Stability Evaluation of Brain Changes in Parkinson's Disease Based on Machine Learning

Chenggang Song [1,2,3,4], Weidong Zhao [4], Hong Jiang [5], Xiaoju Liu [6], Yumei Duan [7], Xiaodong Yu [4], Xi Yu [4], Jian Zhang [8], Jingyue Kui [9], Chang Liu [1,2,3,4*] and Yiqian Tang [4*]

[1] School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China, [2] Center for Information in Medicine, University of Electronic Science and Technology of China, Chengdu, China, [3] MOE Key Lab for Neuroinformation, The Clinical Hospital of Chengdu Brain Science Institute, Chengdu, China, [4] College of Computer, Chengdu University, Chengdu, China, [5] Department of Neurosurgery, Rui-Jin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [6] Department of Abdominal Oncology, Cancer Center, West China Hospital, Sichuan University, Chengdu, China, [7] Department of Computer and Software, Chengdu Jincheng College, Chengdu, China, [8] School of Physics and Electronic Engineering, Sichuan Normal University, Chengdu, China, [9] Department of Urology, Tonghai County People's Hospital, Yuxi, China

Structural MRI (sMRI) has been widely used to examine the cerebral changes that occur in Parkinson's disease (PD). However, previous studies have aimed for brain changes at the group level rather than at the individual level. Additionally, previous studies have been inconsistent regarding the changes they identified. It is difficult to identify which brain regions are the true biomarkers of PD. To overcome these two issues, we employed four different feature selection methods [ReliefF, graph-theory, recursive feature elimination (RFE), and stability selection] to obtain a minimal set of relevant features and nonredundant features from gray matter (GM) and white matter (WM). Then, a support vector machine (SVM) was utilized to learn decision models from selected features. Based on machine learning technique, this study has not only extended group level statistical analysis with identifying group difference to individual level with predicting patients with PD from healthy controls (HCs), but also identified most informative brain regions with feature selection methods. Furthermore, we conducted horizontal and vertical analyses to investigate the stability of the identified brain regions. On the one hand, we compared the brain changes found by different feature selection methods and considered these brain regions found by feature selection methods commonly as the potential biomarkers related to PD. On the other hand, we compared these brain changes with previous findings reported by conventional statistical analysis to evaluate their stability. Our experiments have demonstrated that the proposed machine learning techniques achieve satisfactory and robust classification performance. The highest classification performance was 92.24% (specificity), 92.42% (sensitivity), 89.58% (accuracy), and 89.77% (AUC) for GM and 71.93% (specificity), 74.87% (sensitivity), 71.18% (accuracy), and 71.82% (AUC) for WM. Moreover, most brain regions identified by machine learning were consistent with previous findings, which means that these brain regions are related to the pathological brain changes characteristic of PD and can be regarded as potential biomarkers of PD. Besides, we also found the brain abnormality

of superior frontal gyrus (dorsolateral, SFGdor) and lingual gyrus (LING), which have been confirmed in other studies of PD. This further demonstrates that machine learning models are beneficial for clinicians as a decision support system in diagnosing PD.

1. INTRODUCTION

Parkinson's disease (PD), a serious neurodegenerative disease, is caused by the deterioration of dopaminergic neurons (Group, 2002; Perez-Lloret and Rascol, 2010). It has been reported that disrupted dopamine transmission leads to abnormal motor symptoms and beta oscillations in the subthalamic nucleus (STN) in patient with PD (Mallet et al., 2008). Moreover, the lack of dopamine can damage several areas of the brain, producing a variety of motor and nonmotor symptoms such as resting tremor, bradykinesia, muscle rigidity, depression, and sleep disorders (Koller et al., 1989; Morris, 2000; Chaudhuri and Schapira, 2009; Fox et al., 2011). This disease affects millions of people worldwide and reduces quality of life and happiness. Therefore, the early diagnosis and treatment of PD are particularly important. In general, physicians identify the severity/stage/progression of PD by medical history and neurological examination (Folstein et al., 1975; Fahn et al., 1997). However, the results of these clinical examinations are heavily affected by the knowledge and experience of clinicians, posing a risk to the accurate diagnosis and effective treatment of PD (Fahn et al., 1997; MD et al., 2006; Jankovic, 2008). In order to better diagnose and treat PD, computer-based data analysis methods have been gradually applied to the imaging of neurodegenerative diseases (Hirschauer et al., 2015; Peker et al., 2015; Rana et al., 2015b). The microstructure of gray matter (GM) and white matter (WM) in the brains of patients with PD may change at an early stage of the disease, and structural changes occur earlier than physiological changes (Rektor et al., 2018). Existing studies have shown that neuroimaging analyses can reflect changes in brain microstructure. The development of MRI has made it possible to study the structure and function of the human brain in a noninvasive manner. Structural MRI (sMRI) (Heim et al., 2017) is widely used in the study of neuroimaging due to its advantages of good contrast and high resolution (Duchesne et al., 2009; Ziegler and Augustinack, 2013).

Most existing research studies based on sMRI have used conventional statistical analysis methods to discriminate patients with PD from healthy subjects. It has been reported that patients with PD show GM reductions compared with healthy controls in the bilateral temporal lobe, bilateral occipital lobe, bilateral parietal lobe, bilateral frontal lobe, bilateral insular lobe, bilateral parahippocampal gyrus, bilateral amygdala, right uncus, right precuneus, caudate, and right posterior lobe of the cerebellum (Price et al., 2004; Summerfield et al., 2005; Agosta et al., 2013; Xia et al., 2013; Moccia et al., 2016; Gao et al., 2017; Kikuchi et al., 2017). Price et al. (2004), Agosta et al. (2013), and Moccia et al. (2016) analyzed WM and observed brain changes in the regions of the middle temporal gyrus, right superior longitudinal fasciculus/angular gyrus, and insula. Schwarz et al. (2011) observed changes in the substantia nigra based on manually predefined region(s)-of-interest (ROI(s)). However, conventional statistical analysis often overlooks the correlations among voxels and focuses on group-level differences rather than individual-level diagnosis.

In order to address these issues, machine learning methods have been increasingly applied to neuroimaging in recent years (ChenZhiHong et al., 2020; Liu et al., 2020). Machine learning methods have been widely used for classification studies aiming to distinguish between PD and HC controls (Duchesne et al., 2009; Long et al., 2012; Lei et al., 2018). In addition to being sensitive to subtle differences in the brain, machine learning methods can also be generalized to the diagnosis of individual patients. In neuroimaging studies, the number of features (voxels) is often much larger than the number of subjects, which is a very common problem in machine learning, called the "curse of dimensionality" (Wang et al., 2017, 2019; Altman and Krzywinski, 2018). This problem can easily lead to the overfitting of machine learning models (Guyon and Elisseeff, 2003). Hence, it is essential to utilize feature selection to capture the most relevant features and remove redundant ones.

In general, feature selection includes supervised feature selection and unsupervised feature selection. According to the different attributes of the features, supervised feature selection methods can be divided into three categories: (1) filter methods (Biesiada and Duch, 2005; Sánchezmaroño et al., 2007), which are based on simple statistical parameters (mean value, variation and correlation coefficient, etc.) and rank them in terms of their ability to detect group-level differences; (2) wrapper methods, which are based on the cost function, and sort all features based on their degree of correlation (Kohavi and John, 1997; ChenZhiHong et al., 2020); and (3) embedded methods (Wang et al., 2015), which select relevant features by imposing certain "penalties" to obtain a subset of relevant features. Filtering methods have the benefit of low computational cost, while wrapper methods are superior to filtering methods in performance due to their discriminative ability (Lee and Verleysen, 2007; Chu et al., 2011; Adeli et al., 2016; Cigdem et al., 2018). Considering the interaction among features, embedded methods have shown excellent performance in pattern classification research (Wang et al., 2015). Unsupervised feature selection methods construct related features by means of linear or nonlinear combinations of the original prediction features (Lee and Verleysen, 2007).

An increasing number of feature selection methods are being applied in the field of neuroscience. Chu et al. (2011) used a variety of different feature extraction techniques to construct a classification model, and experiments have revealed that feature

selection methods can enhance the accuracy of classification. Cigdem et al. (2018) classified PD and controls using a probability distribution function based on feature selection methods to build separate decision models for GM and WM, which achieved good classification accuracy of 75.00 and 72.50%, respectively. However, the histogram technology used in that study leads to the loss of information. Adeli et al. (2016) also utilized an unsupervised joint feature-sample selection (JFSS) method to select an optimal subset of samples and features and construct a reliable diagnosis model, which achieved a classification accuracy of 81.9%. Rana et al. (2015a) utilized the filtering feature selection method based on mutual information in conjunction with a support vector machine (SVM) to construct the classification model, and the classification accuracy reached 86.67%. However, these studies were based on region-of-interest (ROI) analysis, which probably neglected some potential brain changes. Rana et al. (2015b) proposed an unsupervised feature selection method based on spectral graph theory and ultimately achieved a classification accuracy of 86.67%. Recently, Babu et al. (2014) employed a meta-cognitive radial basis function network (McRBFN) with recursive feature elimination (RFE) to construct a classification model, which achieved a classification accuracy of 87.21%.

Previous studies have sought to identify brain changes in PD based on voxel-based morphometry (VBM), which has provided the analysis at the level of group to understand the neurobiology of disease, while it has focused on the disease diagnosis of individual patients in a clinical context. Group-level analysis based on statistical analysis fails to implement individual diagnosis, because it only detects brain differences between groups. Furthermore, group-level analysis is sensitive to the acquisition parameters of the sMRI data (Marquand et al., 2013; Salvatore et al., 2013; Cherubini et al., 2014; Rana et al., 2015a,b). Due to the robustness of machine learning techniques, four popular feature selection methods (ReliefF, graph theory, RFE, and stability selection) (Mwangi et al., 2013; Tohka et al., 2016) were used to detect the morphological brain differences characterizing PD on structural brain MRI in our study. These feature selection methods were used to retrieve the minimum sets of relevant and nonredundant features from GM and WM separately. Then, we used SVM to learn the decision model from the selected features. Moreover, due to the excellent properties of the feature selection method, we were able to identify the most discriminative brain regions, and we evaluated their stability both horizontally and vertically. First, we compared the brain changes using different feature selection methods. Then, we conducted a vertical comparison of the brain changes found by machine learning with the results of previously reported traditional statistical analyses. The experimental results showed that machine learning techniques have superior and robust classification performance in distinguishing patients with PD from HCs, and most of the identified brain regions are consistent with previous findings, indicating that they were stable biomarkers of PD. The remaining sections of the present study are organized as follows: Section 2 briefly describes the materials. Section 3 gives a brief description including details on the preprocessing, feature selection, SVM classification, and performance evaluation metrics. Section 4 contains the experimental results and discussion, and section 5 presents the conclusion.

