Towards the identification of Parkinsons Disease using only T1 MR Images

Sara Soltaninejad, Irene Cheng, and Anup Basu

Dept. of Computing Science, University of Alberta, Canada

Abstract. Parkinsons Disease (PD) is one of the most common type of neurological disease caused by progressive degeneration of dopaminergic neurons in the brain. Even though there is no fixed cure for this neurodegenerative disease, earlier diagnosis and resulting earlier treatment can help patients have better quality of life. Magnetic Resonance Imaging (MRI) is one of the most popular means of diagnosis in recent years because it avoids harmful radiations. In this paper, we investigate the plausibility of using MRIs for automatically diagnosing PD. Our proposed method has three main steps : 1) Preprocessing, 2) Feature Extraction, and 3) Classification. The Fressurfer library is used for the first and second steps. For classification, three main types of the classifiers, including Logistic Regression (LR), Random Forest (RF) and Support Vector Machine (SVM) are applied and their classification ability is compared. The Parkinsons Progression Markers Initiative (PPMI) data set is used for evaluation of the proposed method. The proposed system is shown to be promising in assisting the diagnosis of PD.

1 Introduction

Parkinson's Disease (PD) is the second most important neurodegenerative disease after Alzheimer's Disease (AD) that affects middle age and elderly people. The statistical information presented by Parkinsons News Today [1] shows that an estimated seven to ten million people worldwide have Parkinsons disease. PD causes progressive loss of dopamine generating neurons in the brain resulting in two types of symptoms, including motor and non-motor. The motor symptoms are bradykinesia, muscles rigidity, tremor and abnormal gait [2] whereas non-motor symptoms show mental disorders, sleep problems and sensory disturbance [3]. Even though there are some medical methods for diagnosing and progress determination of PD, the results from these experiments are subjective and depends on the expertise of the clinicians. On the other hand, clinicians are expensive and the process is time consuming for the patients [4]. Neuroimaging techniques have significantly improved the diagnosis of neurodegenerative diseases. There are different types of neuro imaging techniques of which Magnetic Resonance Imaging (MRI) is one of the most popular one because it is a cheap and non-invasiveness method. People with PD show their symptoms when they lose almost 80% of their brain dopamine [5]. All of these facts prove the urgent need to have a Computer Aided Diagnosis (CAD) system for automatic detection

of this type of disease. In recent years machine learning has shown remarkable results in the medical image analysis field. The proposed CAD system in neuro disease diagnosis uses different types of imaging data including Single-Photon Emission Computed Tomography (SPECT) (Prashanth et al. [6]), DTI, Positron Emission Tomography (PET)(Loane and Politis [7]) and MRI. In this study, the goal is to utilize structural MRI (sMRI) for developing automated CAD for early diagnosis of PD. Focke et al. [8] proposed a method for PD classification using MR Images. The proposed method in [8] used Gray Matter (GM) and White Matter (WM) individually with an SVM classifier. Voxel-based morphometry (VBM) has been used for preprocessing and feature extraction. The reported results show poor performance 39.53% for GM and 41.86% for WM. In [9] Babu et al. proposed a CAD system for diagnosing PD. Their method has three general steps: feature extraction, feature selection and classification. In the first part the VBM is used over GM to construct feature data. For the feature selection, recursive feature elimination (RFE) was used to select the most discriminative features. In the last step, projection based learning and meta-cognitive radial basis function was used for classification, which resulted in 87.21% accuracy. The potential biomarker for PD is identified as the superior temporal gyrus. The limitation in this work is that VBM is univariate and RFE is computationally expensive. Salvatore et al. [9], proposed a method that used PCA for feature extraction. The PCA was applied on normalized skull stripped MRI data. Then, SVM was used as the classifier, giving 85.8% accuracy. Rana et al. [10] extracted features over the three main tissues of the brain consisting of WM, GM and CSF. Then, they used t-test for feature selection and in the next step SVM for classification. This resulted in 86.67% accuracy for GM and WM and 83.33% accuracy for CSF. In their other work [11], graph-theory based spectral feature selection method was applied to select a set of discriminating features from the whole brain volume. A decision model was built using SVM as a classifier with a leave-one-out cross-validation scheme, giving 86.67% accuracy. The proposed method in [4] was not focused on just individual tissues (GM,WM and CSF) but it considered the relationship between these areas because the morphometric change in one tissue might affect other tissues. $3D$ LBP was used as a feature extraction tool which could produce structural and statistical information. After that, minimum redundancy and maximum relevance with t-test are used as feature selection methods to get the most discriminative and non-redundant features. In the end, SVM is used for classification which gives 89.67% accuracy. In [13], the low level features (GM, cortical volume, etc.) and the high level features (region of interest (ROI) connectivity) are combined to perform a multilevel ROI feature extraction. Then, filter and wrapper feature selection method is followed up with multi kernel SVM to achieve 85.78% accuracy for differentiation of PD and healthy control (HC) data. Adeli et al. [14] proposed a new joint feature-sample selection (JFSS) procedure, which jointly selects the best subset of the most discriminative features and the best sample to build a classification model. Experimental results on both synthetic and publicly available PD datasets show promising results.

