



Modelling the Progression of the Symptoms of Parkinsons Disease Using a Nonlinear Decomposition of 123I FP-CIT SPECT Images

Jose Antonio Simón-Rodríguez, Francisco Jesús Martínez-Murcia^(✉),
Javier Ramírez, Diego Castillo-Barnes, and Juan Manuel Gorriz

Department of Signal Theory, Networking and Communications,
Andalusian Institute on Data Science and Computational Intelligence,
University of Granada, Granada, Spain

josesimon@correo.ugr.es, fjesusmartinez@ugr.es

Abstract. Parkinson's Disease (PD) is one of the most relevant neurodegenerative disorder. It is mainly caused by a loss of dopamine neurons leading to a reduction in the neurotransmitter dopamine, which is essential in the control of movement. While the diagnosis of PD is mainly clinical, new markers are being used with high accuracy in the later stages of the disease, where symptoms are clear. However, the early stages of the disease, when symptoms start to evolve and treatments could potentially be more effective, are yet to be explored. In this work we explore the low-dimensional latent space of the Parkinson's Progression Markers Initiative (PPMI) DaTSCAN imaging dataset, with a twofold objective: to perform an early diagnosis of PD, and to link the low-dimensional representation of the images to symptomatology. Different unsupervised methods have been used to extract the features (ISOMAP and PCA), and the resulting space is evaluated by means of binary or multiclass classification, and linear regression, using Support Vector Machines (SVM). We obtained a diagnosis of PD with an Area Under the ROC Curve (AUC) above 0.94 for three different variables, and a relevant link between the Unified Parkinson's Disease Rating Scale (UPDRS) and the imaging composite features with $R^2 > 0.2$ even for a simple linear model. These results pave the way to explore latent representations in PD and study the progression of the disease and its symptomatology.

Keywords: Parkinson Disease (PD) · ISOMAP · Principal Component Analysis (PCA) · Support Vector Machine (SVM) · ROC Curve

1 Introduction

Parkinson's disease (PD) affects more than 6 million people worldwide [15], being the second most relevant neurodegenerative disorder after Alzheimer's Disease

© Springer Nature Switzerland AG 2022

J. M. Ferrández Vicente et al. (Eds.): IWINAC 2022, LNCS 13258, pp. 104–113, 2022.

https://doi.org/10.1007/978-3-031-06242-1_11

(AD). Parkinson's disease is a severe, progressive and chronic disease of unknown origin which is mainly caused by the progressive loss of dopaminergic neurons of the nigrostriatal pathway [11]. The loss of dopamine neurons leads to a reduction in the neurotransmitter dopamine, which is essential in the control of movement. Nowadays, PD is incurable, and when symptoms appear it means that neuronal destruction is excessive, and any possible treatments that may slow or stop the progression are mainly ineffective. Therefore the greatest challenge in this neurodegenerative disorder is an early diagnostic, even before the apparition of symptoms.

At present, the typical diagnosis of PD is clinical, and based on a subjective view of a patient's symptoms. There are however two main reasons that complicate the diagnosis of PD. First and foremost, because there are many similar movement disorders that conform what is frequently called parkinsonism, of different etiology. Secondly, the severity of these symptoms changes over time. For these reasons, approximately 25% of PD diagnoses are incorrect when compared to autopsy findings [10].

Diagnosis has improved with time, with the increasing use of biomarkers such as tau protein, present in Cerebro-Spinal Fluid (CSF). There are also many studies that aim at diagnosing PD using movement or tremor recorded in wearable sensors [5, 7] or the smartphone data [16]. Or even studies that relate PD with speech faculties: [12].

Neuroimaging modalities yields early diagnosis of PD by providing noninvasive biomarkers. Among them, Single Photon Emission Computed Tomography (SPECT) uses highly specific radiopharmaceuticals –e.g., DaTSCAN– that bind to dopamine transporters in the striatum, making it possible to observe dopaminergic deficits in the brain. However, these images provide high dimensional features that could be better exploited by means of computers than by simple visual inspection. As a consequence, Machine Learning (ML) has thrived in the latest years, mainly to perform a differential diagnosis and enabling Computer Aided Diagnosis (CAD) systems [6, 11, 17]. However, there are a lesser amount of studies that try to make sense of PD progression using these images and other variables.

