IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON NEURAL ENGINEERING INFORMATICS

Received April 24, 2019, accepted May 9, 2019, date of publication May 22, 2019, date of current version June 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2918251

Identification of Early Vascular Dementia Patients With EEG Signal

CHAO WANG[1,2], JIN XU[1,2], SONGZHEN ZHAO[3], AND WUTAO LOU[4]

[1]The Key Laboratory of Biomedical Information of Ministry of Education, Institute of Biomedical Engineering, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, China
[2]National Engineering Research Center for Healthcare Devices, Xi'an Jiaotong University, Xi'an 710049, China
[3]Department of Neurology, The First Affiliated Hospital of Medical College, Xi'an Jiaotong University, Xi'an 710061, China
[4]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong, China

Corresponding author: Jin Xu (xujin@mail.xjtu.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB1300303, in part by the National Natural Science Foundation of China under Grant 31271061, and in part by the grants from the Research Grants Council of the Hong Kong Special Administrative Region, China, under Project CUHK 14113214.

ABSTRACT Effort has been made to find biomarkers for vascular dementia (VaD). Nevertheless, the current findings are typically obtained through statistical tests of group level differences. In clinical practice, however, it is more common to perform individual level inferences, e.g., to determine if a subject is suffering from VaD, which cannot be resolved with statistical analysis. The goal of this study is to develop a method to effectively discriminate early VaD patients from normal controls by combining EEG features with machine learning methods. The EEG signals were recorded from a total of 15 VaD patients and 21 controls during a visual oddball task. Interregional directed connectivity was derived from directed transfer function (DTF) analysis and used as features in classification. Three machine learning methods, including linear discriminant analyais (LDA), error back-propagation (BP) neural network, and support vector machine (SVM) were used as classifiers, and their classification performance was compared. It was found that VaD patients can be effectively identified using the BP and SVM classifiers with high accuracy. In particular, when the SVM classifier was combined with feature selection by Fisher score, it reached an accuracy 86.11%, sensitivity 86.67%, and specificity 85.71%. The area under the curve (AUC, 0.854) indicates a good identification of VaD patients from the normal controls. Since the EEG is noninvasive, inexpensive, and widely available to use, the current study presents a novel clinical application of machine learning methods and could facilitate automatic screening and diagnosis of the VaD at an early stage in future.

INDEX TERMS Electroencephalogram (EEG), information flow, machine learning, vascular dementia.

I. INTRODUCTION

Vascular dementia (VaD) is one of the most common types of dementia worldwide, especially in Asian countries [1]. It is characterized by progressive decline in various cognitive functions, such as memory, attention and execution. Currently, no effective cure method has been found, so it is important to evaluate brain cognitive functions and diagnose patients with VaD at an early stage. With the increasingly larger population recent years, there is a more pressing need to facilitate early diagnosis and intervention of VaD. This has drawn an increasing interest and promoted a number of studies to investigate into VaD [2]–[6]. These studies utilized a wide spectrum of neuroimaging modalities, including electroencephalogram (EEG), magnetoencephalogram, functional magnetic resonance imaging (fMRI) and near-infrared spectroscopy, to find functional alterations in VaD. VaD patients were found to be abnormal in spontaneous [2] and evoked brain activity [5], [6], as well as in interregional brain connectivity patterns [3], [4]. These findings may serve as potential biomarkers for VaD identification.

However, traditionally, identifying abnormalities in VaD patients involves the use of between-group comparisons, typically through statistical tests. In spite of the insightful findings, this strategy has serious limitations: it constrains the obtained conclusions on group level, which may not be straightforward enough to allow to evaluate the diagnostic ability of the identified markers at individual level and thus can be problematic when applied to classify individual subjects in clinical practice [7], [8]. In addition, this approach is

The associate editor coordinating the review of this manuscript and approving it for publication was Zehong Cao.

intrinsically univariate: to extract potential markers that differ significantly between groups, a number of potential markers have to be compared iteratively. Although the univariate method is sensitive to localized effects, the information spatially distributed in the brain tends to be ignored [7], [9], thus making it difficult to evaluate the joint performance of a set of markers.

The shortcomings of traditional group comparison approach can be easily overcome by machine learning technique [7], [9]. Machine learning methods [10], [11] provide a powerful tool and a novel perspective to look into various brain diseases [7], [12], [13]. In the present study, supervised machine learning methods were used. For supervised classifications, the classifier tries to capture the underlying patterns and learn a rule that can maximally differentiate different groups from the training dataset with known class labels. This learned rule is then used to predict the class membership that a new and previously unseen sample belongs to. The more accurately the classifier captures the underlying patterns, the more likely it would distinguish future samples correctly and achieve better performances. Machine learning has been successfully applied to the study of a variety of aging-related brain disorders (for review, see [7]), e.g., mild cognitive impairment (MCI) [14]–[16], Alzheimer's disease (AD) [16]–[18] and Parkinson's disease (PD) [19], as well as normal aging [20].

Since VaD population is large, it will be very important and beneficial if VaD can be screened automatically at an early stage. However, only untill very recently, machine learning methods are combined with neuroimaging markers to identify VaD patients from the normal aging controls. To our knowledge, the only study that applied machine learning technique to VaD identification was performed by [21]. In this study, whole brain connectivity patterns were extracted from resting state fMRI and used as features in SVM classification to identify and differentiate VaD patients from healthy individuals. Compared to fMRI, EEG signal is closely related to postsynaptic potential generated in the cerebral cortex and hence more directly reflects the underlying neuronal electrical activity. EEG is also known for its excellent temporal resolution, which makes it an ideal tool to investigate the dynamic features of neuronal activity. A variety of metrics were proposed to characterize EEG signal (e.g., fuzzy entropy [22]) and used for the investigation of patients in clinical applications, e.g. in migraine patients [23]. Therefore, it would be appealing and of potentially clinical relevance if EEG proves to help distinguish patients with VaD from controls.