In our paper, a CAD system is presented for diagnosing of PD by using only MR T1 images. Our goal is to avoid other modalities, like PET etc., that expose patients to harmful radiation. The general steps of the proposed method is shown in Fig.1 and includes preprocessing, feature extraction and classification.

The remaining sections of this paper are structured as follows: Section 2 and 3 covers materials and methods, which provides details on the dataset, preprocessing and the proposed method for PD classification. Experimental results and discussion are covered in Section 4. Finally, Section 5 presents the conclusion.

2 Data Set

The data used in the preparation of this article is the T1-weighted brain MR images obtained from the PPMI database `www.ppmi-info.org/data`. For up-to-date information on the study please visit `www.ppmiinfo.org`. PPMI is a landmark, large-scale, international and multi-center study to identify PD progression biomarkers [15]. The data that is used in our study contains the original T1 MR image of 598 samples with 411 PD and 187 Control. Furthermore, it contains demographic or clinical information on the age and sex of the subjects. The summary of the data base is presented in Table 1. Based on the demographic information in this table, the balance of dataset is presented for the two type of classes which are Parkinson disease (PD) and healthy control (HC).

Table 1: Demographics of the PPMI

Data Type	**Class**		**Sex**		**Age**		
	PD	**HC**	**F**	**M**	**(25-50)**	**(50-76)**	**(75-100)**
Number of Subjects	411	187	217	381	81	472	45

3 Proposed Method

The overview of our proposed method is presented in Fig.1 which has 3 general steps including: 1- Preprocessing; 2- Feature Extraction; and 3- Classification. Next, each step is explained in detail. The goals of CAD system goals are:

1. Extracting the volume based features from the MR T1 images using the automated surface-based analysis package FreeSurfer.
2. Comparing the capability of different types of classifiers for diagnosing PD.

3.1 Preprocessing

Preprocessing is an essential step in designing the CAD system which provides informative data for the next steps. In this paper, for computing the volumetric information of the MRI subjects, several preprocessing steps are needed.

Fig. 1: The general framework of the proposed methods.

The Freesurfer image analysis suite is used for performing preprocessing over the 3D MRI data. FreeSurfer is a software package for the analysis and visualization of structural and functional neuroimaging data from cross-sectional or longitudinal studies [16]. The FreeSurfer pipeline performs cortical reconstruction and subcortical volumetric segmentation including the removal of non-brain tissue (skull, eyeballs and skin), using an automated algorithm with the ability to successfully segment the whole brain without any user intervention [17]. FreeSurfer is the structural MRI analysis software of choice for the Human Connectome Project which is documented and freely available for download on-line (http://surfer.nmr.mgh.harvard.edu/). In total 31 preprocessing steps has completed using FreeSurfer, of which some are shown in Fig.2.

