Neuroinformatics (2022) 20:765–775
https://doi.org/10.1007/s12021-022-09578-3

ORIGINAL ARTICLE

Design and Application of Automated Algorithms for Diagnosis and Treatment Optimization in Neurodegenerative Diseases

Francisco Estella[1] · Esther Suarez[2] · Beatriz Lozano[2] · Elena Santamarta[2] · Antonio Saiz[2] · Fernando Rojas[1] · Ignacio Rojas[1] · Marta Blazquez[2] · Lydia Nader[2] · Javier Sol[2] · Fernando Seijo[3]

Accepted: 22 February 2022 / Published online: 9 March 2022

Abstract

Neurodegenerative diseases represent a growing healthcare problem, mainly related to an aging population worldwide and thus their increasing prevalence. In particular, Alzheimer's disease (AD) and Parkinson's disease (PD) are leading neurodegenerative diseases. To aid their diagnosis and optimize treatment, we have developed a classification algorithm for AD to manipulate magnetic resonance images (MRI) stored in a large database of patients, containing 1,200 images. The algorithm can predict whether a patient is healthy, has mild cognitive impairment, or already has AD. We then applied this classification algorithm to therapeutic outcomes in PD after treatment with deep brain stimulation (DBS), to assess which stereotactic variables were the most important to consider when performing surgery in this indication. Here, we describe the stereotactic system used for DBS procedures, and compare different planning methods with the gold standard normally used (i.e., neurophysiological coordinates recorded intraoperatively). We used information collected from database of 72 DBS electrodes implanted in PD patients, and assessed the potentially most beneficial ranges of deviation within planning and neurophysiological coordinates from the operating room, to provide neurosurgeons with additional landmarks that may help to optimize outcomes: we observed that x coordinate deviation within CT scan and gold standard intra-operative neurophysiological coordinates is a robust matric to pre-assess positive therapy outcomes- "good therapy" prediction if deviation is higher than 2.5 mm. When being less than 2.5 mm, adding directly calculated variables deviation (on Y and Z axis) would lead to specific assessment of "very good therapy".

Keywords Alzheimer disease · Brain · Classification · Decision trees · Deep brain stimulation · Parkinson disease

Introduction

The evolution and increasing complexity of medical instruments can be associated with greater difficulties in subsequent data interpretation. Thus, biomedical engineering has become one of the main research areas to explore ways to optimize data processing and thus improve diagnosis and treatment options for different diseases.

One focus in healthcare research is on neurodegenerative diseases such as the cognitive disorder Alzheimer's disease (AD) and the movement disorder Parkinson's disease (PD). Both are associated with aging, and their prevalence is increasing worldwide. For example, EURODEM estimates that 53.7% of all dementia cases correspond to AD (Alzheimer's Disease International (ADI), 2018), while about 0.3% of the population are currently suffering from PD (Rocca, 2018). The causes of these diseases are not widely defined, and their evolution directly affects cognitive, functional, psychological and social functions.

The diagnosis of AD is currently based on clinical tests and is only confirmed by anatomopathological exam within the brain. The importance of identifying the type of dementia, in addition to the increase in evidence about earlier stages treatment, makes the identification of diagnose markers key for this field (Livingston et al., 2020). Similarly, PD presents motor difficulties symptoms such as tremor, rigidity, bradykinesia and postural instability. These motor symptoms significantly impair daily living and quality of life and are a high burden on both patients and their caregivers.

✉ Francisco Estella
francisco.estellap@gmail.com

[1] University of Granada, Granada, Spain

[2] Hospital Universitario Central de Asturias, Asturias, Spain

[3] Centro Médico de Asturias, Asturias, Spain

Better classification tools are required in PD to assess which is the best target for therapies such as deep brain stimulation (DBS).

DBS is a surgical procedure used to treat the motor symptoms of various neurological diseases, but mainly movement disorders such as PD (Fisman et al., 2006). Stereotactic techniques are used to implant the leads in the brain, using magnetic resonance imaging (MRI) or computed tomography (CT) to identify the correct area of the brain (thus defining the coordinates) to be targeted for stimulation, usually the subthalamic nucleus (STN), thalamus or globus pallidus. Different methodologies are available to choose the final placement coordinates of the lead, which can add to the already complex intraoperative decision-making regarding the optimal resource to use for imaging, microelectrode recording (MER), intraoperative clinical outcomes, etc.

Both AD and PD affect different structures within the central nervous system, which translates to identifiable neuropsychological patterns that are specific to each disease. The brain alterations do not affect all areas equally. Each brain structure is influenced differently, depending on the neurodegenerative disorder, and can be identified by the different mental patterns using neuroimaging techniques. Some indicators are still unknown, however; therefore, prediagnosis and treatment optimization using automatic algorithms could be key for treatment, or at least help to control its evolution.

Therefore, the first objective of our study was to design and validate an automated classification algorithm to identify different groups in AD according to the disease evolution, through indicators such as anatomical landmarks from MRI images or mathematical variables, using a database of 1,200 patients. Our second objective was to apply this classification algorithm to a database of 72 implanted electrodes for DBS to treat PD using stereotactic systems. We wanted to be able to define in advance whether surgery of this kind would be successful by identifying different patterns of coordinate deviations within the CT scans or MRI in relation to the 'gold standard' currently used in the operating room, i.e., neurophysiological coordinates. We have also specified a maximum deviation range calculated during the planning stage of surgery, to help increase the chances of successful therapy after implantation.

Common framework seems clear when applying to neurodegenerative diseases in which imaging plays a big role for diagnose or treatment. Indeed, AD is also being treated with DBS (Leoutsakos et al., 2018), which would lead us to have an integral approach to this disease thanks to this manuscript, as we are proposing the management of the disease from early pre-diagnosis to therapy outcomes optimization. To assess the feasibility of this method for DBS, Parkinson's Disease was used as it is a more widely used indication; however, as same therapy approach will be applicable to AD, methodology framework gets sense for both indications in this paper.