Actually, EEG has already been widely used and is traditionally considered as a standard tool in the field of brain-computer interface [24], where along with machine learning methods, EEG features are used to identify or discriminate different types of mental activities or states and serve as a control signal to perform spefic tasks, such as, to guide the movement of external mechanical equipments. This is gaining a growing interest and application particularly in rehabilitation engineering (see review [25]). Also, in previous studies, EEG signal was combined with machine learning algorithms and used in a variety of clinical applications, e.g., to identify patients with AD [26]–[30], PD [31], [32] and depression [32]–[34] from normal controls, to detect the cholinergic intervention in healthy adults [27], to predict the response to treatment in subjects with depression [35], [36], to predict seizure onset or discriminate normal, preictal, and seizure in epilepsy patients [37]–[41], and to predict the progression to AD at the MCI stage [42]. They demonstrated the potential power of EEG features in clinical diagnosis and treatment or medication assessment. In most of these studies, however, the features adopted were spectrum- or complexity-related measures, which are based on EEG signal from single channels. Although these measures provide important information regarding local brain activity, the interaction between brain regions is ignored, which is now considered as an basic principle of brain [43]. Brain connectivity has gained a very intense attention recently. Furthermore, since the information in the brain is usually processed in sequential order, the activity of one brain region may be exclusively driven by another region. In this regard, the interaction between spatially distant brain regions is intrinsically directed. Taking these into consideration, EEG and a directed brain connectivity measure (directed transfer function, DTF [44]) are used in this study. Actually, DTF was already adopted to discriminate interictal periods form ictal periods in seizure patients and a high performance was achieved [39], [40]. Instead of resting state EEG, we choose task EEG since the patients in our study are at their early stage and their cognitive decline is relatively mild. For these patients, cognitive deficits are more easily exposed in cognitive tasks. Also, the relatively simple task, a typical oddball paradigm, ensures patients can complete without much difficulty.

In the present study, our goal is to find a way to effectively discriminate patients from the normal controls. This can greatly facilitate the automatic identification and screening of VaD patients in clinical application in the future. We combined brain connectivity and machine learning to identify early VaD patients. Based on DTF, the information flow among brain regions were extracted and served as the feature set in classification. Three supervised machine learning methods, i.e., linear discriminant analysis (LDA), back-propagation neural network (BP) and support vector machine (SVM), were then used as classifiers to differentiate early VaD patients and the controls, and classification performance was evaluated and compared. Primarily, we aimed to test the hypothesis that VaD patients and the controls could be reliably separated using EEG connectivity patterns as features and machine learning methods as classifiers.

II. METHODS

A. SUBJECTS

Thirty-six subjects participated in the study: 15 early VaD patients (VaD group, 13 males) and 21 normal subjects (CTR group, 10 males). VaD patients were enrolled from the First

Affiliated Hospital of Medical College of Xi'an Jiaotong University and control subjects were recruited from Xi'an Jiaotong University. All subjects participated in a series of standardized tests including the Mini-Mental State Examination (MMSE), the Clinical Dementia Rating (CDR) scale, the Hospital Anxiety and Depression Scale (HADS), and the Instrumental Activities of Daily Living (IADL) scale. The patients were diagnosed by expert clinicians according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) criteria, and the MRI revealed evidence of large-vessel stroke, or multiple subcortical lacunar infarcts, and/or extensive white matter lesions in each of them. All patients had experienced the onset of cognitive impairment after a clinical stroke. Patients with a family history of AD were excluded. None of our patients used medication that was expected to influence EEG recordings. Table 1 shows the demographic information of the two groups. Statistical tests showed that there were no significant differences between the two groups in age or education ($p > 0.05$), but the gender distribution was significantly different between patients and the elderly ($p = 0.033$).

TABLE 1. Demographic information of subjects.

	VaD ($n = 15$)	Control ($n = 21$)	p
Age	70.87±9.59	69.90±6.28	0.737[a]
Gender	13/2	10/11	0.033[b]
Education	9.47±4.17	11.38±3.71	0.191[c]
MMSE	24.40±4.24	28.33±1.24	--

Values are presented as mean±SD (standard deviation)
[a]Independent sample t test
[b]Pearson χ^2 test
[c]Mann-Whitney U test

All subjects were right-handed, had normal or corrected-to-normal visual acuity, and no color blindness. After formal approval by the local ethics committee, informed written consent was obtained from all subjects.

B. EEG RECORDING AND PREPROCESSING

A classical visual oddball paradigm was used in the study (Fig. 1). The stimulus sequence contained 300 stimuli and was composed of two types of stimuli, target stimuli (green circles) and nontarget stimuli (red circles), with a probability of 0.20 and 0.80, respectively. Each stimulus randomly appeared at the center of the monitor and lasted for 80 ms with a random inter-stimulus interval of 1000–1200 ms. Each stimulus had a diameter of 3.5 cm, and was presented against a black background on a standard CRT monitor. During inter-stimulus interval, a yellow cross was displayed at the center of the monitor to help subjects keep concentrated. Subjects viewed the stimuli from a distance of about 60 cm. They were instructed to respond to target stimuli by pressing one specified button on the keypad as accurately as possible while ignoring nontarget stimuli. When responding to target

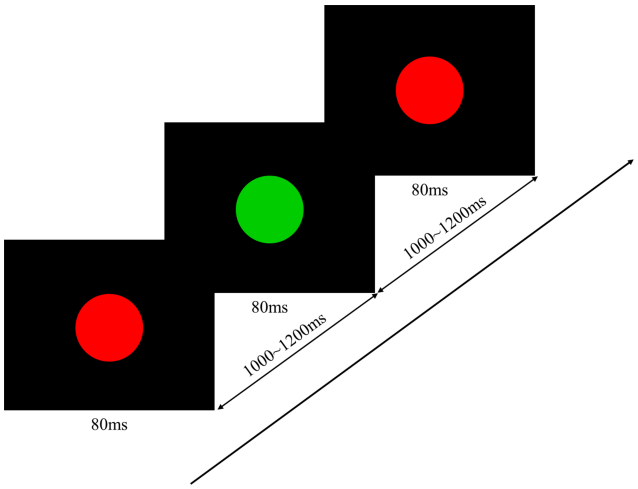

FIGURE 1. Experiment design.

stimuli, subjects were asked to remain quiet at the same time and only move their right index finger to minimize the contribution of muscle artifacts to EEG recordings. Reaction time and accuracy were recorded at the same time.

EEG data were simultaneously recorded from 32 Ag/AgCl electrodes on the scalp using Neuroscan EEG acquisition system (Neuroscan Inc., USA). Electrodes were positioned according to the 10–20 international electrode placement system. A ground electrode was placed on the forehead, and linked mastoid electrodes were used as the reference. Vertical and horizontal electrooculograms (VEOG and HEOG) were also recorded with two pairs of bipolar electrodes in vertical and horizontal directions. During the recording, electrode impedance was kept below 5 kΩ. EEG data were continuously acquired with a sampling rate of 500 Hz and a frequency band of direct current (DC) to 100 Hz.

EEG data were firstly corrected for ocular artifacts using a threshold reduction algorithm (Neuroscan Inc., USA). Then, EEG data were visually scanned for contamination by muscles or other kinds of artifacts, and bad EEG periods were rejected. EEG data were then band-pass filtered at 0.5–30 Hz with a zero-phase shift FIR filter. Finally, each subject's data were segmented into epochs time-locked to stimulus onset. The epoch was defined as the time period from 0 to 400 ms after stimulus onset. Since the information related to the detection of novel stimuli is more relavant for the completion of this task, only correct responses to target stimuli were segmented and used for further analysis. In total, 42.27 ± 7.43 and 45.33 ± 7.64 (mean ± SD) epochs were available for the VaD and CTR groups, respectively. There was no significant difference in the number of epochs between two groups ($p = 0.238$).

