Stage detection of mild cognitive impairment via fMRI using Hilbert Huang transform based classification framework

Jiahao Shi
College of Intelligence and Computing, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin 300350, P. R. China

Baolin Liu[a)]
School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, P. R. China

(Received 1 July 2019; revised 27 March 2020; accepted for publication 6 April 2020;
published xx xxxx xxxx)

Purpose: This work aims to establish a classification framework for the diagnosis of mild cognitive impairment (MCI) at different stages (early MCI and late MCI) through direct analysis of resting-state functional magnetic resonance imaging (rs-fMRI) signals and using the accuracy (total correct rate), specificity (correct rate of late MCI) and sensitivity (correct rate of early MCI) to validate its classification performance.

Methods: All fMR images of subjects were parcellated into 116 regions of interest (ROIs) by applying the Anatomical Automatic Labeling (AAL) template, and the average rs-fMRI signals of each ROI were extracted. The Hilbert-Huang transform (HHT) was introduced into the framework to decompose each rs-fMRI signal into a series of intrinsic mode functions (IMFs) and to analyze these nonstationary and nonlinear time-series from the perspective of multiresolution. After obtaining the instantaneous frequencies and amplitudes of all IMFs of a signal, the Hilbert weighted frequencies (HWFs) were calculated and combined into a vector as the feature of the corresponding ROI. Support Vector Machine (SVM) was implemented to classify MCI at different stages. We used the independent two-sample *t*-test as the feature selection method and measured the classification performance through the leave-one-out cross-validation (LOOCV) method.

Results: Results on 77 early MCI (eMCI) and 64 late MCI (lMCI) with baseline rs-fMRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) yielded 87.94% classification accuracy. Some of the brain regions with significant differences found by previous studies have been confirmed in this work. We found that HWF characteristics exhibited a significant downward trend in all cerebellar regions. The rs-fMRI signals in differential brain regions have not changed completely, but only altered in some narrow frequency bands. The analysis results showed that during the progress of MCI, the main changes of rs-fMRI were concentrated in IMF3, while IMFs with other indexes also contained HWF features with high SVM weights, such as Orbitofrontal superior frontal gyrus in IMF2, Insula in IMF4, and Lobule Ⅲ of vermis in IMF5, indicating that other IMFs provide important information for the diagnosis of MCI as well.

Conclusions: This work confirmed the classification ability of HHT-based classification framework in classification of at different stages of MCI. Through the analysis, we found that during the progress of MCI the main changes of rs-fMRI were concentrated in IMF3, and HWF characteristics showed a significant downward trend in all cerebellar regions. © *2020 American Association of Physicists in Medicine* [https://doi.org/10.1002/mp.14183]

Key words: Alzheimer's disease, mild cognitive impairment, Hilbert-Huang transform, rs-fMRI signal

1. INTRODUCTION

Alzheimer's disease (AD) is a chronic neurodegenerative disease and the most common cause of dementia among older adults. Patients often suffer from a loss of cognitive function and behavioral abilities, which imposes a substantial burden that includes social, psychological, physical, and economic elements both on the patient and caregiver. In developed countries, AD is one of the most financially costly diseases.[1] Mild cognitive impairment (MCI), a prodromal stage of AD with great potential to progress to AD,[2] has recently attracted increasing attention. Moreover, an early treatment for MCI may prevent or at least delay the progression of the disease and preserve some cognitive functions of the brain. Therefore, an accurate diagnosis of MCI is of great importance.

Since the inception of resting-state functional magnetic resonance imaging (rs-fMRI),[3] this advanced medical imaging technique has attracted increasing attention. Spontaneous brain activity has been measured using blood oxygen level-dependent (BOLD) signals obtained from rs-fMRI.[4] Many rs-fMRI studies have explored various approaches to diagnose MCI. Methods combining the functional connectivity network (FCN) with multikernel learning (MKL) technology were proposed to diagnose MCI.[5,6] Wee et al.[7] focused on integrating diffusion tensor imaging (DTI) and resting-state functional magnetic resonance imaging (rs-fMRI) information obtained using a multikernel support vector machine (SVM) to improve the diagnostic accuracy of MCI. By advancing an objective basis for optimizing the functional network representation of general multivariate time-series, Zanin et al.[8] analyzed the functional activity network in the brain of patients with MCI and established a diagnosis. Additionally, dynamic connectivity networks (DCNs) were applied to classify MCI[9] and achieved better classification performance (significant improvements in accuracy, specificity, and sensitivity) than the conventional stationary-based FCNs. Chen et al.[10] introduced the functional correlation tensor (FCT) based on rs-fMRI into the MCI classification to describe the local structural pattern of functional connectivity. Recently, Yu et al.[11] used a new regularized sparse representation framework of a weighted graph to optimize functional brain networks and significantly improved the accuracy of the MCI diagnosis.

However, some aspects remain to be further studied in the field of MCI diagnosis. On the one hand, only a few studies have investigated the diagnosis of MCI at different stages. By conducting human connectome project multimodal parcellation (HCPMMP), Sheng et al.[12] implemented an early MCI (eMCI)/late MCI (lMCI) classification using a small dataset. Jie et al.[13] integrated the temporal and spatial properties of DCNs as features and obtained an accuracy of 78.7% in classifying eMCI/lMCI. Much room for improvement exists in the classification of different stages of MCI. On the other hand, most of these methods focused on functional connectivity in the brain and did not consider the variation in the BOLD signal itself. Previous studies have shown widespread differences in the frequency in some brain regions of patients with MCI, particularly in the low-frequency band.[14–17] The amplitude of low-frequency fluctuations (ALFFs) in some brain regions change with a linear trend during the progression of MCI.[18] This evidence confirms the significant changes in the frequency characteristics of the brain during the progression of MCI, indicating that precise frequency variations in specific brain regions may provide important information for the MCI diagnosis. The Hilbert-Huang transform (HHT) is an effective adaptive signal processing method that is good at analyzing nonstationary and nonlinear signals. By decomposing the signal into a set of narrow-band signals with different time scales, HHT describes the signal from a multiresolution perspective and obtains accurate and instantaneous values for the frequency and amplitude.[19] Therefore, the introduction of HHT into the diagnosis of different stages of MCI may improve the diagnostic performance.

Alterations in the frequency of the BOLD signal is related to cognitive decline.[20] Therefore, in the present study, we hypothesized that the BOLD signal frequencies were altered within a narrow band in distinct brain regions during the progression of MCI. We used HHT to decompose the BOLD signal into a set of intrinsic components and assessed Hilbert weighted frequencies as features to test this hypothesis. Then, we used an SVM to classify eMCI and lMCI and identified the brain regions that were related to the progression of MCI. The experimental results obtained from 141 subjects with baseline rs-fMRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/) confirmed that our method not only achieved satisfactory classification performance but also detected diverse and comprehensive oscillation properties of distinct brain regions. The rest of the paper is organized as described below. In Section 2, we briefly describe the data used in this study, present the methods of the Hilbert-Huang transform and classification framework and describe the method used to compare different combination strategies. In Section 3, we introduce experimental results for the classification performance of different combination strategies, the robustness of the framework, effects of various parameters, important brain regions and major differences in brain regions and frequencies. In Section 4, we present a discussion of different combination strategies, the effects of various parameters, the robustness of the framework, physical meaning of features, and the associations of differential brain regions with disease development. Finally, we describe the conclusions based on the findings in Section 5.

2. MATERIALS AND METHODS

The procedure used in the present study comprises three main steps, including (a) image preprocessing, (b) feature extraction using the Hilbert-Huang transform, and (c) SVM-based classification. In this section, we first introduce the data used in this study and then present a detailed description of each step. Figure 1 shows the steps of the proposed framework.

2.A. Subjects

In this study, we analyzed baseline rs-fMRI data from 141 subjects, including 77 patients with early MCI (eMCI) and 64 patients with late MCI (lMCI), from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The ADNI database is mainly used for the early detection and tracking of Alzheimer's disease, and provides a platform for data sharing among people around the world. According to its own medical classification standard, the ADNI divides MCI into two stages: early MCI (eMCI) and late MCI (lMCI). All rs-fMRI data were acquired at multiple sites on 3.0 Tesla Philips scanners (with various models/systems). The image resolutions for the X and Y dimensions ranged from 2.29 to 3.31 mm and the slice thickness was 3.31 mm. The TR (repetition time) was 3.0 s for all subjects, and the TE (echo time) was 30 ms. The number of volumes for each subject was 140. The demographic and clinical characteristics of the studied subjects are presented in Table I.