Material and Methods

Database Analysis and Algorithm Design

Database

Our first objective was to design a more efficient and accurate classifier to distinguish normal, mild cognitive impairment (MCI) and AD, using MRIs obtained from a large database- Other studies have been exploring the diagnose of AD in relation with classifiers and variables extracted from medical imaging (Sun et al., 2018; Liu et al., 2014). Our work is adding value as it uses not only anatomical variables but also mathematical variables that could be of interest for the pre-diagnosis of the disease, this increases the possibilities of potential correlation within groups of study, and open a new path of study as only pure anatomical variables had been used so far for these purposes. We examined around 2,000 MR images/files of different patients, and 1200 were used in our study (requiring approximately 240 GB of storage). We used the database of the Alzheimer's Disease Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/about/), whose main goal is "to define the progression of Alzheimer's Disease". Each file is a three-dimensional (3D) representation of the information contained within the brain of a patient. Using different cuts/planes or values of variables X, Y, Z to analyze the 3D images, it is possible to obtain three projections of two-dimensional (2D) images in different areas of the brain, which can then be analyzed separately (Fig. 1).

Spatial Normalization

In studies that involve images of many patients, it is often useful (and in our project necessary) to co-register a brain image of a patient to another subject or a standard template – a process known as spatial normalization. To normalize our MR images, we used the "SPM12" toolbox for MATLAB (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Statistical Parametric Mapping (SPM) entails the construction of spatially-extended statistical processes to the test hypotheses regarding regionally specific effects. The maps generated by SPM are image processes with voxel values that are, under the null hypothesis, distributed according to a known probability density function, usually the Student's T or F distributions (i.e., T- or F-maps). The success of statistical parametric mapping is due largely to the simplicity of the idea. Essentially, each voxel is analyzed using any standard (univariate) statistical test. The resulting statistical

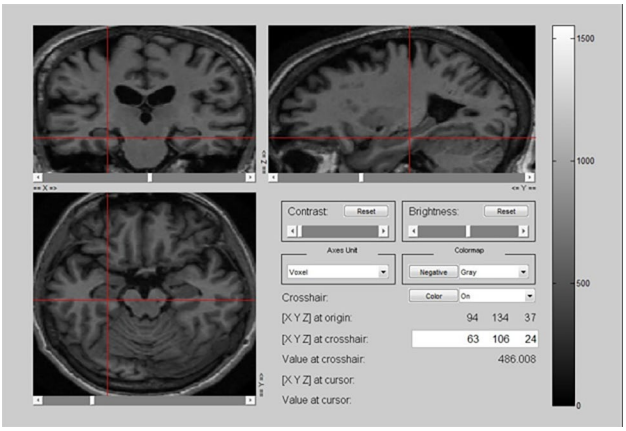

Fig. 1 Use of a graphical user interface to visualize the three planes of the brain using different values of X,Y,Z, obtaining different slides. Example shown: (X,Y,Z) = (63,106,24)

parameters are assembled into an image, the SPM. Due to the considerable number of images involved in our study, we have developed a method to normalize images (Fig. 2).

Medical Images and Slides Selected

15 slides were extracted from every MRI image. The reasons why these specific slides were chosen for the AD diagnosis was because of the information, as hippocampus atrophy, that could be extracted comparing normal and abnormal MRI (Fig. 3). These specific set of slides were agreed within a panel of experts that included neurologist, neuro-radiologist and neurosurgeon. These are the standard slides considered when diagnosing AD in medical practice as standard of care (SoC): the SoC defines hippocampus atrophy as the main marker together with the corresponding "ex-vacuo" dilatation of ventricles.

Images were obtained by repository and stored in specific folders per category as defined by ADNI, these folders where then added to the code to go through images per group and normalized individually using "SPM12" toolbox as defined above.

We could indicate that according to the extracted features, there were two big groups: 322 morphological features per patient and 108 mathematical features per patient (Table 1).

Classification Methods

- Fuzzy Decision Tree: Decision trees based on fuzzy set theory combines the **advantages** of good comprehensibility of decision trees and the ability of fuzzy representation to deal with inexact and uncertain information (Janikow, 1998). There are other classification methods as the ones listed below or more widely used currently such as SVM; however, we decided to choose these ones due to the clarity of the information provided.

- Linear Discriminant Analysis: Classification method that works finding a linear combination of features that characterizes or separates two or more groups to be classified (Hastie et al., 2017).

- Quadratic Discriminant Analysis: More general version of the linear method. Separate measurements of two or more groups by a quadric surface (Hastie et al., 2017).

- Naïve Bayes: Bayes' theorem describes how probable an event is, based on previous conditions related to that event. The Naïve Bayes classifier is a "probabilistic classifiers" based on applying Bayes' theorem with (naive) independence assumptions between the features (Hastie et al., 2017).

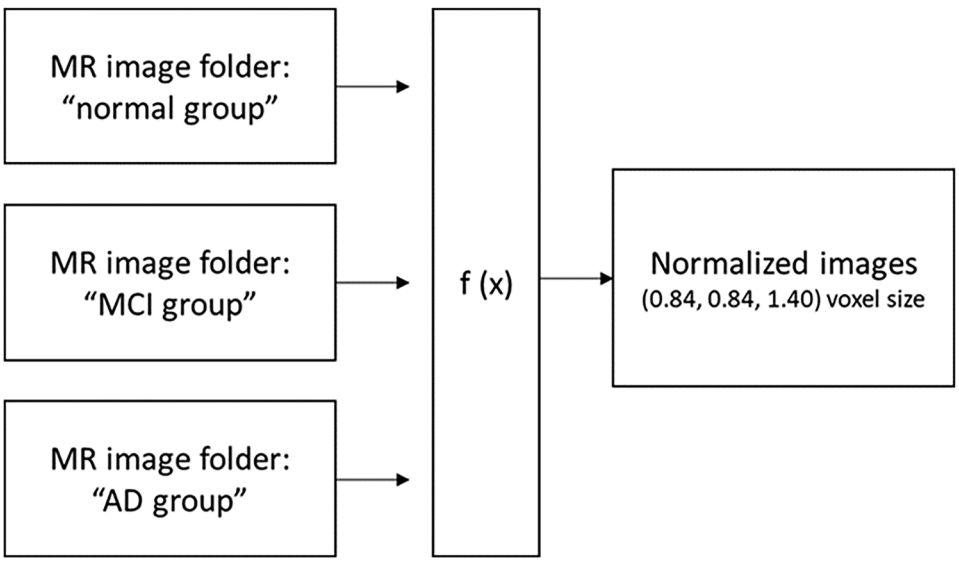

Fig. 2 Normalization method running through the three image folders that normalized all of them to the same format: voxel size was adapted from (2, 2, 2) to (0.84, 0.84, 1.40) for all images