C. FEATURES

In this study, interregional connections served as features and were entered into classifiers for further analysis. We used directed transfer function (DTF) [44] to measure the information flow within the brain. The rationale behind DTF is

sources contain information that could help to better predict the behavior of targets. DTF is based on multivariate autoregressive (MVAR) model, which quantitatively describes the generation and evolution of a signal. The EEG signal from M channels at sampling instant n can be represented by a vector: $X(n) = (X_1(n), X_2(n), \ldots, X_M(n))^{\mathrm{T}}$, in which "T" denotes transposition. Mathematically, a p-order MVAR model has the following form:

$$X(n) = \sum_{k=1}^{p} A_k X(n-k) + E(n) \tag{1}$$

in which A_k $(k = 1, 2, \ldots, p)$ denotes the model coefficient matrix at time lag k, and $E(n)$ is assumed to be an white noise uncorrelated with $X(n)$. In this study, all single-trial EEG data was used to estimate the MVAR model coefficient matrices [45] via Levinson–Robinson–Wiggins algorithm [46], [47]. The optimal model order was chosen as 7 with Akaike Information Criterion [48]. After transforming equation (1) into z-domain and making a substituting: $z^{-1} = e^{-j2\pi f/fs}$, it yields the following:

$$X(f) = E(f)H(f) \tag{2}$$

$$H(f) = (\sum_{k=0}^{p} A_k' e^{-j2\pi kf/f_s})^{-1} \tag{3}$$

where f denotes frequency in Hz, f_s is the sampling frequency, $A_0{}' = \boldsymbol{I}$($\boldsymbol{I}$ is an M-by-M identity matrix), $A_k{}' = -A_k (k = 1, 2, \ldots, p)$. $X(f)$ and $E(f)$ are the representations of $X(n)$ and $E(n)$ in frequency domain, respectively. $H(f)$ is called transfer function. Then, the information flow from channel j to i at frequency f is defined by DTF as follows:

$$DTF_{ij} = \frac{\left|H_{ij}(f)\right|}{\sqrt{\sum_{j=1}^{M} \left|H_{ij}(f)\right|^2}} \tag{4}$$

in which $H_{ij}(f)$ is the (i, j)-th element of matrix $H(f)$. DTF is defined in frequency domain and ranges from 0 to 1: for the information flow from channel j to i, a value of zero implies no flow of information, whereas a value of one indicates that the brain activity at channel i is exclusively driven by the activity at channel j. In the present study, EEG signal from the following 20 electrodes were extracted and used in the subsequent analysis: frontal (Fp1, Fp2, F7, F3, Fz, F4 and F8), left temporal (T7), central (C3, Cz and C4), right temporal (T8), parietal (P7, P3, Pz, P4 and P8), and occipital (O1, Oz and O2). The information flow among these electrodes was evaluated from 0.5 to 30 Hz with a resolution of 0.5 Hz. DTF values were further integrated into five frequency bands, i.e., δ (0.5–3.5 Hz), θ (4–7 Hz), α_1 (8–10 Hz), α_2 (11–13 Hz) and β (14–30 Hz). Since there was a significant difference between VaD patients and the controls in the gender distribution ($p = 0.033$), we regressed out the gender effect from each feature by using a linear regression model before entering into classification.

D. CLASSIFICATION

In the present study, three supervised machine learning methods, i.e., LDA, BP and SVM, served as classifiers respectively. Their classification performance was evaluated and compared. To assess the classifier performance, we combined feature selection and leave one out cross validation (LOOCV). In each fold of LOOCV, one subject serves as testing dataset and the remaining serve as training dataset. Feature selection was performed within LOOCV to select the features that are more relavant to the classification. In this study, an independent sample t-test (two-tailed) was used for the selection of significantly different features between groups ($p < 0.05$). It is worth noting here that feature selection was performed on training data to avoid being contaminated by the information from the testing data, and the same set of features were then extracted from testing data. The main steps in the classification of early VaD patients were illustrated in Fig. 2. In the present study, the classification was implemented with MATLAB and LIBSVM toolbox [49].

1) LDA

LDA is a classical method in machine learning, and is widely idea underlying LDA is to find a weight vector to project original features into a new space, in which it would be easier to separate samples belonging to different classes. The method used by LDA is to make the samples in the same class concentrated enough and the samples that belong to different classes as distant as possible. As such, LDA finds the weight vector by maximizing the ratio of between-class distance to within-class variance.

2) BP

BP is one of the classical learning algorithms in artificial neural network (ANN), which models the behavior of biological neurons in the real world by adopting similar ways of information processing. BP algorithm aims to solve the problems in a multilayer feedforward ANN. In multilayer feedforward ANN, it consists of an input layer, several hidden layers and an output layer. While the input and output layers act as interfaces to receiving the input and producing the final output respectively, hidden layers receive information from the previous layer, process it and relay it to the next layer. Each neuron processes information by modeling it as some specific function of the weighted sum of its input. In addition to this forward propagation of input information, for BP algorithm, it additionally consists of error propagation in a reverse direction in the network. The error is defined as the summed squared deviation between the real output and the expected output. BP algorithm describes the error as a function of all the weights in the network and finds the optimal weights by minimizing the error. In this study, we used a 3-layer network, in which only one hidden layer is included. We set the number of hidden and output neuron as 16 and 1.