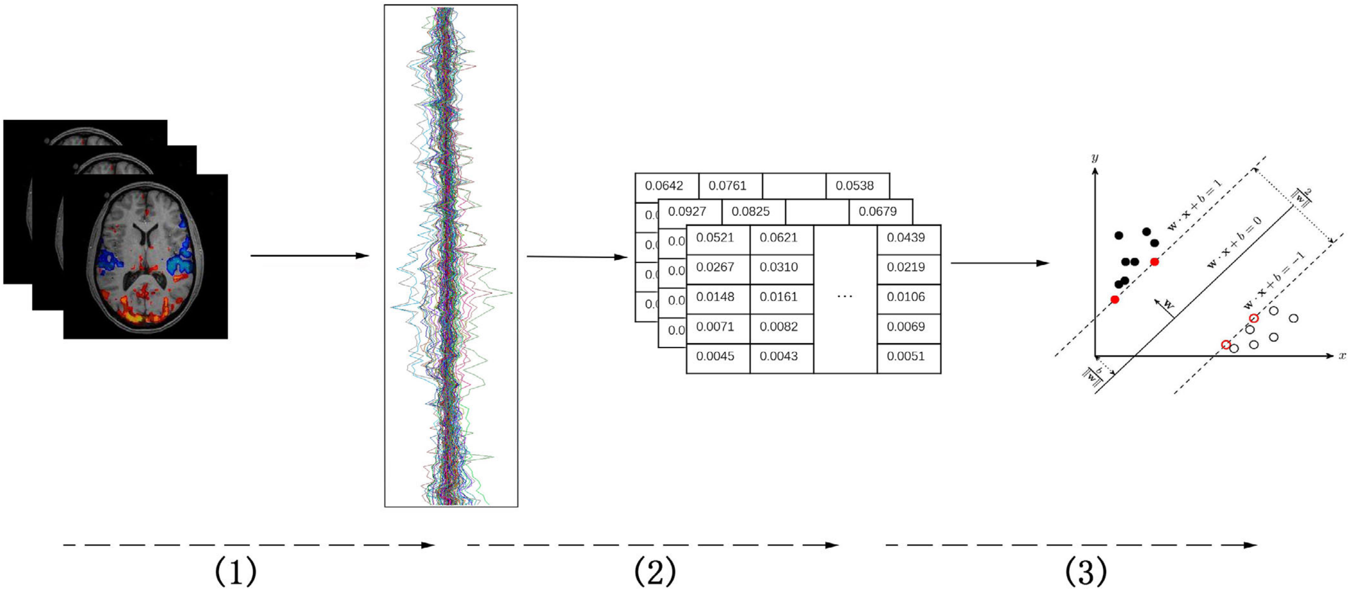

FIG. 1. Illustration of the proposed framework, including (1) image preprocessing, (2) feature extraction using the Hilbert-Huang transform, and (3) support vector machine-based classification

TABLE I. Basic characteristics of the subjects.

Group	Male/female	Age (Mean ± SD)	MMSE (Mean ± SD)
eMCI	32/45	72.4 ± 7.2	28.3 ± 1.4
lMCI	37/27	71.8 ± 7.7	27.9 ± 1.8
p-value	0.055	0.633	0.193

MMSE: mini-mental state examination.

2.B. Image preprocessing

Resting-state fMRI data were preprocessed using DPARSF (http://www.rfmri.org/dpabi), which was implemented in MATLAB software (MathWorks Inc., Sherborn, MA). The fMRI images were corrected for slice timing, realigned and corrected for voxel-specific head motion. All subjects in the study exhibited a maximal displacement of <2.0 mm along each axis, and an angular motion of <2.0° along each axis. Then, the results were smoothed by a kernel with a full width at half maximum (FWHM) of [4 voxels, 4 voxels, 4 voxels] (each voxel represents a 3 mm × 3 mm × 3 mm square in the brain). Thus, the kernel is a three-dimensional Gaussian function with a full width at half maximum of 4 voxels relative to each axis. Finally, the entire brain of each subject in the rs-fMRI space was parcellated into 116 regions of interest (ROIs) by applying the Anatomical Automatic Labeling (AAL) template[21] to the rs-fMRI image of each subject using the DPARSF package. For each of the 116 ROIs, the mean rs-fMRI time-series were calculated by averaging the BOLD signals among all voxels within the specific ROI. Due to the instability of the physical factors such as voltage and magnetic field after starting the machine, the initial fMRI scans were unstable. Therefore, the first 10 data points (30 s) in the BOLD time-series data were discarded, and 130 data points remained in the final data.

2.C. Feature extraction

After pre-processing, we focused on methods designed to extract the frequency feature of BOLD signals. We used the HHT to decompose and demodulate BOLD signals, then calculated the Hilbert weighted frequency as the feature.

The HHT consists of two main parts: empirical mode decomposition to decompose and Hilbert transform to analyze. Empirical mode decomposition (EMD) is the core of the HHT.[19] It utilizes the information from extreme points in the sequence to achieve multiscale decomposition, extracting the natural oscillation of the signal in the time domain, which is called the intrinsic mode function (IMF). It is considered an efficient description of the intrinsic components of time-series and has been widely used in many disciplines.[22] Unlike Fourier-based analyses of time-series data, the HHT does not contain a priori assumptions about the basic function type of the time-series data space, therefore the HHT is an adaptive method that is good at analyzing nonlinear and nonstationary signals, including multiple scale components.[23]

2.C.1. Empirical mode decomposition

Decomposition was performed using a filtering process to decompose the original time-series data into a finite set of IMFs. In general, empirical mode decomposition involves several steps. (a) The cubic spline interpolation is performed on the maximum and minimum values of the time-series, respectively, to obtain the upper and lower envelopes. Then, the upper and lower envelopes are averaged to obtain an average envelope. (b) The difference between the original sequence and the average envelope is set as the first IMF. (c) The average envelope is used as a time-series to perform the steps described above until the remaining time-series is a straight line or a monotonic function. Using the remainder as a residual, the time-series is decomposed into a series of IMFs and calculated using the following equation:

$$X(t) = \sum_{i=1}^{n} S_i + r_n \tag{1}$$

where $X(t)$ is the original signal, S_i is the ith IMF, and r_n is the residual. A typical BOLD signal along with its intrinsic mode function and residual are shown in Fig. 2.

2.C.2. Extension of end points

The EMD calculates the envelope using the local extremum information in the process of decomposing the signal. However, the local extremum is always missing at the end of the signal, and thus a fitting error may be generated at the end point. As the decomposition progresses, the error will become larger and eventually distort the result, which is called the endpoint effect. We extended the time-series using the least squares support vector machine (LS-SVM) as described below to reduce the endpoint effect. (a) The k points at the left (right) endpoint are used as the training set. (b) The predicted characteristics of LS-SVM are used to extend a data point. (c) The steps listed above are repeated N times. The decomposition result was obtained by intercepting the extension after EMD.

To measure the quality of decomposition, The orthogonality index *Ort* was applied to measure the quality of decomposition and was calculated using the following equation:

$$\mathrm{Ort} = \sum_{t=1}^{T} \left(\sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \frac{S_i(t) S_j(t)}{X^2(t)} \right), (i \neq j) \tag{2}$$

where the residual r_n was also viewed as an IMF, in which case $X(t) = \sum_{i=1}^{n+1} S_i$.

Since N was defined as the length of the extension here, N was an integer in a certain section to improve the classification performance. For each BOLD signal, we changed N inside the section to obtain the decomposition result with the minimum *Ort* and used this value as the final decomposition result. We performed a grid search to identify the best parameters. The range of k was set to 4-10 and the range of N was set to [0-10, 10-20, 20-30, 30-40, and 40-50].

