

Adopting Machine Learning Technology for the Classification of Parkinson's Disease

A THESIS SUBMITTED IN
FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

By

Farhan Mohammed

In

Faculty of Engineering and Information Technology

School of Electrical and Data Engineering

UNIVERSITY OF TECHNOLOGY SYDNEY

AUSTRALIA

Submitted JANUARY, 2020

**UNIVERSITY OF TECHNOLOGY SYDNEY
SCHOOL OF ELECTRICAL AND DATA
ENGINEERING**

The undersigned hereby certifies that he has read this thesis entitled “**Adopting Machine Learning Technology for the Classification of Parkinson’s Disease**” by **Farhan Mohammed** and that in his opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy

Production Note:

Signature removed prior to publication.

.....

Prof. Xiangjian (Sean) He

Principal Supervisor

Certificate of Authorship/Originality

I, Farhan Mohammed, declare that this thesis, is submitted in fulfillment of the requirements for the award of Doctor of Philosophy, in the School of Electrical and Data Engineering, Faculty of Engineering and Information Technology here at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Trainig Program.

Production Note:

Signature: Signature removed prior to publication.

Date: 20/07/2020

ABSTRACT

Adopting Machine Learning Technology for the Classification of Parkinson’s Disease

by

Farhan Mohammed

Parkinson’s Disease (PD) is the second most common neuro-degenerative disorder affecting approximately 1% of the population. Major symptoms include tremor, bradykinesia and freezing of gait. The precise diagnosis of PD remains a challenge for clinicians due to the similarity of PD symptoms with other disorders. Although diagnosis is based on clinical symptoms, PD is associated with a plethora of non-motor symptoms adding to its overall disability.

Research into early diagnosis of PD has taken advantage of machine learning-based image analysis. Neuroimaging modalities, like Single Photon Emission Computed Tomography (SPECT), have shown to aid in early diagnosis of PD. SPECT images are powerful tools that depict dopaminergic deficits in brain. Dopamine transporter (DAT) loss is a crucial feature necessary for PD identification. Regular scans are identified by intense and symmetric DAT binding that appear as two “comma-shaped” regions. Any asymmetry of this shape implies an abnormal finding. This thesis proposes three machine learning approaches for the classification of PD.

Firstly, we developed a neural network that classified PD patients from healthy controls. SPECT images were used to train our network. 10-fold cross-validation was used to evaluate our network’s performances. Experimental results indicated that our approach outperformed the benchmark studies.

To classify PD patients into different stages, PDStageNet was implemented which learned several features from the images and their associated stages. The affected re-

gions were enhanced using image segmentation process. Experimental results showed that the proposed approach achieved a very high accuracy in classifying PD patients into five clinical stages of PD progression.

Finally, a classification model was utilised to streamline the process of identifying PD patients for surgical treatment using only clinical data. A feature selection process was used to identify the essential features and determine if the accuracy could be improved. Two experiments were carried out to test this hypothesis. In experiment 1, the best classifier for PD classification was identified. In experiment 2, feature selection process was implemented to determine the essential features. Experimental results indicated that, with only 60% of the features, the accuracy was higher than the benchmark studies.

The significance of these studies is that we propose effective machine learning-based approaches to diagnose PD at its earlier stages, so that the management and prognosis of PD patients can be significantly improved. Given the high performance of our approaches, we believe that the early diagnosis of PD can be done, which will revolutionize PD diagnosis and management.

Acknowledgements

This research would not have been possible without the guidance and the help of many people.

First of all, I would like to thank my supervisor, Prof. Xiangjian He, for his guidance, support and encouragement in my academic career. I have learned so much from him in regards to research and academic writing. I know that Prof. Xiangjian He has always been doing his best for me and believe his knowledge, experience and vision enabled him to help me do my best in my academic career.

I would also like to express my gratitude to my co-supervisors, Dr. Yiguang Lin and Prof. Jinjun Chen. Dr. Yiguang Lin has provided me with valuable and constructive comments that helped me refine my writing skills. I believe that what I have learned from him will be a valuable asset for me in the future. In addition, their guidance, comments and suggestions during the course of my PhD research have been invaluable. Prof. Xiangjian He, Dr. Yiguang Lin and Prof. Jinjun Chen are nice and humble and have a very good reputation in our research community.

Finally, I would like to express my love and gratitude to my parents and my family. I am extremely grateful for their love, prayers and moral support to complete my PhD studies. Being a member of such a wonderful family is the most beautiful thing that happened to me. Thank you very much for everything you have done for me.

Farhan Mohammed

Author's Publications for the PhD

Book chapter

- B-1. **F. Mohammed**, X. He and Y. Lin, “Applications of Machine Learning Techniques in the Diagnosis of Parkinsons Disease: Promises and Challenges,” *mHeath for Belt and Road Initiative: mHealth for Parkinson’s Disease*, (pending publication)

Journal Papers

- J-1. **F. Mohammed**, X. He and Y. Lin, “Easy-to-use Deep Learning Model for Highly Accurate Diagnosis of Parkinson’s Disease using SPECT Images,” *Computerized Medical Imaging and Graphics*, (pending publication).