This study aims to develop an algorithm that can model the progression of PD, offering support for its early diagnosis. The proposed system is based on a non-linear decomposition of the SPECT images using unsupervised machine learning methods and the modelling of the composite variables via support vector machines (SVMs). The system is tested in two differential tasks: differential diagnosis (classification) and disease progression analysis (regression) by means of a longitudinal dataset, both using SVMs. The system is evaluated by means of stratified k -fold cross-validation using a different performance metrics in order to ensure the validity of the decomposition techniques for studying and diagnosing PD.

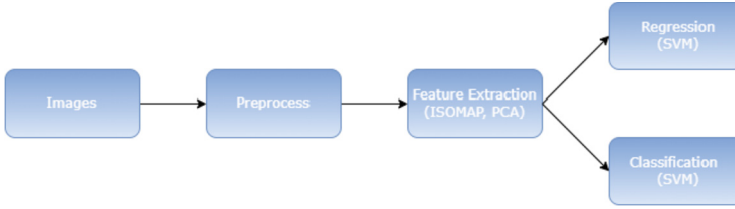


Fig. 1. Flowchart of the study

2 Methodology

2.1 Dataset Description and Image Preprocessing

Data used in the preparation of this article were obtained from the Parkinson’s Progression Markers Initiative (PPMI) database (<https://www.ppmi-info.org/accessdata-specimens/download-data>). For up-to-date information on the study, visit www.ppmi-info.org.

The cohort used for this study is composed by those subjects initially diagnosed either as controls (CTL), with no evidence of neurodegenerative deficits, and Parkinson’s Disease (PD) affected subjects, with different levels of severity. The CTL group consists of 101 males and 53 females, with 2 of them showing mild symptomatology (HY_on=1). The PD group consists of 284 males and 159 females. Table 1 shows the demographic analysis of the dataset. PD subjects and some subjects of the CTL group were clinically followed for up to 5 years, providing data for 1399 sessions, which will be evaluated afterwards to study the progression of imaging biomarkers and their relationship to PD-specific progression indicators.

Table 1. Demographics of the PPMI subjects included in the study.

Prim. Diag.	Sex	HY	N	Age	UPDRS3 on	Cognitive state
PD	M	1.0	124	60.42 [9.96]	14.411 [5.98]	1.091 [0.29]
		2.0	159	64.06 [8.79]	25.346 [8.75]	1.242 [0.44]
		3.0	1	67.74 [0.00]	23.000 [0.00]	1.000 [0.00]
	F	1.0	76	57.00 [9.39]	14.553 [6.51]	1.000 [0.00]
		2.0	81	63.45 [9.26]	24.617 [8.38]	1.045 [0.21]
		3.0	2	63.11 [11.10]	30.000 [8.49]	1.000 [0.00]
CTL	M	0.0	99	61.25 [11.03]	1.133 [2.04]	1.071 [0.26]
		1.0	2	61.07 [9.02]	5.500 [4.95]	1.000 [0.00]
	F	0.0	53	59.21 [12.67]	1.283 [2.33]	1.000 [0.00]

DaTSCAN images were preprocessed using affine registration to the MNI space as in [8]. Intensity normalization was based on a thresholding technique

based on an alpha-stable modelling of the image intensity distribution using the algorithm proposed in [2]. Additionally, three categorical variables will be used for classification: approximate diagnosis at each visit (APPRDX), diagnosis at the first visit (PRIMDIAG) and the value for Hoehn and Yahr (HY) scale for PD, that ranges from 0 (no symptoms) to 3 (severe disability). APPRDX distinguishes between healthy patients, PD's and SWEDD patients, whereas PRIMDIAG just between PD and CTL. The Unified Parkinson's Disease Rating Scale (UPDRS) degree for PD, that measures the degree of symptomatology, is also used as a continuous variable in regression.

2.2 Manifold Learning

Manifold Learning builds on the assumption that real, highly dimensional data such as images lies on a lower-dimensional nonlinear manifold. The group of algorithms aimed at approximating this manifold and how real data is projected to it are known as Manifold Learning algorithms [8].

This work follows the outline at Fig. 1, in which feature extraction via different dimensionality reduction techniques is achieved. After the image preprocessing described at Sect. 2, we apply two algorithms: Principal Component Analysis (PCA) or ISOMAP.