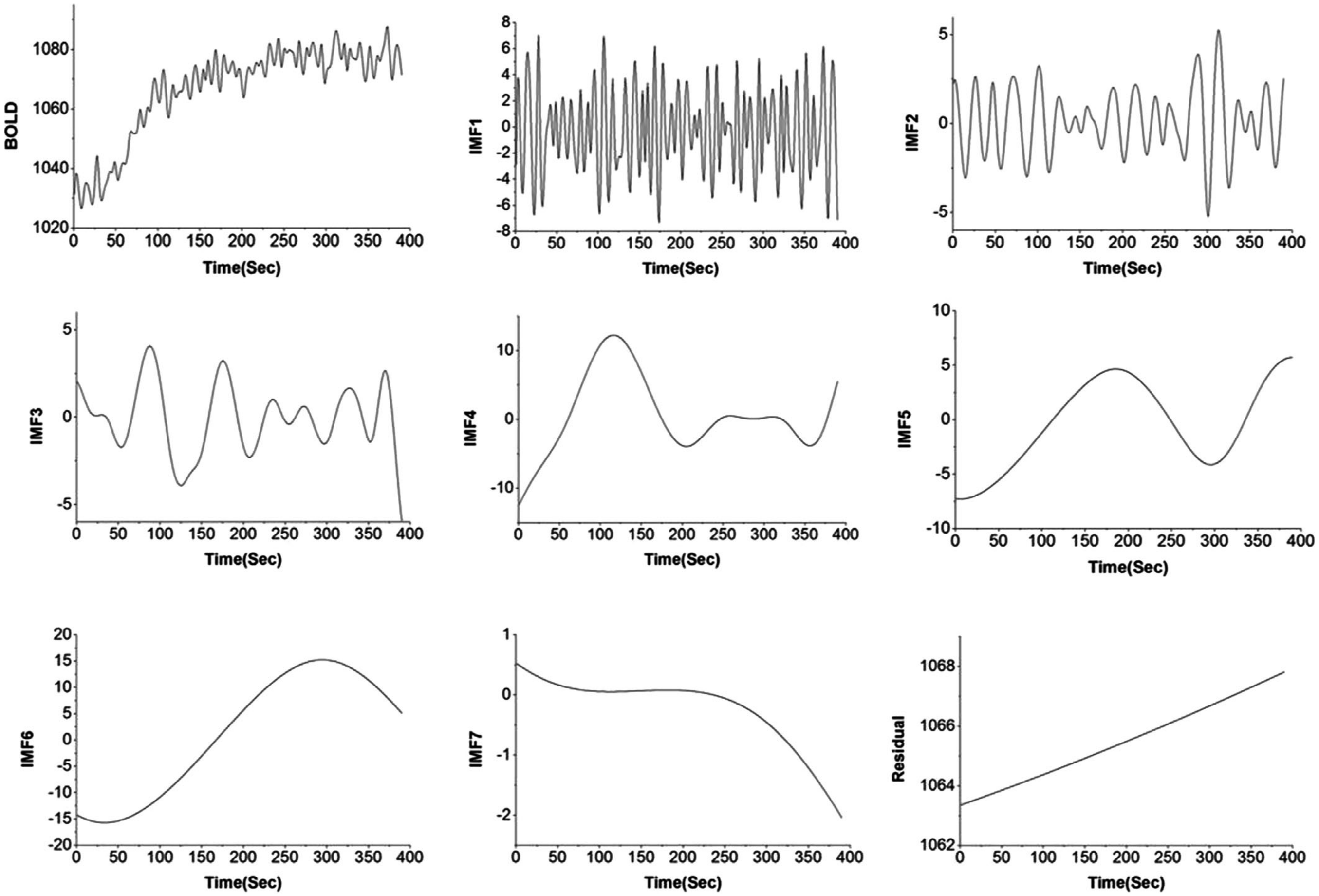

FIG. 2. Empirical mode decomposition of a blood oxygen level-dependent signal. The decomposition yielded seven intrinsic mode functions and a residual trend

2.C.3. Analysis of the Frequency and Amplitude of IMFs

Each IMF was viewed as a narrow-band amplitude and frequency modulated signal represented by $S(t)$, and was calculated using the following equation:

$$\mathrm{S(t)} = \mathrm{A(t)} \cos \varphi(t) \tag{3}$$

where the instantaneous amplitude A and phase φ were obtained using the Hilbert transform of S, defined as

$$S_H(t) = \frac{1}{\pi} \int \frac{S(\tau)}{t-\tau} d\tau \tag{4}$$

$$\mathrm{A(t)} = \sqrt{S^2(t) + S_H^2(t)} \tag{5}$$

$$\varphi(t) = arctan \frac{S_H(t)}{S(t)} \tag{6}$$

The derivative of the phase function is the instantaneous frequency in the usual sense, which is defined as

$$\omega(\mathrm{t}) = \frac{d\varphi(t)}{dt} \tag{7}$$

Thus, the original BOLD signal X was calculated as the sum of all IMFs and residual r

$$\mathrm{X(t)} = \sum_{j=1}^{k} A_j(t) \exp\left(i \int \omega_j(t) dt\right) + r \tag{8}$$

where k is the total number of IMFs, ω_j and A_j represent the instantaneous frequency and the instantaneous amplitude of the jth IMF respectively. Regarding the numerical instability in the Hilbert transform and arctangent function calculation, a direct orthogonal method is always used to replace Hilbert transform and arctangent function operation with cubic spline interpolation and derivative operation, as described in the Appendix.[24,25] We used the instantaneous frequency and amplitude of IMF2 shown in Fig. 1 as an example. The result is presented in Fig. 3.

2.C.4. The Hilbert weighted frequency (HWF)

For a particular IMF with an instantaneous frequency $\omega(i)$ and instantaneous amplitude $a(i)$, we calculated the Hilbert weighted frequency[26] as follows:

$$HWF = \frac{\sum_{i=1}^{m} \omega(i) a^2(i)}{\sum_{i=1}^{m} a^2(i)} \tag{9}$$

where m was the length of time-series. Each signal was represented as a vector consisting of HWF. The first few HWFs of the BOLD signal in all brain regions were combined as features for further classification.

2.D. Classification model for diagnosing MCI

In determining the MCI diagnosis, the amount of data was far less than the feature dimension, and the data were sparsely distributed in the feature space, which was not conducive to improving the classification accuracy. Feature selection was required before the classification to obtain better classification results and more valuable features. We used an independent two-sample t-test for feature selection, and then used support vector machine (SVM) as the classification method, which is described in detail below.

2.D.1. Feature selection

We compared each feature of eMCI and lMCI using an independent two-sample t-test,[27] and selected features with significant differences. We assumed that each feature was selected from the same distribution and calculated the p-values of all features. We performed Kolmogorov-Smirnov tests (K-S tests)[28] on all features to verify their normality, and then calculated p-values using the independent two-sample t-test. We assumed that random variable X displays a normal distribution with mean μ_1 and variance σ_1^2 and random variable Y displays a normal distribution with mean μ_2 and variance σ_2^2. Given the means of the two variables, which are expressed as $\overline{\mathrm{X}}$ and $\overline{\mathrm{Y}}$, then $\overline{\mathrm{X}}$ displays a normal distribution with mean μ_1 and variance σ_1^2/M and $\overline{\mathrm{Y}}$ exhibits a normal distribution with mean μ_2 and variance σ_2^2/L.

$$\overline{X} \sim N\left(\mu_1, \frac{\sigma_1^2}{M}\right) \tag{10}$$

$$\overline{Y} \sim N\left(\mu_2, \frac{\sigma_2^2}{L}\right) \tag{11}$$

where M and L are the number of samples of the two variables respectively. The symbol ~ represents that the random variable exhibits some type of statistical distribution and N indicates a normal distribution. Therefore,

$$\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{M} + \frac{\sigma_2^2}{L}\right) \tag{12}$$

Replace the standard deviations σ_1 and σ_2 with sample standard deviations S_X and S_Y, and then

$$\mathrm{t}' = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{M} + \frac{S_Y^2}{L}}} \sim t(h) \tag{13}$$

Namely, t' exhibits a t-distribution with h degrees of freedom. The degrees of freedom h is calculated using the following formula:

$$\mathrm{h} = \left(\frac{S_X^2}{M} + \frac{S_Y^2}{L}\right)^2 / \left(\frac{S_X^4}{M^2(M-1)} + \frac{S_Y^4}{L^2(L-1)}\right) \tag{14}$$

The P-value is defined as the probability of obtaining a more extreme value than t' under the t-distribution with h degrees of freedom, which is calculated using the following equation:

$$P_{\text{value}} = P_{x \sim t(h)}\left(|x| \geq |t'|\right) \tag{15}$$

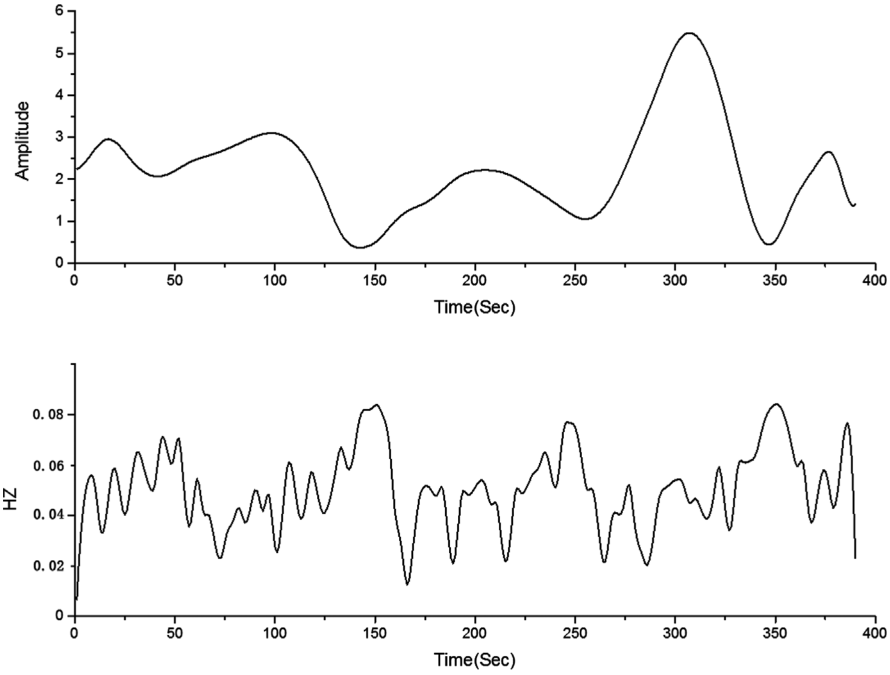

FIG. 3. Instantaneous amplitude and frequency of intrinsic mode function 2

We sorted these *P*-values in ascending order and selected the first few features.