Conference Papers

- C-1. **F. Mohammed**, X. He, Y. Lin and J. Chen, “A Novel Model for Classification of Parkinsons Disease: Accurately Identifying Patients for Surgical Therapy,” *Proc. of 52nd Hawaii International Conference on System Sciences (HICSS-52)*, Hawaii, USA, January, 2019. (ERA Tier A conference)

Contents

Supervisor’s Approval	ii
Certificate	iii
Abstract	iv
Acknowledgments	vi
Author’s Publications for the PhD	vii
List of Figures	xii
List of Tables	xiv
1 Introduction to Parkinson’s Disease	1
1.1 Diagnosing Parkinson’s Disease	1
1.2 Health Expenditures Associated with Parkinson’s Disease in Australia	4
1.3 Challenges in Diagnosing Parkinson’s Disease	6
1.4 Machine Learning Techniques	9
1.4.1 Data Mining Concepts	9
1.4.2 Popular Classification Techniques	11
1.4.3 Deep-Learning Techniques	12
1.5 Parkinson Progression Markers Initiative (PPMI)	15
1.6 Contributions	16
1.7 Organization of this Thesis	17
2 Review of Prior Research Works	18

2.1	Classification of PD Using Vocal Attributes	18
2.2	Classification of PD Using Gait Attributes	23
2.3	Classification of PD Using Striatal Binding Ratio (SBR)	29
2.4	Classification of PD Using Imaging Modalities	33
2.5	Conclusion	34
3	Classification of Parkinson’s Disease using Deep-Learning	35
3.1	Introduction	35
3.2	Proposed Network Approach	39
3.2.1	Network Structure	40
3.2.2	Image Normalisation	43
3.3	Experimental Setup	45
3.3.1	Data Augmentation	46
3.3.2	Training and Testing the Network	47
3.3.3	Evaluation Metrics	47
3.4	Results	48
3.5	Discussion and Conclusion	51
4	Classification of Parkinson’s Disease (PD) into Multiple Stages of Progression	59
4.1	Introduction to PD Stages	59
4.2	Proposed Network Approach	62
4.2.1	Image Segmentation Process	63
4.2.2	Network Architecture	67
4.3	Experimental Setup	68
4.3.1	Training the network	69

4.3.2	Evaluation metrics	73
4.3.3	Testing the network	75
4.4	Results	75
4.5	Discussion and Conclusion	77
4.5.1	Comparison Analysis with Prior Works	81
5	Classification of Parkinson’s Disease Patients for Surgical Treatment	84
5.1	Introduction	84
5.2	Proposed Network Architecture	86
5.2.1	Pre-processing	87
5.2.2	Feature Selection	87
5.2.3	Model	88
5.2.4	Classification	89
5.3	Experiments and Results	90
5.3.1	Experiment 1: General Classification of PD Patients for Surgery Using Different Classifiers	90
5.3.2	Experiment 2: Classification of PD Patients for Surgery after Using Feature Selection	92
5.4	Discussion and Conclusion	93
5.4.1	Comparison of Current Study with Other Research Works	96
6	Conclusion and Future Work	100
6.1	Classification of PD Using Deep-Learning	100
6.2	Multiple Stages Classification of PD	101
6.3	Classification of PD Patients for Surgical Treatment	101

6.4 Future Work 102

Bibliography **104**

List of Figures

1.1	PD deaths in major countries from 1990 - 2017	2
1.2	SPECT scans showing examples of (a) Non-PD patient and (b) PD patient	3
1.3	Health system costs (AU\$) per person in 2014	4
1.4	Health costs (AU\$) associated with PD, in millions	5
1.5	Suggested treatment plan for early PD patient	7
1.6	Suggested treatment plan for late PD patient	8
1.7	Architecture of a CNN model	15
3.1	SPECT scans showing examples of (a) Healthy control and (b) PD patient	36
3.2	Network differences of (a) AlexNet and (b) Our network	42
3.3	SPECT scans showing examples of (a) Normalized PD patient and (b) Normalized healthy control	45
3.4	Performance comparison of our network with AlexNet	53
3.5	Network performance comparison with benchmark studies	55
4.1	Incidence of PD by disease stage	61
4.2	Framework including training and testing for PD stage classification	63
4.3	SPECT images of PD in (a) Stage 1, (b) Stage 2, (c) Stage 3, (d) Stage 4, and (e) Stage 5	64

4.4	Segmented SPECT images of PD in (a) Stage 1, (b) Stage 2, (c) Stage 3, (d) Stage 4, and (e) Stage 5	66
4.5	Network Architectures of (a) VGG-16 and (b) PDStageNet	69
4.6	Example of LabelBinarization process	70
4.7	Examples of (a) TP , (b) TN , (c) FP , and (d) FN for a multi-class classification	74
4.8	Confusion matrix of the study	77
4.9	Example of high confidence classification of PD in (a) Stage 1 (b) Stage 2 (c) Stage 3 (d) Stage 4 and (e) Stage 5	80
4.10	Examples of low confidence classification	82
5.1	Proposed approach architecture.	86
5.2	Flowchart of the model	89
5.3	List of features	89
5.4	Contribution of features to the classification of PD patients for surgery	95
5.5	Classification results using top 7 features and the rest of the features, respectively	97

List of Tables

3.1	Training accuracy results of different architectures	44
3.2	Detailed performance results of the 10-fold using Method 1	49
3.3	Detailed performance results of 10-fold with Method 2	50
3.4	Performance metrics of the network	51
3.5	Computational cost comparison with AlexNet and our network	54
3.6	Comparison table with previous studies	58
4.1	PD Stage characteristics based on HY scale	65
4.2	Results of the network model	77
4.3	Individual stage metrics of the network model	78
4.4	A comparative analysis of the present study and the published studies	83
5.1	General Classification of PD Patients Using Different Classifiers	91
5.2	Classification of PD Patients for Surgery Using Feature Selection	93
5.3	Top seven features' details	95
5.4	Summary and comparisons of previous related works on PD classification	99