2. MATERIALS

2.1. Subjects

All data carried in this study came from the Parkinson's Progression Markers Initiative (PPMI) datasets (https://www. ppmi-info.org/data) (Marek et al., 2011). Briefly, PPMI is a public repository from various centers that provides neuroimaging and associated clinical information of various modes of PD and matched control subjects for data sharing and scientific research. The PPMI cohort includes 600 datasets that comprise 400 participants with PD and 200 healthy subjects. All participants in PPMI have received approval from the Institutional Review Board (IRB). Inclusion criteria of patients with PD in the ppmi datasets were the following: (1) 30 years of age or older. (2) Patients had at least two of the following signs: resting tremor; rigidity; and slowness of movements, with at least one observation of bradykinesia (slowness). (3) Hoehn and Yahr stage 1 or 2 at baseline. Inclusion and exclusion criteria for healthy controls were as following: (1) 30 years of age or older. (2) Healthy subjects whose first-degree relatives had idiopathic PD were excluded. (3) Healthy subjects with a history of neurological disorders, mental illness, head injuries, or substance abuse were excluded.

In this study, we use the MRI data acquired by the PPMI study, in which a T1-weighted, 3D sequence is acquired for each subject using 3T SIEMENS scanners. We have considered 208 subjects (112 healthy subjects and 127 patients with PD) available in the datasets as of September 2019. Among 208 subjects, 18 healthy subjects and 21 patients with PD MR images were excluded due to failure of the segmentation method. Furthermore, all subjects were selected based on the following criteria: (1) All subjects were aged over 50 years and under 80 years; 14 healthy subjects and 17 patients with PD were excluded. (2) All subjects did not have depression, with a Geriatric Depression Scale score of < 5; 8 healthy subjects and 12 patients with PD were excluded. (3) All participant subjects were right-handed; 10 healthy subjects and 9 patients with PD were excluded. (4) patients with PD with Hoehn and Yahr stage < 3 were retained; 8 patients with PD were excluded. (5) The disease duration of patients with PD is greater than 12 months and less than 48 months; 15 patients with PD were excluded. Finally, 44 gender- and age-matched healthy subjects with a similar level of education were selected for comparison.

The disease severity and functional status of each patient were assessed with the Movement Disorder Society–Unified Parkinson's Disease Rating Scale-part III (MDS-UPDRS III) and Hoehn–Yahr (HY) stage. All subjects were assessed with the Montreal Cognitive Assessment (MoCA). The disease duration of patients with PD was defined from the time of symptom onset. The demographic details of the patients whose data were used in our study are shown in **Table 1**.

Variable	PD	HC	p-values
Gender (M/F)	29/15	27/17	0.658[b*]
Age (years)	63.95 ± 8.13	63.89 ± 7.61	0.968[a*]
Education (years)	15.61 ± 2.88	16.07 ± 3.08	0.476[a*]
Disease duration (years)	2.81 ± 1.24	–	–
MDS-UPDRS III score	32.14 ± 9.85	–	–
H&Y score	2.05 ± 0.47	–	–
MoCA score	27.66 ± 2.49	28.23 ± 1.49	0.197[a*]

PD, Parkinson's disease; HC, Healthy controls; M/F, male/female.
MDS-UPDRS III, Movement Disorder Society-Unified Parkinson's.
Disease Rating Scale-part III.
MoCA, Montreal Cognitive Assessment. H&Y = Hoehn-Yahr staging.
[a]Two-sample t-test.
[b]Chi-square test.
[*]Statistical analyses have not shown significant difference between healthy controls and PD according to age, sex, MoCA score, and education year.

2.2. MRI Acquisition

In this study, we used a 3T SIEMENS scanner to obtain T1-weighted 3D sequences for each subject. T1-weighted imaging was acquired with the following parameters: acquisition plane=sagittal, acquisition type = 3D, coil = body, flip angle = 9.0 degrees, matrix X/Y/Z = 240.0/256/176 pixel, manufacturing model = TrioTim, pixel spacing X/Y = 1.0/1.0 mm, pulse sequence = GR/IR, slice thickness = 1 mm, and TE/TI/TR = 2.98/900/2300 ms.

3. MACHINE LEARNING METHODS

The whole classification framework, including data preprocessing, feature selection, and classification is presented in **Figure 1**. As shown in the figure, the whole experiment is divided into three modules: data preprocessing, feature selection, and feature classification. Among these modules, the purpose of data preprocessing is to eliminate the effects of geometric distortion, intensity imbalance, and noise in MRI as a result of human factors or defects in the acquisition equipment itself. Since most studies based on neuroimaging have limited data with very high dimensionality, the classification on original data directly without feature selection will lead to high computational cost and over-fitting problem, which make that machine learning models have poor generalization ability. In order to reduce the risk of overfitting of the classification model and enhance the interpretability of the machine learning model, it is necessary to reduce the dimension of the data and evaluate the importance of features with a feature selection technique. Finally, a linear SVM was used to construct the classification model, which showed significant performance for small data-sets (Vapnik, 1999; Anguita et al., 2011; Peng et al., 2015).

3.1. MRI Data Preprocessing

Voxel-based morphometry (VBM) is an advanced and powerful quantitative MRI procedure designed to detect the brain morphological/volumetric changes in brain diseases. Study has

shown (Farokhian et al., 2017) that a VBM analysis using the Computational Anatomy Toolbox (CAT12) is more robust and accurate in terms of volumetric alterations than using older versions of VBM8 toolbox. Since the purpose of this study is to conduct a horizontal comparison with previous studies based on VBM8, we still considered to adopt VBM8 to exclude the impact of irrelevant factors and keep consistency with previous founding.

The original Digital Imaging and Communications in Medicine (DICOM) images were converted to 3D NIFTI format using MRIcron (https://people.cas.sc.edu/rorden/mricron/index.html). Further image processing was performed based on the voxel-based morphometry (VBM) technique with the VBM8 toolbox (http://dbm.neuro.uni-jena.de) in conjunction with Statistical Parametric Mapping software (SPM8, Wellcome Institute of Neurology, University College London, UK) (http://www.fil.ion.ucl.ac.uk/spm/) running on a Matlab R2013b platform. The preprocessing included the following steps: (1) Spatial normalization was performed by resampling and reorienting 3D T1-weighted MR volume of each subject using T1 templates from SPM8. (2) Gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) were obtained by segmentation of all normalized volumes. Meanwhile, in order to maintain local tissue volume, modulation was performed simultaneously in the segmentation process. (3) Spatial smoothing was performed with a 6 mm isotropic full width at half maximum Gaussian kernel.

3.2. Statistical Analysis

In this study, we applied a two-sample t-test to extract group-level brain changes from GM and WM separately. Whole brain analysis was performed with a threshold of uncorrected $p < 0.05$ and cluster size ≥ 100, which have been verified to effectively balance false positives and negatives (Altamura et al., 2017; Lojowska et al., 2018; Mishra et al., 2019). Finally, the statistical map was utilized as a 3D mask template to extract different voxels from the corresponding smoothed GM and WM.

