UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

DESIGN OF AN IMAGE-BASED SCREENING CLASSIFIER FOR EARLY DETECTION OF PARKINSON'S DISEASE COMPATIBLE PATIENTS USING [18F]PR04.MZ PET TRACER

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES

MARÍA JOSÉ JIMÉNEZ AGUILAR

PROFESOR GUÍA:

SEBASTIÁN ALEJANDRO RÍOS PÉREZ

MIEMBROS DE LA COMISIÓN:

JORGE FELIPE SILVA SANCHEZ

DENIS ROLAND SAURE VALENZUELA

DIEGO ANDRÉS MARTINEZ CEA

SANTIAGO DE CHILE

DICIEMBRE 2021

Abstract

Parkinsonism defines certain symptoms compatible with people who suffer from Parkinson's disease (PD), Essential Tremor, among other neurological disorders. Currently, PD is the second most important neurodegenerative pathology after Alzheimer. Today, the efforts of science are focused on knowing its causes and looking for biomarkers that facilitate an early diagnosis, as well as monitoring the evolutionary process of the disease due to the loss of dopamine in the substantia nigra pars compacta (SNpc) [2].

In this context, clinicians can identify a patient compatible with PD by attending to motor and nonmotor manifestations. According to [5], motor symptoms remain as the core feature for diagnosis, defined as bradykinesia plus rest tremor or rigidity. Since a vast experience is required to give an accurate diagnosis, which still might not be error free, making it early may not always be possible, but certainly necessary. Once first symptoms have slightly appeared on stage one, the progression throughout all five stages of the disease is unstoppable. Although there's no cure, early treatments increase the chances of slowing dopamine deficiency, and therefore, delay the appearance of more severe symptoms.

The possibility of delaying the set of symptoms that characterizes last stages makes more urgent an early diagnosis, but doctors prefer to wait until the symptoms allow an accurate and differential diagnosis. Studies have proved that PD is directly related to the dopamine loss in the dopaminergic system, and therefore, detectable through molecular imaging techniques such as positron emission tomography (PET). The question is: what´s the probability of accuracy of pre-diagnose a patient compatible with PD by detecting the loss of dopamine neurons using machine learning (ML) and deep learning (DL) approaches on neuroimaging, that is, imaging processing?

With this finality, we used a sample of PET/CT images with [18F]PR04.MZ, a new PET imaging tracer, in order to diagnose the illness, and discussed the possibility of presenting a better precision using machine learning versus a movement-disorder specialist. The development of five models, Support Vector Machine (SVM), Random Forest, Logistic Regression, K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN), allowed an extensive comparison between traditional classifiers. Performances were greater than 96% for ML algorithms, and over 98% for the DL model, proving [18F]PR04.MZ tracer allows image-based algorithms have high precision for the identification of PD patterns.

Key Words*: Parkinson Disease, Machine learning, Deep learning, Positron Emission Tomography, [18F]PR04.MZ radiotracer.*

Resumen ejecutivo

El parkinsonismo define aquellos síntomas compatibles con personas que padecen la enfermedad de Parkinson (EP), Temblor Esencial, entre otros trastornos neurológicos. Actualmente, la EP es la segunda enfermedad neurodegenerativa más importante después del Alzheimer. Hoy en día, los esfuerzos de la ciencia buscan conocer sus causas y estudiar biomarcadores que faciliten un diagnóstico precoz, así como en el seguimiento de la pérdida de dopamina en la sustancia nigra pars compacta (SNpc) [2].

En este contexto, los médicos pueden identificar un paciente compatible con EP atendiendo las manifestaciones motoras y no motoras. Según [5], los síntomas motores siguen siendo la característica central del diagnóstico, definida como bradicinesia más temblor de reposo o rigidez. Dado que se requiere una vasta experiencia para dar un diagnóstico preciso, que aún así podría no estar libre de error, no siempre es posible hacerlo en su fase más temprana, pero ciertamente es necesario. Una vez que los primeros síntomas han aparecido levemente en la etapa uno, la progresión a lo largo de las cinco etapas de la enfermedad es imparable. Aunque no existe cura, los tratamientos tempranos aumentan las posibilidades de retrasar la aparición de síntomas más graves.

La posibilidad de retrasar el conjunto de síntomas que caracterizan las últimas etapas hace más urgente un diagnóstico precoz, pero los médicos prefieren esperar hasta que los síntomas permitan un diagnóstico preciso y diferencial. Los estudios han demostrado que la EP está directamente relacionada con la pérdida de dopamina en el sistema dopaminérgico y, por tanto, detectable mediante técnicas de imagen molecular como la tomografía por emisión de positrones (PET). La pregunta es: ¿cuál es la probabilidad de precisión de pre-diagnosticar a un paciente compatible con EP al detectar la pérdida de neuronas dopaminérgicas utilizando enfoques de aprendizaje automático (ML) y aprendizaje profundo (DL) en neuroimagen, es decir, procesamiento de imágenes?

