International Journal of Neural Systems, Vol. 0, No. 0 (0000) 1–16

MORPHOLOGICAL CHARACTERIZATION OF FUNCTIONAL BRAIN IMAGING BY ISOSURFACE ANALYSIS IN PARKINSON'S DISEASE

DIEGO CASTILLO-BARNES[1], FRANCISCO J. MARTINEZ-MURCIA[2], ANDRES ORTIZ[2], DIEGO SALAS-GONZALEZ[1], JAVIER RAMÍREZ[1], JUAN M. GÓRRIZ[1]

[1]*Department of Signal Theory, Telematics and Communications, University of Granada, Periodista Daniel Saucedo Aranda, 18071,Granada, Spain*
E-mail: diegoc@ugr.es
http://sipba.ugr.es/

[2]*Department of Communications Engineering, University of Malaga, Bulevar Louis Pasteur 35, 29071, Malaga, Spain*

Finding new biomarkers to model Parkinson's Disease (PD) is a challenge not only to help discerning between Healthy Control (HC) subjects and patients with potential PD, but also as a way to measure quantitatively the loss of dopaminergic neurons mainly concentrated at substantia nigra. Within this context, the work presented here tries to provide a set of imaging features based on morphological characteristics extracted from $I^{[123]}$-Ioflupane SPECT scans to discern between HC and PD participants in a balanced set of 386 scans from Parkinson's Progression Markers Initiative (PPMI) database. These features, obtained from isosurfaces of each scan at different intensity levels, have been classified through the use of classical Machine Learning classifiers such as Support-Vector-Machines (SVM) or Naïve Bayesian and compared with the results obtained using a Multi-Layer Perceptron (MLP).

The proposed system, based on a Mann-Whitney-Wilcoxon U-Test for feature selection and the SVM approach, yielded a 97.04% balanced accuracy when the performance was evaluated using a 10-fold cross-validation. This proves the reliability of these biomarkers, especially those related to sphericity, center of mass, number of vertices, 2D-projected perimeter or the 2D-projected eccentricity; among others, but including both internal and external isosurfaces.

Keywords: Parkinson's Disease; Neuroimaging; Machine Learning; Isosurfaces; Parkinson's Progression Markers Initiative (PPMI); Single Photon Emission Computed Tomography (SPECT); Computer-Aided-Diagnosis (CAD); Supervised Learning.

1. Introduction

One of the most prevalent neurological disorders, led only by Alzheimer's Disease, is Parkinson's Disease (PD). This pathology is a progressive neurodegenerative disorder characterized by the loss of dopaminergic neurons in the *substantia nigra*.[1] Its most frequent symptoms are: bradykinesia, tremor, ridigity, lack of expressiveness and postural inestability among others.[2,3]

The development of Computer-Aided-Diagnosis (CAD) systems[4–8] based on the analysis of neurological Single Photon Emission Computed Tomography (SPECT) scans is becoming a potential support tool for clinical practice. Among the more popular radiotracers used for PD, the $I^{[123]}$-Ioflupane radioligand is the most frequently used.[9–11] This tracer has a high binding affinity for dopaminergic transporters (DATs) in the brain, so its use results in a grayscale image highly illuminated in those regions where its concentration is higher. As dopaminergic

neurons are mainly located at *substantia nigra*, this region appears to be much brighter than other areas and a quantitative measure of its spatial distribution can be obtained.[12]

It is well-known that morphological information obtained from structural (MRI) and functional (SPECT) scans may be indicative of PD progression:[13,14] when considering a HC subject, its *striatum* appears to be highlighted, symmetrical and with a defined c-sharped portion that corresponds with the caudate and a globular part forming the putamen; on the contrary, patients with PD show the *striatum* area not as highlighted and with a more rounded shape due to the blurring of low intensities.[15] In the current literature there are some works that have dealt with this behaviour from the point of view of 2D projections as carried out by Segovia et al.;[16] delimiting high intensity borders using contour lines as presented in;[17] or just adjusting a non-closed surface to a high intensity threshold level to extract some characteristics from it.[18] These three works have in common that all the morphological features extracted in them represent 2D projections of 3D intensity volumes delimited by a ROI or by the manual selection of some slices in $I^{[123]}$-Ioflupane SPECT scans. Within this context, the work presented here tries to perform a classification of HC participants and PD patients based on four fundamental rules that will simplify the preprocessing of the images: 1) to avoid the need of both intensity and non-affine spatial normalization steps; 2) to avoid the need of a Region Of Interest (ROI) to extract characteristics instead of using the full scan (or at least one of the hemispheres); 3) to make use of 3D volumes that accurately model the spatial distribution of $I^{[123]}$-Ioflupane radioligand uptake; and 4) to present a fully-automated solution. To this end, one approach proposed in recent years has been to model the $I^{[123]}$-Ioflupane SPECT intensity distribution using 3D isosurfaces[19,20] An isosurface can be defined as the three-dimensional analog of an isoline.[21] All the points that make up the surface have the same constant value (intensity). As the $I^{[123]}$-Ioflupane SPECT images capture the presence of dopamine which is mainly concentrated at the *striatum*, isosurfaces are expected to shape adequately this highly illuminated region. In this new scenario, it is necessary to determine an automated way to define the intensity thresholds for the isosurfaces without lowering our classification rates. Decision about which of them should be used for future works will not only depends on their maximum classification rates obtained but also on the potential advantages of each one.

In summary, for this work we have extracted a wide range of morphological features from a balanced dataset of 386 $I^{[123]}$-Ioflupane SPECT scans, classified them using different Machine Learning algorithms including a Multi-Layer Perceptron (MLP) classifier, and validated all results through the use of a 10-fold cross-validation schema. Final list of input features includes (among others): area/volume of each isosurface; its center of mass; 2D projections on the X, Y, Z planes; its sphericity value; its total number of vertices and faces; or the isoperimetric quotient computed for each polyhedron. Note that due to the bilateral nature of the disease, all experiments have been computed considering only: 1) left hemispheres (LH); 2) right hemispheres (RH); 3) the full brain (FULL) and 4) a LH+RH+FULL features list (ALL).

This work is in line with the research proposed in.[19,20] This is a full and improved version of the most important conclusions given in them but including a longer set of morphological features; optimizing the way intensity thresholds should be selected; comparing the classification results obtained using different classifiers and measuring their variability.

2. Materials & Methods

2.1. *Parkinson's Progression Markers Initiative (PPMI)*

Data used in the preparation of this article were obtained from the Parkinsons Progression Markers Initiative (PPMI) database (`www.ppmi-info.org/data`). For up-to-date information on the study, visit `www.ppmi-info.org`. PPMI - a public-private partnership - is funded by the Michael J. Fox Foundation for Parkinsons Research and funding partners, including all partners listed on `www.ppmi-info.org/fundingpartners`.

Informed consents to clinical testing and neuroimaging prior to participation of the PPMI cohort were obtained and approved by the institutional review boards (IRB) of all participating institutions. The PPMI obtained written informed consent from

all study participants before enrolled in the Initiative.

More information are available in `http://www.ppmi-info.org/wp-content/uploads/2014/06/PPMI-Amendment-8-Protocol.pdf`.

2.2. *Participants*

A total of 386 participants after exclusion criteria from PPMI database were selected for this work. This included Healthy Control subjects (HC) and patients diagnosed with Parkinson's Disease (PD).[22]

Participants' demographics have been summarised in Table 1. Note that, to limit the number of results from each subject corresponding with their total amount of visits, only the baseline (BL) visit of each participant has been included for this work. In case of PD participants, this also refers to *de novo* patients[a]. A summary of the UPDRS scale available in the PPMI dataset for PD participants in BL has been depicted in Figure 1. Since the UPDRS scale is sub-grouped into 4 different sets of tests (also called parts),[23] each box in Figure 1 corresponds to one part (first $1-4$ columns) or the average of all parts (last column).

