

**COMPUTER-AIDED EARLY DETECTION OF
PARKINSON'S DISEASE THROUGH MULTIMODAL
DATA ANALYSIS**

R. PRASHANTH



**DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI**

AUGUST 2015

© Indian Institute of Technology Delhi (IITD), New Delhi, 2015

**COMPUTER-AIDED EARLY DETECTION OF
PARKINSON'S DISEASE THROUGH MULTIMODAL
DATA ANALYSIS**

by

R. PRASHANTH

DEPARTMENT OF ELECTRICAL ENGINEERING

Submitted

in fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

AUGUST 2015

CERTIFICATE

This is to certify that the thesis entitled, “**Computer-Aided Early Detection of Parkinson’s Disease through Multimodal Data Analysis**”, being submitted by **Mr. R Prashanth** for the award of the degree of **Doctor of Philosophy** is a record of bonafide research work carried out by him in the Department of Electrical Engineering of the Indian Institute of Technology Delhi.

Mr. R Prashanth has worked under our guidance and supervision and has fulfilled the requirements for the submission of this thesis, which to our knowledge has reached the requisite standard. The results obtained here, have not been submitted to any other University or Institute for the award of any degree.

Dr. Sumantra Dutta Roy	Dr. Shantanu Ghosh	Prof. Pravat K. Mandal
Associate Professor, Dept. of Electrical Engineering, Indian Institute of Technology Delhi	Research Fellow, Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Massachusetts, USA	Professor, National Brain Research Center, India & Associate Professor (Adj.), Dept. of Radiology Johns Hopkins Medicine, USA

Dated:

To my family and friends

Acknowledgements

I express my sincere thanks to my PhD supervisors Dr. Sumantra Dutta Roy, Dr. Shantanu Ghosh and Dr. Pravat Mandal for their able guidance, and also to my other research committee members, Dr. Santanu Chaudhury, Dr. Prem Kalra and Dr. Brejesh Lall, for their valuable comments and feedback which helped to improve the work.

My heartfelt thanks to Dr. Senthil Kumaran of the All India Institute of Medical Sciences (AIIMS), New Delhi, for his most valuable discussion which culminated in the selection of my PhD topic.

My gratitude to Dr. Jyotirmoy Chatterjee and Dr. Pranab Kumar Dutta of IIT Kharagpur, and Dr. Sambuddha Ghosh of Regional of Ophthalmology, Kolkata who were the real inspiration for me to continue in research after my masters.

I thank the reviewers of the manuscripts, that were communicated to peer-reviewed journals, for providing constructive comments that improved the quality of the thesis.

Many thanks to my friends Jitesh Joshi, Pratik Shah, Yogesh T. Piolet, Teesta Naskar, Nikhil Anto Ponnor and Apurva Agrawal who helped me in getting journal papers that were inaccessible at IIT.

I thank my other friends Jeevanand Seshadrinath, Sabyasachi Gupta and other people who were a constant support throughout my PhD study. I thank my fellow lab mates and lab staff in the Multimedia Lab for their help during the study. I thank Dr. Rajesh of the IIT Hospital, Delhi, Dr. Subash Babu and Dr. Bindu of the Ayurveda College

Hospital, Trivandrum, Dr. Pravin Haribakti and Dr. Monal Shah, Ahmedabad, for providing medical help when I got sick. I also thank Drs. Haraja, Aswathy, Maya, Neetu, Parvathy and Vishnu Prasad for their encouragements. My heartfelt thanks to Baiju sir for his help in a spiritual way and my childhood friends Dr.Vishnu and Vineesh for their amazing friendship. I also thank Richard uncle for his important help.

Most importantly, I thank the Parkinson's Progression Marker's Initiative (PPMI) study for publicly sharing their database without which this thesis would not have been possible. I also thank the industry partners and mainly, the Michael J. Fox Foundation (MJFF) for Parkinson's Research for funding the PPMI study. I thank the Data & Publication Committee (DPC) of the PPMI for providing timely administrative reviews of the manuscripts which was a requirement prior to submission to a journal/conference. I thank Vanessa Arnedo, Reena Vanjani, Sohini Chowdhury, Susan Mendick, Shirley Lasch and other members of the PPMI team for their assistance with regard to manuscript submission to DPC, administrative reviews and clearing my queries regarding the PPMI study.