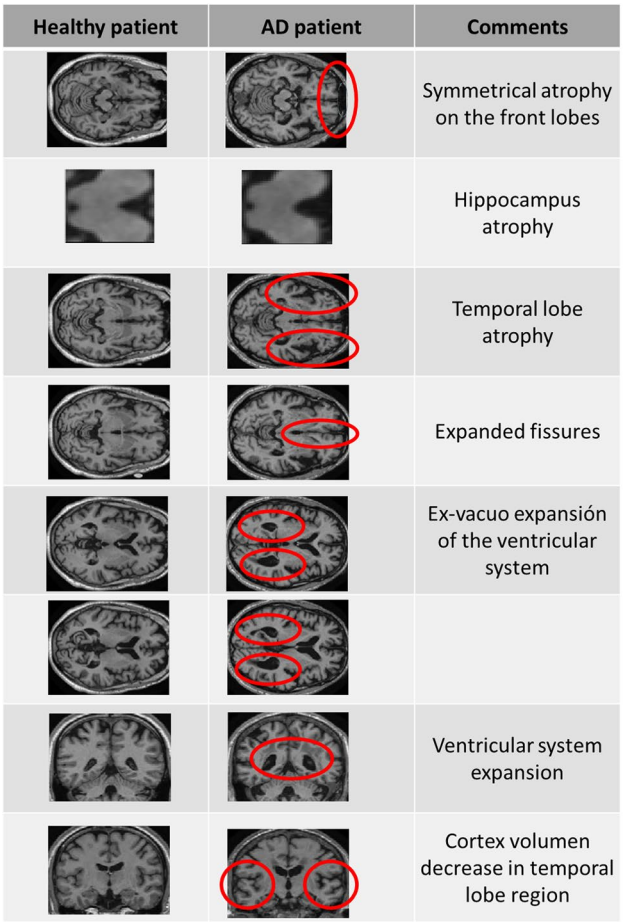

Fig. 3 Medical images used for the classification of patients with Alzheimer's disease (AD). Examples of variables extracted to build the pre-diagnose algorithm built to be used for the therapy assessment algorithm tool for PD

Table 1 Morphological and mathematical features extracted from magnetic resonance images

Morphological features	Mathematical features
Area	Mean
Centroid	Fourier transform coefficients
Bounding box	Cosine transform coefficients
Major axis length	Euclidean distance
Minor axis length	Chebychev distance
Eccentricity	Minkowsky distance
Orientation	Spearman correlation
Convez area	City block distance
Filled area	Mahalanobis distance
Euler number	
Equivalent diameter	
Solidity	
Extent	
Perimeter	
Compactness	
Rectangularity	
Grey white and whole matter volumes	

Application to PD Patients Treated With DBS

The most common brain area for DBS in PD patients is the STN (Fisman et al., 2006), which is a very small brain nucleus (3–4 mm high × 2–3 mm wide). Stimulation of neighboring structures could lead to undesired side effects, and thus poor clinical results. Precision in electrode location is critical during the surgical procedure. Therefore, a stereotactic coordinates system is used to identify the target structure and place the electrode in the right spot. These coordinates are applied using a stereotactic frame located in patient's head. Medical images are then acquired (CT scan and MRI) to identify anterior and posterior commissures (AC-PC) and subsequently the target. These 'target coordinates' are then translated and placed in the frame to perform the external adjustment, to enable the electrodes to be correctly placed deep within the brain.

At the Hospital Central de Asturias (HUCA), 72 electrodes were implanted using planning through CT and MRI scans taken preoperatively (Fig. 4). The stereotactic coordinates were calculated by direct targeting of the STN and by manual calculation from AC-PC landmarks. Neurophysiological coordinates and target length (STN length) were calculated intraoperatively using a MER system, first targeting theoretical coordinates (calculated during planning phase) and then shifting according to the neurophysiological and clinical response (Fig. 5). Patients were then treated with DBS using the stable combination, and medication (levodopa) was adjusted and minimized if possible. The Unified Parkinson's Disease Rating Scale (UPDRS) part III, which assess motor symptoms, was used to assess the baseline level of disease and therapeutic outcomes. Other scales are also studied (e.g., UPDRS II or medication reduction), but we focused on the UPDRS III percentage of improvement because it is the most widely used. A meta-analysis of published articles for DBS in PD has determined the improvement in UPDRS III reported in trials (Table 2). We observed that the maximum improvement in the average UPDRS III score was around 55%; however, recent studies have reported around 60% improvement due to new technologies (Hastie et al., 2017). Therefore, for the purposes of our study, patients with more than 60% improvement in the UPDRS III scale were classed as having received 'very successful therapy' (Table 3). Overall, 33% of the patients in the patient's database analyzed were classified as very successful treatment, 37% as successful treatment, and 14% as not successful treatment.