Con esta finalidad, utilizamos [18F]PR04.MZ PET, un nuevo trazador de imágenes PET, y discutimos la posibilidad de presentar una mejor precisión de diagnóstico usando aprendizaje automático versus la de un especialista en trastornos del movimiento. Para esto desarrollamos cinco modelos: máquina de vectores de soporte (SVM), bosque aleatorio, regresión logística, K-vecinos más cercanos (KNN) y red neuronal artificial (ANN), lo que permitió comparar el desempeño de distintos clasificadores tradicionales. Los rendimientos fueron superiores al 96 % para los algoritmos aprendizaje automático (ML) y superiores al 98 % para el modelo de aprendizaje profundo (DL), lo que demuestra que el trazador [18F]PR04.MZ permite que los algoritmos basados en imágenes tengan una alta precisión para la identificación de patrones compatibles con la EP.

Acknowledgements

I would like to thank God, to whom I owe everything and who is the goal of all my aspirations.

I would like to thank my parents, who encouraged me in every step of this adventure.

I would like to thank my research supervisor Sebastian, who guided me throughout this work; words cannot express my gratitude for all that I learned under your guidance.

I would like to thank AC3E (Centro Basal ANID FB0008) for supporting this thesis development.

I would like to thank PositronMed and its team, who collected and provided the PET scans from patients with written informed consent for this investigation.

Thank you to all the people who accompanied me at this stage of my life.

Table of contents

Chapter 1

Introduction and Context

1.1 Background

Parkinson's Disease (PD) is the second most important neurodegenerative pathology after Alzheimer. A study of the global burden of PD [47] revealed that among Latin American countries, Chile leads the ranking of Parkinson's disease prevalence (Figure 1.1) [45, 47], which may be related with an aging population and an increasing life expectancy.

Considering this, PositronMed's team of technological innovation and clinical research collected two datasets of a balanced group of patients with written informed consent for this investigation. The first one described a slice with the best brain uptake, and the second one included the 3D normalized neuroimaging of a PET/CT scanner (Biograph mCT Flow,

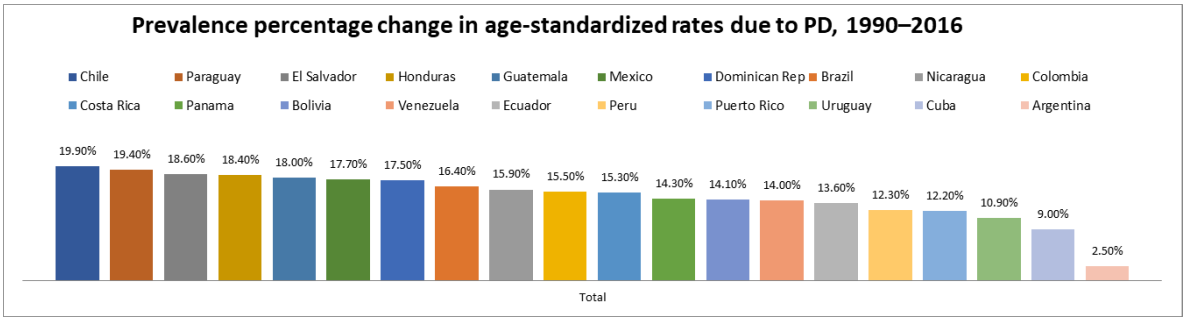


Figure 1.1: Prevalence percentage change due to Parkinson's disease in Latin America between 1990 and 2016. Data adjusted for age, extracted from [47].

Siemens Healthineers, Erlangen, Germany) [1] using [18F]PR04.MZ radiotracer [15].

This dataset was the input used to achieve the final objective: classify [18F]PR04.MZ PET/CT images of 204 subjects (aged 61.54 ± 12.9 years, 118 males, 86 women) between healthy or PD-compatible using machine learning, an artificial intelligence (AI) application widely used in the past decade in the medical imaging field [1]. In order to achieve this, five different classifiers were implemented, and their performance compared. These were Support Vector Machine (SVM), Random Forest, Logistic Regression, k-Nearest Neighbor (k-NN) and Artificial Neural Network (ANN). They were trained with the classification results of the blinded diagnosis of three specialists, who considered the visual and qualitative evaluation of the data to separate the scans in two groups:

- Normal controls (NC), describing symmetric binding in the striatal region.
- PD compatible, where rostro-caudal gradient (RCG) reached high values.

Expect results of this image-based ML to be limited by the conditions and characteristics of the sample previously described.

1.1.1 Diagnosis

Nowadays, the diagnosis of PD is defined with clinical criteria [4]. Bradykinesia, combined with cogwheel rigidity and/or rest tremor, are the main motor symptoms of this disease. Differential diagnosis require both identifying the parkinsonism and defining it's cause, since typical features plus atypical ones, or even non-motor symptoms, can be caused by many other syndromes. While first stage PD patients can present tremor and other movement disorders on one side of the body, fifth, and last, stage patients may require a wheelchair, nursing is needed for every activity, and frequently present non-motor symptoms as dementia.