I am grateful to my family members who have helped me by taking care of my mother during my stay away from home.

Last but not the least, I thank the Almighty who gave me the much necessary endurance to tackle hardships that came upon during the study and life.

Date:

Place: New Delhi

R. Prashanth

Abstract

Early detection of Parkinson's disease (PD) is an important and challenging problem for effective management and treatment. There are no definitive diagnostic tests for PD and the clinical diagnosis relies on presence of classic symptoms, and by the time a patient shows these symptoms, more than 60% of dopaminergic neurons are lost. In this thesis, approaches based on multimodal data analysis and machine learning for early detection of PD are presented.

Recent neuroimaging techniques, such as SPECT, have shown to be sensitive in discriminating PD from healthy normal even at the early stages of the disease. Striatal Binding Ratio (SBR) values are well-established and clinically used quantification measures obtained from SPECT. This work carries out a novel use of these features in developing classification and prediction models for discriminating early PD from healthy controls. A high accuracy is achieved in classification.

Along with discriminating PD from normal, SPECT imaging have also indicated their usefulness in detecting a group of subjects, called Scans Without Evidence of Dopaminergic Deficit (SWEDD) who show normal scans who otherwise were clinically misdiagnosed as PD. This work carries out analysis of SPECT images, for segmenting and quantifying the high activity regions that correspond to striatal uptake, for extracting discriminatory features that eventually are used for classifying degenerative parkinsonism such as PD from healthy normal and non-degenerative conditions such as SWEDD. These computed features show good variation in PD as

compared to normal and SWEDD, and they are statistically significant as well. The classification accuracy is also observed to be very high.

Research studies show that there exist a phase called the premotor or the prodromal phase before the motor phase in PD and that this phase is mostly dominated by non-motor symptoms such as smell loss and REM sleep behavior disorder. This work carries out a novel use of these features along with other relevant features such as imaging and cerebrospinal fluid measurements to classify early PD from healthy normal. A very high accuracy is obtained in classification using an SVM classifier.

As of now, there exists no standard screening procedures for detecting PD, and patient questionnaire are easy tools that might have the potential for a screening instrument. This work carries out the novel use of using the patient questionnaire portion of the widely-used Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) to classify early PD from healthy normal. We observe a very high accuracy in classification.

Therapeutic options in PD depend on the stage and severity of the disease. This work carries out estimation of stage and severity of PD using MDS-UPDRS which evaluates the most pertinent features in PD. The AdaBoost-based classifier ensemble gave a high accuracy in classification. We infer from the study that such models have the potential to aid clinicians in the diagnostic process.

Table of Contents

Certificate	i
Acknowledgement	v
Abstract	vii
Table of Contents	ix
List of Figures	xix
List of Tables	xxiii
List of Key Abbreviations	xxv
1 Introduction	1
1.1 General	1
1.1.1 History	3
1.1.2 Causes	3
1.2 Motivation, Objective and Scope of the work	4
1.2.1 Early PD Detection Based on Striatal Binding Ratio values from SPECT Imaging	4
1.2.2 Detection through Shape- and Surface Model-based Features in SPECT Imaging	5
1.2.3 Early PD Detection Based on Preclinical Non-motor, Imaging and Biological Features	6
1.2.4 Detection Based on Patient Questionnaire parts of MDS-UPDRS (Parts IB and II)	6
1.2.5 Stage and Severity Estimation in PD Based on Complete MDS-UPDRS	7
1.3 Outline of Chapters	9