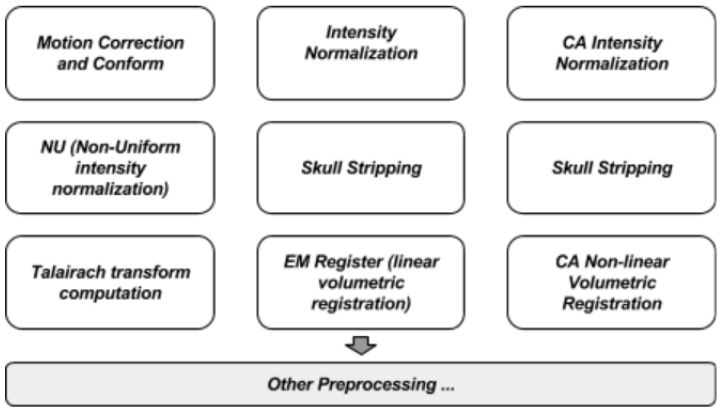


Fig. 2: Preprocessing steps.

There are two types of failures in the preprocessing step which can be categorized into hard failure and soft failure. Hard failures are related to the subjects for whom preprocessing is not successful and the soft failure are related to the subjects that are preprocessed but there are some problem in the results of preprocessing. Out of 568 subjects MRIs, 388 images were successfully preprocessed. Other images were excluded from the dataset due to poor quality of the original images or unknown CDR labels.

3.2 Feature Extraction

After preprocessing using FreeSurfer, a list of volume based features are extracted from different regions of the brain. These features are captured from the regions segmented by brain parcellation using FreeSurfer. Some of the features collected in the left and right hemispheres of the brain are listed below:

1. Left and right lateral ventricle
2. Left and right cerebellum white matter
3. Cerebrospinal fluid (CSF)
4. Left and right hippocampus
5. left and right hemisphere cortex
6. Estimated total intra cranial (eTIV)
7. left and right hemisphere surface holes

The extracted feature data is based on Equation 1.

$$FeatureData = \begin{bmatrix} f_{11} & f_{12}\ f_{13} \dots f_{1n} \\ f_{21} & x_{22}\ x_{23} \dots f_{2n} \\ \dots & \\ f_{s1} & f_{s2}\ f_{s3} \dots f_{sn} \end{bmatrix} \quad (1)$$

Where s is the number of subjects and n is the number of extracted features for that subject. In this study, n is 388 and m is 139.

Furthermore, there are two other type of features which are provided by the PPMI dataset which are age and sex for each subject. Thus, these two biographical information could be added to the extracted feature which gives feature data with the size $(388 * 141)$.

3.3 Classification

In this part the aim is using the extracted volume based features for classifying the MRI data into two classes of PD and HC. In our study, three types of supervised classification algorithm are used. Next, each classification method is described:

- **Logistic Regression (LR):**
  Logistic regression (LR) is a statistical technique which is used in machine learning for binary classification problems. LR belongs to the family of MaxEnt classifiers known as the exponential or log-linear classifiers [18]. Like naive Bayes, it works by extracting some set of weighted features from the input, taking logs, and combining them linearly (meaning that each feature is multiplied by a weight and then added up) [19]. Thus, this model is a suitable binary classifier for our problem.
- **Random Forest (RF):**
  Random forests (RF) is an ensemble learning method for classification, regression and other tasks. This method is presented by Breiman [20], which creates a set of decision trees from a randomly selected subset of training

data. It then aggregates the votes from different decision trees to define final class of the test object. Each tree in a random forest is a weak classifier. A large set of trees trained with randomly chosen data makes a single decision on a majority basis. In the current stage of this research, we tested how accurate decisions can be made by random forests trained by the data coming from a single MRI volume.

