



Combination of Medical Imaging and Demographic Data for Parkinson's Disease Diagnosis

Helena Rico Pereira^{1,2}(✉), José Manuel Fonseca¹,
and Hugo Alexandre Ferreira²

¹ CTS – Center of Technology and Systems, Uninova,
NOVA University of Lisbon, 2829-518 Caparica, Portugal

hri.pereira@campus.fct.unl.pt, jmf@uninova.pt

² Instituto de Biofísica e Engenharia Biomédica da Faculdade de Ciências da,
Universidade de Lisboa, 1749-016 Lisbon, Portugal
hugoferreira@campus.ul.pt

Abstract. The identification of biomarkers to discriminate Parkinson's Disease from other motor diseases is crucial to provide suitable treatment to patients. This study proposes a novel approach for the classification of structural Magnetic Resonance Imaging (MRI), Dopamine Transporter scan data (DaTscan) and demographic information (age and gender) to differentiate PD patients, "Scans Without Evidence for Dopaminergic Deficit" (SWEDD) patients and healthy control subjects using Convolutional Neural Networks (CNN). In Control vs PD, the accuracy of the classifier increased by adding subject gender from 94.5% to 96.0%, while in PD vs SWEDD adding age lead to 88.7% accuracy using slices encompassing the basal ganglia. The CNN was not able to successfully discriminate SWEDD vs Control. Our results suggested that pattern changes in slices encompassing the basal ganglia and the mesencephalon are relevant biomarkers for PD suggesting that this approach may have the potential to aid in PD biomarkers detection.

Keywords: Parkinson's Disease · Convolutional Neural Networks · MRI · DaTscan SPECT · SWEDD

1 Introduction

Parkinson's Disease (PD) is a progressive neurodegenerative disease associated with the degeneration of the dopamine neurons located at the *Substantia Nigra* (SN). This neurodegeneration causes one of the main features of PD patients: the presence of motor symptoms [1]. However, these symptoms only begin to emerge when about 50% of the degeneration of the SN neurons has already occurred [2]. The incidence of PD is higher in the elderly [3] and men [4]. Nowadays, the diagnosis of PD is based on medical history, signs and symptoms, and positive response to PD medication [5]. However, misdiagnosis usually occurs mainly due to symptoms similarity between PD and other motor disorders [6]. Currently, medical imaging techniques, namely Magnetic Resonance Imaging (MRI) and Single Photon Emission Computed Tomography

(SPECT), are not used for diagnosis but reveal useful insights regarding the disease progression [1, 7]. SPECT with Dopamine Transporter imaging (DaTscan) is used to detect presynaptic dopamine dysfunction, which is a biomarker of PD [7]. An interesting finding of DaTscan usage was the identification of a small group of patients (10–20%) [8] diagnosed as having PD but presenting “Scans Without Evidence of Dopaminergic Deficit” (SWEDD), that is, presenting a normal DaTscan. MRI is also of particular interest to differentiate PD from other motor diseases characterized by structural brain changes [1]. In the last few years, Artificial Intelligence (AI) algorithms have become a promising approach for the classification of medical imaging of PD patients [9]. Therefore, the following research question aroused: “Can MRI and SPECT imaging together with subject demographic data (age and gender) provide useful information to aid physicians in reducing diagnosis uncertainty?”. In this paper, a model for PD and SWEDD prediction using MRI and SPECT images in combination with age and gender as additional information using Convolutional Neural Networks (CNN) is proposed.

2 Contribution to Life Improvement

It is possible to find in the literature numerous approaches using AI algorithms to detect diseases and relevant biomarkers [10, 11]. The model proposed in this study aims to address the need for accurate diagnosis, hence, contributing to the improvement of patients’ quality-of life. Usually, these patients have difficulty doing their daily tasks due to the motor and non-motor symptoms such as sleep disturbances, mood disorders and smell impairment [12], which lead to poor quality-of-life. Therefore, early diagnosis at the disease onset may contribute to adjusting the therapeutics and relieve symptoms [13]. Besides, the identification of biomarkers allows PD monitorization and the understanding of the mechanism underlying the disease.