2	Review of the Literature	13
2.1	Nature of the Problem: Early Diagnosis	13
2.2	Clinical Features of PD	13
2.2.1	Motor Features	14
2.2.2	Non-motor Features	16
2.2.2.1	Non-motor Symptoms are often Unrecognized	16
2.3	Diagnostic Criteria for PD	17
2.4	Premotor Stage of PD	18
2.5	Staging of Brain Pathology in PD	19
2.6	Differential Diagnosis	19
2.7	Misdiagnosis in PD	20
2.8	Approaches that Aid in the Detection and Discrimination of PD	21
2.8.1	Neuroimaging in PD	21
2.8.1.1	Magnetic Resonance Imaging (MRI)	21
2.8.1.2	Diffusion-weighted imaging (DWI) and diffusion-tensor MRI (DTI)	22
2.8.1.3	Functional MRI (fMRI)	23
2.8.1.4	Magnetic Resonance Spectroscopy (MRS)	23
2.8.1.5	Positron Emission Tomography (PET)	24
2.8.1.6	Single Photon Emission Computed Tomography (SPECT)	25
2.8.2	Clinical and Behavioral Assessments	28
2.8.2.1	Objective Smell Testing	28
2.8.2.2	REM Sleep Behavior Disorder (RBD) Assessment	29
2.8.3	Genetic Testing	30
2.8.4	Cerebrospinal fluid (CSF) markers	32
2.9	Therapeutic Options in PD	32
2.9.1	Motor Symptoms	32

2.9.2	Non-motor symptoms	33
2.10	Neuroprotective Strategies	34
2.11	Parkinson's Progression Markers Initiative (PPMI)- A Database for Research in PD Biomarkers and Early PD Detection	34
2.11.1	Relevance to other subpopulations such as in Indian context	35
2.12	PD Detection and Characterization through Data Analysis and Machine Learning: A Review	36
2.13	Identified Research Areas	39
2.14	Scope and contribution from the thesis	40
3	Detection and Severity Estimation of Early Parkinson's Disease using Striatal Binding Ratio from SPECT imaging	43
3.1	Introduction	43
3.2	Materials and Methods	46
3.2.1	Database and Study Cohort details	46
3.2.2	Statistical Testing of SBR features	47
3.2.3	Classification and Prediction/Prognostic Modeling based on SBR values to distinguish Early PD from Healthy Controls	49
3.2.3.1	Automatic SVM-based Classification of Early PD from Healthy Controls	50
3.2.3.2	Prediction/Prognostic Modeling for Early PD using Multivariate Logistic Regression	51
3.3	Results and Discussion	51
3.3.1	Statistical Significance of SBR features	51
3.3.2	Automatic Classification and Prediction Modeling for Early PD	52
3.3.2.1	SVM-based Classification to Distinguish Early PD from Healthy Controls	53

3.3.2.2	Prediction Modeling for Early PD using Logistic Regression	54
3.3.2.2.1	Evaluation of the Logistic Model	57
3.3.2.2.2	Comparison with other Similar Models	60
3.3.2.3	Performance comparison through ROC Curves, AUC and Cluster Plots	61
3.3.3	Highlights of the Study	64
3.3.4	Future Work	64
3.3.4.1	Application in Distinguishing SWEDD from Early PD	65
3.3.4.2	Application in Distinguishing Essential Tremor (ET) from Early PD	66
3.3.4.3	Automation of SBR Computation	66
3.4	Conclusion	67
4	Early detection through Shape Analysis and Surface Fitting in SPECT Imaging	69
4.1	Introduction	69
4.2	Materials and Methods	73
4.2.1	Database and Study Cohort details	73
4.2.2	Processing Pipeline	74
4.2.2.1	Preprocessing from PPMI	74
4.2.2.2	Selection of Slices for Processing	74
4.2.2.3	Intensity Normalization	75
4.2.2.4	Shape Analysis through Quantification using Shape-based features	76
4.2.2.5	Surface Fitting based on Intensity Distribution in the Segmented Regions	77
4.2.3	Feature Set	79

