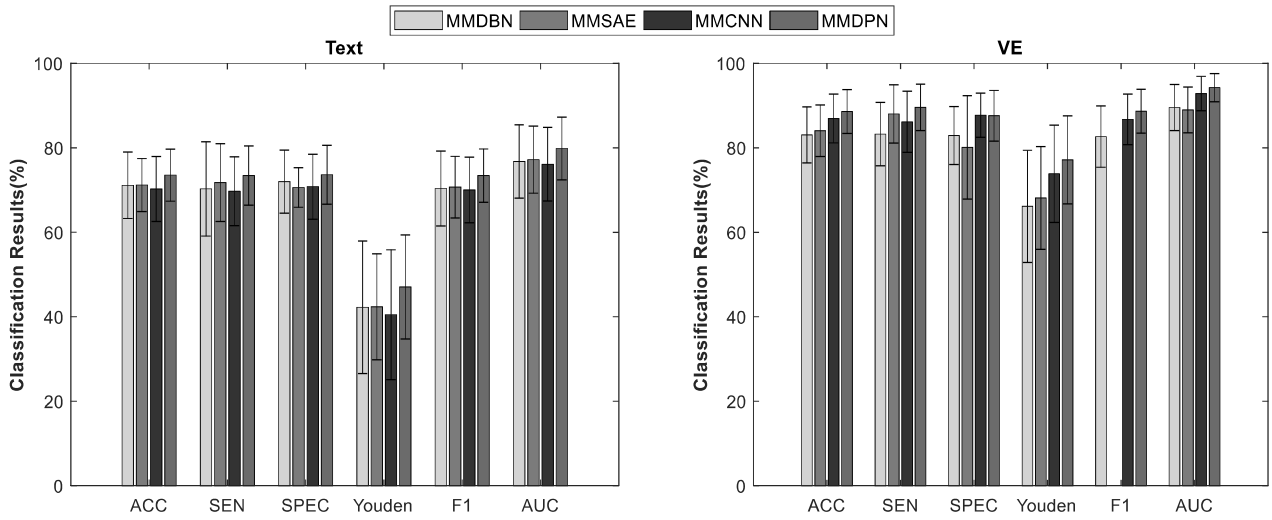

Fig. 6. Performance comparison for MMDBN, MMSAE, MMCNN, and MMDPN for text-based paradigm and VE-based paradigm.

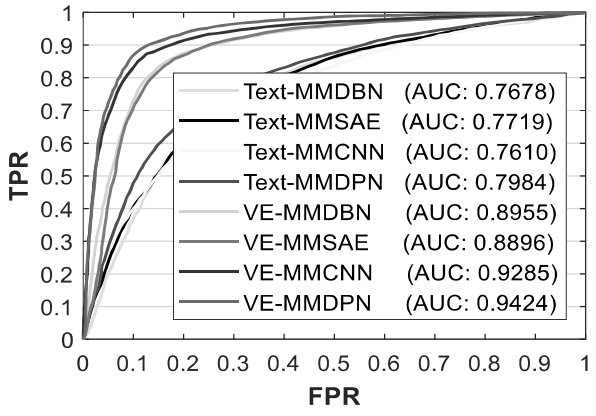

Fig. 7 ROC cures of two different paradigms with different feature encoding algorithms.

Table IV. T-test of VE-based paradigm using MVF, MDPN, and MMDPN.

ρ	ACC	SEN	SPEC	Youden	F1	AUC
MVF vs. MDPN	<0.0001	0.0010	0.0003	<0.0001	<0.0001	0.0009
MVF vs. MMDPN	<0.0001	0.0008	0.0001	<0.0001	<0.0001	0.0002
MDPN vs. MMDPN	0.0011	0.0140	0.0863	0.0011	0.0014	0.0096

TABLE VI T-test of various comparisons in VE-based paradigm.

Method	ACC	SEN	SPEC	Youden	F1	AUC
MMDBN vs. MMDPN	<0.0001	0.0028	0.0051	<0.0001	0.0002	0.0016
MMSAE vs. MMDPN	0.0005	0.4489	0.0514	0.0005	<0.0001	0.0015
MMCNN vs. MMDPN	0.0097	0.0056	0.8900	0.0097	0.0064	0.0159

TABLE VII T-test of various comparisons in text-based paradigm.

Method	ACC	SEN	SPEC	Youden	F1	AUC
MMDBN vs. MMDPN	0.0214	0.1917	0.4185	0.0214	0.0318	0.0055
MMSAE vs. MMDPN	0.0017	0.1345	0.0745	0.0017	0.0007	0.0005
MMCNN vs. MMDPN	0.0140	0.0188	0.1330	0.0140	0.0090	0.0100

TABLE VIII T-test between two paradigms using various methods.