3 State of the Art

Several studies have used medical imaging techniques such as MRI and DaTscan to detect anatomical and physiopathological features of PD [1, 7]. For instance, a review study focused on MRI images and Voxel-Based Morphometry (VBM) reported volume differences in the grey matter of the frontal lobe, olfactory bulb, basal ganglia, and SN [1]. Regarding SPECT images, a study found that performing this exam during a two-year follow-up reduced the period of observation even in ambiguous cases [14] in 90% of the cases. In this regard, each imaging modality provides different biomarkers associated with PD [1]. More recently, some studies made use of AI, namely, machine learning algorithms to differentiate PD from other diseases [10, 11, 15–17]. Using Support Vector Machine (SVM), Singh et al. [18] classified previously extracted intensities from segmented T1-Weighted (T1-W) MRI images of PD, SWEDD and Controls subjects (accuracies above 95%), while Oliveira et al. [19] computed and classified striatal region ratios at voxel level from DaTscan images of PD patients and control subjects (97% accuracy). Adeli et al. [11] classified extracted

features from MRI and SPECT images using a kernel-based SVM. However, studies using SVM have the disadvantage of using features that are known to be related to PD, which leads to lost information that may provide new insights regarding the disease. For instance, Oliveira et al. only considered binding ratio, a PD biomarker, which is expected to lead to higher accuracies. In single-modality classification, Choi et al. [15] used 3D CNN to classify PD and Control images (96% accuracy) and used the trained CNN to classify SWEDD patients, while Esmailzadeh et al. [10] used 3D T1-W images to discriminate PD patients from Control subjects and analyzed the heatmaps, highlighting the most relevant features for the classification (SN and basal ganglia). However, although presenting higher accuracy, the majority of these studies including those that used CNN have the disadvantage of using imbalanced data. Beyond medical imaging, demographic information, such as age [10, 20] and gender [10], have also been considered relevant to differentiate between PD patients from control subjects.

4 Model for Images and Demographic Data Classification

Data used in study were extracted from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. One MRI and one DaTscan image per subject were extracted from PD, SWEDD and Control cohorts. According to PPMI documentation [21], 3D T1-W images were acquired at 1.5T and 3T with MPRAGE GRAPPA or SAG FSPGR sequences, and DaTscan images were obtained 4 h after radiotracer Ioflupane I123 administration. The Mann-Whitney U and the Chi-Squared tests were performed to verify if the data sample used in this study were age and gender-matched using a 5% significance level. MRI pre-processing was performed using SPM12 [22] and included Anterior Commissure (AC) – Posterior Commissure (PC) alignment correction, bias regularization, spatial normalization to Montreal Neurological Institute (MNI) brain template, bias regularization and smoothing. According to PPMI documentation [23], DaTscan images were already pre-processed when extracted from the database. They were reconstructed using an iterative reconstruction algorithm based on hybrid ordered-subset expectation maximization in HERMES (Hermes Medical Solutions, Stockholm, Sweden). Then, a standard 3D Gaussian 6.0 mm filter was applied to the resultant image and normalized to the MNI template. In this work, DaTscan images were co-registered to the corresponding MRI image of each subject. Then, MRI and DaTscan data were divided into 2D axial slices to better adapt our approach and classification models to current clinical practice, and a Contrast Limited Adaptive Histogram Equalization was applied to each MRI slice to improve and uniformize image contrast without overamplify noisy regions [24]. For the classification, three slices of each subject encompassing SN and basal ganglia were averaged, since they are associated with PD. The architecture of CNN used in this work is illustrated in Fig. 1, and it was implemented using the Caffe framework [25]. The CNN, in contrast with other algorithms such as SVM, are able to directly extract features from images and classify their weights, which may result in higher classification accuracies and smaller classification times. The Monte Carlo Cross-Validation (MCCV) method was used to provide an unbiased classification performance estimation. Five batches were created using

MCCV with the proportion 2:1:1 for train, validation and test sets. Besides brain images, age and gender were also provided to the network. Two images were created for each feature to represent this additional information, as illustrated in Fig. 2.

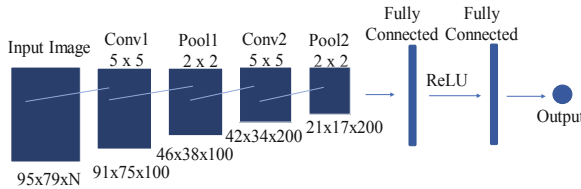


Fig. 1. CNN architecture

Males were defined with horizontal and vertical lines (similar to a chessboard), while vertical lines only represented females. Concerning the age of the subjects, each age value was divided by 100 to be in the range of 0 and 1. Then, for each age, an image with black squares and the background set with the value of the normalized age was created. The Gradient-weighted Class Activation Mapping (Grad-CAM) method proposed by Selvaraju et al. [26] was used to identify relevant regions in the classification. The proposed classification model was computed in a CentOS Linux7 and using the graphics card NVIDIA GeForce GTX TITAN X.