3.3. Feature Selection and Classification

As one of the most important techniques of machine learning, feature selection has been used to analyze neuroimaging because it enables the user to conveniently identify areas that differ among different groups. In this study, we chose four different feature selection methods to construct the learning framework, and we explored their classification performance. The detailed algorithm design is given in the **Supplementary Material**. The four feature selection methods, namely, stability selection (Meinshausen and Bühlmann, 2010), ReliefF (Kira and Rendell, 1992; Robnik-Šikonja and Kononenko, 2003), spectral feature selection (Zhao and Liu, 2007), and recursive feature elimination with linear SVM (Guyon, 2001), are widely used in the field of neuroscience (Mwangi et al., 2013; Tohka et al., 2016). Subsequently, SVM, as a powerful classifier based on the principle of structural risk minimization (Sain, 1997), is used to construct classification models. The main principle of SVM is to find a hyperplane that maximizes the margin between two types of objects to separate measurements.
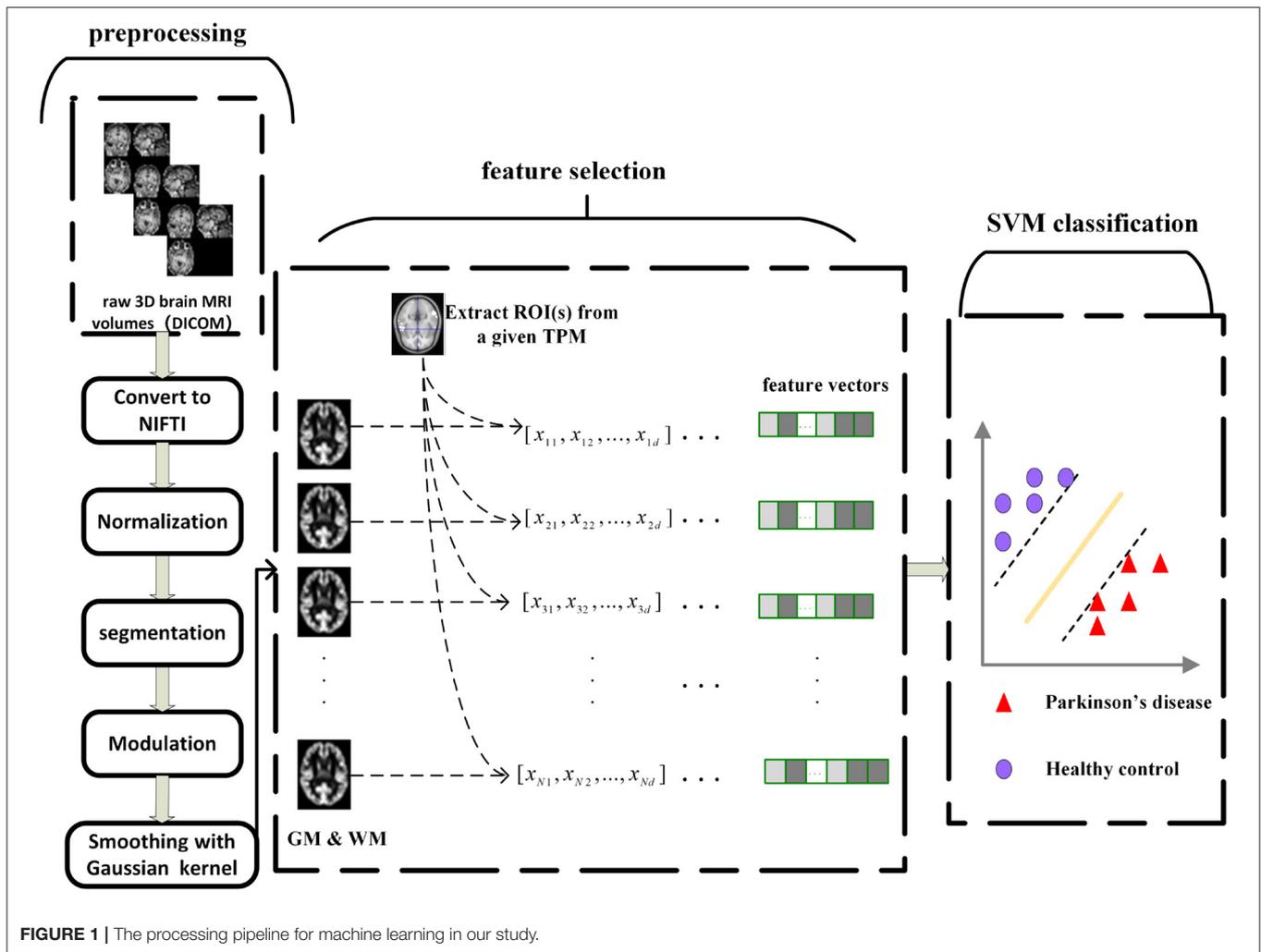
FIGURE 1 | The processing pipeline for machine learning in our study.

3.4. Performance Evaluation

We utilized 5-fold cross-validation to verify the effectiveness of the classification model. We divided the data into five subsets and sequentially designated each one as a test set to test the performance of the model trained on the remaining data. In addition, to further prove the stability of the experimental results, we conducted 10 iterations of each cross-validation. Finally, the performance of the learning model was evaluated by calculating the sensitivity, specificity, and accuracy, which were defined as:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FP + FN}$$

Suppose N represents the number of patients, and M refers to the number of healthy individuals. N' is the number of

patients correctly classified, and M' is the number of healthy controls correctly classified. TP, TN, FP, and FN are represented as follows:

$$TP = \frac{N'}{N}$$

$$TN = \frac{N - N'}{N}$$

$$FP = \frac{M'}{M}$$

$$FN = \frac{M - M'}{M}$$

Accuracy, calculated as an arithmetic mean of sensitivity and specificity, was used to measure the overall performance of the model in PD classification. In general, sensitivity and specificity affect each other, and an increase in one of them will inevitably lead to a decrease in the other.

TABLE 2 | The classification performance (in percentage) of four machine learning methods on GM.

		GM				
	RetainedFeas	OptimalFeas	Sen	Spec	Acc	AUC
SPEC	9,396	8,045	86.93 ± 2.41	91.36 ± 2.94	88.61 ± 1.86	89.15 ± 2.53
ReliefF	9,396	7,809	87.29 ± 3.54	**92.24± 3.69**	**89.58 ± 2.39**	**89.77 ± 3.58**
RFE	5,285	2,841	**92.42 ± 2.62**	85.85 ± 2.46	88.56 ± 1.99	89.14 ± 2.55
STABLASSO	3,547	2368	90.41 ± 3.21	79.55 ± 2.39	84.98 ± 1.99	84.98 ± 2.98

RetainedFeas, The number of features retained by different feature selection methods; OptimalFeas, the optimal feature subsets selected from retained features; Sen, sensitivity; Spec, specificity; Acc, accuracy; AUC, The area under ROC curve. The best performance for each indicator is shown in bold.

TABLE 3 | The classification performance (in percentage) of four machine learning methods on WM.

		WM				
	RetainedFeas	OptimalFeas	Sen	Spec	Acc	AUC
SPEC	9,922	6,048	72.18 ± 3.77	69.35 ± 3.54	70.39 ± 2.33	70.77 ± 3.65
ReliefF	9,922	5,457	**74.87 ± 3.45**	68.77 ± 2.95	**71.18 ± 2.09**	**71.82 ± 3.28**
RFE	7,937	3,360	69.85 ± 4.24	66.31 ± 4.08	68.16 ± 3.49	68.08 ± 4.16
STABLASSO	5,073	3,233	69.30 ± 4.28	**71.93 ± 2.25**	70.11 ± 2.90	70.62 ± 3.87

The best performance for each indicator is shown in bold.

TABLE 4 | The classification performance (in percentage) of four machine learning methods with GM and WM.

		GM + WM				
	RetainedFeas	OptimalFeas	Sen	Spec	Acc	AUC
SPEC	10,740	4,115	89.20 ± 2.31	79.26 ± 4.66	84.20 ± 3.01	84.23 ± 3.74
ReliefF	10,424	3,996	**89.66 ± 3.84**	80.01 ± 3.10	**84.92 ± 1.80**	**84.84 ± 3.48**
RFE	11,372	5,647	86.29 ± 3.07	78.47 ± 4.30	82.37 ± 3.26	82.38 ± 3.96
STABLASSO	4,425	2,753	87.43 ± 2.50	**80.67 ± 1.72**	84.25 ± 1.43	84.05 ± 1.88

The best performance for each indicator is shown in bold.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, we evaluated and compared the classification performance of four different feature selection methods for GM, WM, and GM+WM to investigate the stability of machine learning approaches. For GM or WM, the input data are GM or WM individually, and for GM+WM, the GM and WM have been stacked together and input to training machine learning models. Furthermore, distinguishing brain regions have been identified by feature selection techniques and compared with the findings of conventional statistical analysis.

All machine learning experiments were conducted on Windows with a 3.1GHz Intel Core i5 processor (4 cores) and 12GB RAM. The Python programming language was utilized for the experiments. In this study, spectral feature selection (SPEC) and ReliefF methods were implemented through *scikit-feature* (Li et al., 2018). Stability selection (STABLASSO) was implemented in *stability-selection* (Meinshausen and Bühlmann, 2010). Recursive feature elimination (RFE) was implemented

in *scikit-learn* (Pedregosa et al., 2012). The relevant method parameters were set as follows: (1) ReliefF: The number of neighbors samples k is set to 5. (2) STABLASSO: The LASSO *estimator* was calculated. We used $m = 1,000$ iterations. The experiments were performed with a threshold value of $\pi_{th} = 0.085$. (3) SVM-RFE: Linear SVM is used as the *estimator*, and the *step* parameter is set to 0.1. Leave the rest of the parameters as default values. Notably, the parameter C of the linear-SVM classifier played an important role in all experiments. In this study, we tested different values of this parameter, with $C = \{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1, 10^2, 10^3\}$, and we found that the performance of linear SVM was best when $C = 1$ for both GM and WM. Since each feature selection method requires to assign the number of features (for SPEC, ReliefF, and RFE) or the minimum threshold of stability scores (for STABLASSO) to determine the number of retained features, we retained different numbers of features based on the percentage of voxels in an iterative manner and ultimately selected the number of features with the best classification performance as the optimal number of features for

SPEC, ReliefF, and RFE. For STABLASSO, due to its expensive computational cost, the fixed threshold (threshold = 0.085) of stability scores has been used to determine the optimal number of features in this study. The optimal numbers of features (retained features) selected under different methods are listed in **Tables 2–4**.

4.1. Stability Analysis Based on Classification Performance

Tables 2–4 have listed the classification performance of the four feature selection methods on GM, WM, and GM+WM. As shown in **Tables 2**, **3**, the highest classification performance was 92.24% (specificity), 92.42% (sensitivity), 89.58% (accuracy), and 89.77% (AUC) for GM and 71.93% (specificity), 74.87% (sensitivity), 71.18% (accuracy), and 71.82% (AUC) for WM. Therefore, machine learning methods on GM have achieved better classification performance than those on WM, which has indicated that PD has a greater effect on the brain regions of GM than WM (Feldmann et al., 2010; Ji et al., 2010; Kang et al., 2015). As seen in **Table 4**, ReliefF method achieved the best classification performance: 89.66% (sensitivity), 80.01% (specificity), 84.92% (accuracy), and 84.84% (AUC). In addition, it can be seen that GM+WM was classified more effectively than WM but less effectively than GM. In addition, the variation curves of accuracy with the number of features under the SPEC, ReliefF, and RFE methods are provided in **Supplementary Figures S1**, **S2**.

It should be further noted in **Table 2** that the classification accuracy for GM had a maximum of 89.58% and a minimum of 84.98%, with a difference of approximately 4.6%. For WM, the highest accuracy was 71.18%, and the lowest accuracy was 68.16%, for a small difference of approximately 3%. For GM+WM classification experiment, the highest classification accuracy was 84.92%, and the lowest was 82.37%, for a difference of approximately 2.5%. These small differences have demonstrated that patients with PD can be stably distinguished from HCs with machine learning techniques. Furthermore, the receiver operating characteristic (ROC) curves have been illustrated in **Figure 2**. Generally, the effect of a classifier can also be evaluated by the area under the ROC curve (AUC). The machine learning methods whose AUCs are closest to 1 can be considered to have the best performance. The AUCs are given in **Tables 2–4**, respectively. Similar to the accuracy, the AUCs of different methods on different modalities also have subtle differences, and the ROC curves of different methods in GM/WM/GM+WM were approxiimately the same.