– **Support Vector machine (SVM):**
Support vector machine (SVM) [21] is a well-known supervised machine learning algorithm for classification and regression. It performs classification tasks by constructing optimal hyperplanes in a multidimensional space that separates cases of different class labels. This classification method is more popular because it is easier to use, has higher generalization performance and less tuning comparing to other classifiers. In our case, the kernel SVM is used.

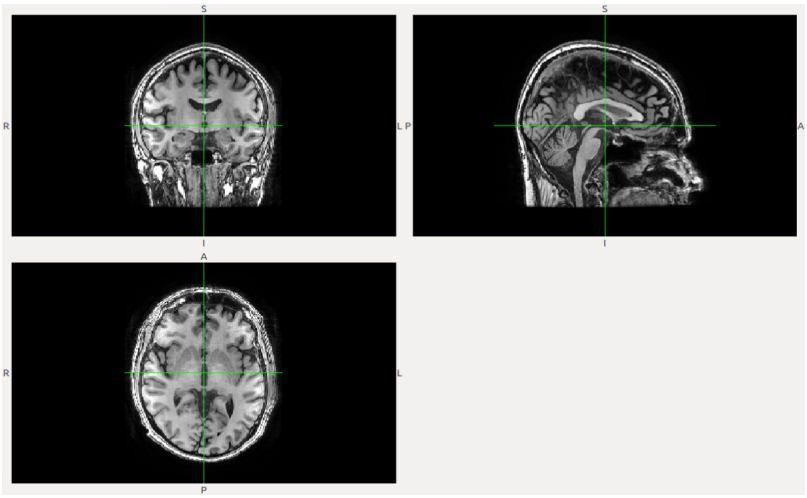
There is a set of parameters for each classifier that needs to be tuned in order to have a fair comparison.

4 Results and Discussion

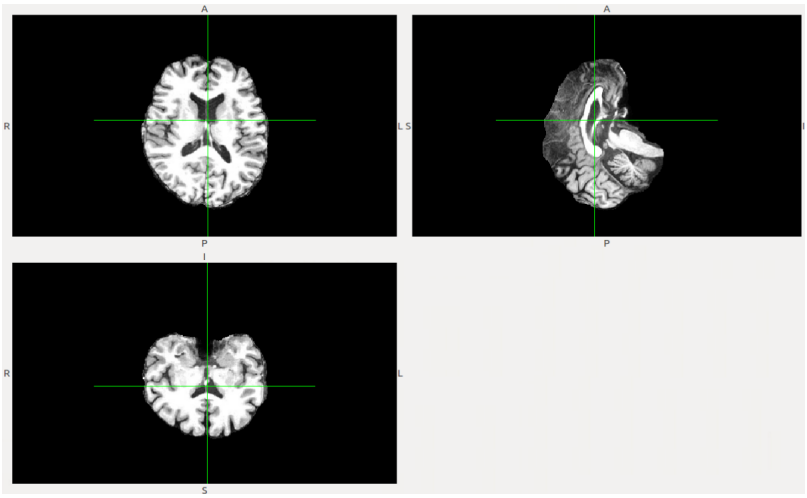
In this section, the experimental results for different steps of the proposed CAD system for diagnosis of PD is presented. First the preprocessing step prepares the MRI data for the next steps using FreeSurfer. Fig.3 shows the MRI for subject 3102 and the resulting image after preprocessing.