4.2.4	Statistical Analysis of Features	80
4.2.5	Classification of Degenerative PS (early PD) from Non-Degenerative Types (Normal/SWEDD)	80
4.2.6	Feature Importance Estimation	81
4.3	Results and Discussion	82
4.3.1	Feature Values, Statistical Analysis and Justification for considering Normal and SWEDD as a Single Group	84
4.3.2	Classification Modeling	89
4.3.2.1	Note on Curse of Dimensionality	90
4.3.3	Estimation of Importance of Features	92
4.3.4	Note on Slice Selection, Threshold Selection and Effect of Segmentation	96
4.3.5	Further Analysis using MRI-A Future Work	101
4.4	Conclusion	103
5	Automatic and High-Accuracy detection of PD based on preclinical (non-motor), biological and imaging features	105
5.1	Introduction	105
5.1.1	Preclinical Non-motor Features	105
5.1.2	Cerebrospinal Fluid (CSF) Markers	106
5.1.3	Neuroimaging Markers	107
5.1.4	Background, Motivation and Outline of the Study	108
5.2	Materials and Methods	109
5.2.1	Database and Study Cohort details	109
5.2.2	Feature Description	110
5.2.2.1	University of Pennsylvania Smell Identification Test (UPSIT)	110
5.2.2.2	REM sleep Behavior Disorder Screening Questionnaire (RBDSQ)	111
5.2.2.3	Biomarkers from Cerebrospinal fluid (CSF)	111

5.2.2.4	Striatal Binding Ratio (SBR) from SPECT Imaging	112
5.2.3	Statistical Analysis of Features	112
5.2.4	Classification of Early PD and Healthy Normal using Automated Models	113
5.3	Results and Discussion	116
5.4	Conclusion	122
6	Easy to use, Low cost and Highly Accurate Detection of PD based on Patient Questionnaire of MDS-UPDRS	125
6.1	Introduction	125
6.2	Materials and Methods	127
6.2.1	Database and Study Cohort details	127
6.2.2	Data Partitioning	130
6.2.3	Predictive Modelling for Distinguishing Early PD from Healthy Normal	130
6.3	Results	132
6.3.1	Logistic Regression Model	132
6.3.1.1	Evaluation of the Logistic Model	135
6.3.2	Classification Trees	135
6.3.3	Boosted Trees, SVM and a Comparison of Classification Performances	136
6.4	Discussion	139
6.4.1	Choosing a Logistic Model and a Classification Tree Model	142
6.4.2	Possible Application as a Screening Instrument for PD Detection	142
6.4.3	Limitations and Future Work	143
6.5	Conclusion	144

7	Parkinson's Disease Stage and Severity Estimation based on MDS-UPDRS (Parts I, II and III)	145
7.1	Introduction	145
7.1.1	Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS)	146
7.1.2	Hoehn and Yahr (HY) Scale	147
7.1.3	Motivation, Background and Outline of the Study	148
7.2	Materials and Methods	149
7.2.1	Database and Study Cohort details	149
7.2.2	Statistical Analysis of Features	150
7.2.3	A Note on the Imbalance in the Dataset used	151
7.2.4	Predictive Modelling for PD Severity and Stage Estimation	152
7.2.5	Performances Measures for Multi-Class Classification Evaluation	153
7.2.6	Two-class (PD/Healthy Normal) Classification and Feature Importance Estimation	154
7.3	Results and Discussion	155
7.3.1	Statistical Testing of Features	155
7.3.2	The Logistic Model	157
7.3.3	Cost and Sampling Proportion Estimation using Genetic algorithm (GA)	161
7.3.4	Cost-sensitive and Sampling based Predictive Modeling, and Performance Comparison of Models	163
7.3.4.1	A Note on Reproducibility of Classifiers	165
7.3.5	Error Analysis	166
7.3.6	Binary Classification and Feature Importance Estimation	167
7.3.7	A Note on Cost-sensitive and Sampling based Imbalanced Data Learning used	170
7.3.8	Limitations and Future Work	171