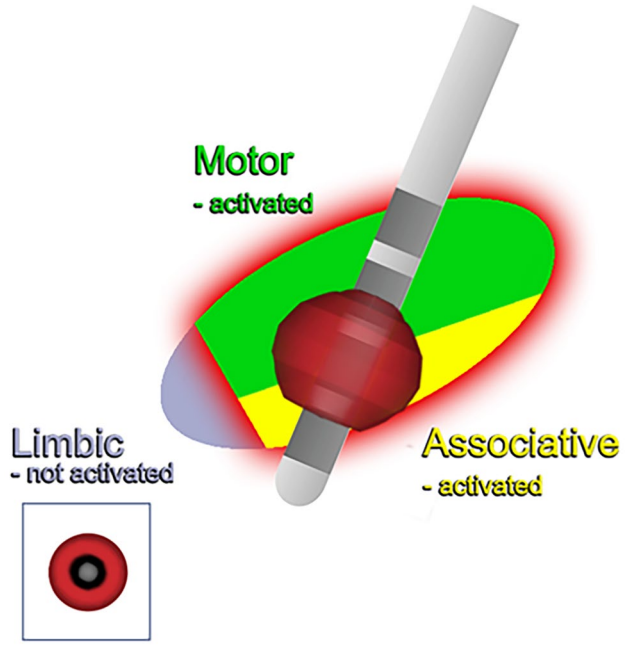

Fig. 4 Deep Brain Stimulation electrode localization most commonly used for PD, using Subthalamic Nucleus (STN) as target. Image shows position of electrode in relation to functional areas of STN and their activation with a traditional monopolar stimulation on those areas

Common Framework

AD and PD are both neurodegenerative diseases that are proven to respond better to therapy when diagnosed at earlier stages (Schuepbach et al., 2019). Also Deep Brain Stimulation has been proven as an indicated therapy for PD, and it is being recently studied for AD (Leoutsakos et al., 2018), which leads to the idea for a common framework when approaching the diagnosis and treatment of these therapies with similar methodologies. The methodology applied in this study presents an early pre-diagnosed classification that could be explored for PD as imaging markers in this kind of neurodegenerative diseases are similar and in any case useful for the diagnosis of the disease; this methodology focused on classification methods to assess the potential outcomes of DBS therapy, which can be applied to AD treated with DBS also, as implant techniques are the same, as well as stereotactic frame coordinate system to be used for both.

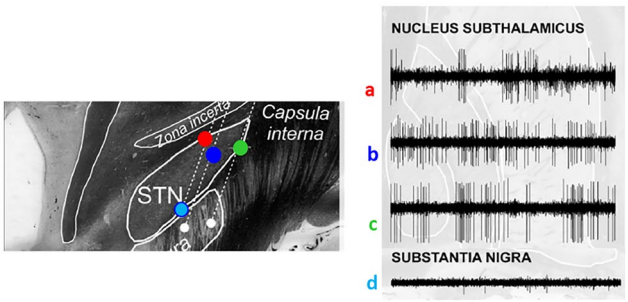

Fig. 5 Subthalamic nucleus microelectrode recording showing different micro-electrode recordings depending on the functional area of the subthalamic nucleus

Feature Selection Using Mutual Information

Feature selection using mutual information (MI) was used (Vergara, 2014). In pattern recognition theory, patterns are represented by a set of variables (features) or measures. Such pattern is a point in a n-dimensional features space. The main goal is to select features that distinguish uniquely between patterns of different classes. Normally, the optimal set of features is unknown and commonly has an irrelevant number or redundant features. Through a pattern recognition process, these irrelevant or redundant features are filtered out and the learning performance of classifiers is greatly improved.

Different criteria have been applied to evaluate the goodness of a feature. In this case, the proposed filter features selection method is based on mutual information as relevance measure and redundancy between the features through minimal-redundancy-maximal-relevance criterion (mRMR). This method was applied folder by folder to go through all datasets.

Let X and Y two random continuous variables with marginal pdfs p(x) y p(y) respectively, and joint probability density function (pdf) p(x,y). The mutual information between X and Y can be represented as:

$$I(X,Y) = \iint p(x,y) log \frac{p(x,y)}{p(x)p(y)} dxdy$$

In the case of discrete variables, the integral operation is reduced to a summation operation. Let X and Y two discrete variables, marginal probabilities p(x) and p(y) respectively and a joint probability mass function p(x,y). The MI between X and Y is expressed as:

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} dxdy$$

The mutual information (MI) has two principal properties that make it different from other dependency measures: 1) the capacity of measure any relationship between variables and 2) its invariance under space transformations.

For mRMR, we considered mutual information-based feature selection for both discrete and continuous data. The MI for continuous variables was estimated using Panzer Gaussian windows. Estimating the mutual information I(C,

Table 2 Summary table of previous clinical trial of deep brain stimulation for Parkinson's disease patients (Hastie et al., 2017; Goetz et al., 2019; Deuschl et al., 2006; Follett et al., 2010; Fraix et al., 2006; Rodriguez-Oroz, 2005; Gervais-Bernard et al., 2009; Moro et al., 2010; Vitek et al., 2020)

Author	N	Design	Time Follow-up	Results
Deuschl et al. (2006)	78	Randomized pairs trial	6 months	41% improvement in UPDRS III
Follett et al. (2010)	299	Multi-center, randomized, blinded	24 months	25,3% improvement in UPDRS III
Fraix et al. (2006)	95	Prospective, multi-center	12 months	57% improvement in UPDRS III
Rodriguez-Oroz (2005)	69	Blinded, multi-center study	3–4 years	50% improvement in UPDRS III with STN and 39% improvement in Gpi
Gervais-Bernard et al. (2009)	42	Prospective, single-center	5 years	55% improvement in UPDRS III
Moro et al. (2010)	51	non randomized prospect, blinded, multi-center study	5–6 years	20% (GPI) to 45% (STN) improvement in UPDRS III scores
Vitek et al. (2020)	157	multicentre, double-blind, randomised, sham-controlled study	1 year	46% (STN) improvement in UPDRS III scores
Timmermann et al. (2015)	40	non-randomised, prospective, multicentre, open-label study	1 year	62,6% (STN) improvement in UPDRS III scores

S) between class variable C and subset of selected features, for minimizing the classification error in the incremental search algorithm, mRMR method is combined with two wrapper schemes. In a first stage, the method is used with the purpose to find the candidate feature set. In a second stage, backward and forward selections were applied in order to find the compact feature set through the candidate feature set that minimizes the classification error.