Method	ACC	SEN	SPEC	Youden	F1	AUC
MMDBN	<0.0001	0.0011	0.0031	<0.0001	<0.0001	<0.0001
MMSAE	<0.0001	0.0014	0.0157	<0.0001	<0.0001	<0.0001
MMCNN	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
MMDPN	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

In Table VI, there are significant differences across all the comparisons, except SEN and SPEC. There are significant differences in terms of Youden, considering the importance of both FPR and FNR. There are significant differences in metrics of ACC, F1, and AUC, which are the same as the text-based paradigm. In general, MMDPN in each paradigm performs better than MMDBN, MMSAE, and MMCNN, and it can improve the performance of BCI based on WI. Table VIII shows that there are significant differences between the two paradigms using the same method. It proves that the performance of WI detection is significantly improved in the VE-based paradigm.

Fig. 7 shows the ROC of two different paradigms with four kinds of feature encoding algorithms. The curves of MMDPN are close to the top left of the diagram, compared with MMDBN, MMSAE and MMCNN. It shows that MMDPN is an effective and robust feature encoding method. The curves of the VE-based paradigm are closer than those of the text-based paradigm, which verifies the effectiveness of VE as well.

IV. DISCUSSIONS

Movement impairment presents great difficulty for patients suffering from neurological disorder due to stroke, accidents or other brain injuries, where the transmission of neuromuscular signals to extremities is interrupted. Disabled patients can modulate EEG rhythms over sensorimotor cortex by imaging or observing motor movements, even though they could not execute the desired movements [26, 27]. Through these processes, the functional network of the brain can be

re-organized and brain plasticity is promoted, thereby yielding positive effect on motor function restoration [21]. Several evidences have indicated that the neuro-rehabilitation approach could be further enhanced by providing the feedback of brain activity [15, 16, 19]. BCIs based neuro-rehabilitation approaches have been developed to promote rehabilitation effect by producing external feedback of movement intentions that are identified using electroencephalogram (EEG) signals acquired non-invasively from the scalp [43-46]. To further enhance active engagement in extremities rehabilitation, virtual environment can be created to display an avatar representing the user, with virtual limbs driven by EEG signals via BCIs. Comparing with exploring this approach to upper-lower rehabilitation, it is not largely used for lower-limb rehabilitation [47]. Because the processing of EEG signals induced from lower-limb movements is more challenging as the signals are generated from deeper brain region and thus more obscure. For this reason, this study focuses on neuro-rehabilitation of lower-limbs with BCI. Through developing advanced EEG feature expression technology, lower-limbs movement intention can be decoded more accurately and thus generate feedback to promote rehabilitation performance. Lately, we have attempted to take this work to real-time rehabilitation robot system for patients to regain walking ability.

Although all subjects are instructed to imagine the kinesthetic experience of lower limb movements, the experimental processes are still accompanied by the observations of walking. To further reduce the effect of visual response on walking imagery, we analyze the EEG signals from the electrodes (e.g. FC1, FC2, FC5, FC6, Cz, C3, C4, T7, T8, CP1, CP2, CP5, and CP6) located approximately above the sensorimotor area. A similar conclusion can be reached. The average accuracy based on CSP is increased by 13.5% when using the VE-based paradigm. We also perform t-test at the 0.05 significance level. The statistical result also shows that the VE-based paradigm significantly outperforms the text-based paradigm (p-value = 0.0002).

In our current study, we only collect EEG data from 9 healthy people and 1 patient. Obviously, the dataset is quite limited, we will recurit more subjects to augment the EEG dataset for a more generalized and robust test. The current VE is only based on the computer display, which needs to improve the immersiveness. The head-mounted display (e.g., HoloLens) will be considered in our future work.

As a future work, we will attemp to enhance the performance via the latest deep learning method. For instance, we can exloit DPN properties and structures for peformance enhancement. Other deep learning methods such as the regularizaed SAE, the stacked deep learning structures can be integrated to furhter improve BCI performance.

V. Conclusions

In this study, we investigate the training paradigms of WI tasks. We evaluate the experimental results of the text-based paradigm and the VE-based paradigm. The experimental results demonstrate that the VE-based paradigm obtains a significant increase on classification results of the single-trial WI task. This paradigm has the potential to improve the reliability and robustness of WI-based BCIs. The preliminary results show that MMDPN has better classification performance, compared with MMSAE, MMDBN, and MMCNN. In our future work, the head-mounted display can be used to further enhance subjects' senses of reality and immersion. In addition, matching subject's characteristics to the training paradigm is likely to improve BCI performance. To optimize the training paradigm and the DPN algorithm for increasing the effectiveness of WI recognition, more EEG data should be collected from additional subjects, as well as more clinical data from the disabled who need lower limb rehabilitation.