After preprocessing with FreeSurfer, for each subject a list of volume-based features is extracted. Also, age and sex are provided for the PPMI data in their website as demographic information of the patients. Some evaluation has been done over the set of extracted features in terms of their discrimination ability. Since PD is an age related disease, the distribution of data in terms of age feature is plotted. Fig.4 shows the distribution of age in the dataset for the subjects with PD and HC labels. The distribution of all the extracted features are plotted in terms of their ability to distinguish the data into two classes of PD and HC. Some of these distributions are shown in Fig.5. As can be seen in Fig.5(a), the subjects with PD have higher brain volume compared to healthy ones. Furthermore, the distribution in Fig.5(b) and (c) illustrate that when people are in the PD category, their CSF and their CC-Anterior volume size is enlarged. Fig.5(d) shows that the surface hole volume in PD is noticeably higher than the normal subjects. Another set of evaluation is done over the extracted features. Data for every two features are plotted versus each other based on the corresponding class. Fig.6 shows the distribution of data based on the two pair of features including $3rd$ ventricles vs lateral ventricles and $3rd$ ventricles vs. left vessels. In both of them, two features tend to have bigger value when the subject is PD.

As explained in the previous section, three types of classifiers are used in this study. These algorithm are run over 388 samples with 141 features. Internal and external cross validation is applied with $K = 10$ for external and $k = 5$ for internal (parameter tunning cross validation). The number of selected samples



(a) Original MR image.



(b) Preprocessed MR image.

Fig. 3: Preprocessing results for one of the subjects.

for the training part is 350 and for the test part is 38. Furthermore, the number of PD and HC in each group is presented in Table 2.

As mentioned before, the classification algorithm needs a set of parameters for tunning which is selected as follow:

– logistic Regression (LR):
Regularization = $[1e-1, 1e-2, 1e-3, 1e-4, 1e-5]$, Tolerance = [1e-1, 1e-2, 1e-3, 1e-4, 1e-5]

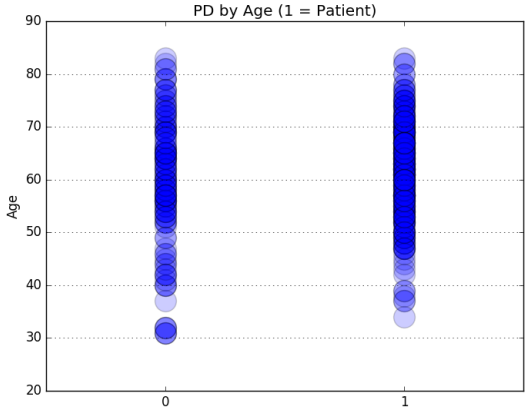


Fig. 4: Distribution of Data in terms of Age feature.

Table 2: Data balance in training and testing parts.

	PD	Hc	Total
Training	236	114	350
Test	26	12	38

- Random Forest (RF):
  Number of estimator =$[5, 10, 15, 20, 25]$, Max depth = $[2 - 10]$
- Support Vector Machine (SVM):
  C = $[0.1, 1, 10, 100, 1000]$, Gamma = $[10, 1, 1e{-}1, 1e{-}2, 1e{-}3, 1e{-}4]$, kernels = $[linear, rbf, poly]$

The evaluation metrics used in this paper for comparing the results of the classification algorithms include accuracy, confusion matrix (recall, precision) and AUC (area under ROC curve). The classification results for LR, RF and SVM are shown in Tables 3, 4, and 5, respectively.

Table 6 shows the general comparison between these methods. The best result is for RF. In the table there are two sets of results related to using age/sex feature or doing the classification based only on the extracted volume based features from FreeSurfer.

Based on the literature review, most papers use VBM for data analysis and feature extraction. In this paper, one of the important goals was evaluating the FreeSurfer features for PD MRIs using machine learning techniques. Generally, the experimental results show that the classification models need more information about the data that should be added to the current features, since these are low-level features and we need a set of high-level features as well. In future research, we are going to determine the useful general features that can be combined with the volume based extracted features.