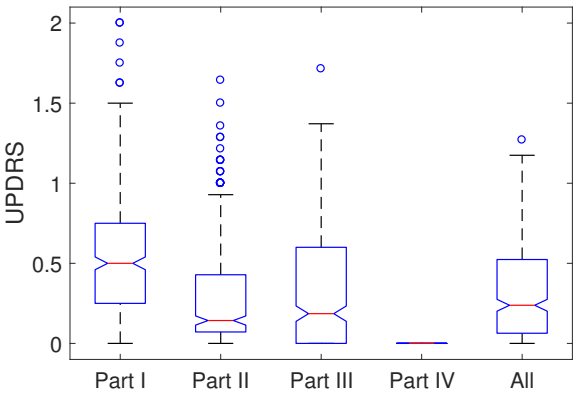

Figure 1. UPDRS scale for PD subjects at BL.

2.3. *Image preprocessing*

2.3.1. *Background and spatial normalization*

The template `Mask_ICV.nii` from the SPM (Statistical Parametric Mapping) software package tool (`SPM12` version available from its website `www.fil.ion.ucl.ac.uk/spm/software/spm12/`) has been spatially normalised to each subject using an affine registration.[24] This results in a mask with ones located at brain voxels and zeros in the rest. Then, each mask is multiplied by its correspondient $I^{[123]}$-Ioflupane scan.

2.3.2. *Generating isosurfaces*

Isosurfaces used in this work are generated using the Marching Cubes algorithm.[25,26] This procedure extracts a 2D surface mesh from a 3D volume. Conceptually, this conforms a 3D generalization of isolines. The version of the algorithm used in this work has been proposed by Lewiner and it is known as the Lewiner Marching Cubes Algorithm (LMCA). This is an improved version of the works presented by Lorensen et al.,[27] and Chernyaev.[28] Originally, this method was designed to generate triangle models of isosurfaces $F(x,y,z)=\alpha$ of a scalar function given by samples over a cuberille grid to modellate a particular volume (2^8 possible arrangements). However, as some of these cases are topologically equivalent (e.g. if considering rotations), the number of configurations can be reduced up to 15 arrangements for Lorensen's proposal and 33 for Chernyaev's. Although this may be sufficient for many scenarios, LMCA is designed for making a heavy use of lookup tables to handle the many different cases of intersections completing the ambiguity resolution of the previous implementations. An isosurface example obtained from the first participant has been depicted in Fig. 2. A second representation but using 8 different intensity levels to generate the isosurfaces map from the first participant has also been included as shown in Fig. 3.

[a]The current definition of *de novo* patients in the PPMI guidelines refers to either newly diagnosed patients with PD or patients not receiving L-dopa.

Table 1. Demographics.

	Number of participants			**Age**		
	Male	Female	Total	Male	Female	Both
HC	128	65	193	61.97 ± 10.72	58.84 ± 11.88	60.90 ± 11.22
PD	127	66	193	62.25 ± 9.92	60.34 ± 9.66	61.60 ± 9.87

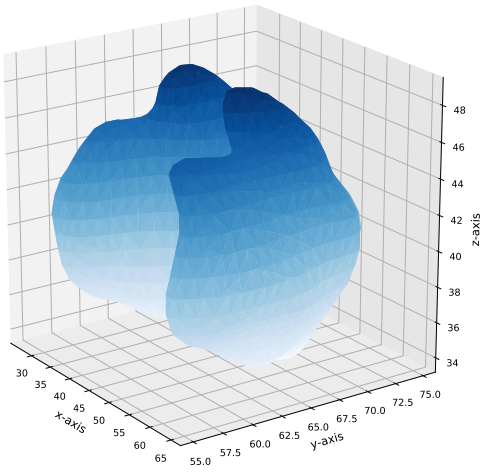

Figure 2. Example showing only the 16$^{\text{th}}$ isosurface (20 in total) obtained from the first participant.

2.3.3. *Obtaining morphological characteristics from SPECT volumes*

Once all the scans have been preprocessed, next step is to calculate the feature set based on the morphological characteristics of each isosurface from the I$^{[123]}$-Ioflupane SPECT volumes. The resulting patterns differ from the anatomical structures and can be used to measure the spatial distribution of the uptake in the brain for different intensity levels.

Given a reference intensity level (I_{ref}) from a range of N possible intensity levels corresponding with the number of isosurfaces[b], $I_{\text{ref}} \in [I_1, I_2, \ldots, I_N]$, the LMCA iterates across the volume, looking for regions which cross this value. If such regions are found, triangulations are generated and added to an output mesh. Then, using this mesh, several features from the delimited volume can be obtained including:

- **Area and Volume** - Of the region within the isosurface at the intensity reference level I_{ref}. For that, `Qhull` library based on the work presented in[29] and available at `www.qhull.org` was used. Documentation about Qhull and applications of this tool (e.g. Delaunay triangulation, intersections of halfspaces or Voronoi diagrams) are also available at this website.
- **Center of Mass (CoM)** - Given a particles system, CoM represents the spatial point where the weighted relative position of the distributed mass sums to zero. Although the feature name has not been changed, this work makes use of intensity values instead of mass. This procedure has been computed for each isosurface using only voxels whose intensity is equal or higher than the reference intensity of the reference isosurface: $I_{\text{Voxel}} \geq I_{\text{ref}}$. Mathematically, Eq. (1) computes the CoM of all M voxels within isosurface I_{Ref} with relative position (x_i, y_i, z_i) and distributed intensities $I(x_i, y_i, z_i)$.

$$\text{CoM}_{I_{\text{Ref}}} = \frac{\sum_{i=1}^{M} (x_i, y_i, z_i) * I(x_i, y_i, z_i)}{\sum_{i=1}^{M} I(x_i, y_i, z_i)} \qquad (1)$$
$$i = 1, 2, \ldots, M$$

- **Projections** - Previous works, like presented in,[16,18] have made use of axial, sagital and coronal projections of the *striatum* region as an intermediate step to obtain reliable features for Parkinson's Disease. The list of these characteristics included: area of the projection, eccentricity, orientation and major/minor axis length of the ellipse that has the same normalised second central moments as the region. Using these ideas, the axial, sagital and coronal projections of each isosurface have been computed to get some features of the projection

[b]The way those intensities thresholds are choosen will be detailed in Section 2.3.4.

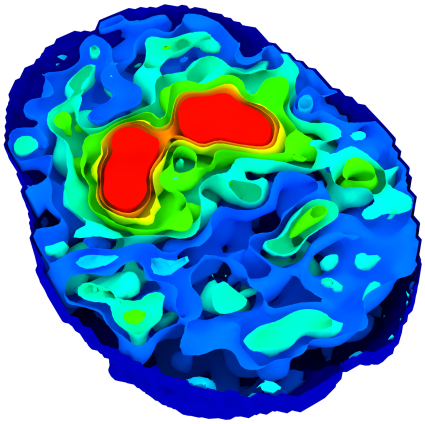

Figure 3. Example showing 8 isosurfaces obtained from the first participant. High intensity parts (red) corresponds with a high uptake of $I^{[123]}$-Ioflupane radiotracer.

including: area, centroid, eccentricity, equivalent diameter, extent, major and minor axis length of the equivalent ellipse, orientation, perimeter and solidity. A representation of the base planes projections (X, Y, Z) of an isosurface computed for a participant is depicted in Fig. 4.

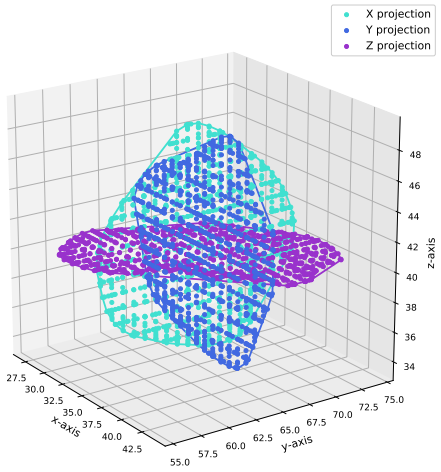

Figure 4. Left 3D-projections of the 16^{th} isosurface obtained from the first participant. Points represent the vertices projections on the base planes X,Y,Z through the point defined by the CoM.