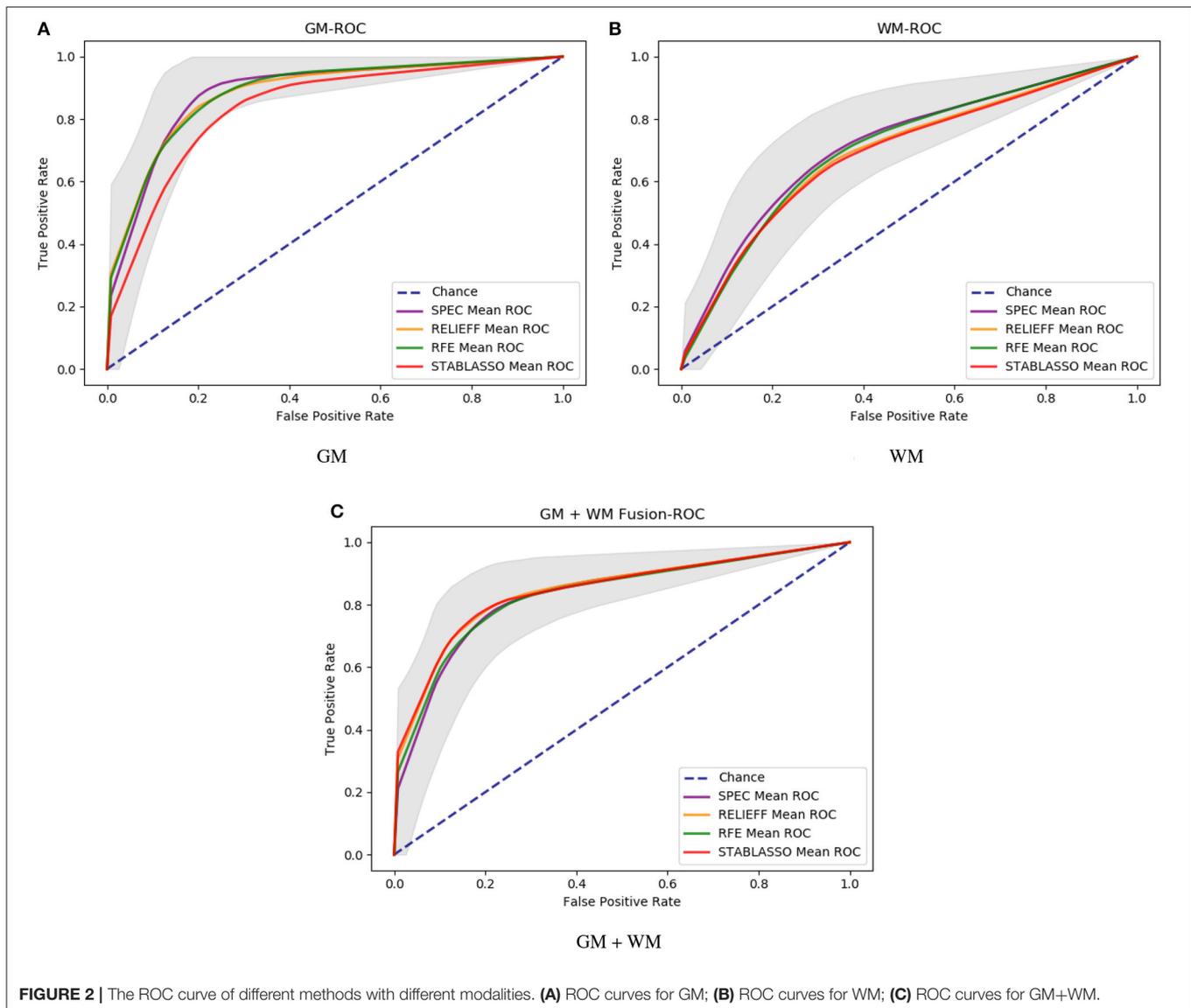## 4.2. Stability Analysis Based on Identified Brain Regions

In this section, we intended to explore the optimal feature subsets from retained features for each feature selection method, which are expected to be PD-associated brain changes in GM/WM. For this purpose, the most distinguished brain regions were defined as the optimal feature subsets if their frequency being selected in cross validation is greater than 80%. For machine learning, the original dimensions of input data for GM and WM are 11,745 and 19,844, respectively. The number of retained features for feature

selection and optimal feature subsets from retained features have been listed in **Tables 2–4**. Moreover, we have compared the distinguished brain regions of GM and WM based on the optimal feature subsets with the findings of previous studies and traditional statistical analysis.

4.2.1. Gray Matter Analysis

There have been numerous prior studies using gray matter to analyze Parkinson's disease. Long et al. (2012) found that the volume of GM was significantly reduced in the Paracentral lobule (PCL) and significantly increased in the Precentral gyrus (PreCG) and the bilateral Posterior cingulate gyrus (PCG) in the PD group. Adeli et al. (2016) found that GM tissue changes play the most important role in PD classification in most regions such as the PreCG, Supplementary motor area (SMA), Superior frontal gyrus (medial, SFGmed), Insula (INS), Median cingulate and paracingulate gyri (DCG), Calcarine fissure and surrounding cortex (CAL), Lingual gyrus (LING), Fusiform gyrus (FFG), Postcentral gyrus (PoCG), Superior parietal gyrus (SPG), Caudate nucleus (CAU), Lenticular nucleus, putamen (PUT), Lenticular nucleus, pallidum (PAL), Thalamus (THA), Superior temporal gyrus (STG), Middle temporal gyrus (MTG), and Inferior temporal gyrus (ITG). Peng et al. (2017) found in their study that the most sensitive GM biomarkers for distinguishing between patients with PD and healthy controls were the SPG, PCL, and Parahippocampal gyrus (PHG). Santos et al. (2013) carried out VBM on GM and observed atrophy in the following brain regions: OLF, Middle frontal gyrus (orbital part, ORBmid), Inferior frontal gyrus (orbital part, ORBinf), Superior frontal gyrus, (medial orbital, ORBsupmed), MTG, STG, INS, and PreCG. On the other hand, Xia et al. (2013) identified the STG, Superior occipital gyrus (SOG), SPG, Middle frontal gyrus (MFG), INS, PHG, and Amygdala (AMYG) as regions with altered GM volume using VBM. Jia et al. (2015) found in their study that GM volume decreased in the Superior frontal gyrus (orbital part, ORBsup), MFG, MTG, ITG, SPG, Inferior parietal, but supramarginal and angular gyri (IPL), Angular gyrus (ANG), and CAU, while GM volume increased in the ORBinf, MOG, Anterior cingulate and paracingulate gyri (ACG), PAL, PUT, and Hippocampus (HIP). Liu et al. (2018) found that the GM/WM brain regions contributing most to PD classification were mainly concentrated in the MOG, PUT, CAU, ORBsup, ACG, PreCG, HIP, Precuneus (PCUN), and PoCG. Rana et al. (2017) further improved the performance of PD classification by defining a fusion feature descriptor (FFD) to capture information and interrelationships between GM and WM tissues simultaneously. Ultimately, they found that the most discriminative brain regions were mainly located in the HIP, DCG, Inferior frontal gyrus (triangular part, IFGtriang), PreCG, MFG, ORBmid, and ACG. These distinguished brain regions have been listed in **Supplementary Table S2**.

In this study, the abnormal brain regions have been illustrated in **Figure 4** and reported based on AAL (Automated Anatomical Labeling atlas) with 116 brain regions (Tzourio-Mazoyer et al., 2002). We found that the SPEC method identified the largest number of features based on brain GM analysis, followed by ReliefF and RFE, and STABLASSO found the fewest, as seen

FIGURE 2 | The ROC curve of different methods with different modalities. **(A)** ROC curves for GM; **(B)** ROC curves for WM; **(C)** ROC curves for GM+WM.
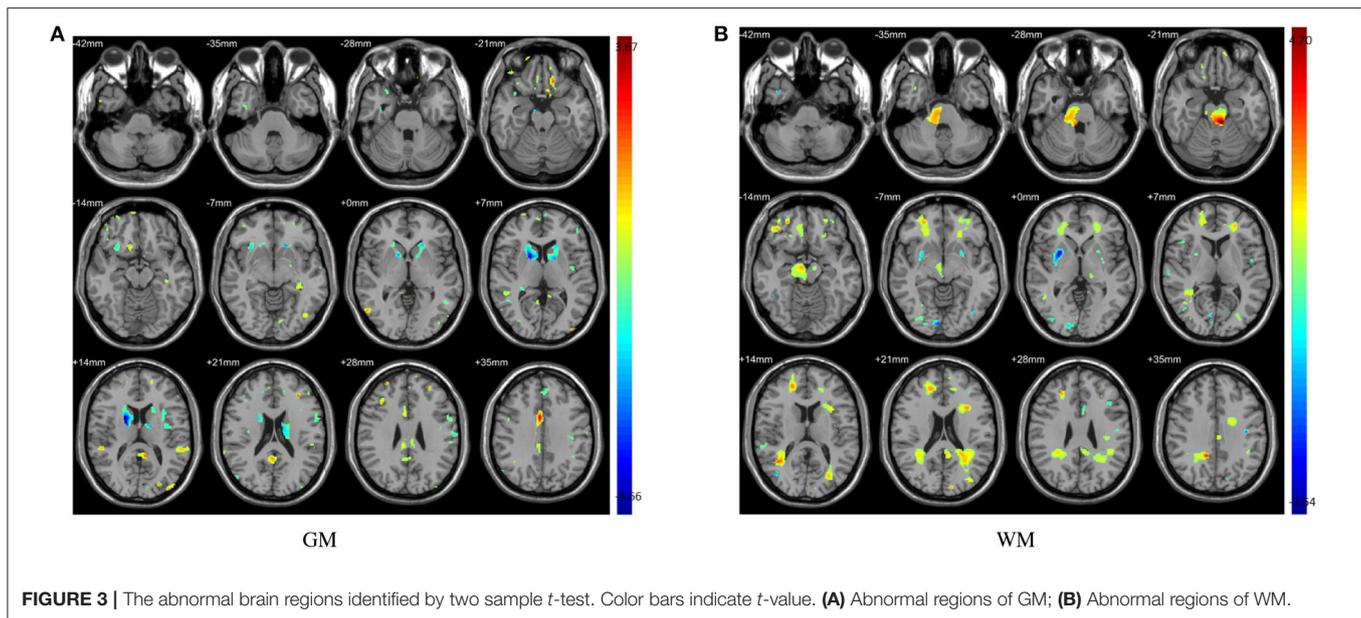
in **Supplementary Table S5**. Furthermore, we note that the significant brain regions found by SPEC included all the brain regions found by the other three methods. The brain regions found commonly by all four methods included the PCUN, CAU, STG, MFG, MOG, MTG, INS, SFGdor, ORBinf, IPL, and ANG. Except SFGdor, other brain regions have been reported in the previous studies. Furthermore, the findings of statistical analysis on GM have been listed in **Supplementary Table S3**. It is easy to find that most brain regions found by statistical analysis also have been identified by machine learning methods, except CAL with the cluster size 113. The possible reason is that the cluster size is too small or the changes in this brain region is subtle.