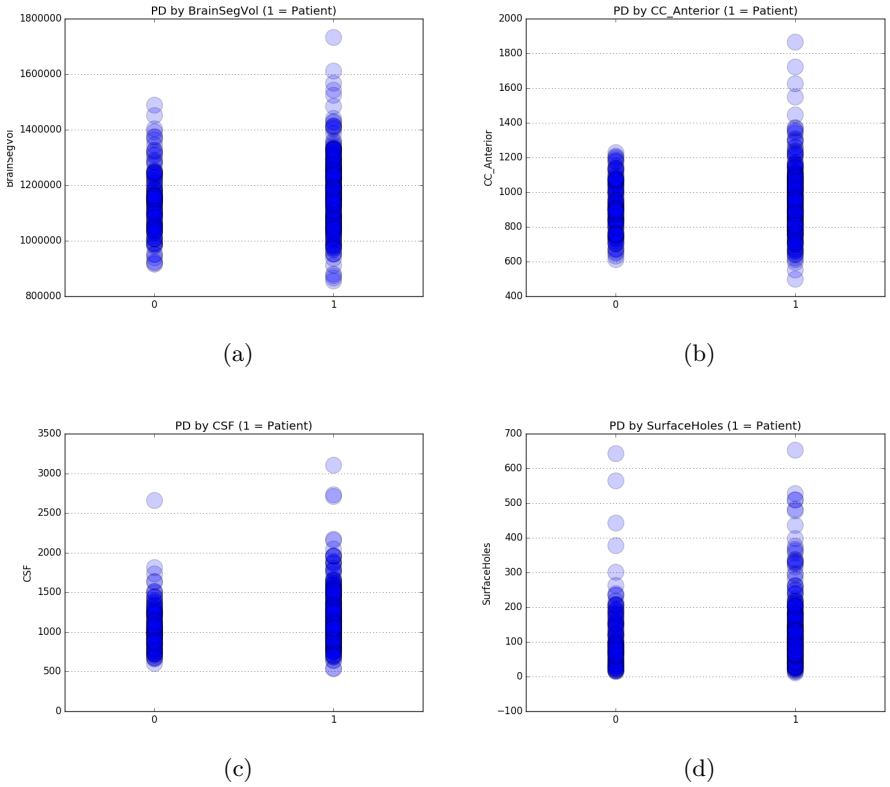
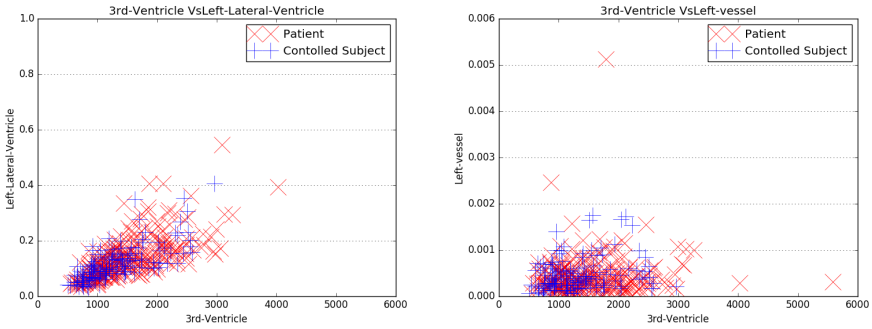


Fig. 5: Data distributions in terms of the class labels and corresponding features; which are: (a) Brain segmented volume. (b) CC- Anterior. (c) CSF. (d) Surface holes.

5 Conclusion

We presented an automatic MRI based CAD system for diagnosing Parkinsons Disease (PD) which is the second common neurodegenerative disease affecting elderly people. This disease is exposed by loss of neuro-transmitters that control body movements and there is no cure other than earlier diagnosis with better and more efficient treatment for patients. MR T1 images from the public PPMI PD data set is used. FreeSurfer is used for feature extraction and preprocessing. The decision model for classification of the extracted feature data is based on LR, RF and SVM methods. In the experimental results, the ability of these three types of classifiers for PD diagnosis are compared to each other. Results show that using only MRI is a potential option for PD diagnosis. This approach will avoid exposing the brain to harmful radiation based scans. In future work, the efficiency of the proposed method could be improved by adding high level features to the



(a) (b)

Fig. 6: Data distribution based on the pair of features: (a)3rd ventricle and left lateral ventricle. (b) 3rd ventricle and left vessel.