Given class variable C, the initial set of features F, an individual feature $f_i \in F$ and a subset of selected features, $S \subset F$, the mRMR criterion for the first order incremental search can be expressed as the optimisation of the following condition:

$$I(C;f) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} dxdy$$

The mRMR criterion, for the first-order incremental search algorithm, tries to optimize the following condition:

$$I\left(C;f_i\right) - \frac{1}{|S|} \sum_{f_S \in S} I(f_S;f_i)$$

where |S| is the cardinality of the selected feature set S, $f_s \in S$. This filter mRMR method is a fast and efficient method because its incremental nature, showing better feature selection and accuracy in classifier including wrapper approach. In this work, mRMR criterion method was used as filter algorithm with the purpose to obtain the relevance of proposed features.

Cross Validation

For the classifier training and validation, 750 patients were used, and then, 450 different patients were chosen to test the efficiency of the method.

Also, we performed several experiments with different number of features selected using the mRMR (75%, 25% and 10% of all the possible features), applying the mRMR methodology.

Also as an important part of the cross-validation scheme followed, we analyzed the confusion matrix of each group to validate the proper selection of features. For example, for class 1 belonging to the best outcomes, the class 2 to the medium ones and the class 3 to the worst outcomes for diagnose or therapy (Class 1 to "healthy group", Class 2 to "MCI group" and Class 3 to "AD group" within the AD task; and Class 1 to "very good therapy", Class 2 to "good therapy" and Class 3 to "no major improvements" within the PD task) we observed the outcome in those 450 patients chosen to assess the robustness of the feature selection process. In total we could get 150 true positives per class, but we got around 100 true positives per class in the scenario in which only 10% of the total variables included in each dataset were chosen, which represents an efficiency of around 70%. This is a good number if we consider that we are doing classification between three different classes and that we are only using few variables (10% of the total as mentioned).

Table 3 Comparison groups, depending on clinical outcomes due to deep brain stimulation versus preoperative scores

Very successful treatment	Successful treatment	Not successful treatment
> 60% UPDRS III improvement	35–60% UPDRS III improvement	< 35% UPDRS III improvement

Table 4 Efficiency obtained for the classification of magnetic resonance images for three classes: healthy, mild cognitive impairment and Alzheimer disease patients

Efficiency	Fuzzy Decision Tree	QDA
Mix Matrix 75%	0.94	LD
Mix Matrix 25%	0.87	0.82
Mix Matrix 10%	0.75	0.67
Morphological Matrix 75%	0.89	0.73
Morphological Matrix 25%	0.68	0.56
Morphological Matrix 10%	0.68	0.48
Mathematical Matrix 25%	0.81	0.71
Mathematical Matrix 10%	0.67	0.64

LD, linear dependence; QDA, Quadratic Discriminant Analysis

Results

Classifier Technique

We decided that the best technique of classification for our project was the Fuzzy Decision Tree, followed by the QDA. The Fuzzy Decision Tree was chosen because of its high efficiency, even when just few features were used for the diagnosis, and because from the point of view of interpretability it can be very useful for human expert diagnosis (Janikow, 1998).

As can be observed in Table 4, the QDA classifier also reaches higher efficiency values, even higher than 80%. However, when the features were reduced to get a diagnosis in a simpler way (to make it easier for the user), the efficiency decreased too much, reaching values under 50% in some cases. For this reason, the Fuzzy Decision Tree classifier was the method that better fit the requirements of our project, because the accuracy obtained using morphological features (that can be easily understood by a neurologist) is sufficient.

Algorithm Application to PD Patients With DBS

Similar classification methods have been applied to other medical applications such as Deep Brain Stimulation for PD. The aim of these publications was to optimize the surgical process by finding the "gold standard" to choose the final coordinates for the procedure (Timmermann et al., 2015; Bermudez et al., 2019). However, none of them classifies and select the most important coordinates calculation to predict the most optimal therapy in advanced.

By applying the algorithm from AD patients to PD patients treated with DBS, we observed that STN length recorded through the MER system was within similar range in all three groups ([4.36, 4.50] mm) (Figs. 6 and 7) and was not of significant value to define better clinical outcomes.

The direct target coordinates calculation showed the highest difference compared with the gold standard neurophysiological coordinates, with major deviation in the Z axis (dev. > 2.3 mm). There was a similar pattern in the calculated coordinates from the CT and MRI scans compared with gold standard for the X and Y axes. However, MRI-scan coordinates differed