References

[1] N. Birbaumer, "Breaking the silence: brain–computer interfaces (BCI) for communication and motor control," *Psychophysiology,* vol. 43, no. 6, pp. 517-532, 2006.

[2] J. R. Wolpaw *et al.*, "Brain-computer interface technology: a review of the first international meeting," *IEEE Trans Rehabil Eng,* vol. 8, no. 2, pp. 164-173, 2000.

[3] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clin Neurophysiol* vol. 113, no. 6, pp. 767-791, 2002.

[4] L. Bonnet, F. Lotte, and A. Lécuyer, "Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery," *IEEE Trans Comput Intell AI Games,* vol. 5, no. 2, pp. 185-198, 2013.

[5] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," *J Neural Eng,* vol. 10, no. 4, p. 046003, 2013.

[6] K. K. Ang *et al.*, "Brain-computer interface-based robotic end effector system for wrist and hand rehabilitation: results of a three-armed randomized controlled trial for chronic stroke," *Front Neuroeng,* vol. 7, p. 30, 2014.

[7] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Trans Biomed Eng,* vol. 63, no. 1, pp. 4-14, 2016.

[8] S. H. Jang, "The recovery of walking in stroke patients: a review," *Int J Rehabil Res,* vol. 33, no. 4, pp. 285-289, 2010.

[9] A. Dunsky, R. Dickstein, E. Marcovitz, S. Levy, and J. Deutsch, "Home-based motor imagery training for gait rehabilitation of people with chronic poststroke hemiparesis," *Arch Phys Med Rehabil,* vol. 89, no. 8, pp. 1580-1588, 2008.

[10] M. Severens, M. Perusquia-Hernandez, B. Nienhuis, J. Farquhar, and J. Duysens, "Using actual and imagined walking related desynchronization features in a BCI," *IEEE Trans Neural Syst Rehabil Eng,* vol. 23, no. 5, pp. 877-886, 2015.

[11] Y. Li, Y. Sun, F. Taya, H. Yu, N. Thakor, and A. Bezerianos, "Single trial EEG classification of lower-limb movements using improved regularized common spatial pattern," in *NER*, 2015, pp. 1056-1059: IEEE.

[12] C. Neuper, A. Schlögl, and G. Pfurtscheller, "Enhancement of left-right sensorimotor EEG differences during feedback-regulated motor imagery," *J Clin Neurophysiol,* vol. 16, no. 4, pp. 373-382, 1999.

[13] C. Neuper and G. Pfurtscheller, "Neurofeedback training for BCI control," in *Brain-computer interfaces*: Springer, 2009, pp. 65-78.

[14] R. Ron-Angevin and A. Díaz-Estrella, "Brain–computer interface: Changes in performance using virtual reality techniques," *Neurosci Lett,* vol. 449, no. 2, pp. 123-127, 2009.

[15] Z. Y. Chin, K. K. Ang, C. Wang, and C. Guan, "Online performance evaluation of motor imagery BCI with augmented-reality virtual hand feedback," in *EMBC*, 2010, pp. 3341-3344: IEEE.

[16] T. Ono, A. Kimura, and J. Ushiba, "Daily training with realistic visual feedback improves reproducibility of event-related desynchronisation following hand motor imagery," *Clin Neurophysiol,* vol. 124, no. 9, pp. 1779-1786, 2013.

[17] T. Sollfrank, D. Hart, R. Goodsell, J. Foster, and T. Tan, "3D visualization of movements can amplify motor cortex activation during subsequent motor imagery," *Front Hum Neurosci,* vol. 9, p. 463, 2015.