- **Sphericity** - Apart from the low intensities obtained in scans of PD patients, the shape of their more highlighted regions also seems to be round. In regard of this, it has also been included the sphericity of each isosurface as a potential feature for Parkinson's Disease diagnosis.
- **Isoperimetric Quotient** - The isoperimetric quotient of a closed curve is defined as the ratio (Q^{ISO}) of the curve area to the area of a circle (A_r) with same perimeter as the curve (p_r):

$$Q_{2D}^{\text{ISO}} = \frac{4\pi A_r}{p_r^2} \tag{2}$$

In case of considering a 3D polyhedron, it can be used both volume (V_r) and surface (S_r) of an sphere as reference instead of the area of a circle and its perimeter. Thus, this quotient remains as follows:

$$Q_{3D}^{\text{ISO}} = \frac{36\pi V_r^2}{S_r^3} \tag{3}$$

- **Number of elements** - When considering healthy subjects, the region with a high DATs concentration is expected to be large and with a defined c-shape. In this case, LMCA will provide a number of vertices and faces greater than the obtained for patients with PD.
- **Other polyhedron features** - Including the Euler characteristic number and the polyhedron orientability. In the case of the Euler characteristic number, χ, this parameter relates the number of vertices V_p, edges E_p, and faces F_p of the polyhedron:

$$\chi = V_p - E_p + F_p \tag{4}$$

Therefore, when concatenating all the characteristics considered for each isosurface, its total number of features is 15 if considering the full brain and 57 when it also includes the information from 2D projections (hemispheres).

2.3.4. *Generating intensity thresholds for each isosurface*

It remains to select the intensity values for each isosurface. For that, this work proposes two alternatives:

(i) Given an image from a subject, a list of N intensity reference values $I_{\mathrm{Ref}j}$ are generated following Eq. (5) where $I_{\mathrm{min}}^{\mathrm{Subject}}$ and $I_{\mathrm{max}}^{\mathrm{Subject}}$ are the respective minimum and maximum intensity levels from each subject.

$$I_{\mathrm{Ref}j} = I_{\mathrm{min}}^{\mathrm{Subject}} + j\frac{I_{\mathrm{max}}^{\mathrm{Subject}} - I_{\mathrm{min}}^{\mathrm{Subject}}}{N} \qquad (5)$$

(ii) Given a reference template, use its minimum/maximum intensity values to obtain a set of $I_{\mathrm{Ref}j}$ levels common to all the subjects. This second alternative requires not only to generate this template but also to decide how maximum/minimum intensity levels from each patient should match with the maximum/minimum intensity levels from the template. Once the template is generated by averaging all the images from a set of HC participants (it will require to apply a non-affine spatial registration procedure followed by an intensity normalization), the difference between its maximum/minimum intensity levels is splited in N parts. Then, each level is used as intensity reference value for each subject image.

Fig. 5 shows an schema of these two approaches. Note that, to avoid excesive highlighted voxels (outliers), maximums can be computed by averaging the 1% of the most intensity values. On the contrary, minimums are always referred to 0.

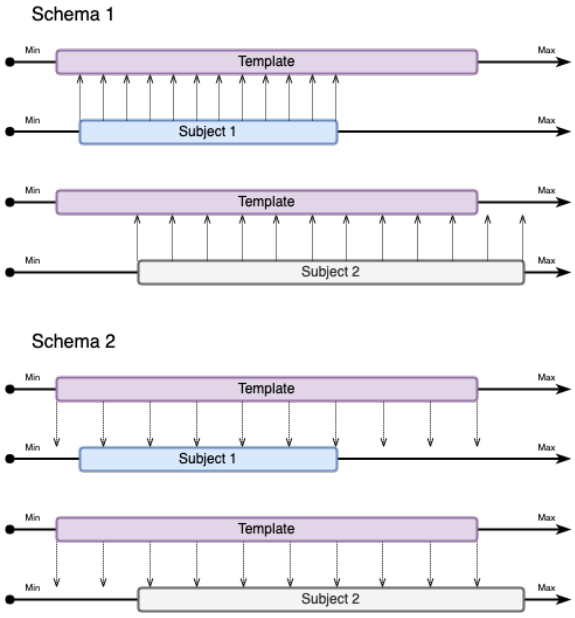

Figure 5. Graphical representation of the two proposed alternatives to determine the isosurfaces intensity thresholds $I_{\mathrm{Ref}j}$.

This allows us to define two experiments, each of them corresponding to the two schemas listed above. It also should be highlighted the importance of using a fixed number of isosurfaces, N, to avoid problems with the dimensionality of the feature vectors. Thus, if a particular isosurface does not fit with any intensity threshold, its feature vector should be filled at least with zeros indicating the no-existence of that layer.

2.4. *Classification algorithms*

For this work, several classification approaches have been used not only as a way to compare their performance when using the morphological features extracted from the I$^{[123]}$-Ioflupane SPECT volumes but also to test the advantages of a multilayer neural network classification over traditional methods.

2.4.1. *Support-Vector-Machines (SVM)*

SVM classifiers are based on linear discriminant functions that define a decision hyperplane in a multidimensional feature space maximizing the separation between different labeled data classes.[30,31]

Given a training set of instance-label pairs $\mathbf{x}_i, y_i$ with $i = 1, 2, \ldots, n$, $\mathbf{x}_i \in \Re^m$ and $y \in \{1, -1\}^n$, a

SVM classifier solves the optimization problem given by Eq. (6) subject to Eq. (7) where $C > 0$ is a penalty factor, and ξ represents the loss function:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i \tag{6}$$

$$y_i\big(\mathbf{w}^T\phi(\mathbf{x}_i) + b\big) \geq 1 - \xi_i \qquad \xi_i \geq 0 \tag{7}$$

In this scenario, the training vectors $\mathbf{x}_i$ are mapped into a higher (maybe infinite) dimensional space by the function ϕ. The form $\phi(\mathbf{x}_i)^T\phi(\mathbf{x}_i)$ is also called the kernel function ($K(\mathbf{x}_i, \mathbf{x}_j)$). According to the kernel methods, for this work we have made use of three commonly used kernels in the extant literature:[4–8]

- **Linear Kernel** - It uses the simplest kernel function, $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i\mathbf{x}_j + C$, where the constant C trades off misclassification of training examples against simplicity of the decision surface. Low values for C smooth the decision surface whereas high values force the classification of all training examples correctly.
- **Radial Basis Function (RBF) kernel** - It also presents the parameter γ which defines the radius of influence of a single training example. Thus, the larger γ value is, the closer other examples must be to be affected. In this case, $K(\mathbf{x}_i, \mathbf{x}_j) = \exp\big(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\big)$.
- **Polynomial kernel** - Where $K(\mathbf{x}_i, \mathbf{x}_j) = \big(\alpha\mathbf{x}_i^T\mathbf{x}_j + C\big)^d$ which makes use of an slope modifier parameter, α, and the polynomial degree d.

In all cases, to tune the parameters, all possible combinations for both $C \in \{0.1, 1, 10, 100\}$ and $\gamma \in \{0.1, 1, 10\}$ have been considered.[32] With respect the polynomial kernel, its degree was fixed to 3.

Related, but not the same, it was also included the called **NuSVC** classifier[33,34] which controls the fraction of training errors and support vectors by its $\nu \in \big(0, 1\big]$ parameter. In our context, its value was finally set to $\nu = 0.1$.

2.4.2. *Naïve Bayesian*

This non-linear classifier is known by its good performance when using a small set of features.[35] In the binary case, it is based on the assumption that the data matrix $\mathbf{X} = \mathbf{x}_0, \mathbf{x}_1 \ldots \mathbf{x}_n$ contains n samples drawn from two different populations with random normal distribution π_c with $c \in [1, 2]$. Using the notation $f_1(\mathbf{x}_i)$ for the probability densities of any of the samples belonging to class 1 or 2 with i the index of any subject of class c, this classifier labels each subject $\mathbf{x}_i$ as label c using the likelihood ratio rule, stating that:

$$\mathbf{x}_i \in \begin{cases} \pi_1 & \text{if } f_1(\mathbf{x}_i)/f_2(\mathbf{x}_i) > 1 \\ \pi_2 & \text{if } f_1(\mathbf{x}_i)/f_2(\mathbf{x}_i) \leq 1 \end{cases} \tag{8}$$