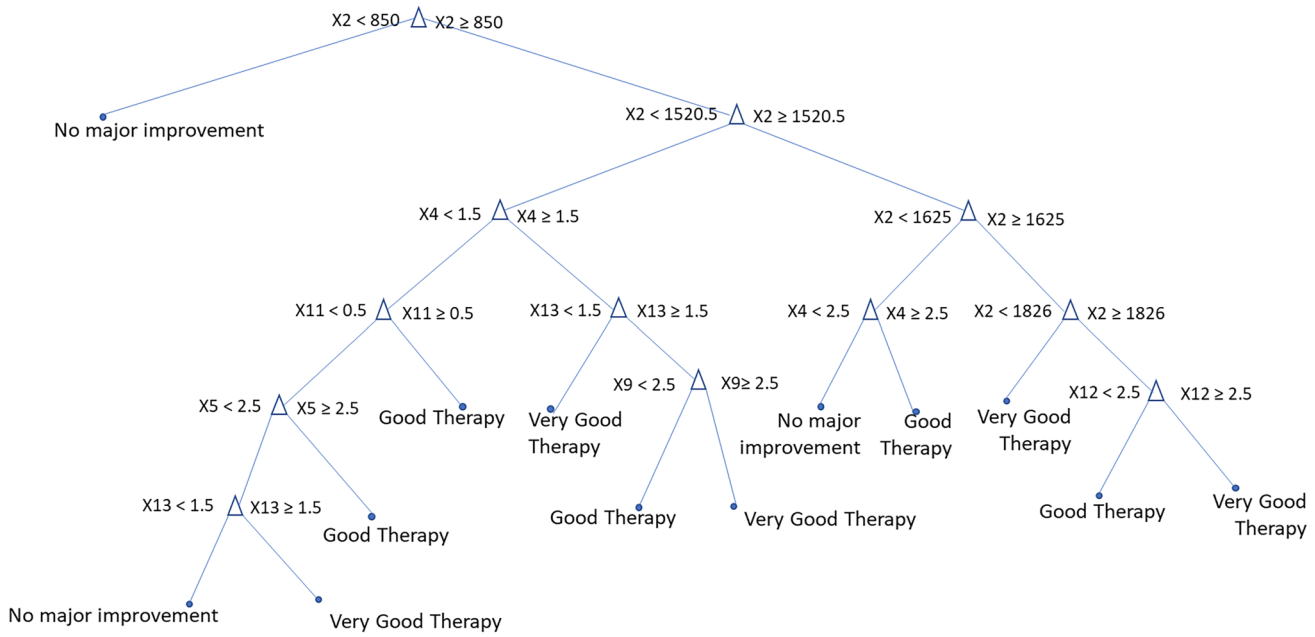

Fig. 6 Decision tree classification including medication as variable of study. It shows how variable 2 (L-dopa medication) and 4 (deviation on X coordinate within MRI planned coordinates and neurophysiological election) are the main variables to assess PD treatment success

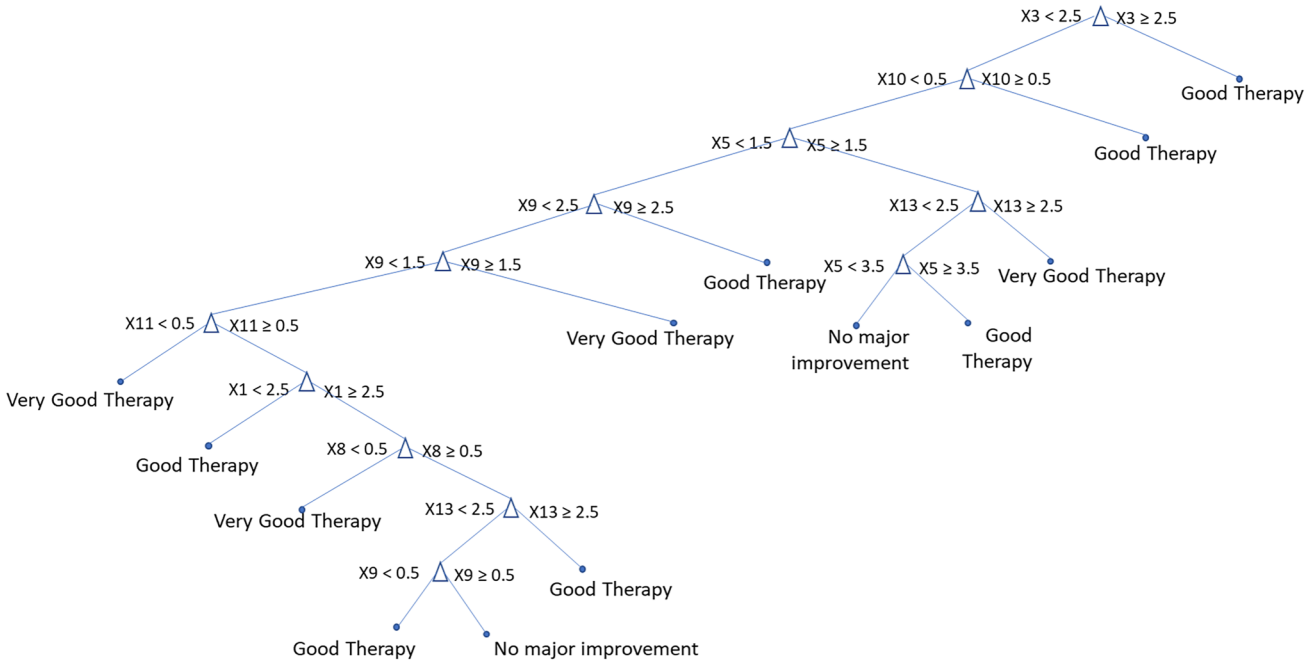

Fig. 7 Decision Tree classification without including medication as variable of study showing variable 3 (deviation on X coordinate within CT planned coordinates and neurophysiological election) as the most important one to assess therapy success and combination of specific ranges of variables (3, 10, 5, 13) or (3, 10, 5, 9) to predict a "Very Good Therapy" response

significantly more than the CT-scan coordinates when comparing the Z axis (Figs. 6 and 7). The deviation average compared to the gold standard neurophysiological coordinates showed no significant value to determine better outcomes, as the three control groups showed similar results for deviations in each coordinate and modality (Figs. 6 and 7).

The variables studied and their corresponding classification assigned (from X1 to X15) are outlined in Table 5. The decision tree classification analysis was performed using these 15 clinical variables as input. Classification output was obtained after training the classification method with the different subject's sub-groups (and 15 clinical variables per subject). The results were studied following two different methodologies, either by considering initial medication or by not taking it into account, and the main highlights are presented below.

Variables Classification Taking Medication Into Account

Medication level is a key indicator for the success of DBS therapy, with 'no major improvement' for patients with preoperative levodopa < 850 mg. This makes sense initially, as most of the eligible patients for DBS therapy are usually advanced PD patients who are already receiving a high dose of medication. In our study, we were able to observe that in cases when DBS was performed even when patients were not yet taking high doses of medication, DBS did not provide major improvement. We observed that the higher the preoperative medication level, the greater the percentage improvement in UPDRS III. We also observed that the deviation between the MRI coordinates and neurophysiological coordinates when comparing the X axis provided better outcomes when the deviation was higher. This means that when trying to find the STN intraoperatively, neurosurgeons should consider that correct deviation should be taken in the X axis rather than the Y axis.

Variable's Classification Not Considering Medication

As we defined previously, direct coordinates are the ones adding higher deviation compared to the gold standard neurophysiological coordinates; therefore, these should not be considered primarily when comparing to MRI and CT scan coordinates and analyzing results.

When we analyzed the decision tree classification after extracting preoperative medication from the variables studied, "very good outcomes" were observed when the X-axis deviation with the CT scan compared to the intraoperative neurophysiological coordinates was less than 2.5 mm and the Z-axis deviation was almost zero (< 0.5 mm).