Time has shown formal diagnostic PD guidelines improves diagnostic accuracy. The International Parkinson and Movement Disorder Society (MDS) published back in 2015 a methodology that enables reproducibility across medical centers and inexperienced clinicians [5]. Studies have validate this criteria [6, 11], proving high performance (accuracy $>$ 90%).

[1] The three-dimensional dataset wasn't sent in its entirety, so augmentation method was used to offset the incomplete dataset.

Despite the MDS-PD criteria relies the diagnosis on the presence of motor syndromes, the challenge of a valid diagnosis at early stages of PD is pushing the use of neuroimaging for a differential diagnosis categorization. Within the variety of imaging techniques finding patterns in presynaptic striatal binding using neuroimaging tracers, studies have shown they can improve diagnosis certainty and reinforce clinicians decisions [7].

Furthermore, the exponential development of machine and deep learning applied in neuroimaging diagnosis has shown advantageous results [8, 9, 10]. Artificial intelligence approaches allows statistical estimation of the resemblance between the training set and an input, leading to predictions based on classification.

1.1.2 Terminology

PD	Parkinson's Disease: progressive neurodegenerative disease without a defined cause, known to respond to dopaminergic drugs.
PET	Positron Emission Tomography: imaging technique which allows to quantify the number of existing dopaminergic neurons in PD patients.
DAT	Dopamine Active Transporter: a transmembrane protein used as tracer in PET imaging [3].
DICOM	Digital Imaging and Communications in Medicine: international standard used to manipulate medical imaging information
$[18F]$PR04.MZ	PET tracer of dopaminergic deficit [15].

1.2 Goals

1.2.1 General Objectives

The objective of this thesis is to discriminate healthy subjects from early compatible PD patients with image-based ML algorithms. This purpose is combined with PositronMed's research work with the [18F]PR04.MZ biomarker as a highly selective PET tracer for imaging dopamine transporters (DAT). Still, both investigations intend to generate a method using artificial intelligence to facilitate the diagnosis for patients in the earliest stage of the disease.

1.2.2 Specific Objectives

- Define and follow the pipeline for image processing, and to guarantee a high-quality dataset. The algorithm created most include preprocessing the images before using them for training and classification.
- Create a first image-based ML approach using a 2D-selected slice; followed by a second approach that automates that selection in a 3D database.
- Construct and train various classification algorithms and compare their results using standardized metrics.
- Select the best model with the help of performance analysis.
- Present an statistical estimation of the resemblance between training and testing sets, that is, the probability of belonging to one group or the other.

1.2.3 Expected Results

For the volumetric data, the research procedure should play the role of a radiologist by using an unsupervised preprocessing image algorithm. The results of this process will submit the input for the image supervised algorithm classifier, intending the role of a neurologist by using the optimized dataset to identify PD subjects.

The investigation on its totality should allow us to discuss the specificity and sensitivity of the machine learning algorithm in comparison to the specialized radiologist-neurologist diagnose, using [18F]PR04.MZ PET images.

1.3 Structure of the thesis

Following Chapter 1, Chapter 2 compile the related work research, setting the AI-gold standard for this thesis. We will also add PositronMed's related publications, pointing out their achievements.

Since the whole investigation was possible thanks to the contribution of PositronMed's team, Chapter 3 will describe to the initial steps of this project. It includes the comparison

between Google Cloud MR Vision algorithm used by one of PositronMed's member, with a set of hand-craft algorithms created for the same finality. This chapter describes the first approach of the procedure that will be used with the tough data.

Motivated by the first approach, Chapter 4 provides the description of the 3D image processing, using as baseline the previous algorithms and adding the corresponding modifications for tough data.

Chapter 5 collects the results of each model, clarifying the performance of accuracy, precision, recall and F1-score achieved in all cases, and pointing the best model according to this indicators.

Finally, Chapter 6 presents this work conclusions, limitations and future work lines.

Chapter 2

State-of-the-art for neuroimaging CADs

2.1 Methodology

Using a guideline may help staying within the bounds proposed for this research. Figure 2.1 provides the general structure through which the investigation was carried out.