When considering that π_c is a multivariate normal population with mean $\boldsymbol{\mu}_c$ and dispersion matrix $\boldsymbol{\Sigma}_c$ for $c \in [1, 2]$, the relation $\frac{f_1(\mathbf{x})}{f_2(\mathbf{x})}$ can be obtained as follows:

$$\begin{gathered} \frac{f_1(\mathbf{x})}{f_2(\mathbf{x})} = \sqrt{\frac{|\boldsymbol{\Sigma}_2|}{|\boldsymbol{\Sigma}_1|}}\exp\big(-\frac{1}{2}(T_1(\mathbf{x}) + T_2(\mathbf{x}) + T_3(\mathbf{x}))\big) \\ T_1(\mathbf{x}) = \mathbf{x}^{\mathbf{T}}(\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1})\mathbf{x} \\ T_2(\mathbf{x}) = -2\mathbf{x}^T(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2) \\ T_3(\mathbf{x}) = \boldsymbol{\mu}_1^T\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2^T\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 \end{gathered} \tag{9}$$

This can be used to define the discrimination function as described in Eq. (10). Since the true parameters $\mu_\mathbf{1}$, $\mu_\mathbf{2}$, $\boldsymbol{\Sigma}_\mathbf{1}$ and $\boldsymbol{\Sigma}_\mathbf{2}$ are not known in reality, they are estimated making use of the maximum likelihood algorithm. Owing to this, the implementation of this classifier did not require any tune.

$$f_c(\mathbf{x}) = \frac{\exp\big(\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_c)^T\boldsymbol{\Sigma}_c^{-1}(\mathbf{x} - \boldsymbol{\mu}_c)\big)}{\sqrt{(2\pi)^p|\boldsymbol{\Sigma}_c|}} \tag{10}$$

2.4.3. *Multi-Layer Perceptron (MLP)*

The last classifier considered for this work was a simple, yet powerful, Multi-Layer Perceptron (MLP) which makes use of three layers of size $L_{1(\text{size})} = 32$, $L_{2(\text{size})} = 32$ and $L_{3(\text{size})} = 2$. This number of neurons has been set to achieve a good trade-off between the number of trainable parameters and the overall accuracy of the model. The general schema of the network is depicted in Fig. 6. Batch normalization and Rectifier Linear Unit (ReLU) activation was used after layers L_1 and L_2, as suggested in.[36] For the output layer, log-softmax activation[37] was used over dimension 1. Dropout[38] was applied during training with a $p = 0.25$ between layers L_1 and L_2, and layers L_2 and L_3 (output layer).

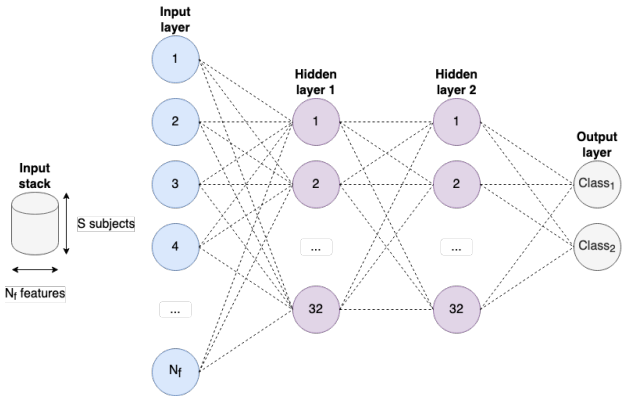

Figure 6. MLP classifier: general schema used for this work. N_f corresponds with the number of input features for each classification.

During the training, it was made use of the Stochastic Gradient Descent (SGD) optimizer, with a fixed learning rate of $\mu_r = 0.1$, providing both a good convergence rate and accuracy. The Negative Log-likelihood loss was computed with a batch size equal to N_b:

$$\begin{aligned} \ell(x, y) = L = \{l_1, \ldots, l_{N_\mathrm{b}}\}^T \\ l_n = -w_{y_n} x_{n,y_n} \\ w_c = \mathrm{weight}[c] \cdot \mathbb{1}\{c \neq \mathrm{ignore\_index}\} \end{aligned} \tag{11}$$

A further mean reduction was applied as general loss for the batch. This resulted in (12) with a batch size set to 16.

$$\ell(x, y) = \sum_{n=1}^{N_\mathrm{b}} \frac{1}{\sum_{n=1}^{N_\mathrm{b}} w_{y_n}} l_n \tag{12}$$

The early stopping technique was applied using an independent validation set. This saved the model which achieved last descending validation loss over the last 30 iterations.

2.5. *Cross-validation strategy*

Classification results have been validated using a stratified K-fold cross-validation strategy schema with $K = 10$ as recommended by Kohavi[39] and Krstajic et al.[40]

Resampling made by the cross-validation ensures that the ratio between classes is preserved in each iteration (fold). This let us to evaluate the generalization capabilities of each generated model maintaining the independence of the results. Moreover, this strategy guarantees that high performance rates do not come from a random selection of input data (avoid bias). Note that the estimation of the generalization error will result in an overstimate of the true prediction error since only using just $K - 1$ folds. This difference in the estimation depends on the learning curve of the classifier and it is reduced as long as K increases.[41]

2.6. *General diagram*

A general diagram including all steps has been depicted in Fig. 7. Note that as Experiment 1 does not require any template to determine the intensity thresholds used to define each isosurface, this step has been removed. In the case of using a template, as necessary for Experiment 2, it can be generated within the CV loop by averaging the training scans of all HC subjects as explained in.[42]

This work was developed using `Python 3.7`. Code for the marching cubes algorithm, isosurfaces and morphological characteristics (Sections 2.3.2-2.3.4) were specifically coded for this project. SVM (Section 2.4.1) and Naïve Bayesian (Section 2.4.2) used the `scikit-learn` library,[43] which in case of SVM is a wrapper of `LIBSVM`.[44] The MLP model (Section 2.4.3) was built in `PyTorch`.[45]

3. Results

Once experiments have been defined as explained in section 2.3.4, the features proposed for this work were extracted from each $I^{[123]}$-Ioflupane SPECT scan and classified through multiple classification schemas variating the number of isosurfaces, N. Following the general diagram depicted in Fig. 7 for experiments 1 and 2, final classification results varying the number of isosurfaces have been calculated. As experiment 1 results were much better than those obtained for experiment 2, we have only included Fig. 8 to show these rates. Moreover, as mentioned in Section 1, owing to the bilateral nature of the disease all the experiments have been performed considering morphological characteristics extracted from left/right hemispheres (LH/RH); the full brain (FULL); and the combination of those three subsets (ALL). Furthermore, regardless classifying only

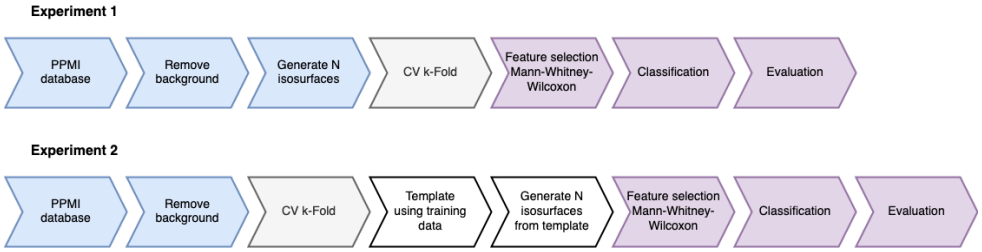

Figure 7. General diagram.

an hemisphere or using the entire brain, as the amount of input features for their classification might be excesive in comparison with the number of input subjects. To identify which features present the major margin separbility between classes, a Mann-Whitney-Wilcoxon U-Test feature selection step was introduced in the cross-validation loop.[46] This non-parametric test of the null hypothesis does not assume that the input distributions follow the normal distribution. Making use of this test, and setting $p_{\text{value}} \leq 0.05$ as threshold, if the amount of features matching this condition did not exceed the number of subjects, all the features are used for feature selection and overfitting is prevented. On the contrary, if this amount is greater than the number of subjects, ns_{train}, the array containing all p_{value} is sorted in increasing order and only the first ns_{train} characteristics are selected so overfitting is also prevented. For example, when $N = 20$ and considering ALL set, a total of $N * (15 + 57 + 57) = 2580$ features are generated. Although most of them presented a $p_{\text{value}} \leq 0.05$ (about 2442 features in both experiments), only ns_{train} of them were finally used for their classification. Final list of the most significative input features included among others the list in Table 2 sorted in increasing order of their p_{value}. The reason why a dimensionality reduction method like PCA was not used is to maintain the physical interpretation of each isosurface feature.