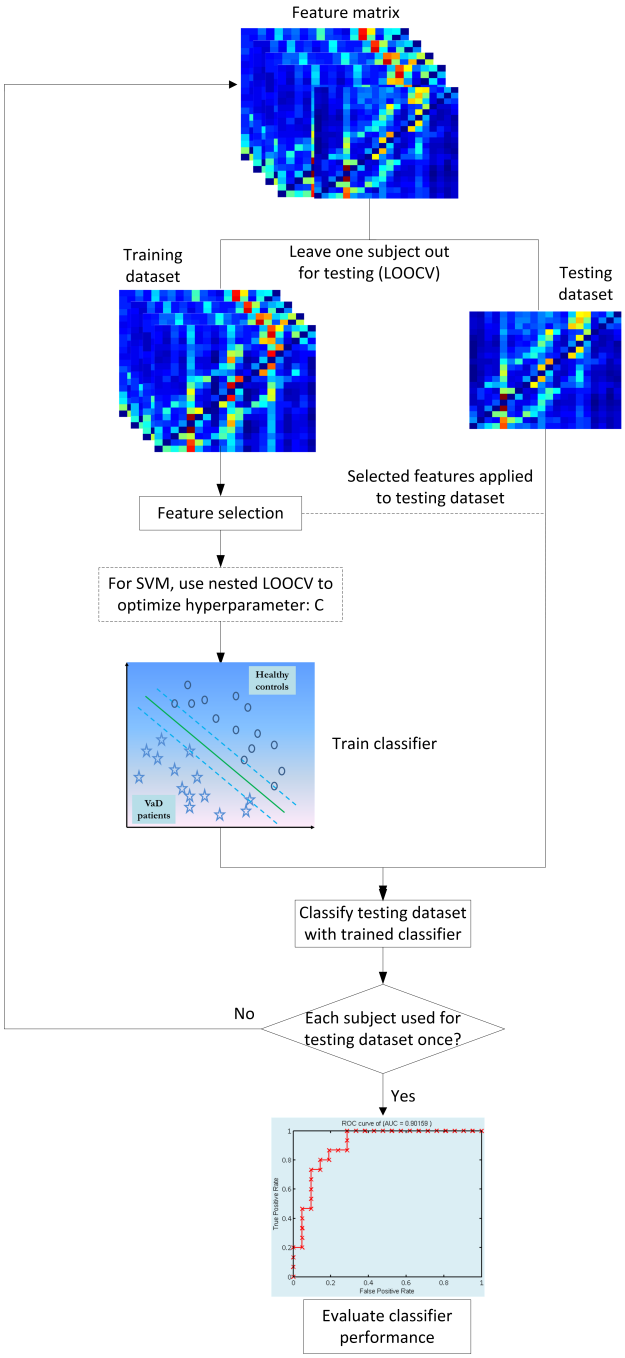

FIGURE 2. The main pipeline for the classification of early VaD patients.

3) SVM

SVM seek to find a hyperplane or a boundary that best separates the samples belonging to different classes according to certain criteria. For SVM, the best hyperplane is defined as the one that achieves maximum margin between two classes; the samples that are closest to the separating hyperplane are referred to as support vectors. Given the large number of features in this study, we used a linear kernel to lower the risk of overfitting. Another advantage of a linear kernel is that, under the settings of this study (binary classification and class labels: VaD: $+1$, CTR: -1), a positive and a negative weight would indicate that the corresponding feature was likely to be strengthened or weakened respectively in VaD patients, compared to the controls [13], [20], [50], [51]. Since SVM has a hyperparameter C, we used an approach known as 'nested leave-one-out cross-validation' to assess the performance [10], [11]. In this approach, within each LOOCV fold, another LOOCV is firstly applied to the leave-one-out training samples to optimize C in a range of parameters C: $[2^{-15}, 2^{-14}, \ldots, 2^{15}]$ (via grid search method); then with this C, a SVM model is learned and tested with training and testing dataset respectively. Before entering into classification, each feature in training dataset was separately scaled to [0, 1] [18], and the parameters obtained from training dataset were then used to normalize the same features in testing dataset [16].

TABLE 2. Performance of different classifiers.

	Acc	*Sens*	*Spec*	*AUC*
LDA	80.56%	80.00%	80.95%	0.8889
BP	83.33%	80.00%	85.71%	0.9397
SVM	72.22%	80%	66.67%	0.7556
SVM[1]	86.11%	86.67%	85.71%	0.8540

[1]Feature selection was performed by using Fisher score.

4) PERFORMANCE EVALUATION

The performance of classifier was evaluated with classification accuracy (*Acc*) and the area under the receiver operating characteristic curve (*AUC*). In this study, classification accuracy is defined as the percent of correctly classified subjects and *AUC* measures the probability that a classifier will assign a higher value to a randomly chosen positive instance than to a randomly chosen negative one [15]. In addition, we also employed sensitivity (Sens) and specificity (Spec):

$$Sens = \frac{TP}{TP + FN} \tag{5}$$

$$Spec = \frac{TN}{TN + FP} \tag{6}$$

where TP and FN denoting true positive and false negative (number of correctly and incorrectly labeled VaD patients), and TN and FP denoting true negative and false positive (number of correctly and incorrectly labeled CTR subjects). They are the percentage of correctly identified VaD and CTR subjects in this study, respectively. Each of these evaluation indexes were calculated as the average across outer LOOCV folds.

III. RESULTS

Table 2 and Fig. 3 showed the classification performance of three classifiers.

SVM is considered to have high generalization ability [10], [40], [52] and good capability to cope with high-dimensional

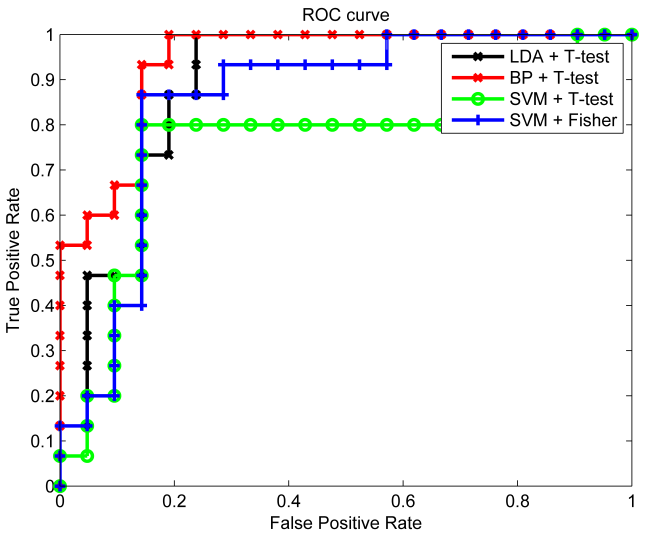

FIGURE 3. ROC of BP and SVM classifiers.

data [10], [53], [54]. Compared to BP, SVM classifier is able to cope with classification problems of unbalanced samples more effectively [40], which is our case. In previous studies, SVM has also been demonstrated to yield superior performance to a number of other approaches, including BP [55]. Therefore, we tried to improve its classification performance. Here, another feature selection method Fisher score (FS) was adopted [17] to select the most informative features within each LOOCV fold. Fisher score is closely related to LDA and is a method to evaluate the potential of a feature to reflect the difference of distinct classes. It is defined as the summed differences between the mean of each class and the mean of all samples devided by the summed within-class deviations across all classes. A larger Fisher score of a feature indicates its high possibility to contain useful information for the discrimanination of different classes. In this study, all features were ranked in descending order according to their Fisher scores and a certain percentage of top-ranking features were selected for use in classification. For a lack of prior knowledge on the number of features, we selected the percent of top-ranking features in a broad range from 0.1% to 100%, with an increment of 0.1%. The result where SVM classifier achieved the highest accuracy was reported: the dimension of features was 931 (49% of the whole feature set). The result (Table 2 and Fig. 3) showed an accuracy 86.11%, a sensitivity 86.67%, a specificity 85.71% and an AUC 0.8540. Yet, we noted that the performance was very stable across a large range after the number of features reached 49%.