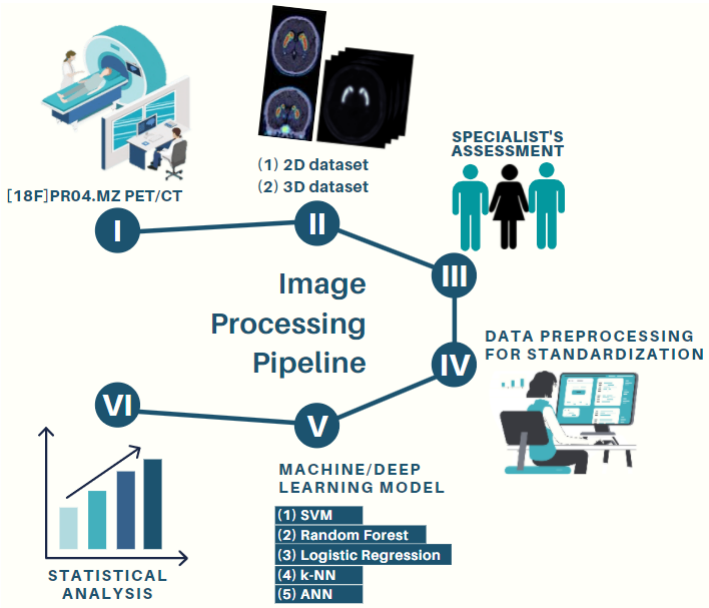


Figure 2.1: Image Processing Pipeline described by: brain scan acquisition through [18F]PR04.MZ PET technique, two different databases (2D/JPG and 3D/DICOM) which will be processed differently, specialists labeling using predetermined criteria, data preprocessing for standardization (unfolded in subsection 2.1.1), ML and DL modeling (unfolded in subsection 2.1.2) and finally the statistical estimation for predictions.

This general framework leads to discriminating healthy subjects from PD, going through the six steps described below:

1. First of all, [18F]PR04.MZ PET/CT images of 204 subjects (aged 61.54 ± 12.9 years, 118 males, 86 women) were collected for the purposes of this research. Between them, 96 were healthy subjects and the rest were PD compatibles (dopamine deficient).

2. There are two different format datasets. The first one includes 2D JPG images per subject, preselected by a radiologist from the raw data, which is the second dataset: 91 slices in rostro-caudal direction per subject. Both of them will be processed differently.

3. Three specialists blindly assessed the images, assigning to each sample two different labels according to three criteria: if the scan is either normal (0) or abnomal (1), if it shows asymmetric binding (1) or not (0) and if the scan shows a high rostro-caudal gradient (1) or not (0). The agreement of at least two of the three clinician's statements set the final label of the scan as Normal (0) or PD compatible (1).

4. The data preprocessing for standardization depends on the dataset used as input. First and second approaches will be destined to chapters 3 and 4 respectively. A more descriptive explanation will be found in subsection 2.1.1.

5. To found out the behavior patterns of the labeled images, the resulting datasets from the previous step will train five different algorithms: Support Vector Machine (SVM), Random Forest, Logistic Regression, k-Nearest Neighbor (KNN) and Artificial Neural Network (ANN). A more descriptive explanation will be found in subsection 2.1.2.

6. Finally the statistical estimation for predictions of binary classes will be done using the following indicators: accuracy, precision, recall and F1-score.

		Predicted	
		0	1
True	0	TN	FP
	1	FN	TP

$$\frac{TP+TN}{TP+TN+FP+FN} \quad \text{(Accuracy)} \quad (2.1)$$

$$\frac{TP}{TP+FP} \quad \text{(Precision)} \quad (2.2)$$

$$\frac{TP}{TP+FN} \quad \text{(Recall)} \quad (2.3)$$

$$2 * \frac{precision * recall}{precision + recall} \quad \text{(F1-score)} \quad (2.4)$$

Where equation (2.1) measures the percentage of samples the model has correctly classified, (2.2) measures the quality of the ML model in classification tasks rated according to the number of TP in a total of PD-compatible predictions, (2.3) reports the amount of PD-compatible subjects the ML model is able to identify and (2.4) combines both (2.2) and (2.3) to measure the performance of precision and recall among various models.

2.1.1 Data pre-processing pipeline

Data preprocessing pipeline is illustrated in Figure 2.2, and seeks the standardization of different datasets. 2D database in the first approach was used as an approximation of what we want to achieve in the second one.

Even though the final goal was the same, the different type of input in each approach marked a different course of operations. In general terms, while dataset (1) seeks the selection of the ROI (region of interest) in 2D images, the striatal region on each image, dataset (2) focuses on the VOI (volume of interest), a volumetric scan of 10 final slices.

Since dataset (2) wasn't sent complete, we used an image augmentation method to increase the number of full-head scans. The data-augmentation technique consisted on a flip of the right and left hemispheres of the brain for each PET, which double the size of the dataset and improved its disadvantageous position against dataset (1). Out of the flip, everything else in the images remain the same. The full process and resulting dataset will be detailed in Chapter 4.