Maximum classification results are obtained when generating the isosurfaces by using the first schema proposed in Section 2.3.4 and classifying using the NuSVC classifier. In fact, for $N_{\text{ISO}} = 20$ isosurfaces, maximum rates reached up to 97.04%. Table 3 and Table 4 have been included to summarize the balanced accuracies obtained for experiment 1 and experiment 2 respectively.

Table 2. Top 10 features with better p_{value} obtained for experiment 1 when using MWW U-Test and $N = 20$ isosurfaces.

Feature name	**Isosurface**	**Region**
2D perimeter (Z)	10	Left hemisphere
2D Eccentricity (Y)	4	Right hemisphere
Sphericity	12	Left hemisphere
Center of Mass	12	Full brain
Num of vertices	13	Right hemisphere
2D Centroid (Y)	3	Right hemisphere
Num of faces	13	Left hemisphere
Volume	13	Left hemisphere
Isoperimetric Q	5	Full brain
Sphericity	13	Right hemisphere

Table 3. Balanced accuracies obtained for experiment 1 when using a total of $N = 20$ isosurfaces.

Classifier	**LEFT**	**RIGHT**	**FULL**	**ALL**
SVM (linear)	90.59%	89.98%	90.91%	92.83%
SVM (rbf)	88.56%	88.42%	88.19%	88.70%
SVM (poly)	90.26%	91.64%	90.72%	93.54%
NuSVC	93.26%	93.75%	94.46%	**97.04**%
Naïve	83.76%	82.01%	84.58%	77.60%
MLP	85.07%	84.96%	83.42%	86.71%

Table 4. Balanced accuracies obtained for experiment 2 when using a total of $N = 20$ isosurfaces.

Classifier	**LEFT**	**RIGHT**	**FULL**	**ALL**
SVM (linear)	56.65%	59.69%	67.32%	59.00%
SVM (rbf)	60.75%	57.52%	58.06%	55.11%
SVM (poly)	64.62%	58.50%	65.50%	60.36%
NuSVC	64.28%	64.79%	64.67%	**65.69**%
Naïve	64.60%	60.19%	61.96%	59.43%
MLP	56.68%	53.03%	57.77%	56.78%

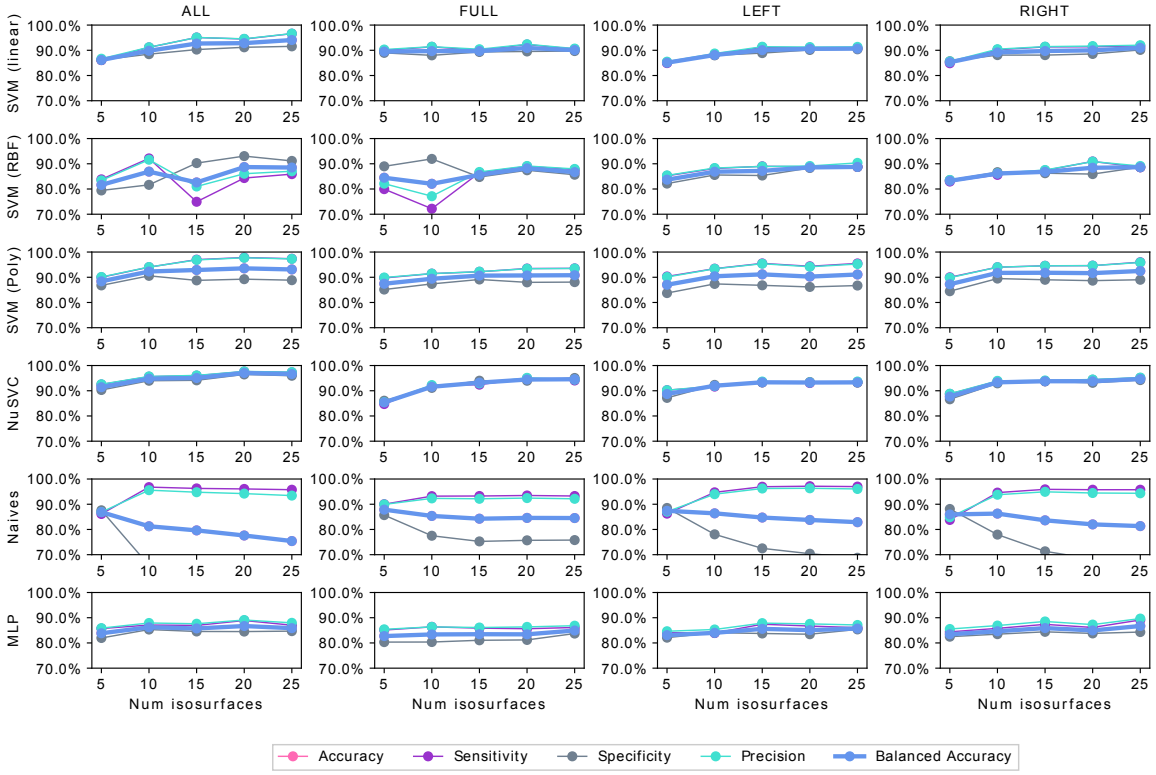

Figure 8. Classification results obtained for all the experiments using different classifiers and following the Schema 1 proposed in Section 2.3.4.

The Receiver Operating Curves (ROC) of each experiment have also been included as depicted in Fig. 9. These representations also include the Area Under the Curve (AUC) obtained for both NuSVC classifier and the MLP for each experiment. Focusing on them, best results are obtained for the NuSVC classifier where AUC parameter for ALL labeled features were greater than in the others features stack ($AUC_{ALL} = 0.9844$). In addition to this, it can been observed clear differences between left and right hemispheres: $AUC_{LEFT} = 0.9792$ and $AUC_{RIGHT} = 0.9718$. On the contrary, the MLP classification resulted in maximum AUC values equal to $AUC_{ALL} = 0.9128$. This is a difference of 0.072 from using the NuSVC.

All the experiments have been repeated 10 times with different randomization in each repetition to determine the variability due to the change in the training/testing data sets. Fig. 10 shows this effect when $N = 20$.

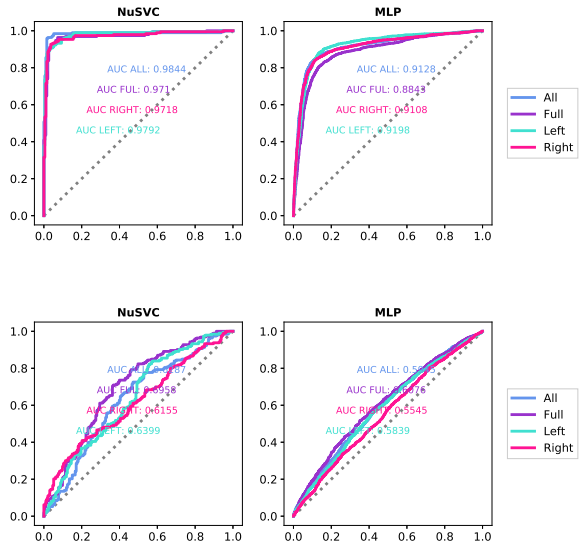

Figure 9. Receiver Operating Curves (ROC) obtained for experiments 1 (top) and 2 (down) using $N = 20$ isosurfaces. Comparison using NuSVC and MLP classifiers. Area Under the Curve (AUC) results have also been included.

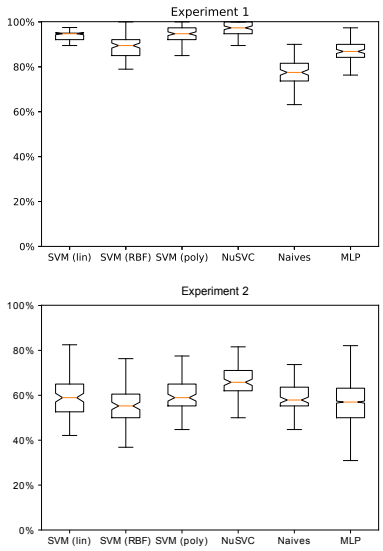

Figure 10. Boxplot representation showing the variability obtained for experiments 1 and 2 when classifying using $N = 20$ isosurfaces and ALL labeled features.