IV. DISCUSSION

In the present study, based on DTF, the directed connections among brain regions were extracted from EEG signals recorded from early VaD patients and healthy controls. They were used as features for the classification of early VaD patients and the controls with different classifiers (LDA, BP and SVM). The results showed a high accuracy for BP and SVM classifiers and the best performance was achieved by SVM combined with FS, which demonstrated the features' power in characterizing and separating the two groups.

A. CLASSIFICATION PERFORMANCE

The exploration of machine learning methods in disease identification and prediction is a relatively new field, but has attracted much attention. Recent years, a number of studies have utilized machine learning methods to investigate into brain disorders. However, VaD, as one of main types of dementia in older adults, is rarely studied. Previously, Zhang *et al.* extracted brain connectivity patterns from resting state fMRI and these features were used to differentiate VaD patients from the controls combined with SVM classifier [21]. Cheveigne *et al.* tried to predict the development of preclinical symptomatic small vessel disease to VaD within 5 years using morphometric measures extracted from structural MRI images and SVM [56]. In our study, EEG signal was analyzed, considering its direct reflection of brain neuronal electric activity. We further evaluated their discriminative ability in the identification of VaD patients and compared the performance of three commonly used machine learning methods: LDA, BP and SVM.

We found the LDA, BP and SVM classifiers could identify VaD patients accurately with the accuracy, sensitivity and specificity >80% and AUC > 0.85, which indicates a good identification of VaD patients. This suggests connectivity patterns can effectively describe and probe into the cognitive changes in VaD patients' brain. In particular, SVM performed the best among three classifiers. In this study, SVM was combined with both two different feature selection methods: *t* test and Fisher score. Although the former displayed a less satisfactory performance, approximately 80% of VaD patients and two thirds of the controls were still correctly identified, resulting in an overall accuracy above 70%. For the latter, it performed even better than LDA and BP: over 85% of VaD patients and over 85% of control subjects were classified correctly, resulting in an overall accuracy over 85% and an AUC larger than 0.85. An inspection of SVM classification with the latter feature selection method further demonstrated the robustness of its performance to the number of features: the performance became almost very stable across a large range of feature dimensions after the reported 49%. This may relect SVM's better capability to capture the abnormalities in the brain function of VaD patients. On the other hand, the superior performance of Fisher score compared to *t* tests also suggests that, the features that could contribute to a successful classification may not be the same as those statistically significant ones. Actually, most of these features (over 2/3) were not of statistical significance in our case. This further validated the rationale for the use of machine learning methods as a complementary way to the traditional statistic methods.

To our knowledge, our study was the first to apply BP algorithm to VaD identification and a comparable classification performance was obtained in VaD patients as compared to

the results in literature in AD and MCI patients, using various neuroimaging features (e.g., [14], [16], [18], [57]–[59]). For example, with resting state fMRI, Dybra *et al.* reported an accuracy of 74%, a sensitivity 82%, a specificity 64% and an AUC 0.80 in discriminating AD patients and the healthy controls [18]; in Cui *et al.*, SVM was combined with MRI features to predict the development of cognition decline within 2 years in MCI subjects and similar results were shown: an accuracy 78.51%, sensitivity 73.33%, specificity 79.75%, and AUC 0.841 [57]; in Zhang *et al.*, SVM was shown to classify AD and MCI patients from normal controls with about 70% $\sim$ 90% accuracy, sensitivity and specificity, using either MRI or other biomarkers [16]; Long *et al.* extracted MRI features based on three different atlas for use in the SVM classification of MCI patients, and all atlas reported accuracies, sensitivities, specificities, and AUC between 80% and 90% [58]. In contrast with SVM, the discriminant power of BP is much less explored. Therefore, we demonstrated a promising potential for VaD screening and automatic identification in real scenario. In addition, in most of these stuides, MRI was used as a imaging method to extract features. Compared to MRI, EEG is well known for its low expense and wide avaliability, which further highlights the future application of our method in reality.

In previous studies, EEG signal combined with machine learning methods displayed its potential in a variety of clinical applications [21], [26]–[35], [37]–[42]. For example, Simpraga et al. used spectral power and amplitude envelope of EEG signal to distinguish AD patients from healthy controls and an accuracy of 73%, sensitivity 73%, and specificity 70% was obtained [27]. Similarly, using EEG spectral mearues, Trambaiolli et al. obtained an accuracy of 91.18%, sensitivity and specificity around 90% for the identification of AD patients with SVM [30]. Lehmann et al. combined multiple spectral measures and machine learning algorithms: for the separation of mild AD and controls, a sensitivity of 85% and a specificity of 78% was reached, and for the separation of moderated AD and controls, a sensitivity of 89% and a specificity of 88% was obtained [29]. Hosseinifard et al. extracted spectral power and nonliear features from EEG signal and they achieved an accuracy of 83.3% in discriminating depression patients and normal controls [33]. Using EEG power at baseline, Cao et al. classified two groups of depression patients: the responders and the non-responders to ketamine, with an average accuracy 81.3%, a sensitivity 82.1% and a specificity 91.9% [36]. In most of these studies, however, spectral information and other single-channel based measures were used. Similar to Wang *et al.* [39], [40], in our study, we used brain connectivity measure (DTF) to take the interactions between brain regions into consideration. The performance was comparable to those in the above studies. Actually, in the studies of Trambaiolli [26] and Khodayari-Rostamabad [35], an improved classification performance was found compared to spectral features for the identification of AD patients [26] and functional connectivity (coherence) between brain regions dominated the most discriminating features including spectral measures in the prediction of their response to medication in depression patients [35].

As regards to the application in VaD patients, there is some methodological difference concerning the use of SVM classifier compared to Zhang *et al.* [21],. While the authors used radial basis function (RBF) as the kernel, which maps the input features nonlinearly into a much higher dimensional space, a linear kernel was employed in our study. A linear kernel is preferred when the dimension of features exceeds the number of samples [7], [17], [51], which is our case. A linear kernel also means a one-to-one correspondence between features and weights, and makes it straightforward to investigate the contributions of individual features to classification performance. In previous studies, based on resting-state fMRI connectivity patterns, linear kernels were used in SVM classifier to identify neuroimaging markers of diseases like stroke [60] and PD [19], as well as typical characteristic like handedness [61]. Comparable classification performance to ours was obtained: an accuracy of 82.6% with a sensitivity of 80% and a specificity of 85% in stroke [60], an accuracy of 93.6% with a sensitivity of 90.5% and a specificity of 96.2% in PD [19], and an accuracy of 86.2% with a sensitivity of 83.3% and a specificity of 88.9% in handness [61].