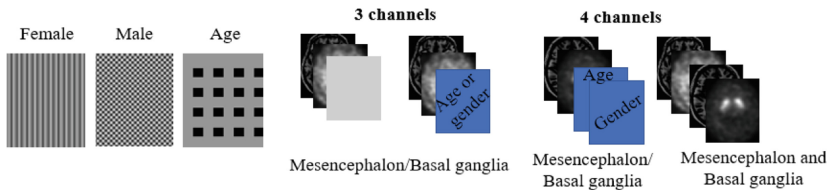


Fig. 2. The 3 channels comprised slices of the mesencephalon or the basal ganglia together with age or gender, while 4 channels comprise that data with both demographic information (age and gender) or alternatively slices of the basal ganglia and the mesencephalon simultaneously.

5 Results and Discussion

The developed model yielded relevant results for the discrimination of PD, SWEDD, and control subjects using a multimodal approach and demographic information. The number of subjects used was 168 control (60 ± 11 years old, 59 females, 109 males), 378 PD (62 ± 10 years old, 136 females, 242 males), 58 SWEDD (62 ± 10 years old, 20 females, 38 males). The results of the Mann-Whitney U and Chi-Square tests suggest that those Control, PD and SWEDD groups do not show statistically significant age and gender differences ($p > 0.05$). The relevant features using by CNN to classify the images are illustrated in Fig. 3. The results of each classification are shown in

Table 1. Overall, the accuracy of the classifier was higher using slices encompassing the basal ganglia in comparison with slices, including the mesencephalon as in the study of Pereira [27]. As expected, the classification Control vs PD using slices encompassing the mesencephalon and the basal ganglia yielded higher accuracy (94.5%) corresponding to a Receiver's Operating Characteristic Area Under Curve (AUC) = 0.98 [0.95–1.00] and Cohen's kappa = 0.98 ± 0.05 . These results are consistent with the study of Adeli et al. [20] in which SN, putamen and caudate were selected as the more relevant region of interest by the SVM model, and with Kollia et al. [28] which achieved 94% using also CNN (55 PD and 33 Control). Although with lower accuracy (79.3%) than SPECT slices, MRI slices encompassing the mesencephalon were able to discriminate Control from PD. Moreover, the analysis of the features extracted from de CNN, illustrated in Fig. 3, suggests the SN as a relevant region for the classification, corroborating other studies in which changes in the SN were found [1]. Concerning the classification of the Control vs SWEDD groups, the CNN was not able to discriminate these two groups, hence supporting the evidence that SWEDD patients present similar brain scans to Control subjects [8].

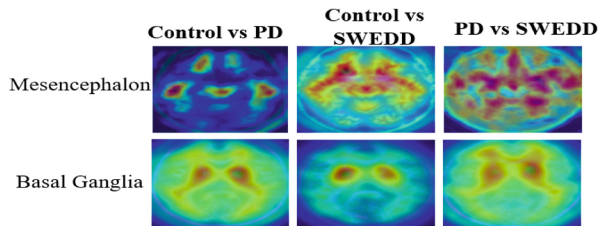


Fig. 3. The heatmaps obtained using the GRAD-Cam algorithm. Regions in red represent the most relevant features for the classification, while blue illustrate the regions less important. (Color figure online)

The classification of PD vs SWEDD using only MRI and DaTscan slices suggest that the performance of our model is better [29] when using slices encompassing the basal ganglia (AUC = 0.93 [0.83–1.00]) in contrast to slices encompassing the *mesencephalon* (AUC = 0.67 [0.82–0.86]), supporting the hypothesis that PD and SWEDD are two distinct groups, in which SWEDD shows no evidence of dopamine deficit [30]. Adding gender to the classification, only increased the accuracy in Control vs PD with slices encompassing the basal ganglia, while decreased in MRI slices encompassing the mesencephalon. These results support the study of Haaxma et al. [31] in which striatal dopamine levels differences were found between men and women. However, for both Control vs SWEDD and PD vs SWEDD, gender decreased the accuracy. It is important to note that in these two comparisons, the number of subjects used in train, validation and test (subjects = 116) was smaller than in Control vs PD (subjects = 336). In contrast to some studies in which age improved the classification of PD vs Control [10, 20], the results obtained in this study showed a decrease in the accuracy value. However, those studies used imbalanced data, which may lead to a biased classification towards the group with more subjects increasing

misclassification of minority classes [32]. Interestingly, our classification of PD vs SWEDD improved when age was included. However, it is important to note that for this case, a small data set (test set = 30 subjects) was used which can potentiate biased estimations. Besides, this CNN model is only able to classify 2D imagens, losing information that 3D imagens can provide, such as volume of brain structures [10].

Table 1. Results of the classification.