4. Discussion

In comparison with the current literature, there are already some works that have measured common differences in shape between HC subjects and patients with PD.[16–18,47] Although most of them have not made use of 3D volumes, they required the use of templates to define a ROI, or used a non-affine spatial normalization preprocessing step; they have aided to improve the knowledge of the disease. Whitin this context, this work was proposed to overcome these limitations by creating a full-automated volumetric approach without losing accuracy in our final results. As included in Table 5 where it was summarised the classification results of these referred works, performance rates obtained by the use of isosurfaces not only has met the target set out, but also they have been exceeded.

The first issue in this work was to determine how to select the intensity thresholds associated to each isosurface. For that purpose, the schemas proposed in Section 2.3.4 should be compared and analysed in terms of their classification rates and their potential advantages/disadvantages. For that, following the first schema (experiment 1), since its intensity thresholds are obtained by dividing the difference between the maximum and the minimum intensity values of the scan, the first schema always yields a fixed number of isosurfaces with non-empty values. This advantage is coupled with the fact that this proposal does require neither a reference template to set up the thresholds nor an intensity procedure which reduces considerably the computational load. On the contrary, the second schema let us compare the same intensities for all subjects. However, this comes at the cost of using a template, which also involves performing not only an intensity normalization preprocessing but also an accurate spatial normalization that allows to math each region of the brain with its corresponding region in the template. At this point, the key is to determine the influence of each approach on the final classification.

If comparing results from experiments 1 and 2 (Fig. 8 and summarised results in Table 3 and Table 4), classification rates using different ML algorithms were good and similar to each other in the case of experiment 1 whereas experiment 2 presented a poor averaged correct rate barely above 67.32% as its best. These results can be explained as the second approach not always finds a match between the template and a subject's scan. Moreover, if the intensity ranges from the template and subject differ in excess, it could happen that the last isosuperface (the one with the highest threshold intensity value) will be too large to model all those voxels with the most information.[12]

Once the first scheme for the intensity threshold was selected, the next step was to determine the number of isosurfaces needed to optimize our classification results. An easy solution for that was to perform a grid search on this number looking at the classification results. As shown in Fig. 8, this max is reached when using $N = 20$ isosurfaces for most of the input regions considered. With fixed scan size of $91 \times 109 \times 91$ voxels and $2 \times 2 \times 2$ mm of voxel-size, using $N = 20$ isosurfaces, this results in outer isosurfaces separated about 10 mm and many other much closer with separation margins about few millimeters. Thus, the use of $N = 20$ isosurfaces also seems to be reasonable in terms of the physical separation between layers.

In relation to the most significant features ($p_{\text{value}} \leq 0.05$) listed at Table 2, what is most noteworthy is the fact that the intermediate isosurfaces will be those with the highest statistical interclass separation in terms of the Mann-Whitney-Wilcoxon U-Test. In fact, when considering $N = 20$, among the most highlighted characteristics, we found fea-

Table 5. Summary of results obtained in previous related works.

Classification	Sample Size	Morphological features	Accuracy	Author	Reference
SVM_{linear} 10 Fold CV	715	Measures obtained from some manually selected slices plus a 3rd degree surface fitting	97.2%	Prashanth et al. (2017)	18
SVM_{linear} LOO CV	388	Measures obtained from 2D projections of the *striatum* area using a ROI	92.3%	Castillo-Barnes et al. (2018)	47
SVM_{linear} 10 Fold CV	147	Measures obtained from a set of contour lines in 4 slices centered at *striatum*	92.5%	Molina et al. (2018)	17
SVM_{linear} 10 Fold CV	189	Measures obtained from 2D projections of the most intense voxels ($0.8I_{max}$)	94.2%	Segovia et al. (2019)	16

tures from layers $12 - 14$ which correspond to the borders of the most intense voxels. This result coincides with that observed in[48] when using the gradient magnitude for a set of I[123]-Ioflupane SPECT images.

To quantitatively demonstrate the advantages of the isosurface methodology, it has been compared to well known methodologies in the literature. This is the case of the Voxels As Features (VAF) approach, which directly uses voxels intensities as features, or the composite features obtained using Principal Component Analysis (PCA).[49] For that, two pipelines similar to Fig 7 were defined to evaluate both approaches:

- VAF - It includes a non-affine spatial normalization; an intensity normalization to the maximum intensity values and a threshold to delimit the boundaries of the ROI. To obtain this region, the training set is averaged following the work presented in[42] and the threshold is set up to 50% of the maximum template intensity[20] as shown in Fig 11.
- PCA - Apart from the non-affine spatial normalization and the intensity normalization to the maximum intensity values, within the cross-validation loop, the PCA algorithm is calculated for a number of components, *nComp* for their posterior evaluation.

Both methods use an SVM classifier with Linear Kernel classification. Evaluation of the stratified 10-fold cross-validation schema with 10 repetitions yieled maximum balanced accuracy rates equal to 85.67% for VAF and 88.89% for PCA when using $nComp = 24$ (PCA results in function of *nComp* are depicted in Fig 12).

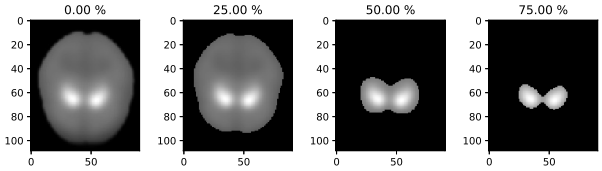

Figure 11. Regions Of Interest (ROIs) generated from the $(0\%, 25\%, 50\%, 75\%)$ percentages of the maximum intensity value of the template.

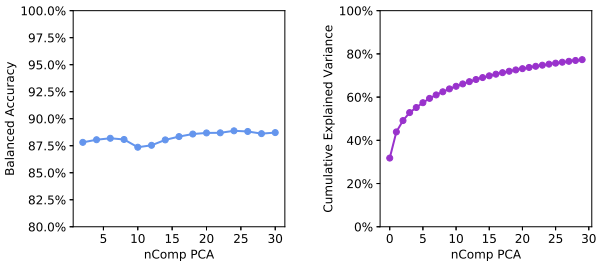

Figure 12. (left) Balanced accuracy obtained when applying PCA to the preprocessed intensity values from $I^{[123]}$-Ioflupane SPECT scans instead of using isosurfaces. (right) Cummulative explained variance. All results are referred to the SVM with Linear Kernel classification.

For this work, significant differences in classification rates were found when comparing isolated LEFT and RIGHT hemispheres. When taking as reference the best classification scenario with $N = 20$ isosurfaces and using a SVM classifier with Linear Kernel, classification results considering only left hemispheres were aproximately 0.6% better than using right hemispheres. In the case of the NuSVC classifier, this difference was similar (only using the right

hemispheres gives an improvement of 0.5% with respect only using left hemispheres). Although this result is noteworthy, it is well-known that Parkinson's Disease do not always presents a symmetric behaviour[50,51] so these differences were expected.

The classification performance, even in the case of SVM with linear kernel, has proven the high reliability of using isosurfaces instead of classical VAF analysis or PCA. This approach reduces strongly the number of input features but maintaining all its physical interpretation. Differences between the SVM and NuSVC lies in the selection of the NuSVC hyperparameter $\nu = 0.1$ which penalises incorrect classifications. As this term can been rewritten in terms of the C parameter from SVM, it can been stated that both schemas are equivalent.

In the last few years, the use of Neural Networks (NN) with special emphasis on Deep Learning (DL)[52–54] has become very popular in neuroscience.[55–60] For example, a recent publication consisted of solid volumes obtained using a single threshold and directly passed to a Convolutional Neural Network (CNN).[20] In that scenario, the features were computed using the NN but without including neither a physical interpretation nor analysing how this threshold should be calculated but only the use of large amounts of volumetric intensities which could be increasing the bias of the solution proposed. In comparison with our MLP, as no more features than the available number of subjects were selected, this problem has been to a large extent resolved.