7.4	Conclusion	172
8	Conclusions	173
8.1	Summary of contributions	175
8.2	Future Work	176
8.2.1	Future Work in terms of Analysis, in the Current Framework	178
8.2.1.1	The Role of Classifiers, Relative Utility, and Possible Combinations	178
8.2.1.2	The Role of Features, Feature Combination, Feature Selection, Orthogonality Issues	179
8.2.1.3	A Closer Look at the Classifier Data itself: Datasets from Different Sites	179
8.2.1.4	A Staged/Cascaded Approach for Diagnosis in a Screening Population, and Subsequent Validation through large scale field tests	180
8.2.1.5	Developing a mobile app for pre-screening, based on low-cost self-administered instruments	180
8.2.1.6	The Effect of Segmentation on Image-Based Features	181
8.2.2	Future work in terms of examining characteristics related to PD itself	181
8.2.2.1	Differentiating PD from other forms of Parkinsonism	181
8.2.2.2	Differentiating PD from SWEDD	182
8.2.2.3	Further analysis of the prodromal phase	182
8.2.2.4	The use of other modalities in PD analysis	183
8.2.2.5	The role of genetics in PD	183

References	185
Appendices	217
A.1 Subject Selection Criteria by PPMI	217
A.2 DaTSCAN Imaging	217
A.3 Striatal Binding Ratio (SBR) Computation	218
A.4 Description of Different Statistical Tests and Techniques used in the Study	219
A.5 Description of Shape-based Features	222
A.6 A Review of Machine Learning based Classifiers	224
A.6.1 Logistic Regression	224
A.6.2 Support Vector Machine (SVM)	225
A.6.3 Naïve Bayes	225
A.6.4 Boosted Trees	226
A.6.5 Random Forests	227
A.7 Classifiers used for Multi-class Classification	228
A.7.1 Ordinal Logistic Regression	228
A.7.2 Cost-sensitive Support Vector Machine (SVM)	228
A.7.3 Cost-sensitive AdaBoost-based Classifier Ensemble	230
A.7.4 RUSBoost-based Classifier Ensemble	231
A.8 Genetic Algorithm for Optimal Cost and Sampling Proportion Estimation for Multi-class Imbalance Data Learning	232
List of Publications	235
Bio-Data	237

List of Figures

1.1	Illustration of brain regions such as the striatum and substantia nigra, and dopaminergic pathways.	2
2.1	^{123}I -FP-CIT SPECT scan images for a (a) Healthy Control (HC), (b) Scan Without Evidence of Dopaminergic Deficit (SWEDD), and (c) Early Parkinson's Disease (PD)	26
3.1	Flowchart of the analysis carried out	46
3.2	Histograms and notched box plots of the SBR values of (a, b) right caudate (c, d) left caudate (e, f) right putamen (g, h) left putamen for normal and early PD populations	48
3.3	Box plot of SBR values of (a) right caudate (b) left caudate (c) right putamen (d) left putamen for the misclassified samples from the SVM classifier using RBF kernel.	56
3.4	Histogram of the predicted risk probabilities.	61
3.5	Plot of predicted risk probabilities of PD vs. the lowest putamenal SBR	61
3.6	ROC curves of the different classifiers used	62
3.7	Cluster plot of data points (both correctly classified and misclassified) from (a) SVM_{RBF} , (b) $\text{SVM}_{\text{Linear}}$, (c) Logistic regression	63
4.1	Flowchart of the analysis carried out	77
4.2	Slice averaging, image segmentation and surface fitting for healthy normal (a, b, c, d), SWEDD (e, f, g, h), and early PD (i, j, k, l).	85
4.3	Box plots of the shape-based features (a-p), surface fitting-based features (q-ad) and SBR-based features (ae-ah).	87-88
4.4	Plot of accuracies obtained after adding one feature at a time to the classifier.	94
4.5	Plot of cross-validation error of the boosted trees ensemble with number of trees in the ensemble.	95