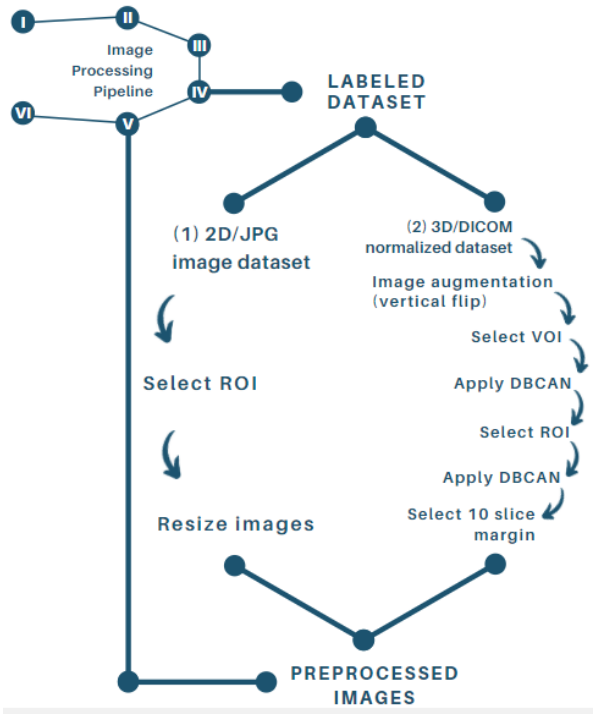


Figure 2.2: Data preprocessing for standardization pipeline

2.1.2 Machine Learning model pipeline

ML and DL model pipelines are illustrated in Figure 2.3. Starting with the output of step 4, a selected percentage of the preprocessed images flow through the pipeline to build a trained model. The proposed image-based algorithms include Support Vector Machine (SVM), Random Forest, Logistic Regression, k-Nearest Neighbor (KNN) and Artificial Neural Network (ANN), all of them reinforced with validation techniques such as ten-fold cross validation over a split dataset.

The methods and techniques used for each dataset will be made explicit in Chapters 3 and 4. All of these models were finally used to classify the remaining images in order to test the model performance in step 6. Results were presented in Chapter 5 of this thesis.

However, this framework remains in general lines, since machine learning and deep learning PET image processing workflows differ in feature extraction, selection and classification techniques. While ML algorithms clearly separate this steps, in ANN feature extraction and selection is an automatic one-step process [10] and will be clarified later.

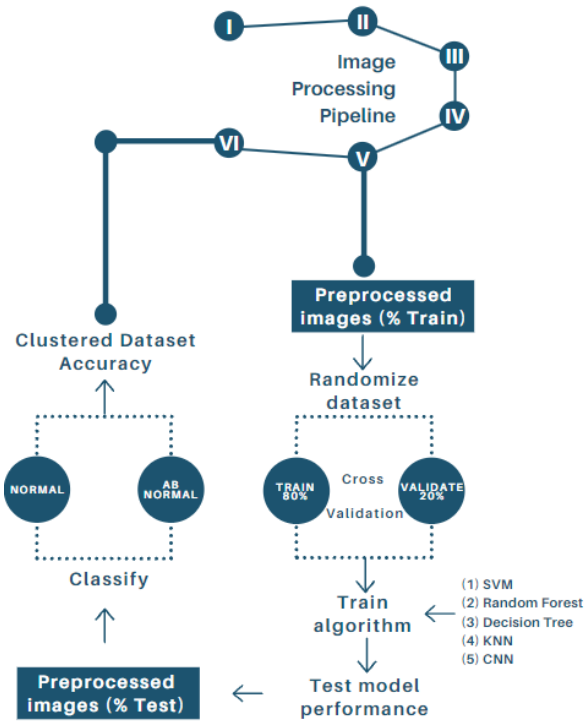


Figure 2.3: Machine and Deep Leaning model pipeline

2.2 Related machine learning algorithms

The research work of this section aims to provide a general view of the artificial intelligence implementation as a potential resource of pathology diagnosis. There's plenty of studies that support the diagnosis and analysis of diseases through image processing with machine and deep learning. That's why in this section we summarize the state of art in computer aided diagnosis (CAD) systems in the latest years. To target the publications related with this research, we used the boolean search strings shown in Table 2.1, in the advanced search engine of four journals: PubMed, Springer, ScienceDirect and IEEE Xplore.

Table 2.1: Boolean search strings used to retrieve the publications on four journals: PubMed, Springer, ScienceDirect and IEEE Xplore.

Journal database	*Boolean search string*
PubMed Springer ScienceDirect IEEE Xplore	'machine' **AND** 'deep' **AND** 'learning' **AND** 'neuroimaging' **AND** 'machine learning' **AND** 'deep learning'

Although the use of artificial intelligence for medical applications began several years ago, we will focus on finding those that were published in the last 10 years. There are many symptoms that may describe PD, that's why studies use different types of database to make an accurate diagnosis: speech data, video recordings, force sensors, handwriting dynamics, among others [12]. As said before, we are interested in CADs, but above all, those who use neuroimaging as source for the diagnosis of diseases. Publications that did not include the use of neuroimaging for computational methods were taken out of the count. The result of this review is illustrated in Figure 2.4.