5. Conclusions

In summary, this work proposes an advanced and automatic method to obtain reliable morphological features from $I^{[123]}$-Ioflupane SPECT scans making use of isosurfaces obtained from each scan and the posterior processing of their characteristics including, among others: volume, area, sphericity, centers of mass, or the total number of vertices/faces from each isosurface. This has let us to classify a dataset of 386 subjects from PPMI database with a balanced accuracy of 97.04% following an intensity threshold selection schema with advantages such as: 1) no need of applying a non-affine spatial normalization procedure; 2) no need of an intensity normalization of the $I^{[123]}$-Ioflupane SPECT images; 3) being interpretable from the point of view of the physiology of the brain. In this sense, the main limitation of our proposal could be the computational costs associated to the calculation of each isosurface. It is expected that conclusions obtained here will be used in future publications helping to interpret the longitudinal evolution of these morphological patterns in potential patients with Parkinson's Disease and will help the community in their finding of new markers for PD,[61] and its mathematical model.[62]

Acknowledgments

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.
This work was supported by the MINECO/FEDER under the RTI2018-098913-B-I00 and PGC2018-098813-B-C32 projects and the General Secretariat of Universities, Research and Technology, Junta de Andalucía under the Excellence FEDER Project A-TIC-117-UGR18.

Data Availability Statement

The dataset analysed for this study can be found in the Parkinsons Progression Markers Initiative (PPMI) database available at the website `www.ppmi-info.org/data`.

References

Bibliography

1. A. S. Harms, C. J. Barnum, K. A. Ruhn, S. Varghese, I. Trevino, A. Blesch and M. G. Tansey, Delayed dominant-negative TNF gene therapy halts progressive loss of nigral dopaminergic neurons in a rat model of parkinson's disease, *Molecular Therapy* **19**(1) (2011) 46–52.
2. J. Jankovic, Parkinson's disease: clinical features and diagnosis, *Journal of Neurology, Neurosurgery & Psychiatry* **79**(4) (2008) 368–376.
3. S. Bhat, U. R. Acharya, Y. Hagiwara, N. Dadmehr and H. Adeli, Parkinson's disease: Cause factors, measurable indicators, and early diagnosis, *Computers in Biology and Medicine* **102** (2018) 234–241.
4. I. A. Illán, J. M. Górriz, J. Ramírez, F. Segovia, J. M. Jiménez-Hoyuela and S. J. Ortega Lozano, Automatic assistance to Parkinson's disease diagnosis in DaTSCAN SPECT imaging, *Medical Physics* **39**(10) (2012) 5971–5980.
5. F. Martinez-Murcia, J. Górriz, J. Ramírez, I. Illán and A. Ortiz, Automatic detection of Parkinsonism using significance measures and component analysis

in DaTSCAN imaging, *Neurocomputing* **126**(Supplement C) (2014) 58–70.
6. A. Augimeri, A. Cherubini, G. L. Cascini, D. Galea, M. E. Caligiuri, G. Barbagallo, G. Arabia and A. Quattrone, CADA–computer-aided DaTSCAN analysis, *EJNMMI Physics* **3**(1) (2016) p. 4.
7. S. Badoud, D. V. D. Ville, N. Nicastro, V. Garibotto, P. R. Burkhard and S. Haller, Discriminating among degenerative parkinsonisms using advanced 123I-ioflupane SPECT analyses, *NeuroImage: Clinical* **12**(Supplement C) (2016) 234–240.
8. N. Nicastro, J. Wegrzyk, M. G. Preti, V. Fleury, D. Van de Ville, V. Garibotto and P. R. Burkhard, Classification of degenerative parkinsonism subtypes by support-vector-machine analysis and striatal 123I-FP-CIT indices, *Journal of Neurology* (2019).
9. F. Sixel-Döring, K. Liepe, B. Mollenhauer, E. Trautmann and C. Trenkwalder, The role of 123I-FP-CIT-SPECT in the differential diagnosis of Parkinson and tremor syndromes: a critical assessment of 125 cases, *Journal of Neurology* **258**(12) (2011) 2147–2154.
10. R. A. Hauser and D. G. Grosset, [123I]FP-CIT (DaTscan) SPECT Brain Imaging in Patients with Suspected Parkinsonian Syndromes, *Journal of Neuroimaging* **22**(3) (2011) 225–230.
11. T. Booth, M. Nathan, A. Waldman, A.-M. Quigley, A. Schapira and J. Buscombe, The Role of Functional Dopamine-Transporter SPECT Imaging in Parkinsonian Syndromes, Part 2, *American Journal of Neuroradiology* **36**(2) (2015) 236–244.
12. K. L. Marek, R. B. Innis, C. H. Van Dyck, B. Fussell, M. Y. Early, S. W. Eberly, D. Oakes and J. P. Seibyl, [123I] -CIT SPECT imaging assessment of the rate of Parkinson's disease progression, *Neurology* **57**(Issue 11) (2001) 2089–2094.
13. A. Garg, S. Appel-Cresswell, K. Popuri, M. J. McKeown and M. F. Beg, Morphological alterations in the caudate, putamen, pallidum, and thalamus in Parkinson's disease, *Frontiers in Neuroscience* **9** (2015) p. 101.
14. C. Owens-Walton, D. Jakabek, X. Li, F. A. Wilkes, M. Walterfang, D. Velakoulis, D. van Westen, J. C. Looi and O. Hansson, Striatal changes in Parkinson disease: An investigation of morphology, functional connectivity and their relationship to clinical symptoms, *Psychiatry Research: Neuroimaging* **275** (2018) 5–13.
15. J. Dukart, F. Sambataro and A. Bertolino, Distinct Role of Striatal Functional Connectivity and Dopaminergic Loss in Parkinson's Symptoms, *Frontiers in Aging Neuroscience* **9** (2017) p. 151.
16. F. Segovia, J. M. Górriz, J. Ramírez, F. J. Martinez-Murcia and D. Castillo-Barnes, Assisted diagnosis of Parkinsonism based on the striatal morphology, *International Journal of Neural Systems* (2019).
17. M. Molina-Molina, D. Castillo-Barnes, F. J. Martinez-Murcia, F. Segovia, J. M. Górriz, J. Ramírez and D. Salas-Gonzalez, [123i]FP-CIT SPECT brain imaging for parkinson's diagnosis using contour lines, *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, (IEEE, 2018).
18. R. Prashanth, S. D. Roy, P. K. Mandal and S. Ghosh, High-accuracy classification of parkinson's disease through shape analysis and surface fitting in 123i-ioflupane SPECT imaging, *IEEE Journal of Biomedical and Health Informatics* **21**(3) (2017) 794–802.
19. D. Castillo-Barnes, D. Salas-Gonzalez, J. Ramírez, F. J. Martinez-Murcia, F. Segovia and J. M. Górriz, Analysis of i[123]-ioflupane SPECT intensity isosurfaces to assist the diagnosis of parkinsonism, *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, (IEEE, 2018).
20. M. Martínez-Ibañez, A. Ortiz, J. Munilla, D. Salas-Gonzalez, J. M. Górriz and J. Ramírez, *Isosurface Modelling of DatSCAN Images for Parkinson Disease Diagnosis*, *Understanding the Brain Function and Emotions* eds. J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo and H. Adeli (Springer International Publishing, 2019), pp. 360–368.
21. M. Imre, J. Tao and C. Wang, Identifying nearly equally spaced isosurfaces for volumetric data sets, *Computers & Graphics* **72** (2018) 82–97.
22. K. Wyman-Chick, P. Martin, M. Minár and R. Schroeder, Cognition in Patients with a Clinical Diagnosis of Parkinson Disease and Scans Without Evidence of Dopaminergic Deficit (SWEDD): 2-Year Follow-Up, *Cognitive and Behavioral Neurology* **29**(Issue 4) (2016) 190–196.
23. L. J. Evers, J. H. Krijthe, M. J. Meinders, B. R. Bloem and T. M. Heskes, Measuring parkinson's disease over time: The real-world within-subject reliability of the MDS-UPDRS, *Movement Disorders* **34** (jul 2019) 1480–1487.
24. D. Castillo-Barnes, F. J. Martinez-Murcia, F. Segovia, I. A. Illán, D. Salas-Gonzalez, J. M. Górriz and J. Ramírez, Comparison between affine and non-affine transformations applied to i[123]-fp-cit spect images used for parkinson's disease diagnosis, *Understanding the Brain Function and Emotions*, eds. J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo and H. Adeli (Springer International Publishing, 2019), pp. 379–388.
25. T. Lewiner, H. Lopes, A. W. Vieira and G. Tavares, Efficient Implementation of Marching Cubes' Cases with Topological Guarantees, *Journal of Graphics Tools* **8**(2) (2003) 1–15.
26. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart and T. Yu, scikit-image: image processing in python, *PeerJ* **2** (2014) p. e453.
27. W. E. Lorensen and H. E. Cline, Marching Cubes: A High Resolution 3D Surface Construction Algo-