Considering both of these analyses (i.e., with or without preoperative medication), we observed that potentially a deviation of > 1.5 mm from the MRI coordinates calculated should be taken into account when trying to find the target intraoperatively, but also that neurosurgeons should

Table 5 Clinical variables studied with their corresponding classification variable assigned for the study

Variable	Definition
X1	Neurophysiological STN length (mm.)
X2	L-Dopa medication in mg. (pre-op)
X3	Deviation coordinates Xct vs Xnph (mm.)
X4	Deviation coordinates Xmri vs Xnph (mm.)
X5	Deviation coordinates Xdir vs Xnph (mm.)
X6	Deviation coordinates Xcontrol vs Xnph (mm.)
X7	Deviation coordinates Yct vs Ynph (mm.)
X8	Deviation coordinates Ymri vs Ynph (mm.)
X9	Deviation coordinates Ydir vs Ynph (mm.)
X10	Deviation coordinates Ycontrol vs Ynph (mm.)
X11	Deviation coordinates Zct vs Znph (mm.)
X12	Deviation coordinates Zmri vs Znph (mm.)
X13	Deviation coordinates Zdir vs znph (mm.)
X14	Deviation coordinates Zcontrol vs Znph (mm.)
X15	L-Dopa medication in mg. (post-op, stim ON)

control, postoperative computed tomography (CT) scan image-based coordinate; ct, CT scan image-based coordinate; dir, direct coordinate decided by specialist by pointing it directly in the image; mri, MRI scan image-based coordinate; nph, intra-operative micro-electrode recording validated coordinates ('gold standard' neurophysiological recording); STN, subthalamic nucleus

not deviate further than 2.5 mm from the CT scan X axis coordinates calculated during the planning step.

Discussion

We present a new methodology based on a large data set that is able to classify three different categories of MR brain images using features than can be easily understood by a human expert (i.e., a neurologist rather than a computer) and can aid decision making, thus being of added value for neurosurgeons and neurologist for the final placement of the leads, critical for a positive outcome from DBS therapy. This powerful tool uses the same methodology as other classification processes related to imaging with any other variables (such as dystonia, essential tremor or any other indication with its specific targets defined). However, it is important to remember the variability within patients; thus, it is impossible to assess a general and standard methodology for all treated cases.

As we could see, classification methodology was validated and optimized by checking the minimum sufficient number of variables to be selected following mRMR feature selection. Selected feature selection methodology was applicable and tested for both tasks, in which three different classes were used (very good, good or no major improvement for therapy and three groups also for the diagnostics task). For both, 10% of total size of features were selected being anatomical/morphological variables the ones with higher weight; thus, also confirming the proper use of these imaging techniques in SoC for medical practice nowadays.

For application in PD, it is important to note that neurophysiological length of the STN recorded does not affect clinical outcomes if it is within range (> 4 mm). Our results indicate that CT-scan calculations are the most accurate when planning for DBS. Direct calculation should be avoided, as major deviation in the Z axis was observed, which could be even dangerous during surgical procedure, leading to hemorrhages or undesired stimulation.

Limitations of these dataset need to be mentioned, as having a limited dataset for the PD algorithm to assess potential outcomes of DBS therapy focus on data from one center only might be a bias for general application in other centers, in which the DBS procedure might vary in specific steps of the process. We could assess that this algorithm is valid for centers with the standardized DBS procedure flow followed in the center studied: use of stereotactic planning through direct/anatomical planning of coordinates using imaging, neurophysiological intraoperative assessment of coordinates and finales verification using computer tomography (CT) scan. Also, although comparison and common framework within algorithm application has been demonstrated above, a potential limitation on the common framework might be the fact that additional biomarkers in addition to the main anatomical and mathematical ones considered in this study are being used for each indication, such as behavioral patterns or cognitive scales; further studies using all these datasets per indication could potentially led to different weights on the different variables studied when creating decision trees.

The main advantage of comparing our data to others in the same line of investigation is the big database available. Also, there are no similar studies applying this kind of decision algorithms to compare DBS coordinates and assess the most important deviations when assessing the success of the therapy. The fact that only one center data was studied when applying the algorithm to DBS outcomes assessment is also positive as no deviations in the surgical workflow are to be considered, which could potentially bias the results; thus allowing for a proper validation or comparison within the different techniques when deciding the best coordinates to be used within the DBS procedure.

As further advantages of the use of these datasets we could consider that similarities within both indications in terms of brain atrophy as one of the main anatomical biomarkers for diagnostic could lead us to potential use of the diagnostic algorithm of AD for PD patients with cognitive impairment related to PD, or to future use of the algorithm when treating AD patients with DBS as mentioned above.

Other positive point of these datasets is the fact that the standard of care of both indications for healthcare professionals currently use anatomical biomarkers through

imaging as one of the most relevant indicators for diagnosis and therapy outcomes validation; this big database confirms that hypothesis and current medical practice, and adds a more clear and automatized pattern to consider and filter the main landmarks when taking clinical decisions for both diagnose and assessment of potential therapy outcomes, key when deciding on potential surgical interventions for the treatment of a specific pathology.

Further analysis is needed to assess a specific landmark as key indicator for successful therapy, meanwhile the use of these classification methods could result in more insights about the minimum variables needed to correctly place the lead, thus optimizing the current DBS methodology in centers where micro-electrode recording is being used as gold standard, as well as proving this algorithm as useful when diagnosing these neurodegenerative diseases, confirming that anatomical biomarkers, as currently use in standard medical practice, have enough wight or importance to allow for clinical decisions considering them.

Conclusion

As technology advances, the need for algorithms to optimize the time of the clinical team is becoming more important. Preoperative algorithms will lead to optimization for operating room timings, and allow functional units using DBS to better understand and estimate potential positive outcomes for many patients. It is also of great utility for pre-diagnosis of other indications, which is key in neurodegenerative diseases such as AD and PD. Our study shows promising results in classification for pre-diagnosis of AD and for DBS cases in PD, allowing physicians to start treating neurodegenerative diseases earlier to avoid a faster evolution of the disease and to optimize the decision of final placement of leads in DBS. More research and development on preoperative algorithms to potentially assess DBS for PD outcomes are needed, but our study demonstrates that CT scan images are the most accurate and should be chosen as the initial method when manually planning coordinates for DBS. This is increasingly changing as automatic fusion and planning algorithms are being introduced to calculate planning using fewer calculations. In contrast to previous beliefs, our results indicate that the recorded STN length in DBS for PD is not the best indicator for a better trajectory and shows less weight with better clinical outcomes.