4.2.2. White Matter Analysis

Similarly, a great number of studies have been conducted to analyze the potential value of WM in the classification of

PD. Ding et al. (2011) found that the changes in WM tissue volume in most regions, including SOG, Inferior occipital gyrus (IOG), PCG, and PCL, contributed to the improvement of overall classification performance. Peng et al. (2017) also found that the most sensitive WM biomarkers for distinguishing between patients with PD and healthy controls were the ITG, FFG and MFG. Long et al. (2012) found that brain regions showing WM volume changes were mainly located in the PreCG, ORBinf, Rolandic operculum (ROL), OLF, HIP, AMYG, PoCG, and CAU. On the other hand, the MOG, IPL, PAL, THA, ITG, CAU, and PUT were identified by Adeli et al. (2016) as regions whose WM volume was affected. Rana et al. (2017) further improved the performance of PD classification by defining a fusion feature descriptor (FFD) to capture information and interrelationships between GM and WM tissues simultaneously. Ultimately, they found that the most discriminative brain regions were mainly concentrated in

FIGURE 3 | The abnormal brain regions identified by two sample *t*-test. Color bars indicate *t*-value. **(A)** Abnormal regions of GM; **(B)** Abnormal regions of WM.

the HIP, DCG, IFGtriang, PreCG, MFG, ORBmid, and ACG. Liu et al. (2018) found that the GM/WM brain regions most contributing to PD classification were mainly concentrated in the MOG, PUT, CAU, ORBsup, ACG, PreCG, HIP, PCUN, and PoCG. These distinguished brain regions have been listed in **Supplementary Table S2**.

The distinguished brain regions on WM identified by machine learning have been shown in **Figure 5** and reported based on AAL listed in **Supplementary Table S6**. The common regions found by four feature selection include PoCG, ORBmid, SFGdor, IOG, and LING. Except for SFGdor and LING, other regions have been reported in previous findings. The results detected by statistical analysis have been listed in **Supplementary Table S4** and shown in **Figure 3**. Similarly, most regions also have been found by machine learning methods, while some small regions have not been found, such as INS (cluster size = 99), IPL (cluster size = 109), ANG (cluster size = 47), IFGoperc (cluster size = 39), and SPG (cluster size = 57).
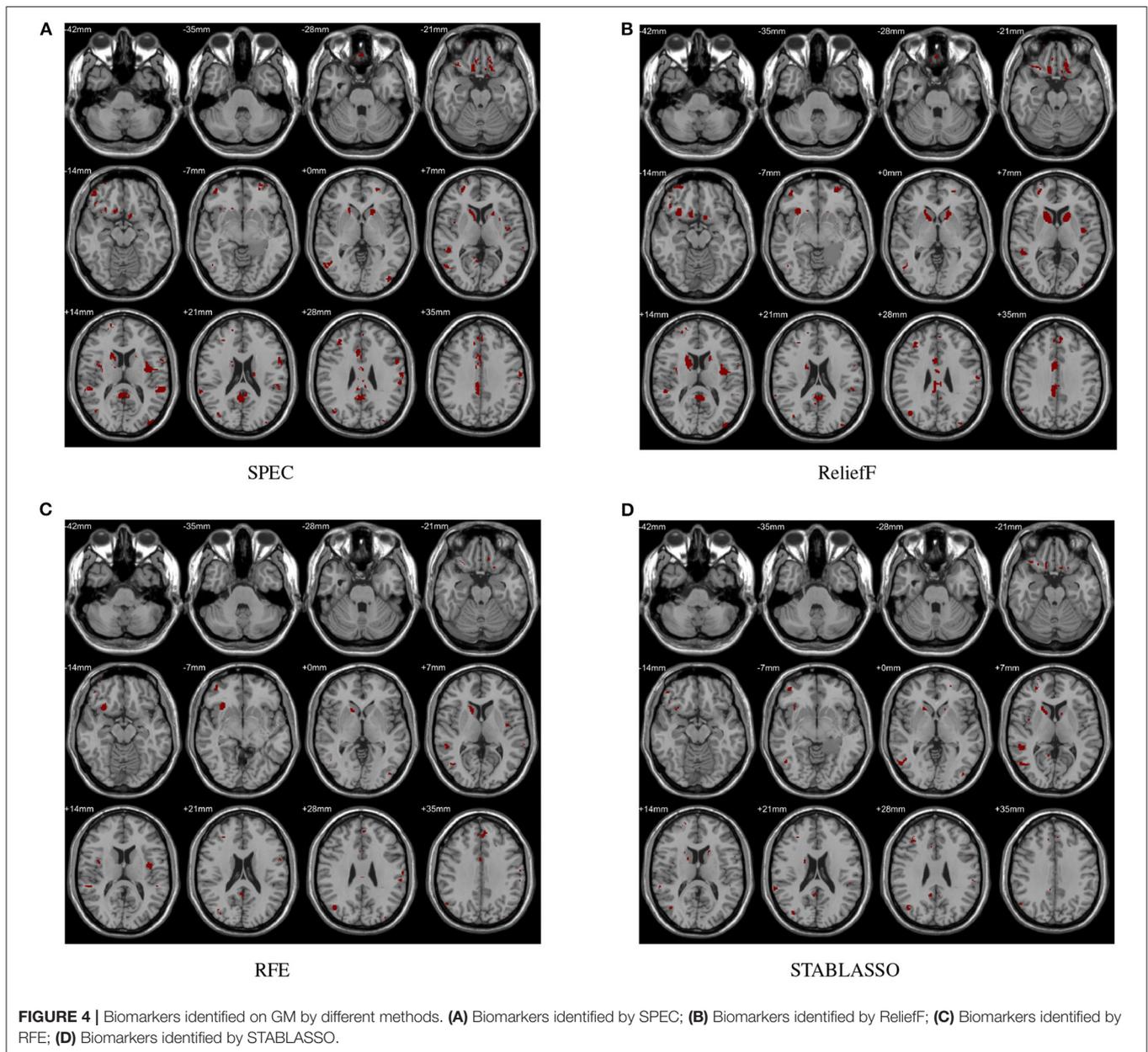
4.3. Discussion

Parkinson's disease (PD) is the most common neurodegenerative disorder (Kalia and Lang, 2015). At its core, it is a movement disease with early prominent death of dopaminergic neurons in the substantia nigra pars compacta (SNpc) (Kalia and Lang, 2015). Pathologically, it is also characterized by accumulation of misfolded –synuclein, which is found in intra-cyto-plasmic inclusions called Lewy bodies (Balestrino and Schapira, 2020). The gold standard for diagnosis of PD has been the presence of SNpc degeneration and Lewy pathology at post-mortem pathological examination (Kalia and Lang, 2015). It has been reported that the damage of substantia nigra correlated with basal ganglia dysfunction for patients with PD (Vitali et al., 2020). In our study, the changes of the caudate nucleus (CAU) in basal ganglia have also been found in GM, which have emphasized