Principal Component Analysis (PCA) is a linear dimensionality reduction technique which allows to identify new linear subspaces. The principal components (PC) are each one of the spatial directions which maximizes the variance of the data while being orthogonal to each other, therefore they are uncorrelated variables [9].

In contrast to the linear nature of PCA, we have applied an isometric feature mapping, or Isomap, an algorithm that performs nonlinear modelling of a manifold extending metric multidimensional scaling via geodesic distances. The aim is to find a linear, lower dimensional subspace in which to embed the data in a high dimensional space, preserving the geodesic distance between the data [3].

We used the algorithm for computing Isomap as defined in [14], that is composed of three steps:

1. First, K-nearest neighbors are applied to construct a neighborhood graph.
2. Second, the shortest path between all pairs of point is calculated by estimating geodesic distances, usually using the Dijkstra or Floyd-Warshall algorithms.
3. Finally, a d-dimensional Euclidean embedding is constructed by a partial eigenvalue decomposition (i.e., taking the d largest eigenvalues of the kernel).

2.3 Classification and Regression Experiments

Last step in our analysis is the application of machine learning modelling for predicting scores. We have used an algorithm that has proven robust in many applications: Support Vector Machines (SVM), which have been widely used in

Alzheimer’s [13] or cancer [1]. We have used the implementation of LIBLINEAR for SVM classifiers (SVC) and SVM regression (SVR) [4].

Linear SVC is a particular case of SVR, as both try to predict a target variable Y from a set of data X , which in our case are the coordinates of the projections of the images in the manifold. SVMs in general try to model the curve inherent to the trend of data, creating a linear hyperplane that in the case of classification, separates (theoretically) the data. In the case of SVR, it becomes a predictor of certain variables.

5-fold stratified cross-validation (CV) is used to obtain performance measures both for regression and classification. For the classification approach, the average and standard deviation of accuracy, sensitivity, specificity and balanced accuracy are provided, along with the ROC curve in each CV loop. In the case of regression, Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and R^2 are used as performance measures.

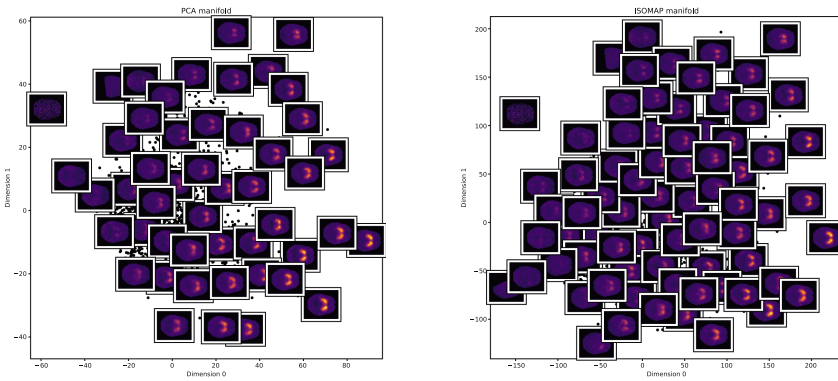


Fig. 2. DaTSCAN central slice placed in the corresponding coordinates of the first 2 dimensions of PCA (left) and Isomap (right).

3 Results and Discussion

First of all, we observe the differences between the linear decomposition with PCA and non-linear decomposition with Isomap in Fig. 2. First, two dimensions are shown for each methodology. Both approaches show similar trends in the first dimension (dimension 0), spanning from negative to positive values that are related to the intensity of the striatal region. However, the Isomap decomposes in a more uniformly distributed coordinate space, thanks to its nonlinear modelling, whereas PCA is more influenced by extreme values and outliers. Dimension 1, however, differs. In Isomap it is related to the intensity of the tails of the striatal region (putamen), whereas in PCA seems to measure roughly the asymmetry of the image.

With the aim of predicting the symptomatology of a patient just from image composite values, we trained a SVR with the projections of both Isomap and ICA. The performance of the SVR in predicting the variable UPDRS is shown at Table 2. There, the Isomap is shown a more accurate decomposition than the PCA, with no big differences between the two and three component decomposition, and higher R2 (above 0.2) than PCA.

Table 2. Performance of the SVR in predicting UPDRS from the 2 and 3-component PCA and Isomap projections.