[18] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design," *Front Hum Neurosci,* vol. 7, p. 568, 2013.
[19] T. Kondo, M. Saeki, Y. Hayashi, K. Nakayashiki, and Y. Takata, "Effect of instructive visual stimuli on neurofeedback training for motor imagery-based brain–computer interface," *Hum Mov Sci,* vol. 43, pp. 239-249, 2015.
[20] S. Liang, K.-S. Choi, J. Qin, W.-M. Pang, Q. Wang, and P.-A. Heng, "Improving the discrimination of hand motor imagery via virtual reality based visual guidance," *Comput Methods Programs Biomed,* vol. 132, pp. 63-74, 2016.
[21] Z. Qiu *et al.*, "Optimized motor imagery paradigm based on imagining Chinese characters writing movement," *IEEE Trans Neural Syst Rehabil Eng,* vol. 25, no. 7, pp. 1009-1017, 2017.
[22] X. Liu *et al.*, "Performance Evaluation of Walking Imagery Training Based on Virtual Environment in Brain-Computer Interfaces," in *ISM*, 2017, pp. 25-30: IEEE.
[23] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection," *Expert Syst Appl,* vol. 38, no. 11, pp. 14314-14320, 2011.
[24] J. F. D. Saa and M. S. Gutierrez, "EEG signal classification using power spectral features and linear discriminant analysis: A brain computer interface application," in *Eighth Latin American and Caribbean Conference for Engineering and Technology*, 2010, pp. 1-7.
[25] D. Devlaminck, B. Wyns, M. Grosse-Wentrup, G. Otte, and P. Santens, "Multisubject learning for common spatial patterns in motor-imagery BCI," *Comput Intell Neurosci,* vol. 2011, p. 8, 2011.
[26] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans Rehabil Eng,* vol. 8, no. 4, pp. 441-446, 2000.
[27] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clin Neurophysiol,* vol. 110, no. 5, pp. 787-798, 1999.
[28] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Front Neurosci,* vol. 6, p. 39, 2012.
[29] R. Ortner, J. Scharinger, A. Lechner, and C. Guger, "How many people can control a motor imagery based BCI using common spatial patterns?," in *NER*, 2015, pp. 202-205: IEEE.
[30] Z. J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG," *Electroencephalogr Clin Neurophysiol,* vol. 79, no. 6, pp. 440-447, 1991.
[31] L. Yun, Z. Lifeng, and Z. Shujun, "A hand gesture recognition method based on multi-feature fusion and template matching," *Procedia Eng,* vol. 29, pp. 1678-1684, 2012.
[32] N. Robinson, A. P. Vinod, K. K. Ang, K. P. Tee, and C. T. Guan, "EEG-based classification of fast and slow hand movements using wavelet-CSP algorithm," *IEEE Trans Biomed Eng,* vol. 60, no. 8, pp. 2123-2132, 2013.
[33] B. Yang, H. Li, Q. Wang, and Y. Zhang, "Subject-based feature extraction by using fisher WPD-CSP in brain–computer interfaces," *Comput Methods Programs Biomed,* vol. 129, pp. 21-28, 2016.
[34] N. Lu, T. Li, X. Ren, and H. Miao, "A deep learning scheme for motor imagery classification based on restricted boltzmann machines," *IEEE Trans Neural Syst Rehabil Eng,* vol. 25, no. 6, pp. 566-576, 2017.
[35] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J Neural Eng,* vol. 14, no. 1, p. 016003, 2016.
[36] M. Veres, M. Moussa, and G. W. Taylor, "Modeling grasp motor imagery through deep conditional generative models," *IEEE Robot Autom Lett,* vol. 2, no. 2, pp. 757-764, 2017.
[37] Z. Yu *et al.*, "A Deep Convolutional Neural Network Based Framework for Automatic Fetal Facial Standard Plane Recognition," *IEEE J Biomed Health Inform,* 2017.
[38] R. Livni, S. Shalev-Shwartz, and O. Shamir, "An algorithm for training polynomial networks," *arXiv preprint arXiv:1304.7045,* 2013.
[39] R. Livni, D. Lehavi, S. Schein, H. Nachliely, S. Shalev-Shwartz, and A. Globerson, "Vanishing component analysis," in *ICML*, 2013, pp. 597-605.
[40] J. Shi, S. Zhou, X. Liu, Q. Zhang, M. Lu, and T. Wang, "Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset," *Neurocomputing,* vol. 194, pp. 87-94, 2016.
[41] Y. Li *et al.*, "An EEG-based BCI system for 2-D cursor control by combining Mu/Beta rhythm and P300 potential," *IEEE Trans Biomed Eng,* vol. 57, no. 10, pp. 2495-2505, 2010.
[42] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *IEEE Trans Neural Syst Rehabil Eng,* vol. 12, no. 2, pp. 258-265, 2004.
[43] M. Gomez-Rodriguez, J. Peters, J. Hill, B. Schölkopf, A. Gharabaghi, and M. Grosse-Wentrup, "Closing the sensorimotor loop: haptic feedback facilitates decoding of motor imagery," *Journal of neural engineering,* vol. 8, no. 3, p. 036005, 2011.
[44] K. K. Ang and C. Guan, "Brain–computer interface for neurorehabilitation of upper limb after stroke," *Proceedings of the IEEE,* vol. 103, no. 6, pp. 944-953, 2015.
[45] F. Pichiorri *et al.*, "Brain–computer interface boosts motor imagery practice during stroke recovery," *Annals of neurology,* vol. 77, no. 5, pp. 851-865, 2015.
[46] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Transactions on Biomedical Engineering,* vol. 63, no. 1, pp. 4-14, 2016.
[47] D. T. Jeffery, J. A. Norton, F. D. Roy, and M. A. Gorassini, "Effects of transcranial direct current stimulation on the excitability of the leg motor cortex," *Experimental brain research,* vol. 182, no. 2, pp. 281-287, 2007.