An additional control study was performed by comparing with the classification performance of network parameters derived from brain network analysis. The following network topological parameters were used due to their wide use in previous brain network studies, including global (clustering coefficient, characteristic path length and their normalized version, global and local efficiency, and smallworldness) and local (the indegree, outdegree, global and local efficiency, and betweenness centality of each node) ones [62]. It was found that the classification performance of the features in our study was superior to that of network features (accuracy 80.56%, sensitivity 80%, specificity 80.95% and AUC 0.8476). Therefore, the connectivity features used in our study may indeed better capture the patological changes underlying early VaD patients' brain, and when combined with machine learning methods, can identify early VaD patients in a more effective way.

Overall, the results in this study indicate that connectivity measures extracted from EEG signals can very well reveal important neurophysiological information of altered brain of VaD patients and this information can be further utilized to characterize VaD patients and distinguish them from normally aging brains.

B. LIMITATIONS AND FUTURE WORK

In the analysis of EEG signal, the choice of reference is an important issue to consider. The reference provides a baseline for the measurement of EEG signal. Currently, many reference methods exist, e.g., average reference and the so-called 'reference electrode standadization technique' (REST) [63]. Average reference assumes the head as a sphere and proposes to use the average potential of all electrodes on the sphere

as the baseline. However, this is not the case in reality since the head is not spherical and it is usually only sampled on the upper surface (i.e., scalp) by EEG electrodes. For REST, accurate head models incorporating brain tissues' biophysical properties would be very crucial [64], which so far are still poorly understood. It has been demonstrated that estimation of brain connectivity can be remarkably impacted by the reference choice [64]. For DTF, since it is based on the estimation of covariance matrix, Blinowska (the author who proposed DTF) suggests that introducing additional correlations between EEG channels by reference methods should be avoided, e.g., by average referencing [65]. When using DTF, it is recommend that EEG signals be referenced with respect to the channels not involved in the model estimation [65]. Therefore, in this study, we used the commonly used linked mastoids as reference.

Despite the insightful findings, there are several limitations in the current study. Firstly, due to practical reasons, the sample size (the number of subjects) is small. However, the leave-one-out cross-validation we employed is considered to be able to provide a relatively less biased estimate of classification performance [10]. Future studies on larger samples of dataset or independent datasets [66], [67] will be needed to confirm the findings of our study. A related issue is the unbalanced sample sizes between two classes of subjects. Unbalance is likely to make the classifer tend to learn the characteristics of the class with larger size and therefore bias the classification performance. However, it can be found in our study that in all classifiers, the sensitivity is comparable and even higher than specificity, which means the classifiers may indeed learn the rules to separate VaD patients from controls. Reaching a balance in sample size would greatly benefit the study of VaD. Secondly, VaD is quite heterogeneous, containing some different subtypes, such as small-vessel and large-vessel disease [68]. We were not able to further divide them due to the relatively small samples. Again, it would be advantageous with larger samples, when classifiers can be constructed to further differentiate between subtypes of diseased populations. This will be appealing and of great interest to clinicians.

V. CONCLUSIONS

This study investigated the power of EEG connectivity patterns in discriminating early VaD patients from healthy controls with multiple machine learning methods, i.e., LDA, BP and SVM. Using either BP or SVM as classifers, especially SVM, the connectivity features achieved a high classification performance, suggesting they have good capability in characterizing the pathological changes in VaD patients' brain. Since EEG is noninvasive, inexpensive and widely avaliable to use, the method proposed in this study is of practical significance for the application in VaD patients, and may have important implications both for the automatic screening and for the clinical diagnosis of VaD patients at an early stage in the future.

REFERENCES

[1] D. W. Desmond, "Vascular dementia," *Clin. Neurosci. Res.*, vol. 3, pp. 437–448, Mar./Apr. 2004.
[2] E. C. W. van Straaten, W. de Haan, H. de Waal, P. Scheltens, W. M. van der Flier, F. Barkhof, T. Koene, and C. J. Stam, "Disturbed oscillatory brain dynamics in subcortical ischemic vascular dementia," *BMC Neurosci.*, vol. 13, p. 85, Jul. 2012.
[3] E. C. W. van Straaten, J. den Haan, H. de Waal, W. M. van der Flier, F. Barkhof, N. D. Prins, and C. J. Stam, "Disturbed phase relations in white matter hyperintensity based vascular dementia: An EEG directed connectivity study," *Clin. Neurophysiol.*, vol. 126, pp. 497–504, Mar. 2015.
[4] J. Xu, W. Lou, S. Zhao, and C. Wang, "Altered directed connectivity in patients with early vascular dementia during a visual oddball task," *Brain Topogr.*, vol. 28, pp. 330–339, Mar. 2015.
[5] S. Tak, S. J. Yoon, J. Jang, K. Yoo, Y. Jeong, and J. C. Ye, "Quantitative analysis of hemodynamic and metabolic changes in subcortical vascular dementia using simultaneous near-infrared spectroscopy and fMRI measurements," *NeuroImage*, vol. 55, pp. 176–184, Mar. 2011.
[6] Z.-Y. Sun, J.-H. Wang, J.-L. Sun, T.-J. Wang, L. Li, Y.-H. Dong, J. Wu, W.-Z. Cui, Y.-J. Wu, and P.-Y. Lu, "Magnetoencephalography assessment of evoked magnetic fields and cognitive function in subcortical ischemic vascular dementia patients," *Neurosci. Lett.*, vol. 532, pp. 17–22, Jan. 2013.
[7] G. Orrù, W. Pettersson-Yeo, A. F. Marquand, G. Sartori, and A. Mechelli, "Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review," *Neurosci. Biobehav. Rev.*, vol. 36, pp. 1140–1152, Apr. 2012.
[8] L.-L. Zeng, H. Shen, L. Liu, L. Wang, B. Li, P. Fang, Z. Zhou, Y. Li, and D. Hu, "Identifying major depression using whole-brain functional connectivity: A multivariate pattern analysis," *Brain A, J. Neurol.*, vol. 135, no. 5, pp. 1498–1507, 2012.
[9] R. C. Craddock, P. E. Holtzheimer, III, X. P. Hu, and H. S. Mayberg, "Disease state prediction from resting state functional connectivity," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1619–1628, 2009.
[10] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *NeuroImage*, vol. 56, no. 2, pp. 387–399, 2011.
[11] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: A tutorial overview," *NeuroImage*, vol. 45, pp. S199–S209, Mar. 2009.
[12] S. Klöppel, A. Abdulkadir, C. R. Jack, Jr., N. Koutsouleris, J. Mourão-Miranda, and P. Vemuri, "Diagnostic neuroimaging across diseases," *NeuroImage*, vol. 61, pp. 457–463, Jun. 2012.
[13] F. X. Castellanos, A. Di Martino, R. C. Craddock, A. D. Mehta, and M. P. Milham, "Clinical applications of the functional connectome," *NeuroImage*, vol. 80, pp. 527–540, Oct. 2013.
[14] C.-Y. Wee, P.-T. Yap, W. Li, K. Denny, J. N. Browndyke, G. G. Potter, K. A. Welsh-Bohmer, L. Wang, and D. Shen, "Enriched white matter connectivity networks for accurate identification of MCI patients," *NeuroImage*, vol. 54, no. 3, pp. 1812–1822, 2011.
[15] C.-Y. Wee, P.-T. Yap, D. Zhang, K. Denny, J. N. Browndyke, G. G. Potter, K. A. Welsh-Bohmer, L. Wang, and D. Shen, "Identification of MCI individuals using structural and functional connectivity networks," *NeuroImage*, vol. 59, no. 3, pp. 2045–2056, 2012.
[16] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 55, no. 3, pp. 856–867, 2011.
[17] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory," *Clin. Neurophysiol.*, vol. 126, pp. 2132–2141, Nov. 2015.
[18] M. Dyrba, M. Grothe, T. Kirste, and S. J. Teipel, "Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM," *Hum. Brain Mapping*, vol. 36, pp. 2118–2131, Jun. 2015.
[19] Y. Chen, W. Yang, J. Long, Y. Zhang, J. Feng, Y. Li, and B. Huang, "Discriminative analysis of Parkinson's disease based on whole-brain functional connectivity," *PLoS ONE*, vol. 10, Apr. 2015, Art. no. e0124153.
[20] T. B. Meier, A. S. Desphande, S. Vergun, V. A. Nair, J. Song, B. B. Biswal, M. E. Meyerand, R. M. Birn, and V. Prabhakaran, "Support vector machine classification and characterization of age-related reorganization of functional brain networks," *NeuroImage*, vol. 60, pp. 601–613, Mar. 2012.
[21] D. Zhang, B. Liu, J. Chen, X. Peng, X. Liu, Y. Fan, M. Liu, and R. Huang, "Determination of vascular dementia brain in distinct frequency bands with whole brain functional connectivity patterns," *PLoS ONE*, vol. 8, Jan. 2013, Art. no. e54512.