	ISOMAP 2	ISOMAP 3	PCA 2	PCA 3
MAE	12.317 [0.749]	12.318 [0.746]	12.692 [0.730]	12.692 [0,733]
R2	0.207 [0.063]	0.206 [0.064]	0.190 [0.0557]	0.191 [0.056]
MSE	261.594 [29.299]	261.468 [29.310]	266.494 [27.564]	266.374 [27.516]
RMSE	16.148 [0.917]	16.144 [0.918]	16.302 [0.849]	16.299 [0.847]

The resulting SVR model for Isomap and PCA is shown at Fig. 3. There, the black lines represent a perfect linear reconstruction. We can observe that the predictions using Isomap decomposition are more linear with respect to the real UPDRS score than those of the PCA. The error encountered is due mainly to larger UPDRS values, which involves a more severe symptomatology.

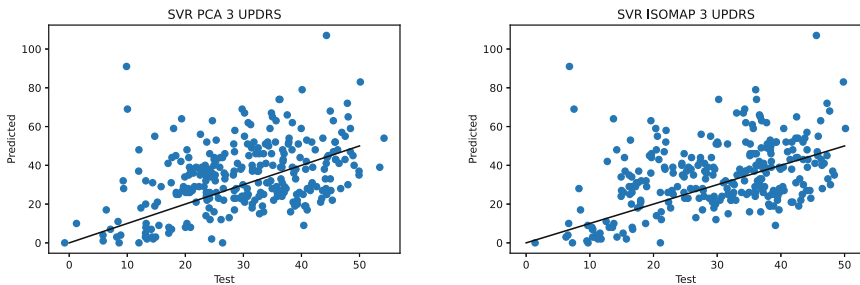


Fig. 3. Comparison of the predictions obtained by a SVR using the PCA and the Isomap decomposition with two components.

The composite imaging features obtained from PCA and Isomap are also used to predict three different targets: PRIMDIAG, APPRDX, and HY scale (see Sect. 2 for a description of these variables). To do so, a SVC is trained, and measures of Accuracy, Sensitivity, Specificity, and Balanced Accuracy (average of Sensitivity and Specificity) are reported in Table 3.

Table 3. Performance of the SVC when using the two- and three-dimensional projections of Isomap and PCA.

Variable	Performance	ISOMAP 3	PCA 3	PCA 2	ISOMAP 2
APPRDX	Accuracy	0.879 [0.028]	0.856 [0.043]	0.855 [0.041]	0.871 [0.034]
	Bal. Accuracy	0.882 [0.043]	0.847 [0.052]	0.847 [0.054]	0.877 [0.048]
	Sensitivity	0.886 [0.088]	0.835 [0.091]	0.835 [0.091]	0.886 [0.097]
	Specificity	0.878 [0.032]	0.859 [0.046]	0.858 [0.043]	0.869 [0.040]
PRIMDIAG	Accuracy	0.880 [0.031]	0.857 [0.046]	0.856 [0.044]	0.873 [0.036]
	Bal. Accuracy	0.751 [0.046]	0.724 [0.054]	0.722 [0.054]	0.745 [0.048]
	Specificity	0.520 [0.086]	0.474 [0.099]	0.470 [0.099]	0.507 [0.091]
	Sensitivity	0.982 [0.014]	0.973 [0.015]	0.973 [0.015]	0.982 [0.015]
HY	Accuracy	0.611 [0.038]	0.594 [0.020]	0.593 [0.022]	0.606 [0.021]
	Bal. Accuracy	0.466 [0.021]	0.441 [0.019]	0.438 [0.013]	0.457 [0.009]
	Class 0	0.897 [0.066]	0.817 [0.108]	0.829 [0.086]	0.903 [0.066]
	Class 1	0.143 [0.060]	0.126 [0.039]	0.086 [0.038]	0.072 [0.040]
	Class 2	0.824 [0.065]	0.821 [0.061]	0.838 [0.059]	0.852 [0.057]
	Class 3	0.000 [0.000]	0.000 [0.000]	0.000 [0.000]	0.000 [0.000]

It can be observed that the best performing decomposition is again Isomap with three dimensions. In this cases, it is able to approximate each subject's and session diagnosis with high accuracy. It is also capable of predicting primary diagnosis with high sensitivity (>0.98). For its part, the symptomatology as measured by HY, was predicted with a Multiclass one-vs-all SVC. The results were good for classes 0 and 2, but not for classes 1 and 3. This is relevant for interpreting this result, as class 0 are controls and class 2 has severe symptoms. Class 3 subjects were interpreted as class 2, probably because they show similar dopaminergic deficit, in contrast to controls. Class 1 is so close to class 0 that some subjects with primary diagnosis as CTL have class 1 assigned. However, it is very likely that some of these subjects are incorrectly interpreted as controls and other as class 2 because they show dopaminergic deficit.