the relationship between substantia nigra and basal ganglia again. However, the relationship between the nigral anatomical changes, evaluated as structural alterations or neuromelanin signal decrease and the dopaminergic nigro-striatal function needs to be further clarified (Prange et al., 2019).
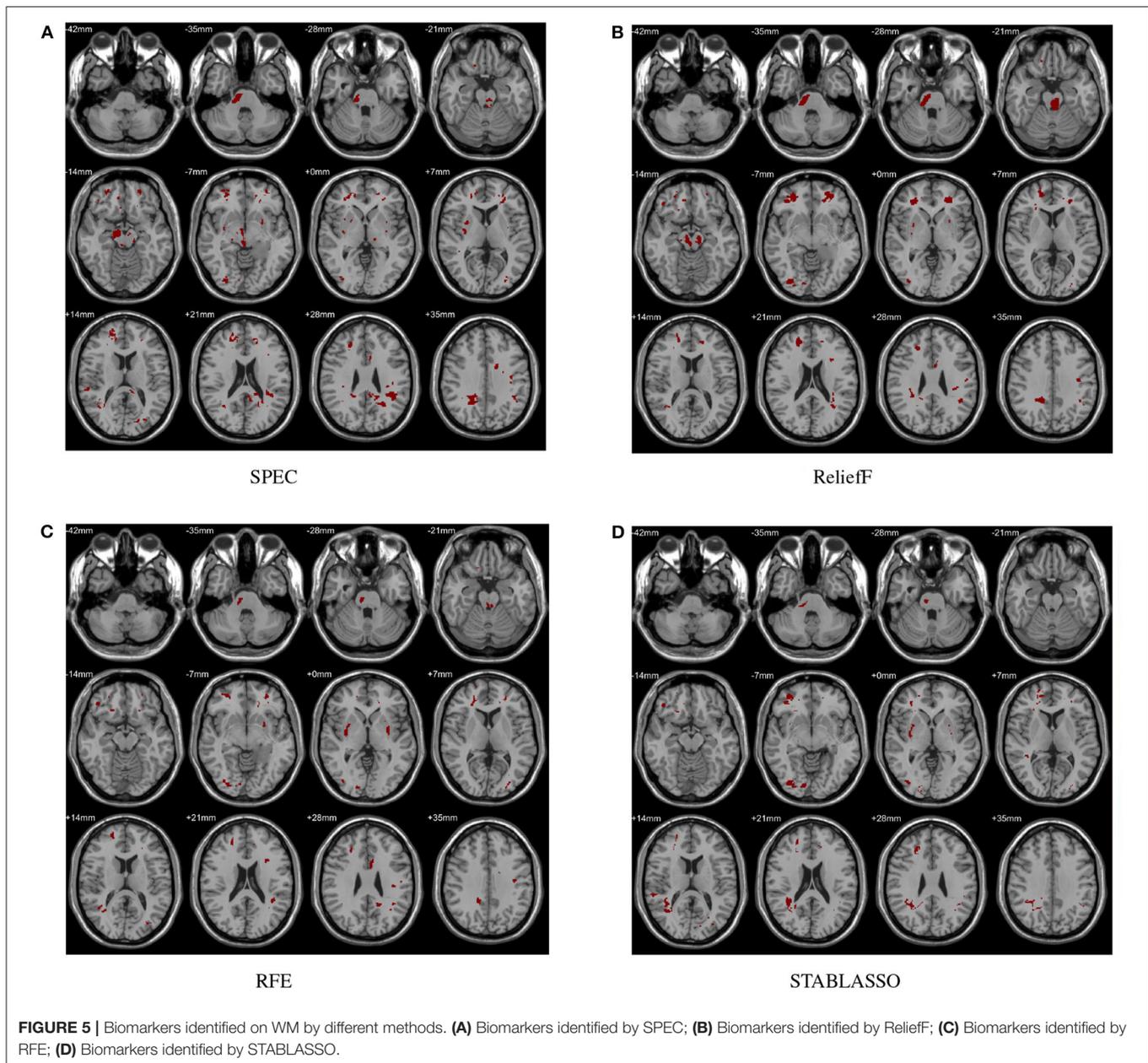
Recently, the detection and diagnosis of neurodegenerative diseases have attracted a great deal of research interest. There have been a large number of studies on GM and WM changes in patients with PD. They also identified several brain regions significantly affected by PD. Zheng et al. (2019) found that brain GM changes in the left anterior insula were associated with mild cognitive impairment in PD. In González Redondo et al. (2014), bilateral areas of atrophy in the middle frontal gyrus and bilateral GM loss in the medial-superior frontal gyrus were reported in patients with PD. Significantly increased WM density in the occipital lobes, posterior cingulate gyrus, and paracentral lobule was found in PD with olfactory dysfunction compared with healthy controls (Ding et al., 2011). As reported in Niethammer et al. (2013), a significant relationship between the loss of dopaminergic input to the caudate nucleus and the expression of a cognition-related disease network was found in unmedicated patients with PD. Gao et al. (2017) revealed that subcortical atrophy (for example, in the limbic lobe) was associated with impaired memory in patients with PD.

Although machine learning techniques have attracted great interest in the field of neuroscience, some researchers still remain skeptical about the stability and interpretability of machine learning outputs. As a result, instead of developing a new machine learning framework, this study focuses on horizontal comparison of different machine learning methods and vertical comparison between machine learning techniques and statistical analysis to enhance the credibility of machine learning. For this purpose, we applied a typical machine learning framework with feature selection and feature classification to distinguish PD from HCs. Four popular feature selection methods were utilized to find

FIGURE 4 | Biomarkers identified on GM by different methods. **(A)** Biomarkers identified by SPEC; **(B)** Biomarkers identified by ReliefF; **(C)** Biomarkers identified by RFE; **(D)** Biomarkers identified by STABLASSO.

a set of relevant and nonredundant features. A support vector machine was employed to build the classification models. An sMRI dataset was constructed to evaluate the performance of machine learning methods. Experimental results showed that the maximum accuracy levels based on GM and WM were 89.58 and 71.18%, respectively, which highlight the role of feature information in improving classification accuracy. Machine learning techniques are more effective than traditional statistical methods at extracting meaningful information from high-dimensional structural images and achieving stable classification results. Moreover, machine learning methods overcome the shortcomings of traditional statistical methods and extend simple analysis of differences from the group level to the individual

diagnostic level. In addition, machine learning techniques identified many GM and WM brain changes that could effectively distinguish patients with PD from controls. Through further comparison, most brain changes identified by machine learning techniques were found to be consistent with previous studies of structural MRI. At the same time, we also found some brain regions have not been reported in previous studies on structural MRI, such as SFGdor for GM and LING, SFGdor for WM, while SFGdor and LING have been confirmed in other modality studies. Li et al. (2020) has pointed out the reduced betweenness centrality in right SFGdor for PD. Auning et al. (2014) has recovered higher radical diffusivity in LING on DTI when compared PD and NC. Our findings support

FIGURE 5 | Biomarkers identified on WM by different methods. **(A)** Biomarkers identified by SPEC; **(B)** Biomarkers identified by ReliefF; **(C)** Biomarkers identified by RFE; **(D)** Biomarkers identified by STABLASSO.

that SFGdor and LING may be the important biomarkers in PD.

In light of the above analysis, the major contributions to the field by this study concern the following issues: (i) Stability—The horizontal comparison of different machine learning methods has demonstrated that machine learning techniques can robustly classify PD and HCs; (ii) Interpretability—In terms of feature selection, machine learning techniques are helpful in identifying the distinct brain changes that can be used as potential biomarkers of PD; (iii) Consistency—Based on the vertical comparison, the brain changes found by machine learning techniques are partly consistent with those found by statistical analysis. Despite superior performance and ease of utilization, the machine learning method used in the paper also has a few

disadvantages: (i) Although this study investigated classification performance on GM and WM by simply concatenating the GM and WM data, the correlation between the two types of tissues was ignored; such information could be helpful in improving the performance of machine learning. (ii) The dataset used in this study is limited and needs to be further expanded to make the results more generalizable. (iii) Machine learning techniques were used to classify PD solely on the basis of sMRI, and its needs; the same techniques need to be applied to more neuroimaging modalities to prove their effectiveness. (iv) Machine learning techniques need to be extended to other disease investigations to prove their universality. (v) To assess the stability of the discriminative brain regions, we have focused only on the brain regions that were identified by all four machine

learning methods and reported by previous research work at the same time. However, the biomarkers of PD may also include some brain regions identified by our machine learning methods but not reported by previous findings based on conventional statistical analysis.

In future research work, this study should be expanded upon in the following directions: (i) It cannot be denied that the pre-processing of MRI data is crucial for the subsequent data analysis and VBM analysis with different softwares and versions are controversial until now. As pointed out in Farokhian et al. (2017), CAT12 has shown great advantage to detect the volumetric alterations compared with VBM8, so it necessitates to conduct further investigation based on CAT12 in the future. (ii) In order to fully consider the correlation of different modalities, it is necessary to integrate multi-modal features to capture the information and interrelationship simultaneously. (iii) In order to further demonstrate the universality of machine learning, it might be worthwhile to study the consistency of the results of machine learning techniques on a larger dataset.

5. CONCLUSION

This study was designed to analyze the changes in GM and WM brain tissue in PD. In this study, we adopted four popular feature selection methods to extract a group of the most relevant and nonredundant features, and we then used linear support vector machines to construct classification models for WM and GM. In the experiment, we have achieved stable and satisfactory classification performance on both GM and WM. The brain changes identified by feature selection partially corresponded to previous findings, such as PCUN, CAU, STG, MFG, MOG, MTG, INS, ORBinf, IPL, ANG in GM and PoCG, ORBmid, and IOG in WM, which proved the stable performance of machine learning. Furthermore, machine learning also found commonly the brain changes in SFGdor and LING, which have been proved in other studies. These findings have not only demonstrated the good stability of machine learning technique, but also the superior ability to explore the brain changes, and these brain changes can be regarded as potential biomarkers of PD. These potential biomarkers can help clarify the neurological mechanisms of PD. Moreover, machine learning techniques can provide a decision system for the diagnosis of PD and can be generalized to the diagnosis of other neurodegenerative diseases.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Parkinson's Progression Markers Initiative (PPMI), https://www.ppmi-info.org/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Michael J. Fox Foundation for Parkinson's Research. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CS, WZ, HJ and XL have contributed equally to this study. CS, WZ, CL and YT conceived and designed this study, HJ, XL and JK participated in the analysis of MRI dataset, YD, XDY, XY, JZ helped to draft the manuscript. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

The authors express their gratitude to the editor and the reviewers for their useful comments and suggestions that helped in improving the quality of the research work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom. 2021.735991/full#supplementary-material

REFERENCES

Adeli, E., Shi, F., An, L., Wee, C.-Y., Wu, G., and Wang, T. (2016). Joint feature-sample selection and robust diagnosis of Parkinson's disease from mri data. *Neuroimage* 141, 206–219. doi: 10.1016/j.neuroimage.2016.05.054

Agosta, F., Canu, E., Stojkovic, T., Pievani, M., Tomic, A., Sarro, L., et al. (2013). The topography of brain damage at different stages of Parkinson's disease. *Hum. Brain Mapp.* 34, 2798–2807. doi: 10.1002/hbm. 22101

Altamura, A., Maggioni, E., Dhanoa, T., Ciappolino, V., Paoli, R. A., Cremaschi, L., et al. (2017). The impact of psychosis on brain anatomy in bipolar disorder: a structural mri study. *J. Affect. Disord.* 233, 100–109. doi: 10.1016/j.jad.2017.11.092

Altman, N., and Krzywinski, M. (2018). The curse(s) of dimensionality. *Nat. Methods* 15, 397. doi: 10.1038/s41592-018-0019-x

Anguita, D., Ghio, A., and Ridella, S. (2011). Maximal discrepancy for support vector machines. *Neurocomputing* 74, 1436–1443. doi: 10.1016/j.neucom.2010.12.009

Auning, E., Kjærvik, V. K., Selnes, P., Aarsland, D., Haram, A., Bjørnerud, A., et al. (2014). White matter integrity and cognition in Parkinson's disease: a cross-sectional study. *BMJ Open* 4:e003976. doi: 10.1136/bmjopen-2013-003976

Babu, G. S., Suresh, S., and Mahanand, B. S. (2014). A novel pbl-mcrbfn-rfe approach for identification of critical brain regions responsible for Parkinson's disease. *Expert Syst. Appl.* 41, 478–488. doi: 10.1016/j.eswa.2013.07.073

Balestrino, R., and Schapira, A. V. H. (2020). Parkinson disease. *Eur. J. Neurol.* 27, 27–42. doi: 10.1111/ene.14108

Biesiada, J., and Duch, W. (2005). "Feature selection for high-dimensional data: a kolmogorov-smirnov correlation-based filter," in *International Conference on Computer Recognition Systems* (Berlin; Heidelberg: Springer).