[22] Z. Cao and C.-T. Lin, "Inherent fuzzy entropy for the improvement of EEG complexity evaluation," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 1032–1035, Apr. 2018.
[23] Z. Cao, C.-T. Lin, C.-H. Chuang, K.-L. Lai, A. C. Yang, J.-L. Fuh, and S.-J. Wang, "Resting-state EEG power and coherence vary between migraine phases," *J. Headache Pain*, vol. 17, p. 102, Nov. 2016.
[24] L. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update," *J. Neural Eng.*, vol. 15, no. 3, p. 031005, 2018.
[25] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nature Rev. Neurol.*, vol. 12, p. 513, Aug. 2016.
[26] L. R. Trambaiolli, A. C. Lorena, F. J. Fraga, P. A. M. Kanda, R. Anghinah, and R. Nitrini, "Improving Alzheimer's disease diagnosis with machine learning techniques," *Clin. EEG Neurosci.*, vol. 42, no. 3, pp. 160–165, Jul. 2011.
[27] S. Simpraga, R. Alvarez-Jimenez, H. D. Mansvelder, J. M. A. van Gerven, G. J. Groeneveld, S. S. Poil, and K. Linkenkaer-Hansen, "EEG machine learning for accurate detection of cholinergic intervention and Alzheimer's disease," *Sci. Rep.*, vol. 7, no. 1, p. 5775, Jul. 2017.
[28] N. N. Kulkarni and V. K. Bairagi, "Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier," *IETE J. Res.*, vol. 63, pp. 11–22, Oct. 2017.
[29] C. Lehmann *et al.*, "Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG)," *J. Neurosci. Methods*, vol. 161, no. 2, pp. 342–350, Apr. 2007.
[30] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the diagnosis of Alzheimer's disease," *Clin. Neurophysiol.*, vol. 128, no. 10, pp. 2058–2067, Oct. 2017.
[31] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, and U. R. Acharya, "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Comput. Appl.*, to be published.
[32] S. Vanneste, J.-J. Song, and D. De Ridder, "Thalamocortical dysrhythmia detected by machine learning," *Nature Commun.*, vol. 9, Mar. 2018, Art. no. 1103.
[33] B. Hosseinifard, M. H. Moradi, and R. Rostami, "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal," *Comput. Methods Programs Biomed.*, vol. 109, no. 3, pp. 339–345, 2013.
[34] W. Mumtaz, S. S. A. Ali, M. A. M. Yasin, and A. S. Malik, "A machine learning framework involving EEG-based functional connectivity to diagnose major depressive disorder (MDD)," *Med. Biol. Eng. Comput.*, vol. 56, pp. 233–246, Feb. 2018.
[35] A. Khodayari-Rostamabad, J. P. Reilly, G. M. Hasey, H. de Bruin, and D. J. MacCrimmon, "A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder," *Clin. Neurophysiol.*, vol. 124, no. 10, pp. 1975–1985, 2013.
[36] Z. Cao, C.-T. Lin, W. Ding, M.-H. Chen, C.-T. Li, and T.-P. Su, "Identifying ketamine responses in treatment-resistant depression using a wearable forehead EEG," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1668–1679, Jun. 2019.
[37] N. Wang and M. R. Lyu, "Extracting and selecting distinctive EEG features for efficient epileptic seizure prediction," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 5, pp. 1648–1659, Sep. 2015.
[38] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Comput. Biol. Med.*, vol. 100, pp. 270–278, Sep. 2017.
[39] D. Wang, D. Ren, K. Li, Y. Feng, D. Ma, X. Yan, and G. Wang, "Epileptic seizure detection in long-term EEG recordings by using wavelet-based directed transfer function," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2591–2599, Nov. 2018.
[40] G. Wang, D. Ren, K. Li, D. Wang, M. Wang, and X. Yan, "EEG-based detection of epileptic seizures through the use of a directed transfer function method," *IEEE Access*, vol. 6, pp. 47189–47198, 2018.
[41] G. Wang, Z. Sun, R. Tao, K. Li, G. Bao, and X. Yan, "Epileptic seizure detection based on partial directed coherence analysis," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 3, pp. 873–879, May 2016.
[42] S.-S. Poil, W. de Haan, W. M. van der Flier, H. D. Mansvelder, P. Scheltens, and K. Linkenkaer-Hansen, "Integrative EEG biomarkers predict progression to Alzheimer's disease at the MCI stage," *Frontiers Aging Neurosci.*, vol. 5, p. 58, Oct. 2013.
[43] O. Sporns, *Networks of the Brain*. Cambridge, MA, USA: MIT Press, 2011.
[44] M. J. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biol. Cybern.*, vol. 65, no. 3, pp. 203–210, 1991.
[45] M. Ding, S. L. Bressler, W. Yang, and H. Liang, "Short-window spectral analysis of cortical event-related potentials by adaptive multivariate autoregressive modeling: Data preprocessing, model validation, and variability assessment," *Biological*, vol. 83, pp. 35–45, Jun. 01 2000.
[46] M. Morf, A. Vieira, and T. Kailath, "Covariance characterization by partial autocorrelation matrices," *Ann. Statist.*, vol. 6, no. 3, pp. 643–648, 1978.
[47] M. Morf, A. Vieira, D. T. L. Lee, and T. Kailath, "Recursive multichannel maximum entropy spectral estimation," *IEEE Trans. Geosci. Electron.*, vol. 16, no. 2, pp. 85–94, Apr. 1978.
[48] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.
[49] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
[50] D.-L. Feis, K. H. Brodersen, D. Y. von Cramon, E. Luders, and M. Tittgemeyer, "Decoding gender dimorphism of the human brain using multimodal anatomical and diffusion MRI data," *NeuroImage*, vol. 70, pp. 250–257, Apr. 2013.
[51] J. Mourao-Miranda, A. L. W. Bokde, C. Born, H. Hampel, and M. Stetter, "Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data," *NeuroImage*, vol. 28, pp. 980–995, Nov. 2005.
[52] C. Davatzikos, "Machine learning in neuroimaging: Progress and challenges," *NeuroImage*, to be published.
[53] A. Ben-Hur, C. S. Ong, S. Sonnenburg, B. Schölkopf, and G. Rätsch, "Support vector machines and kernels for computational biology," *PLoS Comput. Biol.*, vol. 4, no. 10, 2008, Art. no. e1000173.
[54] A. Serra, P. Galdi, and R. Tagliaferri, "Machine learning for bioinformatics and neuroimaging," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 8, p. e1248, Sep./Oct. 2018.
[55] P. Sajda, "Machine learning for detection and diagnosis of disease," *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 537–565, Aug. 2006.
[56] C. Lambert, E. Zeestraten, O. Williams, P. Benjamin, A. J. Lawrence, R. G. Morris, A. D. Mackinnon, T. R. Barrick, and H. S. Markus, "Identifying preclinical vascular dementia in symptomatic small vessel disease using MRI," *NeuroImage, Clin.*, vol. 19, pp. 925–938, Jan. 2018.
[57] Y. Cui, P. S. Sachdev, D. M. Lipnicki, J. S. Jin, S. Luo, W. Zhu, N. A. Kochan, S. Reppermund, T. Liu, J. N. Trollor, H. Brodaty, and W. Wen, "Predicting the development of mild cognitive impairment: A new use of pattern recognition," *NeuroImage*, vol. 60, pp. 894–901, Apr. 2012.
[58] Z. Long, J. Huang, B. Li, Z. Li, Z. Li, H. Chen, and B. Jing, "A comparative atlas-based recognition of mild cognitive impairment with voxel-based morphometry," *Frontiers Neurosci.*, vol. 12, p. 916, Dec. 2018.
[59] S. Kloppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack, Jr., J. Ashburner, and R. S. J. Frackowiak, "Automatic classification of MR scans in Alzheimer's disease," *Brain*, vol. 131, no. 3, pp. 681–689, Mar. 2008.
[60] A. K. Rehme, L. J. Volz, D.-L. Feis, I. Bomilcar-Focke, T. Liebig, S. B. Eickhoff, G. R. Fink, and C. Grefkes, "Identifying neuroimaging markers of motor disability in acute stroke by machine learning techniques," *Cerebral Cortex*, vol. 25, pp. 3046–3056, May 2014.
[61] E.-M. Pool, A. K. Rehme, S. B. Eickhoff, G. R. Fink, and C. Grefkes, "Functional resting-state connectivity of the human motor network: Differences between right- and left-handers," *NeuroImage*, vol. 109, pp. 298–306, Apr. 2015.
[62] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.
[63] D. Yao, "A method to standardize a reference of scalp EEG recordings to a point at infinity," *Physiol. Meas.*, vol. 22, no. 4, pp. 693–711, 2001.
[64] F. Chella, V. Pizzella, F. Zappasodi, and L. Marzetti, "Impact of the reference choice on scalp EEG connectivity estimation," *J. Neural Eng.*, vol. 13, no. 3, 2016, Art. no. 036016.