rithm, *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, *SIGGRAPH '87*, (ACM, New York, NY, USA, 1987), pp. 163–169.
28. E. V. Chernyaev, Marching Cubes 33: Construction of Topologically Correct Isosurfaces. Technical Report CERN CN 95-17, tech. report, CERN (1995). Presented at GRAPHI CON '95, Saint-Petersburg, Russia, 3-7 July, 1995.
29. C. B. Barber, D. P. Dobkin and H. Huhdanpaa, The quickhull algorithm for convex hulls, *ACM Transactions on Mathematical Software* **22**(4) (1996) 469–483.
30. V. Vapnik, *Statistical Learning Theory*, 1 edn. (John Wiley and Sons, 1998).
31. T. Joachims, Text categorization with Support Vector Machines: Learning with many relevant features, *Machine Learning: ECML-98*, (Springer Berlin Heidelberg, 1998), pp. 137–142.
32. J. Bergstra and Y. Bengio, Random search for hyper-parameter optimization, *Journal of machine learning research* **13**(Feb) (2012) 281–305.
33. S. Knerr, L. Personnaz and G. Dreyfus, Neurocomputing: algorithms, architectures and applications, *by Françoise Fogelman Soulié and Jeanny Hérault. Berlin, Heidelberg: Springer Berlin Heidelberg. Chap. Single-layer learning revisited: a stepwise procedure for building and training a neural network* (1990) 41–50.
34. P.-H. Chen, C.-J. Lin and B. Schlkopf, A tutorial on ν-support vector machines, *Applied Stochastic Models in Business and Industry* **21** (mar 2005) 111–136.
35. T. F. Chan, G. H. Golub and R. J. LeVeque, Updating formulae and a pairwise algorithm for computing sample variances, *COMPSTAT 1982 5th Symposium held at Toulouse 1982*, Springer1982, pp. 30–41.
36. S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift. (2015).
37. J. Chung, C. Gulcehre, K. Cho and Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling (2014).
38. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, Dropout: A simple way to prevent neural networks from overfitting, *Journal of Machine Learning Research* **15** (2014) 1929–1958.
39. R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection., *IJCAI'95 Proceedings of the 14th international joint conference on Artificial intelligence*, *IJCAI'95* **2**, (Quebec, Canada, 1995), pp. 1137–1145.
40. D. Krstajic, L. J. Buturovic, D. E. Leahy and S. Thomas, Cross-validation pitfalls when selecting and assessing regression and classification models, *Journal of Cheminformatics* **6**(1) (2014).
41. T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning* (Springer-Verlag GmbH, 2009).
42. D. Salas-Gonzalez, J. M. Górriz, J. Ramírez, I. A. Illán, P. Padilla, F. J. Martínez-Murcia and E. W. Lang, Building a FP-CIT SPECT brain template using a posterization approach, *Neuroinformatics* **13** (mar 2015) 391–402.
43. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, Scikit-learn: Machine learning in python, *Journal of machine learning research* **12**(Oct) (2011) 2825–2830.
44. R.-E. Fan, P.-H. Chen and C.-J. Lin, Working set selection using second order information for training support vector machines, *Journal of machine learning research* **6**(Dec) (2005) 1889–1918.
45. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, Pytorch: An imperative style, high-performance deep learning library, *Advances in Neural Information Processing Systems*, 2019, pp. 8024–8035.
46. H. Mann and D. Whitney, On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other, *The Annals of Mathematical Statistics* **18**(1) (1947) 50–60.
47. D. Castillo-Barnes, J. Ramírez, F. Segovia, F. J. Martinez-Murcia, D. Salas-Gonzalez and J. M. Górriz, Robust Ensemble Classification Methodology for I123-Ioflupane SPECT Images and Multiple Heterogeneous Biomarkers in the Diagnosis of Parkinson's Disease, *Frontiers in Neuroinformatics* **12** (2018) p. 53.
48. D. Castillo-Barnes, F. Segovia, F. J. Martinez-Murcia, D. Salas-Gonzalez, J. Ramírez and J. M. Górriz, Classification Improvement for Parkinson's Disease Diagnosis Using the Gradient Magnitude in DaTSCAN SPECT Images, *Advances in Intelligent Systems and Computing*, **771** (Springer International Publishing, 2018), pp. 100–109.
49. M. López, J. Ramírez, J. Górriz, I. Illán, D. Salas-Gonzalez, F. Segovia and R. Chaves, Svm-based cad system for early detection of the alzheimer's disease using kernel pca and lda, *Neuroscience Letters* **464**(3) (2009) 233–238.
50. J. Blesa, Inter-hemispheric asymmetry of nigrostriatal dopaminergic lesion: a possible compensatory mechanism in Parkinson's disease, *Frontiers in Systems Neuroscience* **5** (2011).
51. M. C. A. Santos, L. S. Campos, R. P. Guimarães, C. C. Piccinin, P. C. Azevedo, L. G. Piovesana, B. M. D. Campos, A. C. S. Amato-Filho, F. Cendes and A. D'Abreu, Does Side of Onset Influence the Pattern of Cerebral Atrophy in Parkinson's Disease?, *Frontiers in Neurology* **7** (2016).
52. J. Torres, A. Galicia, A. Troncoso and F. Martínez-Álvarez, A scalable approach based on deep learning for big data time series forecasting, *Integrated Computer-Aided Engineering* **25**(4) (2018) 335–348.
53. F. Vera-Olmos, E. Pardo, H. Melero and N. Malpica,

Deepeye: Deep convolutional network for pupil detection in real environments, *Integrated Computer-Aided Engineering* **26**(1) (2018) 85–95.
54. T. Yang, C. Cappelle, Y. Ruichek and M. E. Bagdouri, Multi-object tracking with discriminant correlation filter based deep learning tracker, *Integrated Computer-Aided Engineering* **26**(3) (2019) 273–284.
55. U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan and H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals, *Computers in Biology and Medicine* **100** (2017).
56. A. Antoniades, L. Spyrou, D. Martin-Lopez, A. Valentin, G. Alarcon, S. Sanei and C. C. Took, Deep neural architectures for mapping scalp to intracranial EEG, *International Journal of Neural Systems* **28**(08) (2018) p. 1850009.
57. F. J. Martinez-Murcia, J. M. Górriz, J. Ramírez and A. Ortiz, Convolutional neural networks for neuroimaging in parkinson's disease: Is preprocessing needed?, *International Journal of Neural Systems* **28**(10) (2018) p. 1850035.
58. O. M. Manzanera, S. K. Meles, K. L. Leenders, R. J. Renken, M. Pagani, D. Arnaldi, F. Nobili, J. Obeso, M. R. Oroz, S. Morbelli and N. M. Maurits, Scaled subprofile modeling and convolutional neural networks for the identification of parkinson's disease in 3d nuclear imaging data, *International Journal of Neural Systems* **29**(09) (2019) p. 1950010.
59. C. Hua, H. Wang, H. Wang, S. Lu, C. Liu and S. M. Khalid, A novel method of building functional brain network using deep learning algorithm with application in proficiency detection, *International Journal of Neural Systems* **29**(01) (2019) p. 1850015.
60. A. H. Ansari, P. J. Cherian, A. Caicedo, G. Naulaers, M. D. Vos and S. V. Huffel, Neonatal seizure detection using deep convolutional neural networks, *International Journal of Neural Systems* **29**(04) (2019) p. 1850011.
61. P. Gómez-Vilda, Z. Galaz, J. Mekyska, J. M. F. Vicente, A. Gómez-Rodellar, D. Palacios-Alonso, Z. Smekal, I. Eliasova, M. Kostalova and I. Rektorova, Vowel articulation dynamic stability related to parkinson's disease rating features: Male dataset, *International Journal of Neural Systems* **29**(02) (2019) p. 1850037.
62. M. Caiola and M. H. Holmes, Model and analysis for the onset of parkinsonian firing patterns in a simplified basal ganglia, *International Journal of Neural Systems* **29**(01) (2019) p. 1850021.