ROC curves for the differential binary diagnosis are shown at Fig. 4. Individual curves for each CV fold are shown, and the mean ROC curve is shown in blue. There it can be seen that the AUC of our methodology is very high for providing a good differential diagnosis of PD based solely on imaging markers.

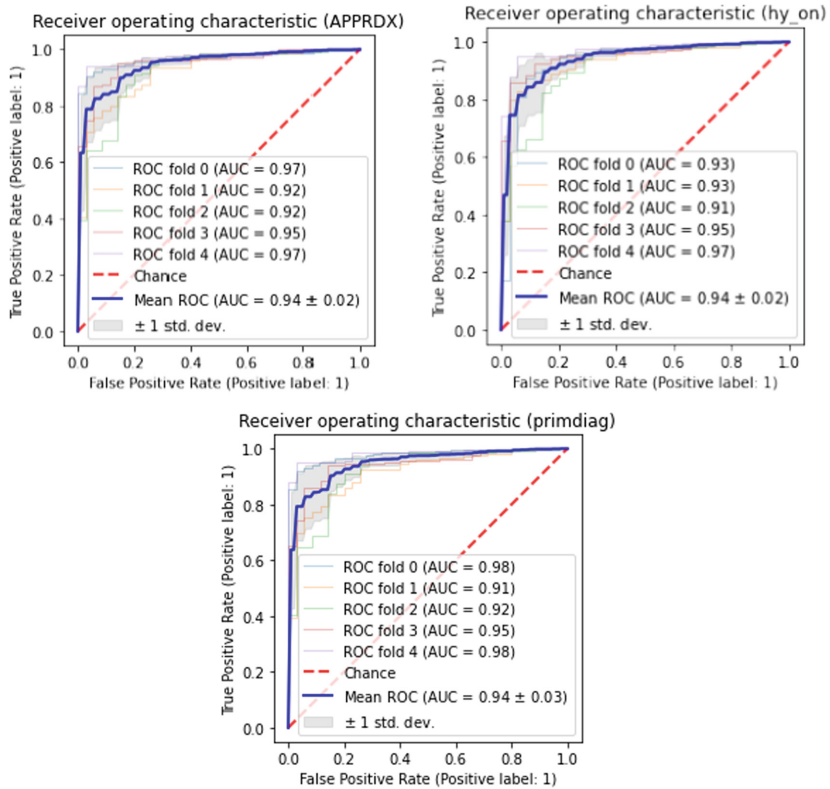


Fig. 4. ROC curves obtained with SVC applied to ISOMAP 3 output with the approximate diagnosis, HY scale and primary diagnosis.

4 Conclusions

Due to the critical importance and the difficulty of diagnosing Parkinson's disease, it is necessary to find reliable and accurate methods of early diagnosis. The aim of this work is to explore the usefulness of nonlinear manifold learning methodologies to inform a new space in which imaging features relate to symptomatology of PD, allowing for the creation of longitudinal PD progression models. The proposed methodology, using Isomap as a manifold learning method, and Support Vector Machines (SVM) as machine learning models for classification and regression, yielded high performance in regression, binary classification, and multi-class classification, as in the case of the HY score.

Each of the dimensions of the nonlinear subspace found by Isomap can be related to relevant changes in the brain such as the concentration of dopamine transporters (DaT) in the striatum or DaT concentration at the putamen (as it is the case of dimension 1). This builds a machine learning model that, unlike

many in the literature, is fully interpretable by healthcare professionals, paving the way for a more informed diagnosis.

The latent space of Isomap and other manifold learning algorithms in DaTSCAN images is yet to be fully explored. Other decomposition methods based on self-supervised neural networks could be of help here, along with more complex regression models that account for differences in distribution and hierarchical models with covariates, paving the way for new interpretable Computer Aided Diagnosis systems to understand the diagnosis and progression of PD.