2.D.2. Support vector machine

Since Cortes and Vapnik proposed the Support Vector Machine (SVM) technique,[29] it has been widely used in brain science research to solve the binary classification problem.[27,30–38] The decision function of SVM is shown as follows:

$$Y(x_k) = \text{sign}\left\{\sum_{i=1}^{K} \alpha_i y_i.K(x_k, x_i) + b\right\}$$

where *K* represents the kernel function. The solution of SVM aims to minimize the objective function of the following optimization problems and calculate α_i:

$$\begin{aligned}&\min_{\alpha}\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j.y_iy_j.K\left(x_i,x_j\right)-\sum_{i=1}^{N}\alpha_i\\&\text{s.t.}0\le\alpha_i\le\text{C}\\&\sum_{i=1}^{N}\alpha_iY_i=0\end{aligned} \tag{17}$$

where *C* is a regularization parameter that represents the importance of outliers. The tradeoff between margin and misclassification error is controlled by this parameter.

In the experiment, we used the leave-one-out cross-validation (LOOCV) approach to measure the classification effect. In each fold, one subject was used as the test set, and the others were used as the training set for feature selection and classification. The procedure was repeated until every subject participated in the experiment as the test set. Then, we calculated the classification accuracy.

2.E. Comparison of different combinations of strategies

We used three methods to analyze the BOLD signal and to compare the performance of different signal processing methods and classifiers: the short-time Fourier transform,[39] wavelet transform,[40] and Hilbert-Huang transform.

Specifically, we applied the short-time Fourier transform by dividing the BOLD signal into segments with a length of 2^M to capture multiscale information. Because the length of the BOLD signal was set to 130, $M = 1, 2, 3...7$, and the last two points of BOLD signal were discarded.

After obtaining multiscale spectrum information by using the short-time Fourier transform and wavelet transform, the weighted frequency was extracted as a feature. Similar to HWF, the weighted frequency for spectral information at each scale was defined as the sum of the product of frequency and the square of the amplitude divided by the sum of the square of the amplitude.

We applied four classification methods to classify features: support vector machine, logical regression,[41] random forest,[42] and back propagation neural network.[43] By combining these three signal processing methods with the four classification methods we obtained 12 different combination strategies. Then, we compared the classification performance of these combination strategies.

3. RESULTS

We calculated the accuracy (total correct rate), specificity (correct rate of lMCI), and sensitivity (correct rate of eMCI) to validate the classification performance. The three numbers in each cell of the tables below represent the sensitivity, specificity, and accuracy respectively.

3.A. Classification performance of different combination strategies

Table II shows the classification results of 12 different combination strategies using three signal processing methods (short-time Fourier transform, wavelet transform, and Hilbert-Huang transform) and four classification methods (support vector machine, logical regression, random forest, and back propagation neural network). Since classification is considered as a Bernoulli process with a probability p, where p is the true accuracy rate, the classification of the sample set is viewed as a binomial distribution experiment with an expected value of Np and a variance of $Np\ (1 - p)$, where N is the number of samples. Therefore, for a single classification accuracy rate acc, $acc\ (1 - acc)$ approximates the variance of the classification. We assumed that other strategies have the same classification accuracy as HHT + SVM and used independent two-sample t-tests to obtain p-values. We used the false discovery rate (FDR) correction here to compare the classification performance of other strategies with HHT + SVM and avoid possible errors associated with multiple comparisons.

3.B. Effects of various parameters

In this section, we mainly discuss the effects of method parameters on the results. Three parameters are able to be adjusted in the experiment: the number of features, the number of IMFs, and the SVM kernel function parameter. We used three indicators of the classification, accuracy, specificity, and sensitivity, to evaluate the classification results.

Table II. Classification results of different combination strategies.

	SVM	Logistic regression	Random forest	Back propagation neural network
Short time Fourier transform	59.74%	57.14%	71.43%	58.44%
	43.75%	39.06%	54.69%	45.31%
	52.48% (P = 0.0005)	48.94% (P = 0.0004)	63.83% (P = 0.0006)	52.48% (P = 0.0006)
Wavelet transform	61.04%	61.04%	71.43%	68.83%
	53.13%	42.19%	56.25%	56.25%
	57.45% (P = 0.0005)	52.48% (P = 0.0005)	64.54% (P = 0.0007)	63.12% (P = 0.0006)
Hilbert-Huang transform	**92.21%**	87.01%	79.22%	80.02%
	82.81%	78.12%	71.88%	65.62%
	87.94%	71.88% (P = 0.0011)	75.89% (P = 0.0013)	73.62% (P = 0.0015)

The method combining the Hilbert-Huang transform with SVM is superior to the other strategies.
Bold indicate that the Hilbert-Huang transform and SVM combination strategy achieved the best classification performance.

3.B.1. Number of features

Figure 4 shows the effect of the number of selected features on the classification results of eMCI compared with lMCI. At a confidence level of 0.05, all features exhibited normal distributions. Therefore, we sorted all P-values in ascending order and selected the first few features for classification. We set the number of IMFs to 5 and used the linear kernel function in SVM to investigate the effects of the number of selected features. As shown in Fig. 4, the classification performance presented a bell-shaped trend as the number of features changed. When the number of features was set to 30, the method achieved the best classification performance, obtaining results with an accuracy, specificity, and sensitivity of 87.94%, 82.81%, and 92.21%, respectively, in distinguishing eMCI from lMCI.

3.B.2. Number of IMFs

Figure 5 shows the effect of the number of IMFs on the results of the classification of eMCI and lMCI. Here, the number of features was set to 30 and the linear kernel function was used in SVM. As shown in Fig. 5, the accuracy varied substantially with the number of IMFs, indicating that the method is sensitive to this parameter. The selection of the top five IMFs as the best choice for the classification of eMCI and lMCI.

3.B.3. SVM kernel function parameter

Figure 6 shows the effect of the Radial Basis Function (RBF) parameters on the results of the classification of eMCI and lMCI. We used an RBF with different full width at half maximum (FWHM) values as the kernel function in SVM.

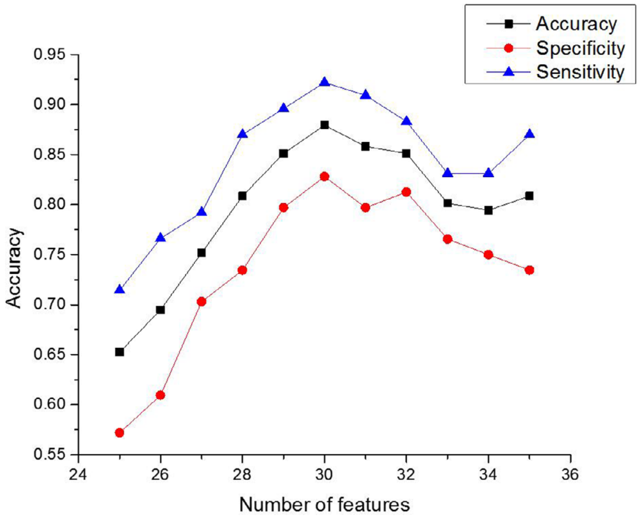

Fig. 4. The relationship between the classification and the number of selected features in the early mild cognitive impairment (eMCI) vs lMCI classification. Accuracy: total correct rate; specificity: correct rate of lMCI; sensitivity: correct rate of eMCI

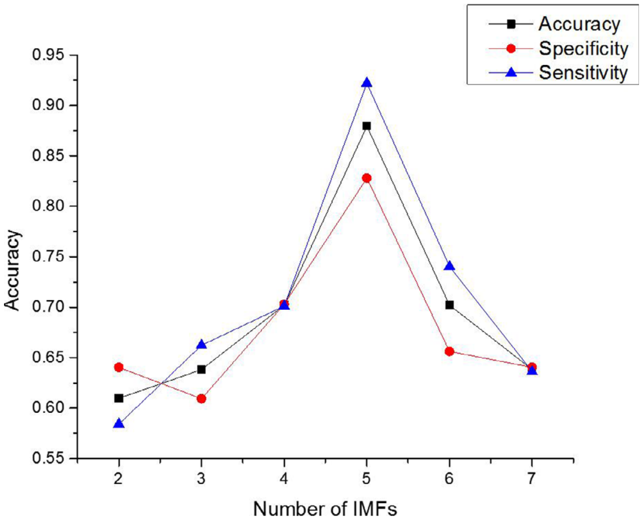

FIG. 5. The relationship between the classification and the number of intrinsic mode functions in the early mild cognitive impairment (eMCI) vs lMCI classification. Accuracy: total correct rate; specificity: correct rate of lMCI; sensitivity: correct rate of eMCI

Notably, even the best classification performance of SVM obtained using the RBF kernel function was worse than the SVM with the linear function, indicating that the features are complex and sufficiently separable using a linear function to effectively calculate the classification surface using only the linear kernel function.