[65] K. J. Blinowska, "Review of the methods of determination of directed connectivity from multichannel data," *Med. Biol. Eng. Comput.*, vol. 49, pp. 521–529, May 2011.
[66] H. G. Schnack, M. Nieuwenhuis, N. E. M. van Haren, L. Abramovic, T. W. Scheewe, R. M. Brouwer, H. E. H. Pol, and R. S. Kahn, "Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects," *NeuroImage*, vol. 84, pp. 299–306, Jan. 2014.
[67] M. Nieuwenhuis, N. E. van Haren, H. E. H. Pol, W. Cahn, R. S. Kahn, and H. G. Schnack, "Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples," *NeuroImage*, vol. 61, pp. 606–612, Jul. 2012.
[68] L.-O. Wahlund, T. Erkinjuntti, and S. Gauthier, *Vascular Cognitive Impairment in Clinical Practice*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

CHAO WANG received the B.Eng. degree in biomedical engineering from Nanchang Hangkong University, Nanchang, China, in 2011, and the M.Eng. degree in biomedical engineering from Xi'an Jiaotong University, Xi'an, China, in 2014, where he is currently pursuing the Ph.D. degree in biomedical engineering. His research interests include the early detection of vascular dementia patients with EEG and the multimodal neuroimaging fusion analysis of normal aging with EEG and fMRI.

JIN XU received the B.S. degree in precision instrument from Xi'an Technological University, Xi'an, China, in 1991, the M.S. degree in optical engineering from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, in 1999, and the Ph.D. degree in biomedical engineering from Xi'an Jiaotong University, Xi'an, in 2006. She held a postdoctoral position with the Memory Lab in Carnegie Mellon, Pittsburgh, PA, USA. She is currently a Professor with the School of Life Science and Technology, Xi'an Jiaotong University. Her main research interests are biomedical signal processing, brain functional information detection, and neural engineering.

SONGZHEN ZHAO received the Bachelor of Medicine degree in clinical medicine from Xi'an Medical University, Xi'an, China, in 1984. She is currently an Associate Chief Physician and also an Associate Professor with the Department of Neurology, First Affiliated Hospital of Xi'an Jiaotong University, School of Medicine, Xi'an. Her research interest mainly focuses on the basic and clinical research of cerebrovascular diseases, including vascular dementia and epilepsy.

WUTAO LOU received the B.Eng. degree in medical information engineering from Sichuan University, Chengdu, China, in 2009, the M.Eng. degree in biomedical engineering from Xi'an Jiaotong University, Xi'an, China, in 2012, and the Ph.D. degree in imaging and interventional radiology from The Chinese University of Hong Kong, Hong Kong, in 2016. He is currently a Postdoctoral Fellow with the Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong. His research interests include EEG/fMRI signal processing, brain network analysis, and cognitive decline.