Chaudhuri, K. R., and Schapira, A. H. (2009). Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment. *Lancet Neurol.* 8, 464–474. doi: 10.1016/S1474-4422(09)70068-7

Chen, Z. H., Yan, T., Wang, E., Jiang, H., Tang, Y., Yu, X., et al. (2020). Detecting abnormal brain regions in schizophrenia using structural mri via machine learning. *Comput. Intell. Neurosci.* 2020: 6405930 doi: 10.1155/2020/6405930

Cherubini, A., Morelli, M., Nisticó, R., Salsone, M., Arabia, G., Vasta, R., et al. (2014). Magnetic resonance support vector machine discriminates between Parkinson disease and progressive supranuclear palsy. *Mov. Disord.* 29, 266–269. doi: 10.1002/mds.25737

Chu, C., Hsu, A.-L., Chou, K.-H., and Bandettini, P. (2011). Does feature selection improve classification accuracy? impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage* 60, 59–70. doi: 10.1016/j.neuroimage.2011.11.066

Cigdem, O., Yilmaz, A., Beheshti, I., and Demirel, H. (2018). "Comparing the performances of pdf and pca on Parkinson's disease classification using structural MRI images," in *2018 26th Signal Processing and Communications Applications Conference (SIU)* (Izmir: IEEE), 1–4.

Ding, H., Wu, X., Zhang, K., Fan, F., Zhu, C., Li, K., et al. (2011). Change of olfactory function associated structures in Parkinson's disease: a voxel - based morphometry study. *Chin. J. Contemporary Neurol. Neurosurg.* 11, 54–59. doi: 10.3969/j.issn.1672-6731.2011.01.012

Duchesne, S., Rolland, Y., and Vérin, M. (2009). Automated computer differential classification in parkinsonian syndromes via pattern analysis on mri. *Acad. Radiol.* 16, 61–70. doi: 10.1016/j.acra.2008.05.024

Fahn, S., Elton, R., and Committee, U. (1997). *The Unified Parkinson's Disease Rating Scale (University of Liverpool Library)*, Vol. 2, 153–163. doi: 10.1017/CBO9780511763281.013

Farokhian, F., Beheshti, I., Sone, D., and Matsuda, H. (2017). Comparing cat12 and vbm8 for detecting brain morphological abnormalities in temporal lobe epilepsy. *Front. Neurol.* 8:428. doi: 10.3389/fneur.2017.00428

Feldmann, A., Illes, Z., Kosztolanyi, P., Illes, E., and Nagy, F. (2010). Morphometric changes of gray matter in Parkinson's disease with depression: a voxel-based morphometry study. *Movement Disord.* 23, 42–46. doi: 10.1002/mds.21765

Folstein, M., Folstein, S., and McHugh, P. (1975). Mini-mental state-practical method for grading cognitive state of patients for clinicians. *J. Psychiatr. Res.* 12, 89–98. doi: 10.1016/0022-3956(75)90026-6

Fox, S. H., Regina, K., Shen-Yang, L., Bernard, R., Klaus, S., Miguel, C., et al. (2011). The movement disorder society evidence-based medicine review update: treatments for the motor symptoms of Parkinson's disease. *Mov. Disord.* 26, S2–S41. doi: 10.1002/mds.23829

Gao, Y., Nie, K., Huang, B., Mei, M., Guo, M., Xie, S., et al. (2017). Changes of brain structure in Parkinson's disease patients with mild cognitive impairment analyzed via vbm technology. *Neurosci. Lett.* 658, 121–132. doi: 10.1016/j.neulet.2017.08.028

González Redondo, R., Garcia-Garcia, D., Clavero, P., Gasca-Salas, C., Garcia-Eulate, R., Zubieta, J., et al. (2014). Grey matter hypometabolism and atrophy in Parkinson's disease with cognitive impairment: a two-step process. *Brain* 137(Pt 8), 2356–2367. doi: 10.1093/brain/awu159

Group, P. S. (2002). Dopamine transporter brain imaging to assess the effects of pramipexole vs levodopa on Parkinson's disease progression. *JAMA* 287, 1653–1661. doi: 10.1001/jama.287.13.1653

Guyon, I. (2001). Erratum: gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422. doi: 10.1023/A:1012487302797

Guyon, I., and Elisseeff, A. (2003). An introduction of variable and feature selection. *J. Mach. Learn. Res.* 3, 1157–1182. doi: 10.1162/508153244303322753616

Heim, B., Krismer, F., Marzi, R. D., and Seppi, K. (2017). Magnetic resonance imaging for the diagnosis of Parkinson's disease. *J. Neural Transm.* 124, 1–50. doi: 10.1007/s00702-017-1717-8

Hirschauer, T. J., Adeli, H., and Buford, J. A. (2015). Computer-aided diagnosis of Parkinson's disease using enhanced probabilistic neural network. *J. Med. Syst.* 39, 179. doi: 10.1007/s10916-015-0353-9

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *J. Neurol Neurosurg. Psychiatry* 79, 368–376. doi: 10.1136/jnnp.2007.131045

Ji, E. L., Park, B., Song, S. K., Sohn, Y. H., Park, H. J., and Lee, P. H. (2010). A comparison of gray and white matter density in patients with Parkinson's disease dementia and dementia with lewy bodies using voxel-based morphometry. *Mov. Disord.* 25, 28–34. doi: 10.1002/mds.22858

Jia, X., Liang, P., Li, Y., Shi, L., Wang, D., and Li, K. (2015). Longitudinal study of gray matter changes in Parkinson disease. *AJNR Am. J. Neuroradiol.* 36, 2219–2226. doi: 10.3174/ajnr.A4447

Kalia, L. V., and Lang, A. E. (2015). Parkinson's disease. *Lancet* 386, 896–912. doi: 10.1016/S0140-6736(14)61393-3

Kang, D.-Z., Chen, F.-Y., Wang, F.-Y., Wu, G.-R., Liu, Y., Wu, G., et al. (2015). Brain gray matter volume changes associated with motor symptoms in patients with Parkinson's disease. *Chin. Neurosurg. J.* 1, 9. doi: 10.1186/s41016-015-0003-6

Kikuchi, K., Hiwatashi, A., Togao, O., Yamashita, K., Somehara, R., Kamei, R., et al. (2017). Structural changes in Parkinson's disease: voxel-based morphometry and diffusion tensor imaging analyses based on 123i-mibg uptake. *Eur. Radiol.* 27, 5073–5079. doi: 10.1007/s00330-017-4941-6

Kira, K., and Rendell, L. A. (1992). A practical approach to feature selection. *Mach. Learn. Proc.* 48, 249–256. doi: 10.1016/B978-1-55860-247-2.50037-1

Kohavi, R., and John, G. H. (1997). Wrappers for feature subset selection. *Artif. Intell.* 97, 273–324. doi: 10.1016/S0004-3702(97)00043-X

Koller, W. C., Glatt, S., Vetere-Overfield, B., and Hassanein, R. (1989). Falls and Parkinson's disease. *Clin. Neuropharmacol.* 12, 98. doi: 10.1097/00002826-198908000-00006

Lee, J., and Verleysen, M. (2007). *Nonlinear Dimensionality Reduction, Vol. 8226.* New York, NY: Springer. doi: 10.1007/978-0-387-39351-3

Lei, H., Zhao, Y., Wen, Y., Luo, Q., Cai, Y., Liu, G., et al. (2018). Sparse feature learning for multi-class Parkinson's disease classification. *Technol. Health Care* 26, 193–203. doi: 10.3233/THC-174548

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., et al. (2018). Feature selection: a data perspective. *ACM Comput. Surveys* 50, 94. doi: 10.1145/3136625

Li, J., Zeng, Q., Zhou, W., Zhai, X., and Cheng, G. (2020). Altered brain functional network in Parkinson disease with rapid eye movement sleep behavior disorder. *Front. Neurol.* 11:563624. doi: 10.3389/fneur.2020.563624

Liu, G.-D., Li, Y.-C., Zhang, W., and Zhang, L. (2020). A brief review of artificial intelligence applications and algorithms for psychiatric disorders. *Engineering* 6, 462–467. doi: 10.1016/j.eng.2019.06.008

Liu, L., Wang, Q., Adeli, E., Zhang, L., and Zhang, H. (2018). Exploring diagnosis and imaging biomarkers of Parkinson's disease via iterative canonical correlation analysis based feature selection. *Comput. Med. Imaging Graphics* 67, 21–29. doi: 10.1016/j.compmedimag.2018.04.002