3.B.4. Other parameters

Three other parameters were included in the experiment, two in end point extension, training set length *k,* and extension length *N*, and one in SVM, the regularization parameter *C*. We performed a grid search to identify the best values of *k* and *N*. When *k* was set to 7 and *N* was set to 20–30, we obtained the best classification performance. The value of *C* ranged from 0.1 to 0.7 to explore its effect on the classification results. Table III shows the effect of parameter *C*, and the best choice of *C* in this experiment was 0.4.

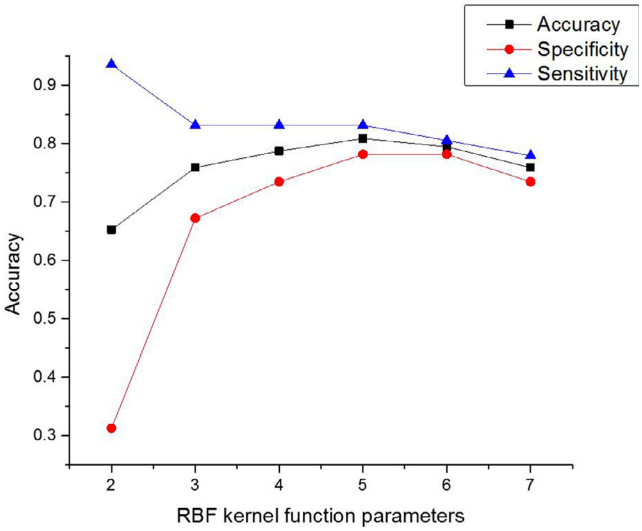

FIG. 6. The relationship between classification and radial basis function (RBF) kernel function parameters in the early mild cognitive impairment (eMCI) vs lMCI classification. RBF kernel function parameter: the full width at half maximum of RBF. Accuracy: total correct rate; specificity: correct rate of lMCI; sensitivity: correct rate of eMCI

In particular, we varied the number of hidden layers of neural network from 1 to 6 to explore the relationship between the depth of the neural network and classification results. Table IV shows the changes in the classification results with the increase in the neural network depth.

3.C. Robustness of the framework

We added white noise to the BOLD signal to test the robustness of the framework and obtained the classification performance of the framework. The standard deviation of the white noise was set to *d* times the BOLD signal. We explored the change in the classification results of the framework when *d* increased from 0 to 0.45. The results are shown in Table V.

3.D. Important brain regions

In this section, we explore the important features of the eMCI vs lMCI classification in our proposed method. Since we used the accuracy of the leave-one-out cross-validation (LOOCV) approach, we defined the brain regions whose IMF features were selected by all LOO folds as important brain regions. First, the features of each brain region were inspected using independent two-sample t-tests in all LOO folds. Then, the features with significant differences were selected under the condition of a confidence level of 0.05. In the linear kernel SVM classification, the averaged weights of the SVM corresponding to these features were used to measure their importance. Table VI shows the importance of these IMFs in important brain regions.

As shown in Table VI and Fig. 7, the important IMFs were concentrated in IMF3, which includes the orbitofrontal superior frontal gyrus, opercular inferior frontal gyrus, triangular inferior frontal gyrus, olfactory cortex, cuneus, lingual gyrus, middle occipital gyrus, postcentral gyrus, crus I and lobule VIII of the cerebellar hemisphere. Additionally, IMFs that were very important in other indexes were identified, including IMF2 in the orbitofrontal superior frontal gyrus (0.9907), IMF4 in the insula (1.1703) and IMF5 in lobule III of the vermis (0.7594). The locations of ROIs corresponding to the important IMFs are shown in Fig. 8.

3.E. Major differences in brain regions and frequencies

We inspected the IMF features of each brain region using an independent two-sample t-test and obtained the ROIs with significantly different IMF features between eMCI and lMCI along with their Hilbert weighted frequencies to identify the

Table III. The effect of C on classification results.

Value of C	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Classification results	87.01%	88.31%	90.91%	**92.21%**	89.61%	89.61%	89.61%
	79.69%	81.25%	82.81%	**82.81%**	82.81%	84.38%	82.81%
	83.69%	85.11%	87.23%	**87.94%**	86.52%	87.23%	86.52%

When C was set to 0.4, the best classification result was achieved.
Bold indicate that the framework achieved the best classification performance when C was set to 0.4.

Table IV. The relationship between the number of hidden layers and classification results.

Number of hidden layers	1	2	3	4	5	6
Classification results	80.02%	74.03%	75.32%	72.73%	76.62%	75.32%
	65.62%	68.75%	70.31%	67.19%	60.94%	56.25%
	73.62%	71.63%	73.05%	70.12%	69.50%	66.67%

The classification performance decreases as the number of neural network layers increases.

Table V. The relationship between the classification and the added white noise.

White noise	0	0.05	0.1	0.15	0.2
Classification results	92.21%	85.71%	84.42%	79.22%	67.53%
	82.81%	81.25%	78.13%	68.75%	67.19%
	87.94%	83.69%	81.56%	74.47%	67.38%

White noise	0.25	0.3	0.35	0.4	0.45
Classification results	63.63%	61.04%	57.14%	55.84%	45.45%
	65.63%	64.06%	48.44%	40.62%	43.75%
	64.54%	62.41%	53.19%	48.94%	44.68%

The unit of noise is the standard deviation (SD) of the BOLD signal. As the standard deviation of the white noise increases, the classification ability of the framework decreases, and ultimately it will fail.

major brain regions showing differences. Table VII shows the results.

As shown in Table VII, most of the important IMFs of distinct brain regions exhibited obvious differences in two categories with relatively small p-values, which verified the important role of these IMFs in classification. As shown in Table VII, during the progression of MCI, the Hilbert weighted frequencies were increased in ROIs such as the precentral gyrus, opercular superior frontal gyrus, triangular inferior frontal gyrus, hippocampus, parahippocampal gyrus, cuneus, middle occipital gyrus, fusiform gyrus, inferior parietal angular gyrus, angular gyrus, precuneus and caudate nucleus, but were decreased in the orbitofrontal superior frontal gyrus, rolandic operculum, olfactory cortex, insula, lingual gyrus, inferior occipital gyrus, postcentral gyrus, and temporal pole (superior). Remarkably, the HWF features of all cerebellar ROIs, such as R.crus I, crus II, lobule III, lobule VIII, lobule X of the cerebellar hemisphere, and lobule III of

Table VI. Important IMFs of distinct brain regions and their averaged SVM weights in the classification of eMCI and lMCI.