4.6	Plot of cross-validation classification error and out-of-bag error of the Random forests ensemble with number of trees in the ensemble.	95
4.7	Plot of estimates of feature importance for all shape-based (features from 1 to 16), surface fitting-based (features from 17 to 30) and SBR-based features (features from 31 to 34).	96
4.8	Plot of cross-validation classification error and out-of-bag error of the Random forests ensemble for selecting the number of trees in the ensemble for estimating feature importance	97
4.9	SPECT images of different slices from a healthy normal.	98-99
4.10	Plot of the areas of the segmented striatal regions from slices numbered from 33 to 50.	100
4.11	Histogram of the thresholds used for segmenting the striatal region in (a) healthy normal, (b) SWEDD and (c) Early PD subjects	101-102
4.12	T2-weighted MRI image of a normal control and a early PD patient.	104
5.1	Flowchart of the proposed analysis	115
5.2	Box plots of features for normal and early PD populations (a) RBDSQ score, (b) UPSIT score, (c) A β 1-42, (d) α -syn, (e) P-tau ₁₈₁ , (f) T-tau, (g) T- tau/A β 1-42, (h) P-tau ₁₈₁ /A β 1-42, (i) P-tau ₁₈₁ /T-tau, (j) SBR _{RC} , (k) SBR _{LC} , (l) SBR _{RP} , and (m) SBR _{LP}	115-116
5.3	ROC plots for (a) Naïve Bayes, (b) SVM (c) Boosted Trees, and (d) Random forests classifiers.	120
5.4	Plot of 10-fold cross-validation error of boosted trees classifier with number of trees (b) plot of 10-fold cross-validation error and out of bag error of Random forests classifier with number of trees	121
5.5	Plot of features importance scores computed using Random forests technique.	123
6.1	Stacked bar charts shows the severity of PQ features in MDS-UPDRS for (a) normal population (b) early PD population	129
6.2	Classification tree model	136
6.3	Plot of feature importance scores obtained using classification trees	137

6.4	ROC plots for classification using (a) logistic regression (b) classification trees (c) boosted trees, and (d) SVM	138
6.5	(a) Plot of deviance of fit of models obtained for 100 different partitions (iterations) of input data into training and testing sets, (b) Plot of 10-fold cross-validation classification error and the number of learners (trees) in the boosted tree model	139
7.1	Stacked bar plots showing the severity of symptoms in (a) normal stage, (b) early stage and (c) moderate stage PD	150
7.2	Flowchart for predictive modeling and feature importance estimation	151
7.3	Plot of the expected value of the class computed using the probabilities obtained from the ordinal logistic model	159
7.4	Plot of Spearman's rank correlation coefficients between a feature and the true stage (R), and between a feature and the expected value of the class (R').	159
7.5	Plots of the fitness values (mean and best fitness, on the left side), and the best individual (on the right side) obtained from GA for the 3 classifiers used: (a) cost-sensitive SVM, (b) cost-sensitive AdaBoost-based ensemble, and (c) RUSBoost-based ensemble	163
7.6	Box plots of the misclassified instances from the classifiers used in the study (a) ordinal logistic regression, (b) cost-sensitive SVM, (c) cost-sensitive AdaBoost-based ensemble, and (d) RUSBoost-based ensemble	167
7.7	Estimates of importance scores for all 59 features.	169

List of Tables

3.1	SBR values (mean $\pm$ standard deviation (SD)) of right caudate, left caudate, right putamen and left putamen for healthy normal and early PD populations	47
3.2	Result of statistical testing of SBR features using univariate logistic regression	52
3.3	Accuracies (in %) obtained for SVM classifiers using RBF and linear kernels for each fold during 10-fold cross-validation	54
3.4	Confusion matrix and performance measures for the SVM classifier with RBF kernel (SVM_{RBF}) and SVM classifier with linear kernel (SVM_{Linear})	55
3.5	Four-predictor logistic model developed using the SBR values of left putamen, right putamen, left caudate and right caudate as predictors	56
3.6	Three-predictor logistic model obtained using left putamen SBR, right putamen SBR and the interaction term (left caudate SBR*right caudate SBR) as predictors.	57
3.7	Omnibus Test of Model Coefficients	57
3.8	Goodness-of-fit tests	59
3.9	Classification Table	59
4.1	List of shape-based features	78
4.2	List of Striatal Binding Ratio (SBR)-based features	81
4.3	Shape-based, surface fitting-based and SBR-based feature values (mean $\pm$ SD), and statistical testing for comparing healthy normal vs. SWEDD and early PD vs. Normal/SWEDD	89-90
4.4	Performance measures obtained for the classifiers used	91
4.5	Training and testing accuracy obtained for various classifiers used	93
5.1	Feature values (in terms of mean $\pm$ SD) for normal and early PD groups	118
5.2	Performance measures of various classifiers used in the study	119