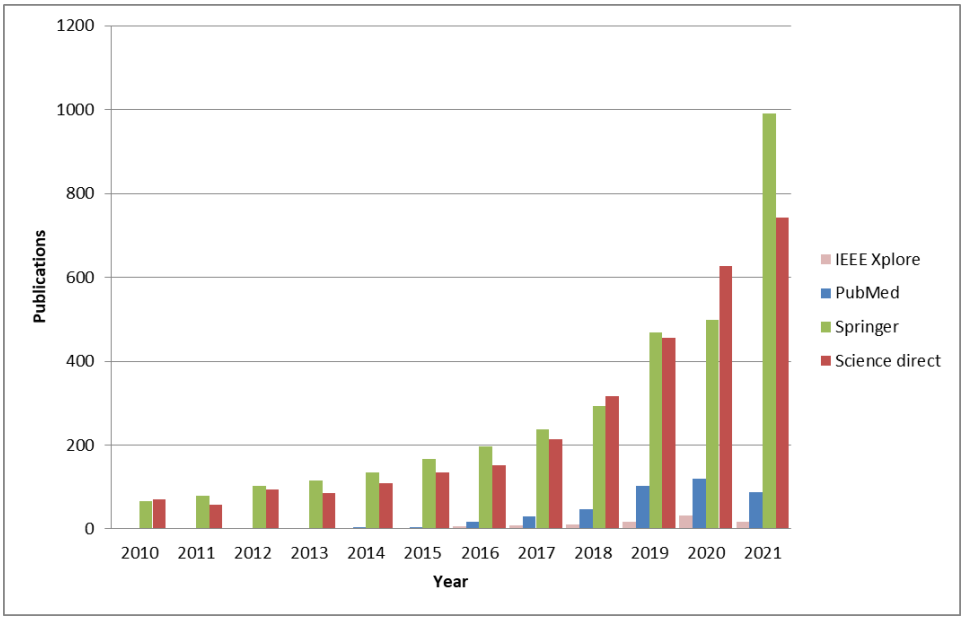


Figure 2.4: Number of publications in PubMed, Springer, ScienceDirect and IEEE Xplore per year (January 2010 to October 2021) using the boolean search strings shown in Table 2.1.

The graph demonstrates the exponential growth of this research area, specially in the last 5 to 6 years. This prompted us to select those studies, in the range of 2015-2021, whose aims were similar to ours. Table 2.2 summarizes the results of this research, presenting approaches that seek to discriminate PD patients from normal controls (NC) using artificial intelligence models to neuroimage data. In order to compare them with the achievements and contributions of the present thesis, we extract the dataset size, radiotracer applied, machine or deep learning classifier and best results. The table also includes a column of comments and achievements, which expand other research proposals for future lines of work.

Table 2.2 puts on evidence the uniqueness of this study, since we use a new radiotracer: [18F]PR04.MZ. The research also helped to detect the classifiers that perform the best with this type of dataset, as well as to give an idea of which is the gold-standard in algorithmic terms. Throughout this research, we found out accuracy, sensitivity, specificity, among other parameters may achieve high levels using complex algorithmic structures such as artificial neural networks. Still, there are classifiers reliable enough that may provide interpretability in an algorithmic approach. The combination of simple and complex classifiers here presented validates this thesis as a complete research work for the previously established objectives.

2.3 [18F]PR04.MZ PET tracer

Improving tools that support the diagnosis of neurodegenerative diseases is a challenge that include scientific advancement in radiotracers for the visualization and identification of structural and functional changes in the brain, which may describe neurodegenerative features. Traditional radiotracers used in neuroimaging techniques, such as SPECT and PET, are employed to assessed diagnostic decisions, quantify dopaminergic activity, track disease progression and treatment monitoring [7, 14].

Among other neurodegenerative diseases, Parkinson's disease (PD) is below Alzheimer's disease (AD), Huntington's disease (HD) and Multiple System Atrophy (MSA), as a focus for investigation. Compared to AD, studies on PD brain imaging tracers are barely moving forward [13]. The benefits of finding appropriate radiotracers can be decisive early-stage diagnosis of PD and avoiding misdiagnoses.

A new PET tracer, [18F]PR04.MZ, has shown very high selectivity, specificity and affinity for DAT uptake [15]. On the present investigation, the reader will find the first approaches on computer-aided diagnosis of PD-compatible using [18F]PR04.MZ PET tracer. [18F]PR04.MZ has shown in previous studies excellent properties of imaging and quantification outcomes [15, 16, 17]. The purpose of the following chapters is to validate [18F]PR04.MZ PET as a new radiotracer which, combined with computational methods, can provide an accurate diagnosis.

Table 2.2: State of art reference results of PD detection comparing dataset size, radiotracer applied, ML classifier and best metric results. As baseline, all studies aim to classify PD patients from NC using AI models to neuroimage data (n_databases = 21) between 2015-2021.