ROI	IMF index	Weight
L.Orbitofrontal superior frontal gyrus	**IMF2**	**0.9907**
L.Orbitofrontal superior frontal gyrus	IMF3	0.4076
L.Opercularis inferior frontal gyrus	IMF3	-0.1329
L.Triangular inferior frontal gyrus	IMF3	-0.7773
R..Olfactory cortex	IMF3	0.7629
R. Insula	**IMF4**	**1.1703**
R.Hippocampus	IMF4	-0.6426
R. Parahippocampal gyrus	IMF2	-0.5533
R. Cuneus	IMF3	-0.5156
L.Lingual gyrus	IMF3	-0.0167
L. Middle occipital gyrus	IMF1	-0.3658
L. Middle occipital gyrus	IMF3	-0.4505
L.Inferior occipital gyrus	IMF2	0.1307
L. Postcentral gyrus	IMF3	0.0839
R.Angular gyrus	IMF4	-0.5715
R.Crus Ⅰ of cerebellar hemisphere	IMF3	0.6111
L.Lobule Ⅲ of cerebellar hemisphere	IMF5	0.5460
L.Lobule Ⅷ of cerebellar hemisphere	IMF3	0.2879
Lobule Ⅲ of vermis	**IMF5**	**0.7594**

L. = Left, R. = Right.
Bold indicate that highlights ROIs and IMFs with high weight in the SVM classification.

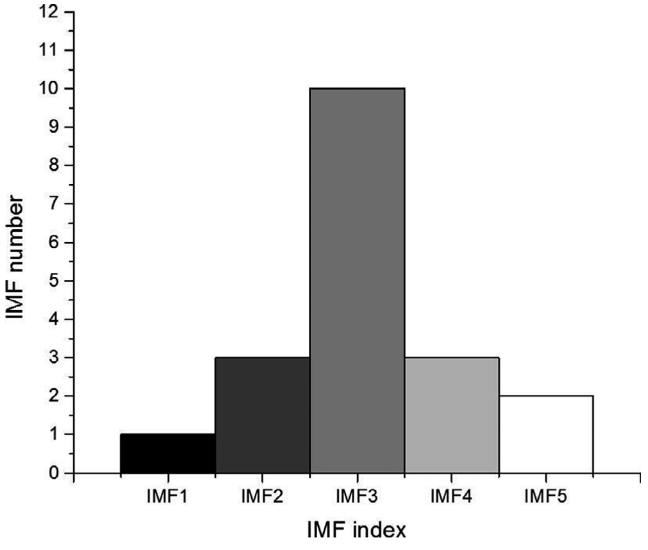

Fig. 7. The distribution of important intrinsic mode functions (IMFs) in different IMF indexes

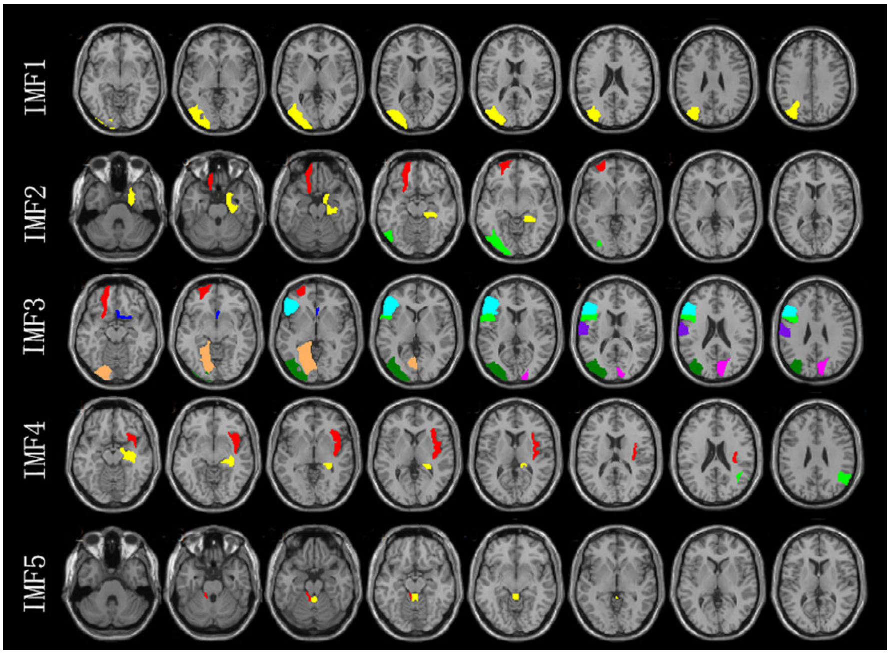

Fig. 8. The locations of regions of interests (ROIs) corresponding to the important intrinsic mode function (IMFs) from IMF1 to IMF5. Each row shows the locations of ROIs corresponding to important IMFs of different indexes, and different colors represent different ROIs

the vermis, significant decreased during the progression of MCI.

4. DISCUSSION

4.A. Different combination strategies

Fourier transform is a global transform method, and it is unable to easily transform nonstationary signals. Short-time Fourier transform divides the signal into small segments, in which the nonstationary signal is approximately regarded as the stationary signal. However, this approximation is unable to characterize the features of nonstationary signals. Wavelet transform sufficiently describes the local characteristics of the signal, but it is still based on the Fourier transform, which is unable to simultaneously improve the temporal resolution and frequency resolution. The Hilbert-Huang transform is an adaptive method. Based on the instantaneous frequency and amplitude, the Hilbert-Huang transform simultaneously improve the temporal resolution and frequency resolution.[25] Therefore, the Hilbert-Huang transform more accurately describes the characteristics of nonstationary signal. As shown in Table II, the classification results obtained using the Hilbert-Huang transform are significantly better than the other two methods. Regarding the classification method, SVM shows a higher classification accuracy than other methods. These findings confirm the effectiveness of the classification framework combining the Hilbert-Huang transform with SVM.

4.B Effects of various parameters

In this section, we mainly discuss the effects of different parameters on the results. In the experiments, three parameters, including the number of features, the number of IMFs and SVM kernel parameters, substantially affected the results. When we set the number of features to 30, the number of IMFs to 5 and used the linear kernel function, we obtained the best classification performance for the comparison of eMCI with lMCI, and the accuracy, specificity, and sensitivity were 87.94%, 82.81%, and 92.21% respectively. We used the average change rate, namely, the quotient of the total change in classification accuracy and the total change in the parameters, to quantify the effect of changes in these parameters on the classification results. As shown in Figs. 4–6, the classification performance presents a relatively stable trend with the change in feature number (average change rate: 3.48%), while it changes rapidly with the number of IMFs (average change rate: 10.35%). A potential explanation for this finding is that changes in the number of IMFs may lead to tremendous changes in the total number of features. For this experiment, when the number of IMFs changes by 1, the feature number would increase or decrease by 116 before feature selection. The use of an inappropriate number of IMFs will result in a substantial loss of valid information or serious concealment, decreasing the classification performance. Therefore, the choice of the appropriate number of IMFs is very important. According to Fig. 5, the best strategy in this experiment was to select the first five IMFs. In theory, the classification performance of nonlinear kernels is similar to linear kernels.[29] However, as shown in Fig. 6, we were unable to obtain the best classification results with parameters that have not been carefully selected. Thus, the features we extracted are complex and sufficiently separable using a linear kernel to effectively calculate the classification surface using only the linear kernel function. In particular, as shown

TABLE VII. Different IMF features of distinct brain regions and p-values in the classification of eMCI vs. lMCI.

ROI	IMF index	eMCI	lMCI	*P*-value
R.Precentral gyrus	IMF2	0.03	0.0328	0.028
L.Orbitofrontal superior frontal gyrus	IMF2	0.0389	0.036	0.022
L.Orbitofrontal superior frontal gyrus	IMF3	0.0168	0.0147	0.009
L.Opercularis superior frontal gyrus	IMF3	0.0142	0.0159	0.018
L.Triangular inferior frontal gyrus	IMF3	0.0144	0.0162	0.014
L.Rolandic operculum	IMF3	0.0162	0.0144	0.028
R.Olfactory cortex	IMF3	0.0185	0.0163	0.008
R.Insula	IMF4	0.0062	0.0052	0.026
R.Hippocampus	IMF4	0.0056	0.0067	0.02
R. Parahippocampal gyrus	IMF2	0.0336	0.0369	0.021
R. Cuneus	IMF2	0.0336	0.0361	0.031
R. Cuneus	IMF3	0.0148	0.017	0.003
L.Lingual gyrus	IMF3	0.0154	0.0138	0.015
R.Lingual gyrus	IMF2	0.0349	0.0336	0.035
L.Middle occipital gyrus	IMF1	0.0484	0.052	0.006
L.Middle occipital gyrus	IMF3	0.0139	0.0156	0.028
L.Inferior occipital gyrus	IMF2	0.0384	0.0344	0.002
L.Fusiform gyrus	IMF5	0.0009	0.0015	0.026
L. Postcentral gyrus	IMF3	0.0137	0.0121	0.021
L. Postcentral gyrus	IMF5	0.0016	0.001	0.045
L.Inferior parietal angular gyrus	IMF1	0.0490	0.0516	0.03
R. Angular gyrus	IMF4	0.0049	0.0061	0.003
R. Precuneus	IMF1	0.0491	0.0514	0.038
L.Caudate nucleus	IMF5	0.0014	0.0021	0.048
L.Temporal pole(superior)	IMF4	0.0059	0.005	0.028
R.crus Ⅰ of cerebellar hemisphere	**IMF3**	**0.0160**	**0.0131**	**0.004**
L.crus Ⅱ of cerebellar hemisphere	**IMF2**	**0.0373**	**0.0343**	**0.028**
L.Lobule Ⅲ of cerebellar hemisphere	**IMF5**	**0.0020**	**0.0013**	**0.023**
L.Lobule Ⅷ of cerebellar hemisphere	**IMF3**	**0.0149**	**0.0128**	**0.005**
L.Lobule Ⅹ of cerebellar hemisphere	**IMF3**	**0.0160**	**0.0143**	**0.027**
Lobule Ⅲ of vermis	**IMF5**	**0.0024**	**0.0014**	**0.008**

L. = Left, R. = Right.
Bold indicate the ROI of the cerebellar region has a significant decrease in frequency.

in Table IV, classification results became worse as the depth of the neural network increased, potentially because the connection weights in the neural network to be trained increase as the depth of the neural network increases. However, the amount of data is unable to meet the requirements of training, and thus many weights lead to a decrease in classification accuracy.