Table 3: Logistic regression performance

Logit	Train Accuracy	Test Accuracy	TN	FP	FN	TP	AUC
K_0	0.6867	0.6500	3	10	4	23	0.5413
K_1	0.6781	0.6750	0	13	0	27	0.5000
K_2	0.6819	0.7179	2	11	0	26	0.5769
K_3	0.6733	0.6666	3	10	3	23	0.5576
K_4	0.6991	0.5384	2	11	7	19	0.4423
K_5	0.6618	0.6153	0	13	2	24	0.4615
K_6	0.6657	0.6842	1	11	1	25	0.5224
K_7	0.7057	0.6578	2	10	3	23	0.5256
K_8	0.6857	0.6052	1	11	4	22	0.4647
K_9	0.6742	0.6315	0	12	2	24	0.4615

current ones. The classification rate with MRI needs to be improved to get close to those using raditation based scanning.

References

1. https://parkinsonsnewstoday.com/parkinsons-disease-statistics/.
2. S.H. Fox, R. Katzenschlager, S.Y. Lim, B. Ravina, K. Seppi, M. Coelho, W. Poewe, O. Rascol, C.G. Goetz, C. Sampaio, The movement disorder society evidence-based medicine review update: treatments for the motor symptoms of Parkinsons disease, Mov. Disord. 26 (S3) (2011) S2S41.
3. K.R. Chaudhuri, A.H. Schapira, Non-motor symptoms of Parkinsons disease: dopaminergic pathophysiology and treatment, The Lancet Neurology 8 (5) (2009) 464474.
4. Rana, B., Juneja, A., Saxena, M., Gudwani, S., Kumaran, S., Behari, M. and Agrawal, R. (2017). Relevant 3D local binary pattern based features from fused

Table 4: Random forests performance

RandomFrest	Train Accuracy	Test Accuracy	TN	FP	FN	TP	AUC
K_0	0.7270	0.6750	0	13	0	27	0.5000
K_1	0.7327	0.6500	1	12	2	25	0.5014
K_2	0.7507	0.6923	1	12	0	26	0.5384
K_3	0.7392	0.6923	1	12	0	26	0.5384
K_4	0.7335	0.6666	2	11	2	24	0.5384
K_5	0.7277	0.7179	2	11	0	26	0.5769
K_6	0.7399	0.6578	0	12	1	25	0.4807
K_7	0.7171	0.6578	0	12	1	25	0.4807
K_8	0.7257	0.6315	1	11	3	23	0.4839
K_9	0.7199	0.7105	1	11	0	26	0.5416

Table 5: Support Vector Machine performance

SVM	Train Accuracy	Test Accuracy	TN	FP	FN	TP	AUC
K_0	0.7528	0.6000	3	10	6	21	0.5042
K_1	0.7471	0.6000	7	6	10	17	0.5840
K_2	0.7134	0.6923	5	8	4	22	0.6153
K_3	0.7335	0.5128	2	11	8	18	0.4230
K_4	0.7478	0.6666	3	10	3	23	0.5576
K_5	0.7449	0.5897	4	9	7	19	0.5192
K_6	0.7228	0.7105	4	8	3	23	0.6089
K_7	0.7400	0.4473	3	9	12	14	0.3942
K_8	0.7171	0.5789	4	8	8	18	0.5128
K_9	0.7428	0.5789	3	9	7	19	0.4903

feature descriptor for differential diagnosis of Parkinsons disease using structural MRI. Biomedical Signal Processing and Control, 34, pp.134-143.