Information Sharing Statement

The datasets generated during and/or analysed during the current study are available in the figshare repository, https://figshare.com/s/69483c46308c2f3fe675 https://figshare.com/s/e3ed61cc0bec6f08b259 https://figshare.com/s/449aa9a8975d0ef4e777. Other datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of Interest The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Alzheimer's Disease International (ADI). (2018). World Alzheimer Report 2018: The state of the art of dementia research: New frontiers. *Alzheimer's Disease International (ADI).*

Bermudez, C., Rodriguez, W., Huo, Y., Hainline, A. E., Li, R., Shults, R., & Landman, B. A. (2019). Towards Machine Learning Prediction of Deep Brain Stimulation (DBS) Intra-operative Efficacy Maps. *Proceedings of SPIE - The International Society for Optical Engineering.*

Deuschl, G., Schade-Brittinger, C., et al. (2006). A Randomized Trial of Deep-Brain Stimulation for Parkinson's Disease. *New England Journal of Medicine, 355*, 896–908.

Fisman, G. K., Herzog, J., Fisman, D., Lang, A. E., & Deuschl, G. (2006). Subthalamic Nucleus Deep Brain Stimulation: Summary and Meta-Analysis of Outcomes. *Movement Disorders, 21*(Suppl 14), S290–S304.

Follett, K. A., Weaver, F. M., Stern, M., et al. (2010). Pallidal versus Subthalamic Deep-Brain Stimulation for Parkinson's Disease. *New England Journal of Medicine, 362*, 2077–2091.

Fraix, V., Houeto, J. L., Lagrange, C., et al. (2006). Clinical and economic results of bilateral subthalamic nucleus stimulation in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 77*, 443–449.

Gervais-Bernard, H., Xie-Brustolin, J., Mertens, P., et al. (2009). Bilateral subthalamic nucleus stimulation in advanced Parkinson's disease: Five-year follow-up. *Journal of Neurology, 256*, 225–233.

Goetz, L., Bhattacharjee, M., Ferraye, M. U., Fraix, V., Maineri, C., Nosko, D., Fenoy, A. J., Piallat, B., Torres, N., Krainik, A., Seigneuret, E., David, O., Parent, M., Parent, A., Pollak, P., Benabid, A. L., Debu, B., & Chabardès, S. (2019). Deep Brain Stimulation of the Pedunculopontine Nucleus Area in Parkinson Disease: MRI-Based Anatomoclinical Correlations and Optimal Target. *Neurosurgery, 84*(2), 506–518. https://doi.org/10.1093/neuros/nyy151

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *The Elements of Statistical Learning*. Springer.

Janikow, C. Z. (1998). Fuzzy decision trees: Issues and methods Man and Cybernetics. IEEE Press.

Krüger, R., Schuepbach, M., Deuschl, G., & EARLYSTIM study group. (2019). Quality of life predicts outcome of deep brain stimulation in early Parkinson disease. *Neurology, 92*(10), e1109–e1120. https://doi.org/10.1212/WNL.0000000000007037

Leoutsakos, J. M. S., Yan, H., Anderson, W. S., Asaad, W. F., Baltuch, G., Burke, A., & Lyketsos, C. G. (2018). Deep brain stimulation targeting the fornix for mild alzheimer dementia (the ADvance Trial): A two year follow-up including results of delayed

activation. *Journal of Alzheimer's Disease, 64*(2), 597–606. https://doi.org/10.3233/JAD-180121

Liu, M., Zhang, D., & Shen, D. (2014). Alzheimer's Disease Neuroimaging Initiative. Identifying informative imaging biomarkers via tree structured sparse learning for AD diagnosis. *Neuroinformatics, 12*(3), 381–94. https://doi.org/10.1007/s12021-013-9218-x

Livingston, G., Huntley, J., Sommerlad, A., et al. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet, 396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Moro, E., Lozano, A., Pollak, P., et al. (2010). Long-Term Results of a Multicenter Study on Subthalamic and Pallidal Stimulation in Parkinson's Disease. *Movement Disorders, 25*, 578–586.

Rocca, W. A. (2018). The burden of Parkinson's disease: a worldwide perspective. *Elsevier open access*.

Rodriguez-Oroz, M. C., Obeso, J. A., Lang, A. E., Houeto, J. L., Pollak, P., Rehncrona, S., & Van Blercom, N. (2005). Bilateral deep brain stimulation in Parkinson's disease: A multicentre study with 4 years follow-up. *Brain, 128,* 2240–2249.

Sun, Z., Qiao, Y., Lelieveldt, B. P. F., & Staring, M. (2018). Alzheimer's Disease NeuroImaging Initiative. Integrating spatial-anatomical regularization and structure sparsity into SVM: Improving interpretation of Alzheimer's disease classification. *NeuroImage, 178*, 445–460. https://doi.org/10.1016/j.neuroimage.2018.05.051

Timmermann, L., Seijo, F., et al. (2015). Multiple-source current steering in subthalamic nucleus deep brain stimulation for Parkinson's disease (the VANTAGE study): A non-randomised, prospective, multicentre, open-label study. *Lancet Neurology, 14*(7), 693–701.

Vergara, J. (2014). A review of feature selection methods based on mutual information. *Neural Computing and Applications, 24*, 175–186. https://doi.org/10.1007/s00521-013-1368-0

Vitek, J. L., Jain, R., Chen, L., Tröster, A. I., Schrock, L. E., House, P. A., & Starr, P. A. (2020). Subthalamic nucleus deep brain stimulation with a multiple independent constant current-controlled device in Parkinson's disease (INTREPID): a multicentre, double-blind, randomised, sham-controlled study. *The Lancet, 19*.