4.C. Robustness of the framework

As shown in Table V, when the noise increased from 0 to 0.1 SD, the classification accuracy decreased from 87.94% to 81.56%, but still maintained the effectiveness of classification. When the noise further increased to 0.45 SD, the classification accuracy decreased to 44.68%, and the framework totally lost the classification ability. Thus, this framework has a certain anti-noise ability and still maintains a relatively strong classification ability when the noise is <0.1 SD. However, under conditions with a high level of noise, the classification ability of the framework will continue to decline and ultimately the framework will fail.

4.D. Physical meaning of features

As an adaptive method, the HHT decomposes multicomponent signals into single-component signal components, that is, IMF, with time scale ranging from small to large values. IMF has the characteristics of a local narrow band, which more accurately describes the time-frequency characteristics of multicomponent signals.[25] Each IMF has its own clear physical meaning, but further studies are needed to identify these physical meanings.[25] The Hilbert weighted frequency effectively reflects the average frequency fluctuation characteristics of local narrow-band signals, and it is a robust feature.[26] The eigenvector composed of the Hilbert weighted frequencies of multiple IMFs with a time scale ranging from small to large values shows the amplitude-frequency distribution characteristics of fMRI signals in specific brain regions in each frequency band, avoiding overlap and confusion between frequency bands.[26] It highlights the role of frequency components in the main part of energy in a specific frequency band; therefore, it is an accurate and complete description of signals. As shown in Table VII, specific HWF features were increased in cerebral ROIs, such as the precentral gyrus, opercular superior frontal gyrus, triangular inferior frontal gyrus, hippocampus, and parahippocampal gyrus, but were decreased in the orbitofrontal superior frontal gyrus, rolandic operculum, olfactory cortex, and insula. Frequency features of all cerebellar ROIs, such as R.crus I, crus II, lobule III, lobule VIII, lobule X of the cerebellar hemisphere, and lobule III of the vermis, decreased significantly. By combining the Hilbert weighted frequency vectors of all ROIs as features and comparing them, we identified the differences in the narrow frequency bands of distinct brain regions between different stages of MCI, which is more conducive to further analysis and research.

4.E. The association of different brain regions with disease development

Few studies have investigated the changes in brain regions at different stages of MCI. In this experiment, we identified the brain regions displaying significant differences at different stages of MCI, which were consistent with previous studies, including the cuneus, middle occipital gyrus,[44] lingual gyrus, insula, olfactory cortex, and inferior occipital gyrus.[45] Differences in the parahippocampal gyrus, postcentral gyrus, angular gyrus, and temporal pole (superior) have also been reported in previous studies.[13] We also observed differences in other brain regions at different stages of MCI, such as the precentral gyrus, orbitofrontal superior frontal gyrus, opercular superior frontal gyrus, triangular inferior frontal gyrus, rolandic operculum, hippocampus, fusiform gyrus, inferior parietal angular gyrus, precuneus and caudate nucleus.

In addition, some cerebellar areas have also been confirmed to display significant differences that were consistent with previous studies, including lobule X of the cerebellar hemisphere and lobule III of the vermis.[13] Crus I, crus II, lobule III, and lobule VIII of the cerebellar hemisphere also exhibited significant changes during the progression of MCI. As shown in Table VII, frequency features of all cerebellar ROIs were significantly decreased. These results are consistent with other neuroscience studies,[46–48] confirming that the cerebellum may be involved in human cognitive processes and may provide useful information for the diagnosis of MCI.

Notably, the BOLD signals in these brain regions have not changed completely but were only altered in some narrow frequency bands of the signals, which were accurately detected using the HHT method. As shown in Fig. 7, the differences in the whole brain range from IMF1 to IMF5 in different stages of MCI, but the distributions of altered signals were more concentrated in IMF3. Thus, the changes in frequency observed during the progression of MCI are mainly attributed to IMF3. On the other hand, as shown in Table VI, IMFs with other indexes also contain HWF features with high weights, such as the orbitofrontal superior frontal gyrus in IMF2, insula in IMF4, and lobule III of the vermis in IMF5, indicating that other IMFs also provide important information for the diagnosis of MCI. Based on these results, the frequency features differ in distinct brain regions, and the method described here will be more conductive to analyzing these frequency features in actual patients in future studies.

4.F. Limitations

This study was subject to some limitations. First, the limited antinoise ability made this method relatively vulnerable to the impact of noise. In future studies, denoizing methods should be combined to improve the robustness of this method. Second, the amount of data was relatively small and the length of each time-series was relatively short (130 data points). The TR (repetition time) of the dataset was 3.0 s, and thus we were unable to examine the activities during a shorter period. The lack of data limited the effect of the combination of this method and deep learning. Larger datasets and longer time-series may be needed to obtain more accurate diagnostic results in future studies.

5. CONCLUSIONS

In this paper, we proposed a learning framework based on the HHT for the classification of different stages of MCI. Specifically, we first decomposed BOLD signals from each brain region into a series of IMFs using the HHT, and demodulated them to obtain the instantaneous amplitude and frequency. Then, the Hilbert weighted frequency was calculated as a feature, and feature selection was performed using the t-test. Finally, SVM was implemented for classification. The proposed method effectively classified eMCI and lMCI. Using this analysis, the main changes in rs-fMRI were located in IMF3, and HWF characteristics showed a significant decreasing trend in all cerebellar regions during the progression of MCI.

FUNDING INFORMATION

This work was supported by the National Key Research and Development Program of China (No. 2018YFB0204304) and National Natural Science Foundation of China (No. U1736219 and No. 61571327).

ETHICAL APPROVAL

All procedures performed in studies involving human participants were in accordance with the ethical standards of the national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

INFORMED CONSENT

Informed consent was obtained from all individual participants included in this study.

CONFLICT OF INTEREST

We declare that we have no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations that can inappropriately influence our work.

APPENDIX 1

The direct orthogonal method first estimates the amplitude by local extremum information, and then uses the frequency modulation portion of the original signal and its orthogonal component to estimate the frequency. Specifically, the following steps are included:

1. For a single-component signal $s(t)$, determine all the maxima points of $|s(t)|$, mark as (t_k, s_k), $k = 1, 2, \cdots, M$
2. All the extreme points are fitted with cubic splines, and the resulting envelope function is recorded as $a_{11}(t)$. Then $s(t)$ can be standardized by $a_{11}(t)$, that is.

$$s_1(t) = \frac{s(t)}{a_{11}(t)} \tag{A1}$$

3. If $s_1(t)$ is not a pure FM signal, repeat the above process n times until $s_n(t)$ is an FM signal, denoted as $F(t)$.

4. The instantaneous amplitude of the signal is defined as

$$A(t) = \frac{s(t)}{F(t)} = a_{11}(t)a_{12}(t)\ldots a_{1n}(t) \tag{A2}$$

After removing the amplitude modulation portion of the original signal, the frequency modulation portion F(t) is obtained, and the frequency is estimated by the following steps.

1. For the single-component signal $x(t)$, the FM portion $F(t)$ and its orthogonal component $G(t)$ can be expressed as

$$\begin{cases} F(t) = cos\varphi(t) \\ G(t) = sin\varphi(t) \end{cases} \tag{A3}$$

2. Deriving on both sides,

$$\begin{cases} F'(t) = -\varphi^{'}(t)sin\varphi(t) \\ G'(t) = \varphi'(t)cos\varphi(t) \end{cases} \tag{A4}$$

1. Since the instantaneous frequency is non-negative, the instantaneous frequency ω(t) can be calculated as follows:

$$\omega(t) = \varphi'(t) = \sqrt{F'(t)^2 + G'(t)^2} \tag{A5}$$

Jiahao Shi and Baolin Liu equally contributed as co-first authors.

[a]Author to whom correspondence should be addressed. Electronic mail: liubaolin@tsinghua.edu.cn; Telephone: +86-10-62781789; Fax: +86-10-6278 1789.