Lojowska, M., Ling, S., Roelofs, K., and Hermans, E. (2018). Visuocortical changes during a freezing-like state in humans. *Neuroimage* 179, 313–325. doi: 10.1016/j.neuroimage.2018.06.013

Long, D., Wang, J., Xuan, M., Gu, Q., Xu, X., Kong, D., et al. (2012). Automatic classification of early Parkinson's disease with multi-modal mr imaging. *PLoS ONE* 7:e47714. doi: 10.1371/journal.pone.0047714

Machado, A., Rezai, A., Kopell, B., Gross, R., Sharan, A., and Benabid, A. (2006). Deep brain stimulation for Parkinson's disease: surgical technique and perioperative management. *Mov. Disord.* 21, S247–S258. doi: 10.1002/mds.20959

Mallet, N., Pogosyan, A., Sharott, A., Csicsvari, J., and Magill, P. (2008). Disrupted dopamine transmission and the emergence of exaggerated beta oscillations in subthalamic nucleus and cerebral cortex. *J. Neurosci.* 28, 4795–4806. doi: 10.1523/JNEUROSCI.0123-08.2008

Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., et al. (2011). The Parkinson Progression Marker Initiative (PPMI). *Prog. Neurobiol.* 95, 629–635. doi: 10.1016/j.pneurobio.5562011.09.005

Marquand, A., Filippone, M., Ashburner, J., Girolami, M., Mourão–Miranda, J., Barker, G., et al. (2013). Automated, high accuracy classification of parkinsonian disorders: a pattern recognition approach. *PLoS ONE* 8:e69237. doi: 10.1371/journal.pone.0069237

Meinshausen, N., and Bühlmann, P. (2010). Stability selection. *J. R. Stat. Soc.* 72, 417–473. doi: 10.1111/j.1467-9868.2010.00740.x

Mishra, V., Sreenivasan, K., Zhuang, X., Yang, Z., Cordes, D., and Walsh, R. (2019). Influence of analytic techniques on comparing dti-derived measurements in early stage Parkinson's disease. *Heliyon* 5:e01481. doi: 10.1016/j.heliyon.2019.e01481

Moccia, M., Tedeschi, E., Ugga, L., Erro, R., Picillo, M., Caranci, F., et al. (2016). White matter changes and the development of motor phenotypes in *de novo* Parkinson's disease. *J. Neurol. Sci.* 367, 215–219. doi: 10.1016/j.jns.2016.06.015

Morris, M. E. (2000). Movement disorders in people with Parkinson disease: a model for physical therapy. *Phys. Ther.* 80, 578–597. doi: 10.1093/ptj/80.6.578

Mwangi, B., Tian, T., and Soares, J. (2013). A review of feature reduction techniques in neuroimaging. *Neuroinformatics* 12, 229–244. doi: 10.1007/s12021-013-9204-3

Niethammer, M., Tang, C., Ma, Y., Mattis, P., Ko, J. H., Dhawan, V., et al. (2013). Parkinson's disease cognitive network correlates with caudate dopamine. *Neuroimage* 78, 204–209. doi: 10.1016/j.neuroimage.2013.03.070

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2012). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Peker, M., Sen, B., and Delen, D. (2015). Computer-aided diagnosis of Parkinson's disease disease using complex-valued neural networks and mrmr feature selection algorithm. *J. Healthc Eng.* 6, 281. doi: 10.1260/2040-2295.6.3.281

Peng, B., Wang, S., Zhou, Z., Liu, Y., Tong, B., Zhang, T., et al. (2017). A multilevel-roi-features-based machine learning method for detection of morphometric biomarkers in Parkinson's disease. *Neurosci. Lett.* 651, 88–94. doi: 10.1016/j.neulet.2017.04.034

Peng, J., Zhou, Y., and Chen, C. (2015). Region-kernel-based support vector machines for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* 53, 1–15. doi: 10.1109/TGRS.2015.2410991

Perez-Lloret, S., and Rascol, O. (2010). Dopamine receptor agonists for the treatment of early or advanced Parkinson's disease. *CNS Drugs* 24, 941. doi: 10.2165/11537810-000000000-00000

Prange, S., Metereau, E., and Thobois, S. (2019). Structural imaging in Parkinson's disease: new developments. *Curr. Neurol. Neurosci. Rep.* 19, 50. doi: 10.1007/s11910-019-0964-5

Price, S., Paviour, D., Scahill, R., Stevens, J., Rossor, M., Lees, A., et al. (2004). Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and Parkinson's disease. *Neuroimage* 23, 663–669. doi: 10.1016/j.neuroimage.2004.06.013

Rana, B., Juneja, A., Saxena, M., Gudwani, S., Kumaran, S., Agrawal, R., et al. (2015a). Regions-of-interest based automated diagnosis of Parkinson's disease using t1-weighted mri. *Expert Syst. Appl.* 42, 4506–4516. doi: 10.1016/j.eswa.2015.01.062

Rana, B., Juneja, A., Saxena, M., Gudwani, S., Kumaran, S., Behari, M., et al. (2015b). Graph theory based spectral feature selection for computer aided diagnosis of Parkinson's disease using t1-weighted mri. *Int. J. Imaging Syst. Technol.* 25, 245–255. doi: 10.1002/ima.22141

Rana, B., Juneja, A., Saxena, M., Gudwani, S., Kumaranc, S., Agrawal, R., et al. (2017). Relevant 3d local binary pattern based features from fused feature descriptor for differential diagnosis of Parkinson's disease using structural mri. *Biomed. Signal Process Control.* 34, 134–143. doi: 10.1016/j.bspc.2017.01.007

Rektor, I., Svatkova, A., Vojtisek, L., Zikmundová, I., Vanícek, J., Király, A., et al. (2018). White matter alterations in Parkinson's disease with normal cognition precede grey matter atrophy. *Parkinson. Relat. Disord.* 46, e9–e10. doi: 10.1016/j.parkreldis.2017.11.030

Robnik-Šikonja, M., and Kononenko, I. (2003). Theoretical and empirical analysis of relieff and rrelieff. *Mach. Learn.* 53, 23–69. doi: 10.1023/A:1025667309714

Sain, S. R. (1997). The nature of statistical learning theory. *Technometrics* 38, 409–409. doi: 10.1080/00401706.1996.10484565

Salvatore, C., Cerasa, A., Castiglioni, I., Gallivanone, F., Augimeri, A., López, M., et al. (2013). Machine learning on brain mri data for differential diagnosis of Parkinson's disease and progressive supranuclear palsy. *J. Neurosci. Methods* 222, 230–237. doi: 10.1016/j.jneumeth.2013.11.016

Sánchezmaroño, N., Alonsobetanzos, A., and Tombillasanromán, M. (2007). "Filter methods for feature selection–a comparative study," in *Intelligent Data Engineering Automated Learning-ideal, International Conference* (Birmingham).

Santos, M., Campos, L., Guimarães, R., Piccinin, C., Azevedo, P., and Piovesana, L. (2013). Whole brain voxel based morphometry analysis in idiopathic Parkinson's disease [abstract]. *Mov. Disord.* 28.

Schwarz, S., Rittman, T., Gontu, V., Morgan, P. S., Bajaj, N., D. and Auer, P. (2011). T1-weighted mri shows stage-dependent substantia nigra signal loss in Parkinson's disease. *Mov. Disord.* 26, 1633–1638. doi: 10.1002/mds.23722

Summerfield, C., Junqué, C., Tolosa, E., Salgado-Pineda, P., Gómez-Ansón, B., Martí, M., et al. (2005). Structural brain changes in Parkinson disease with dementia. *Arch. Neurol.* 62, 281–285. doi: 10.1001/archneur.62.2.281

Tohka, J., Moradi, E., and Huttunen, H. (2016). Comparison of feature selection techniques in machine learning for anatomical brain mri in dementia. *Neuroinformatics* 14, 279–296. doi: 10.1007/s12021-015-9292-3

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Vapnik, V. (1999). Vapnik vn. an overview of statistical learning theory. *IEEE Trans. Neural. Netw.* 10, 988–999. doi: 10.1109/72.788640

Vitali, P., Pan, M. I., Palesi, F., Germani, G., Faggioli, A., Anzalone, N., et al. (2020). Substantia nigra volumetry with 3-T MRI in *de novo* and advanced parkinson disease. *Radiology* 296, 401–410. doi: 10.1148/radiol.2020191235

Wang, S., Tang, J., and Liu, H. (2015). "Embedded unsupervised feature selection". in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, AAAI'15*. (AAAI Press), 470–476.

Wang, Y., Tang, Y. Y., Li, L., Hong, C., and Pan, J. (2017). Atomic representation-based classification: theory, algorithm and applications. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 6–19. doi: 10.1109/TPAMI.2017.2780094

Wang, Y., Tang, Y. Y., Li, L., and Zheng, X. (2019). Block sparse representation for pattern classification: theory, extensions and applications. *Pattern Recognit.* 88, 198–209. doi: 10.1016/j.patcog.2018.11.026

Xia, J., Wang, J., Tian, W., Ding, H., Wei, Q., Huang, H., et al. (2013). Magnetic resonance morphometry of the loss of gray matter volume in Parkinson's disease patients. *Neural Regen. Res.* 8, 2557–2565. doi: 10.3969/j.issn.1673-5374.2013.27.007

Zhao, Z., and Liu, H. (2007). "Spectral feature selection for supervised and unsupervised learning," in *Proceedings of the 24th International Conference on Machine Learning, ICML '07* (New York, NY: Association for Computing Machinery), 1151–1157. doi: 10.1145/1273496.1273641

Zheng, D., Chen, C., Song, W., Yi, Z., Zhao, P., Zhong, J., et al. (2019). Regional gray matter reductions associated with mild cognitive impairment in Parkinson's disease: a meta-analysis of voxel-based morphometry studies. *Behav. Brain Res.* 371:111973. doi: 10.1016/j.bbr.2019.111973

Ziegler, D., and Augustinack, J. (2013). Harnessing advances in structural mri to enhance research on Parkinson's disease. *Imaging Med.* 5, 91–94. doi: 10.2217/iim.13.8