Reference	*Year*	*Dataset size*	*Radiotracer*	*ML classifier*	*Best result*	*Comments and achievements*
[26]	2021	42	[11C]raclopride PET	SVM	Accuracy=0.967	Included patients with multiple system atrophy (MSA n = 8), corticobasal degeneration (CBD n = 6), and dementia with Lewy bodies (DLB n = 5), vs HC and achieved Acc=0.921.
[27]	2020	202	99mTc-TRODAT-1 SPECT	AlexNet	Accuracy=0.825, Recall=0.753 Precision=0.874 F1-score=0.809	For each subject, CNNs analyzed and compared two input image types: grayscale and pseudo color, both with five selected slices. Other CNNs used: GoogLeNet, Residual Neural Network, VGG, and DenseNet.
[28]	2020	642	[123I]FP-CIT SPECT	CNN (VGG16)	Accuracy=0.952 Sensitivity=0.975 Specificity=0.909	Used Local Interpretable Model-Agnostic Explainer (LIME) to provide a plausible reason for the prediction of PD patients, making the model more interpretable.
[29]	2020	406	3T T1-weighted MRI	3D CNN	Accuracy=0.9529 Recall=0.943 Precision=0.927	MRI may be used with or without a biomarker to capture internal organs, but if it does, it is different from the one used for computed tomography (CT).
[36]	2019	408	T2-weighted MRI	ResNet50	Accuracy=0.886	Idem. Used a slice-based subset of randomly picked slices.
[37]	2019	204	MRI	SSAE	Accuracy=0.8524 Sensitivity=0.681 Precision=0.9583 AUC=0.8604 F1-score=0.7064	This study extracted features of brain regions from MRI data to obtain 116-dimensional feature matrix of gray matter (GM), white matter (WM) and mean diffusivity (MD). It also used SSAE to identify scans without evidence of dopamine deficit (SWEDD).
[30]	2019	189	[123I]FP-CIT SPECT	SVM	Accuracy=0.94	Used morphological feature extraction through parcellation of the striatum before applying SVM.
[31]	2019	550	[123I]FP-CIT SPECT	SVM	Accuracy=0.784	This study also aimed to identify SWEDD patients as a PD with motor symptoms or PD with nonmotor symptoms or just a control lookalike. Other ML classifiers used: RF, multilayer perceptron (MLP) NN, LR and K-NN.
[32]	2019	578	[123I]FP-CIT SPECT	SVM	Accuracy=0.929	Studied general degenerative parkinsonian syndromes: PD, multiple system atrophy-parkinsonian type, progressive supranuclear palsy and corticobasal syndrome. Also used striatal VOIs uptake, VOIs asymmetry indices and caudate/putamen (C/P) ratio as input for SVM.
[33]	2019	350	18F-(FDG) PET	GLS-DBN	Accuracy=0.90 Sensitivity=0.96 Specificity=0.84 AUC=0.912	Applied Locally linear embedding (LLE) to reduce the dimensionality of pre-processed PET data in all subjects and used two different batches of data to validate the model's performance and test model's robustness on other datasets.
[34]	2019	645	[123I]FP-CIT SPECT	CNN	Accuracy=0.972	Proved that a high-dimensional CNN can be trained to deal with variable image characteristics.
[35]	2019	182	18F-(FDG) PET	SVM radial basis	Accuracy=0.9126 Specificity=0.9327	Used 3 kernel (linear, sigmoid, and radial basis) functions and random forest classifier to detect feature generalization ability and classification reliability. It proved that radiomic features provide additional classification information for the traditional voxel values information.
		48	18F-(FDG) PET	SVM radial basis	Accuracy=0.9018 Specificity=0.9205	
[38]	2018	642	[123I]FP-CIT SPECT	ALEXNET3D	Accuracy=0.941 ROC-AUC score=0.984	Showed that sufficiently complex model, i.e. three-dimensional version of the ALEXNET, can effectively account for spatial differences.
[39]	2018	652	[123I]FP-CIT SPECT	SVM	Accuracy=0.979 Sensitivity=0.98 Specificity=0.976	Showed that, individually, the length of the striatal region, the putaminal binding potential and the striatal binding potential, generated high Acc. Other classifiers used: KNN and LogReg.
[40]	2017	624	[123I]FP-CIT SPECT	CNN-PD Net	Accuracy=0.96 Sensitivity=0.942 Specificity=1.0	The study compared CNN-PD Net versus visual analysis combined with conventional quantification. The quantification method calculated regional DAT binding ratio (BR) in target regions (putamen/caudate and occipital cortex), segmented through automated anatomical labeling (AAL) template.
[41]	2017	715	[123I]FP-CIT SPECT	SVM	Accuracy=0.9729 Sensitivity=0.9737 Specificity=0.9718	The study also included SWEDD subjects to compute shape- and surface-fitting-based features. Other classifiers used: NB, boosted trees and RF.
[42]	2017	304 local DB	[123I]FP-CIT SPECT	SVM	Accuracy=0.88-0.92	Compared ML methods with semi-quantification approaches based on striatal binding ratios (SBRs) from both putamen, with and without consideration of the caudates.
		657 PPMI DB	[123I]FP-CIT SPECT	SVM	Accuracy=0.95-0.97	
[43]	2017	172	T1-weighted MRI	Multi-kernel SVM	Accuracy=0.8578 Sensitivity=0.8764 Specificity=0.8779	Multilevel-ROI-features-based machine learning method proved to be superior in characterizing the brain structural alterations and brain connections for PD patients.
[44]	2015	604	[123I]FP-CIT SPECT	EPNN	Accuracy=0.986	It also classified PD patients vrs SWEDDs and obtained a classification accuracy of 92.5% using EPNN. Other classifiers used: probabilistic neural network (PNN), SVM, k-NN and DT.