REFERENCES

1. Sylvie BG, Dina Z, Ezio G, Gabriel G, Jean-Pierre M. The economical impact of dementia. *Press Medicale*. 2005;34:35–41.
2. Reiman EM, Langbaum JB, Fleisher AS, et al. Alzheimer's prevention initiative: a plan to accelerate the evaluation of presymptomatic treatments. *J Alzheimer's Dis*. 2011;26:321–329.
3. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med*. 2010;34:537–541.
4. Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. 2007;8:700–711.
5. Biao J, Daoqiang Z, Wei G, Qian W, Chong-Yaw W, Dinggang S. Integration of network topological and connectivity properties for neuroimaging classification. *IEEE Trans Biomed Eng*. 2014;61:576–589.
6. Chen G, Ward BD, Xie C, et al. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. *Int J Med Radiol*. 2011;259:213–221.
7. Wee C-Y, Yap P-T, Zhang D, et al. Identification of MCI individuals using structural and functional connectivity networks. *NeuroImage*. 2012;59:2045–2056.
8. Zanin M, Sousa P, Papo D, et al. Optimizing functional network representation of multivariate time series. *Sci Rep*. 2012;2:630.
9. Wee CY, Yang S, Yap PT, Shen D. Sparse temporally dynamic resting-state functional connectivity networks for early MCI identification. *Brain Imaging Behav*. 2015;10:1–15.
10. Chen X, Zhang H, Zhang L, Shen C, Lee SW, Shen D. Extraction of dynamic functional connectivity from brain grey matter and white matter for MCI classification. *Hum Brain Mapp*. 2017;38:5019.
11. Yu R, Qiao L, Chen M, Lee S-W, Fei X, Shen D. Weighted graph regularized sparse brain network construction for MCI identification. *Pattern Recognit*. 2019;90:220–231.
12. Sheng J, Wang B, Zhang Q, et al. A novel joint HCPMMP method for automatically classifying Alzheimer's and different stage MCI patients. *Behav Brain Res*. 2019;365:210–221.
13. Jie B, Liu M, Shen D. Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease. *Med Image Anal*. 2018;47:81–94.
14. Ying H, Jinhui W, Zhilian Z, et al. Frequency-dependent changes in the amplitude of low-frequency fluctuations in amnestic mild cognitive impairment: a resting-state fMRI study. *NeuroImage*. 2011;55:287–295.
15. Xi Q, Zhao X, Wang P, et al. Spontaneous brain activity in mild cognitive impairment revealed by amplitude of low-frequency fluctuation analysis: a resting-state fMRI study. *Radiol Med*. 2012;117:865–871.
16. Zhao Z-L, Fan F-M, Lu J, et al. Changes of gray matter volume and amplitude of low-frequency oscillations in amnestic MCI: an integrative multi-modal MRI study. *Acta radiol*. 2015;56:614–621.
17. Yang L, Yan Y, Wang Y, et al. Gradual disturbances of the amplitude of low-frequency fluctuations (ALFF) and fractional ALFF in Alzheimer spectrum. *Front Neurosci*. 2018;12:975.
18. Liang P, Xiang J, Liang H, Qi Z, Li K, Alzheimer's Disease NeuroImaging Initiative. Altered amplitude of low-frequency fluctuations in early and late mild cognitive impairment and Alzheimer's disease. *Curr Alzheimer Res*. 2014;11(4):389–398.
19. Huang NE, Zheng S, Long SR, et al. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc Math Phys Eng Sci*. 1971;1998:903–995.
20. Yang AC, Tsai SJ, Lin CP, Peng CK, Huang NE. Frequency and amplitude modulation of resting-state fMRI signals and their functional relevance in normal aging. *Neurobiol Aging*. 2018;70:59–69.
21. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002;15:273–289.
22. Niazy RK, Xie J, Miller K, Beckmann CF, Smith SM. Spectral characteristics of resting state networks. In: VanSomeren EJW, VanDerWerf YD, Roelfsema PR, Mansvelder HD, DaSilva FHL, eds. *Slow Brain Oscillations of Sleep, Resting State and Vigilance*. Vol 193. Progress in Brain Research. Amsterdam, Netherlands: Elsevier Science; 2011:259–276. 1016/B978-0-444-53839-0.00017-X
23. Wu Z, Huang NE, Long SR, Peng C-K. On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc Natl Acad Sci*. 2007;104:14889–14894.
24. Zheng J, Cheng J. Improved Hilbert-Huang transform and its applications to rolling bearing fault diagnosis. *J Mech Eng*. 2015;51:138.
25. Huang NE, Wu Z, Long SR, Arnold KC, Chen X, Blank K. On instantaneous frequency. *Adv Adapt Data Anal*. 2009;01:177–229.
26. Wu X, Wu T, Liu C, Wen X, Yao L. Frequency clustering analysis for resting state functional magnetic resonance imaging based on Hilbert-Huang transform. *Front Hum Neurosci*. 2017;11:61.
27. Zhang Z, Xu J, Tang J, Zou Q, Guo F. Diagnosis of brain diseases via multi-scale time-series model. *Front Neurosci*. 2019;13:1–8.
28. Shiryayev AN15. On The Empirical Determination of A Distribution Law. Sel Work A N Kolmogorov. January 1992:139–146. https://doi.org/10.1007/978-94-011-2260-3_15
29. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273–297.
30. Li J, Qiu L, Xu L, Pedapati EV, Erickson CA, Sunar U. Characterization of autism spectrum disorder with spontaneous hemodynamic activity. *Biomed Opt Express*. 2016;7:3871.
31. Zeng P, Huang J, Wu S, et al. Characterizing the structural pattern predicting medication response in herpes zoster patients using multivoxel pattern analysis. *Front Neurosci*. 2019;13:534.
32. Mikolas P, Hlinka J, Skoch A, et al. 07-Connectivity of the anterior insula differentiates participants with first-episode schizophrenia spectrum disorders from controls: a machine-learning study. *Psychol Med*. 2016;46:2695–2704.

33. Bi X, Xu Q, Luo X, Sun Q, Wang Z. Weighted random support vector machine clusters analysis of resting-state fMRI in mild cognitive impairment. *Front Psychiatry*. 2018;9:340.
34. Memarian N, Torre JB, Haltom KE, Stanton AL, Lieberman MD. Neural activity during affect labeling predicts expressive writing effects on well-being: GLM and SVM approaches. *Soc Cogn Affect Neurosci*. 2017;12:1437–1447.
35. Guo H, Zhang F, Chen J, Xu Y, Xiang J. Machine learning classification combining multiple features of a hyper-network of fMRI data in Alzheimer's disease. *Front Neurosci*. 2017;11:1–22.
36. Cui X, Xiang J, Wang B, Xiao J, Niu Y, Chen J. Integrating the local property and topological structure in the minimum spanning tree brain functional network for classification of early mild cognitive impairment. *Front Neurosci*. 2018;12:1–11.
37. Zago L, Hervé P-Y, Genuer R,, et al. Predicting hemispheric dominance for language production in healthy individuals using support vector machine. *Hum Brain Mapp*. 2017;38:5871–5889.
38. Bi XA, Chen J, Sun Q, Liu Y, Wang Y, Luo X. Analysis of asperger syndrome using genetic-evolutionary random support vector machine cluster. *Front Physiol*. 2018;9:1–10.
39. Hannaford B, Short LS. Time Fourier Analysis of the Electromyogram: Fast Movements and Constant Contraction; 1986.
40. Kikuchi H. Wavelet and signal processing;2001.
41. Hosmer DW Jr., Lemeshow S, Sturdivant RX.Applied Logistic Regression. 2013;34(3):358–359.
42. Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
43. Rumelhart D, Mcclelland J. Learning and Relearning in Boltzmann Machines; 1986.
44. Wang B, Niu Y, Miao L, et al. Decreased complexity in Alzheimer's disease: resting-state fMRI evidence of brain entropy mapping. *Front Aging Neurosci*. 2017;9:378.
45. Niu Y, Wang B, Zhou M, et al. Dynamic complexity of spontaneous bold activity in Alzheimer's disease and mild cognitive impairment using multiscale entropy analysis. *Front Neurosci*. 2018;12:677.
46. Schmahmann JD. The cerebellum and cognition. *Neurosci Lett*. 2019;688:62–75.
47. Anteraper SA, Guell X, D'Mello A, Joshi N, Whitfield-Gabrieli S, Joshi G. Disrupted cerebrocerebellar intrinsic functional connectivity in young adults with high-functioning autism spectrum disorder: A data-driven, whole-brain, high-temporal resolution functional magnetic resonance imaging study. *BRAIN Connect*. 2019;9:48–59.
48. Zacharia TT, Eslinger PJ. Functional MRI activation patterns of cerebellum in patients with epilepsy and brain tumors. *Clin Anat*. 2019;32:1053–1060.