6.1	Components in Part-IB and Part-II of MDS-UPDRS	128
6.2	A typical partition of input dataset into training and testing set. 70% is used for training and rest for testing	130
6.3	Multivariate logistic model obtained using all the 20 predictors	133
6.4	Logistic model developed using stepwise variable selection	134
6.5	Predictors that were removed during stepwise selection process	134
6.6	Performance measures obtained for the classifiers used	137
6.7	Comparison with related works	140
7.1	Statistical testing for comparing 3 groups (normal, early and moderate stage PD)	156
7.2	The logistic model	158
7.3	Spearman's rank correlation of features with true stage (R), and with the expected value of the class (R')	160
7.4	Confusion matrices obtained for the classifiers used	164
7.5	Performance measures of the classifiers used	165
7.6	Training and testing accuracies obtained for different classifiers used in the study, from 10-fold cross validation	166
7.7	Statistical testing of features for 2 class(Normal/PD) comparison	168
7.8	Confusion matrix for 2-class classification (PD/Normal) using Random forests	169
7.9	Performance measures obtained for the classifiers using optimal and default costs/sampling proportion settings	171
8.1	Comparison with related works	178

List of Important Abbreviations

^{123}I -FP-CIT	^{123}I -2-carbomethoxy-3-(4-iodophenyl)-N-(3-fluoropropyl)nortropane
^{18}F FDG-PET	^{18}F -fluoro-2-deoxyglucose PET
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AUC	Area Under the ROC curve
A β 1-42	amyloid beta peptide 1-42
BOLD	Blood Oxygenation Level-Dependent
CSF	Cerebrospinal Fluid
DAT	Dopamine Transporter
DLB	Dementia with Lewy bodies
DRT	Dopamine Replacement Therapy
DTI	Diffusion-Tensor Imaging
DWI	Diffusion-Weighted imaging
ET	Essential Tremor
FA	Fractional Anisotropy
fMRI	Functional MRI
FN	False Negative
FP	False Positive
GA	Genetic Algorithm
HC	Healthy Control
HY scale	Hoehn and Yahr scale
ICA	Independent Component Analysis
LRRK2	leucine-rich repeat kinase 2
MAO-B	Monoamine Oxidase B
MCI	Mild Cognitive Impairment
MD	Mean Diffusivity
MDS-UPDRS	Movement Disorder Society- Unified Parkinson's Disease Rating Scale
MLR	Multivariate Logistic Regression

MRI	Magnetic Resonance Imaging
MRS	Magnetic Resonance Spectroscopy
MSA	Multiple System Atrophy
NMS	Non-Motor Symptoms
PCA	Principal Component Analysis
PD	Parkinson's Disease
PET	Positron Emission Tomography
PPMI	Parkinson's Progression Markers Initiative
PQ	Patient Questionnaire
PS	Parkinsonian Syndrome
P-tau181	tau phosphorylated at threonine 181
RBD	REM sleep Behavior Disorder
RBDSQ	RBD Screening Questionnaire
RBF	Radial Basis Function
REM	Rapid Eye Movement
RF	Random Forests
ROC	Region Operating Characteristic
ROI	Region of Interest
RUS	Random Undersampling
SBR	Striatal Binding Ratio
SBR _{LC}	SBR value of the left caudate
SBR _{LP}	SBR value of the left putamen
SBR _{RC}	SBR value of the right caudate
SBR _{RP}	SBR value of the right putamen
SPECT	Single Photon Emission Computed Tomography
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SWEDD	Scans Without Evidence of Dopaminergic Deficit
TN	True Negative
TP	True Positive
T-tau	Total tau
UPDRS	Unified Parkinson's Disease Rating Scale
UPSIT	University of Pennsylvania Smell Identification Test

VMAT2

α -syn

Vesicle Monoamine Transporter

α -synuclein