Acknowledgements. This work was supported by the MCIN/ AEI/10.13039/501100011033/ and FEDER “Una manera de hacer Europa” under the RTI2018-098913-B100 project, by the Consejería de Economía, Innovación, Ciencia y Empleo (Junta de Andalucía) and FEDER under CV20-45250, A-TIC-080-UGR18, B-TIC-586-UGR20 and P20-00525 projects. Work by F.J.M.M. is supported by the MCIN AEI IJC2019-038835-I ‘Juan de la Cierva - Incorporación’ fellowship.

References

1. Artan, Y., et al.: Prostate cancer localization with multispectral MRI using cost-sensitive support vector machines and conditional random fields. *IEEE Trans. Image Process.* **19**(9), 2444–2455 (2010)
2. Castillo-Barnes, D., Martínez-Murcia, F.J., Ramírez, J., Górriz, J., Salas-Gonzalez, D.: Expectation-maximization algorithm for finite mixture of α -stable distributions. *Neurocomputing* **413**, 210–216 (2020)
3. Cayton, L.: Algorithms for manifold learning. Univ. of California at San Diego Technical report 12(1–17), 1 (2005)
4. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J.: Liblinear: a library for large linear classification. *J. Mach. Learn. Res.* **9**, 1871–1874 (2008)
5. Górriz, J.M., et al.: Artificial intelligence within the interplay between natural and artificial computation: advances in data science, trends and applications. *Neurocomputing* **410**, 237–270 (2020). <https://doi.org/10.1016/j.neucom.2020.05.078>
6. Illán, I., Górriz, J., Ramírez, J., Segovia, F., Jiménez-Hoyuela, J., Ortega Lozano, S.: Automatic assistance to Parkinson’s disease diagnosis in datscan spect imaging. *Med. Phys.* **39**(10), 5971–5980 (2012)
7. Kubota, K.J., Chen, J.A., Little, M.A.: Machine learning for large-scale wearable sensor data in Parkinson’s disease: Concepts, promises, pitfalls, and futures. *Mov. Disord.* **31**(9), 1314–1326 (2016)
8. Martínez-Murcia, F.J., Ortiz, A., Górriz, J.M., Ramírez, J., Castillo-Barnes, D.: Studying the manifold structure of Alzheimer’s disease: a deep learning approach using convolutional autoencoders. *IEEE J. Biomed. Health Inform.* **24**(1), 17–26 (2019)
9. Murcia, F.J.M.: Statistical neuroimage modeling, processing and synthesis based on texture and component analysis: tackling the small sample size problem. Ph.D. thesis, Universidad de Granada (2017)
10. Pahwa, R., Lyons, K.E.: Early diagnosis of Parkinson’s disease: recommendations from diagnostic clinical guidelines. *Am. J. Manag. Care* **16**(4), 94–99 (2010)

11. Rojas, A., et al.: Application of empirical mode decomposition (EMD) on datscan spect images to explore Parkinson disease. *Expert Syst. Appl.* **40**(7), 2756–2766 (2013). <https://doi.org/10.1016/j.eswa.2012.11.017>, <https://www.sciencedirect.com/science/article/pii/S0957417412012274>
12. Sakar, B.E., et al.: Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE J. Biomed. Health Inform.* **17**(4), 828–834 (2013)
13. Stoeckel, J., Fung, G.: Svm feature selection for classification of spect images of alzheimer’s disease using spatial information. In: Fifth IEEE International Conference on Data Mining (ICDM 2005). IEEE (2005). 8-pp
14. Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. *Science* **290**(5500), 2319–2323 (2000)
15. Vos, T., et al.: Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *The lancet* **388**(10053), 1545–1602 (2016)
16. Zhan, A., et al.: Using smartphones and machine learning to quantify Parkinson disease severity: the mobile Parkinson disease score. *JAMA Neurol.* **75**(7), 876–880 (2018)
17. Zubal, I.G., Early, M., Yuan, O., Jennings, D., Marek, K., Seibyl, J.P.: Optimized, automated striatal uptake analysis applied to spect brain scans of Parkinson’s disease patients. *J. Nucl. Med.* **48**(6), 857–864 (2007)