Group	Mesencephalon			Basal ganglia		
	Acc (%)	Sen (%)	Spe (%)	Acc (%)	Sen (%)	Spe (%)
Control vs PD	79.3	82.9	75.7	93.1	93.8	92.7
Control vs SWEDD	52.7	56.0	49.3	57.0	48.0	57.3
PD vs SWEDD	63.3	66.7	60.0	87.3	90.1	84.0
Control vs PD + age	76.9	83.4	71.4	93.3	93.8	92.9
Control vs SWEDD + age	62.7	62.7	62.7	54.0	60.0	48.0
PD vs SWEDD + age	65.3	64.0	66.7	88.7	86.7	90.7
Control vs PD + gender	71.4	72.3	70.0	96.0	97.3	94.3
Control vs SWEDD + gender	52.0	65.3	39.7	56.7	56.0	57.3
PD vs SWEDD + gender	61.3	62.7	60.0	86.7	93.3	80.0
Control vs PD + age + gender	72.4	70.0	74.7	92.3	91.9	89.3
Control vs SWEDD + age + gender	53.3	46.7	60.0	53.3	53.3	57.3
PD vs SWEDD + age + gender	67.6	68.0	65.3	85.7	85.3	89.3
	Mesencephalon + Basal ganglia					
	Acc (%)		Sen (%)		Spe (%)	
Control vs PD	94.5		92.9		96.1	
Control vs SWEDD	53.3		57.3		49.3	
PD vs SWEDD	84.7		84.0		85.3	

Acc = Accuracy, Sen = Sensitivity, Spe = Specificity

6 Conclusion and Future Work

A multi-modality approach using CNN and additional demographic information is proposed to differentiate between PD, SWEDD and Control subjects. Notably, this work extends previous studies by adding age and gender as additional information for the CNN classifier, being the first study that combines MRI, SPECT, age and gender to distinguish between PD and SWEDD patients. The results support the relevance of biomarkers, such as the basal ganglia and the SN for the classification of images of PD and SWEDD patients. However, it is essential to note some limitations of this study. Although previous studies support the results, the model was only internally validated. Moreover, the images that represent age and gender were created arbitrarily. Therefore, future work should include an external validation of the proposed approach as well as

optimization of network parameters or improve the CNN architecture to classify also 3D images. Additional techniques to translate and integrate the demographic information in the network should also be investigated.

Acknowledgments. The authors would like to thank the financial support from Fundação para a Ciência e Tecnologia (FCT) under the project UID/BIO/00645/2019 and UIDB/00066/2020 (CTS – Center of Technology and Systems), Programa Operacional Temático Competitividade e Internacionalização under the project POCI-01-0145-FEDER-016428, to the NVIDIA GPU Grant Program and to work by PPMI personnel that went into accumulating the data, as well as funding of the study. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and industry partners: <https://www.ppmi-info.org/about-ppmi/who-we-are/study-sponsors/>.

References

1. Saeed, U., Compagnone, J., Aviv, R.I., et al.: Imaging biomarkers in Parkinson's disease and Parkinsonian syndromes: current and emerging concepts. *Transl Neurodegener.* **6**, 8 (2017)
2. Cheng, H.-C., Ulane, C.M., Burke, R.E.: Clinical progression in Parkinson disease and the neurobiology of axons. *Ann. Neurol.* **67**(6), 715–725 (2010)
3. Reeve, A., Simcox, E., Turnbull, D.: Ageing and Parkinson's disease: why is advancing age the biggest risk factor? *Ageing Res. Rev.* **14**(100), 19–30 (2014)
4. Wooten, G.F., Currie, L.J., Bovbjerg, V.E., et al.: Are men at greater risk for Parkinson's disease than women? *J. Neurol. Neurosurg. Psychiatry* **75**(4), 637–639 (2004)
5. Williams, D.R., Litvan, I.: Parkinsonian syndromes. *Contin. Lifelong Learn. Neurol.* **19**(5), 1189–1212 (2013)
6. Fearnley, J.M., Lees, A.J.: Striatonigral degeneration. A clinicopathological study. *Brain* **113**(6), 1823–1842 (1990)
7. Brooks, D.J.: Imaging approaches to Parkinson disease. *J. Nuclear Med.* **51**(4), 596–609 (2010)
8. Erro, R., Schneider, S.A., Stamelou, M., et al.: What do patients with scans without evidence of dopaminergic deficit (SWEDD) have? New evidence and continuing controversies. *J. Neurol. Neurosurg. Psychiatry* **87**(3), 319–323 (2016)
9. Xu, J., Zhang, M.: Use of magnetic resonance imaging and artificial intelligence in studies of diagnosis of Parkinson's disease. *ACS Chem. Neurosci.* **10**(6), 2658–2667 (2019)
10. Esmaeilzadeh, S., Yang, Y., Adeli, E.: End-to-end Parkinson disease diagnosis using brain MR-images by 3D-CNN. *CoRR*. abs/1806.05233 (2018)
11. Adeli, E., Wu, G., Saghafi, B., et al.: Kernel-based joint feature selection and max-margin classification for early diagnosis of Parkinson's disease. *Sci. Rep.* **7**(1), 41069 (2017)
12. Olanow, C.W., Schapira, A.H.V., Obeso, J.A.: Parkinson's disease and other movement disorders in Harrison's principles of internal medicine, 19th edn, pp. 2609–2626. McGraw-Hill Education, New York (2015)
13. Lang, A.E., Lozano, A.M.: Parkinson's disease. *N. Engl. J. Med.* **339**(16), 1130–1143 (1998)
14. Tolosa, E., Vander, B.T., Moreno, E.: Accuracy of DaTSCAN (123I-Ioflupane) SPECT in diagnosis of patients with clinically uncertain parkinsonism: 2-year follow-up of an open-label study. *Mov. Disord.* **22**(16), 2346–2351 (2007)