5. Adeli, E. et al. Joint feature-sample selection and robust diagnosis of parkinsons disease from MRI data. NeuroImage 141, 206219 (2016).
6. Prashanth, R., Roy, S. D., Mandal, P. K., and Ghosh, S., Automatic classification and prediction models for early Parkinsons disease diagnosis from SPECT imaging. Expert Syst. Appl. 41:33333342, 2014.
7. Marios Politis and Clare Loane, Serotonergic Dysfunction in Parkinson's Disease and Its Relevance to Disability, TheScientificWorldJOURNAL, vol. 11, Article ID 172893, 9 pages, 2011.
8. N.K. Focke, G. Helms, S. Scheewe, P.M. Pantel, C.G. Bachmann, P. Dechent, J. Ebentheuer, A. Mohr, W. Paulus, C. Trenkwalder, Individual voxel-base subtype prediction can differentiate progressive supranuclear palsy from idiopathic parkinson syndrome and healthy controls, Hum. Brain Mapp. 32 (11) (2011) 19051915.
9. G.S. Babu, S. Suresh, B.S. Mahanand, A novel PBL-McRBFN-RFE approach for identification of critical brain regions responsible for Parkinsons disease, Expert Syst. Appl. 41 (2) (2014) 478488.
10. C. Salvatore, A. Cerasa, I. Castiglioni, F. Gallivanone, A. Augimeri, M. Lopez, G. Arabia, M. Morellie, M.C. Gilardic, A. Quattrone, Machine learning on brain MRI data for differential diagnosis of Parkinsons disease and Progressive Supranuclear Palsy, J. Neurosci. Methods 222 (2014) 230237.

Table 6: Comparing performance of different classifiers

Methods/Criteria	Age/Sex Feature	Train Accuracy	Test Accuracy	AUC
LR	No	0.6806	0.6467	0.4912
LR	Yes	0.6858	0.6313	0.4794
RF	No	0.7425	0.67	0.4647
RF	Yes	0.7396	0.6673	0.5086
SVM with rbf kernel	No	0.6752	0.6753	0.500
SVM with rbf kernel	Yes	0.6752	0.6753	0.500
SVM with linear kernel	No	0.7362	0.5977	0.521
SVM with linear kernel	Yes	0.7259	0.5927	0.5273

11. B. Rana, A. Juneja, M. Saxena, S. Gudwani, S.K. Senthil, R. Agrawal, M. Behari, Regions-of-interest based automated diagnosis of Parkinsons disease using T1-weighted MRI, Expert Syst. Appl. 42 (9) (2015) 45064516.
12. B. Rana, A. Juneja, M. Saxena, S. Gudwani, S.K. Senthil, M. Behari, R.K. Agrawal, Graph-theory-based spectral feature selection for computer aided diagnosis of Parkinsons disease using T1-weighted MRI, Int. J. Imaging Syst. Technol. 25 (3) (2015) 245255.
13. Bo Peng, Suhong Wang, Zhiyong Zhou, Yan Liu, Baotong Tong, Tao Zhang, Yakang Dai, A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinsons disease, Neuroscience Letters, Volume 651,2017, Pages 88-94, ISSN 0304-3940.
14. Adeli, E., Shi, F., An, L., Wee, C. Y., Wu, G., Wang, T., & Shen, D. (2016). Joint feature-sample selection and robust diagnosis of Parkinson's disease from MRI data. NeuroImage, 141, 206-219.
15. https://ida.loni.usc.edu/home.
16. https://surfer.nmr.mgh.harvard.edu/fswiki.
17. A. Worker and et al. Cortical thickness, surface area and volume measures in parkinsons disease, multiple system atrophy and progressive supranuclear palsy. PLOS ONE, 9(12), 2014.
18. Peng, Chao-Ying Joanne, et al. An Introduction to Logistic Regression Analysis and Reporting. The Journal of Educational Research, vol. 96, no. 1, 2002, pp. 314.
19. J. Martin and D. Jurafsky, Speech and language processing.Prentice, Hall, 2000.
20. L. Breiman. Random forests. Machine learning, 45(1):532, 2001.
21. V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.