15. Choi, H., Ha, S., Im, H.J., et al.: Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. *NeuroImage Clin.* **16** (September), 586–594 (2017)
16. Singh, G., Samavedham, L.: Algorithm for image-based biomarker detection for differential diagnosis of Parkinson's disease. *IFAC-PapersOnLine* **48**(8), 918–923 (2015)
17. Kollia, I., Stafylopatis, A.-G., Kollias, S.: Predicting Parkinson's disease using latent information extracted from deep neural networks. *CoRR*. abs/1901.07822 (2019)
18. Singh, G., Samavedham, L.: Unsupervised learning based feature extraction for differential diagnosis of neurodegenerative diseases: a case study on early-stage diagnosis of Parkinson disease. *J. Neurosci. Methods* **256**, 30–40 (2015)
19. Oliveira, F.P.M., Castelo-Branco, M.: Computer-aided diagnosis of Parkinson's disease based on [(123)I]FP-CIT SPECT binding potential images, using the voxels-as-features approach and support vector machines. *J. Neural Eng.* **12**(2), 026008 (2015)
20. Palumbo, B., Fravolini, M.L., Buresta, T., et al.: Diagnostic accuracy of Parkinson disease by support vector machine (SVM) analysis of 123I-FP-CIT brain SPECT data. *Med. (Baltimore)* **93**(27), e228 (2014)
21. Parkinson's Progression Markers Initiative. MRI - Technical Operations, PPMI (2015)
22. Friston, K.J.: Statistical Parametric Mapping: The Analysis of Functional Brain Images. *Functional Neuroimaging: Technical*, p. 656. Elsevier/Academic Press, Amsterdam (2006)
23. Marek, K., Chowdhury, S., Siderowf, A., et al.: The Parkinson's progression markers initiative (PPMI) – establishing a PD biomarker cohort. *Ann. Clin. Transl. Neurol.* **5**(12), 1460–1477 (2018)
24. Zuiderveld, K.: Contrast limited adaptive histogram equalization. In: *Graphics Gems*, pp. 474–485. Elsevier (1994)
25. Jia, Y., et al.: Caffe: convolutional architecture for fast feature embedding. In: *Proceedings of ACM International Conference on Multimedia - MM14*, pp. 675–678 (2014)
26. Selvaraju, R.R., Cogswell, M., Das, A., et al.: Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *IEEE Conference on Computer Vision*, October 2017, pp. 618–626 (2017)
27. Pereira, H.R., Ferreira, H.A.: Classification of patients with Parkinson's disease using medical imaging and artificial intelligence algorithms. In: Henriques, J., Neves, N., de Carvalho, P. (eds.) *MEDICON 2019*. IP, vol. 76, pp. 2043–2056. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-31635-8_241
28. Tagaris, A., Kollias, D., Stafylopatis, A., et al.: Machine learning for neurodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset. *Int. J. Artif. Intell. Tools* **27**(03), 1850011 (2018)
29. Hajian-Tilaki, K.: Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* **4**(2), 627–635 (2013)
30. Bajaj, N.: SWEDD for the general neurologist. *ACNRSO* **10**(4), 30–31 (2010)
31. Haaxma, C.A., Bloem, B.R., Borm, G.F., et al.: Gender differences in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* **78**(8), 819–824 (2007)
32. Longadge, R., Dongre, S.: Class imbalance problem in data mining review. *Eur. J. Intern. Med.* **24**(1), e256 (2013)