Chapter 3

First ML and DL approach

3.1 Dataset description

A total of 204 patients (118 men, 86 women, 61.6 ± 12.9 years) referred for routine WB PET/CT scanner (Biograph mCT Flow, Siemens Healthineers, Erlangen, Germany) with [18F]PR04.MZ were included for this first approach (Table 3.1). Using the same split from a previous experiment performed by PositronMed, the training set had a size of 129 2-dimensional PET/CT images, each one of them selected by a radiologist from a 3-dimensional full scan. The remaining 75 images were saved for the testing set. This dataset was given in JPG image format.

For the purposes of this research, three specialists blindly assessed the images, assigning to each scan two different labels according to three criteria:

(a) If the scan was either normal (0) or abnormal (1).

(b) If the scan showed asymmetric binding (1) or not (0).

(c) If the scan showed a rostro-caudal gradient (1) or not (0).

According to the previous criteria, each scan was classified as compatible with PD (1) or not (0). Comparing the results of all three specialists, each scan was finally labeled as normal (0) or abnormal (1), requiring the agreement of at least two of them. This visual and quantitative imaging interpretation is going to be the gold standard of the further analysis.

Table 3.1: Group composition and demographic details of the data used in this work.

	Normal Training	*Abnormal Training*	*Normal Testing*	*Abnormal Testing*
Amount	54	75	39	36
Age	63.3±13.9	59.0±13.9	60.9±11.5	65.3±9.5

3.1.1 [18F]PR04.MZ PET images

No studies have been made about machine learning algorithms for neuroimages using [18F]PR04.MZ PET tracer. PositronMed research group has worked to deepen the qualities of this new radiotracer, and the present investigation contains the first approaches of creating an effective automated computer-aided diagnosis of PD.

After the tracer-injection and the PET scan, the raw projection data is transferred and reconstructed by a medical technologist, who then uses PMOD software for stereotactic normalization using a normal space template image. After the transformations are applied to the PET, final pre-processed images are exported in the two formats used for this investigation:

(a) 2D image, which is a pseudo-color image pre-selected among the full head scan.

(b) 3D image, a full head scan in DICOM format with 91 x 109 x 91 resolution.

This chapter describes the approach that focused on the first dataset, while the procedures in Chapter 4 were performed using the second one. The region of interest (ROI), putamen and caudate, was outlined before exporting for image processing, as seen in Figure 3.1.

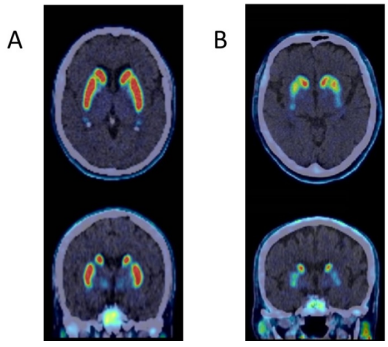


Figure 3.1: Axial and coronal section images at the level of the normalized striatum in A. Normal Control. B. Dopamine deficit patient.

3.2 Google Cloud MR Vision AI platform

The trial version of Google platform was used to train a model that classifies the images of the dataset previously described. The procedure taken for the task didn't exceed uploading the 129 2D images, respectively labeled, and start running the software for training.

The training time was of about approximately thirteen hours, and the information about the model used very limited. The results of the Google platform algorithm were used to compare the ones obtained from hand-craft computer methods.

3.3 Machine learning classification models development

Entering into the specifications of the procedure illustrated in Figure 2.2, extra margins from each pseudo-color image were removed, automating the cropping for the whole set, to select the target-region. Then, we resized the scans into a resolution of 40 x 80, which homogenizes the input of the model. Since this set contains only 2D images, the preprocess only contemplates these steps.

For the development of this investigation, a total of 129 PET normalized images were used to train five of the most common ML algorithms, and detect the behavior patterns between the labeled images of each group: Normal Control (NC) or PD compatible. These were: Support Vector Machine (SVM), Random Forest, Logistic Regression, K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN). Concretely, we used scikit-learn functions (http://scikit-learn.org) for ML models and tensorflow-keras (https://www.tensorflow.org/guide/keras/sequential_model) for the ANN deep learning model, and implemented them in Python. The hyper-parameter optimizer applied to each model will be describe respectively. From the total training data, and using the same distribution as in the Google Cloud platform algorithm, we separate 63% of the data for training and 37% for verification.

Unfolding the ML/DL Model, Step 5 in Figure 2.1, the pipeline is illustrated in Figure 2.3. Firstly, 63% of the preprocessed images were input. Then the set was randomized, and split into training (80%) and validation (20%). Applying cross validation method, each of the five train models were executed. Details of each algorithm will be described in the next section.

We measured the performance of each model, and then predicted the probability of the testing dataset to be NC or PD compatible. The final step of the ML model pipeline match with step 6 of the main pipeline methodology: the statistical analysis.

3.3.1 Support Vector Machine (SVM)

Support Vector Machine (SVM), first proposed in 1995 [22], is a technique that seeks to find a hyperplane that allows a binary classification of data, maximizing the margin between the hyperplane and the classes. [20, 21] are some examples of how SVM has been used to achieve effective diagnosis. The results there shown encourage us to try its performance on this experiment.

We selected GridSearchCV to optimize the parameters of our classifier estimator, including C, kernel and gamma, to find the best parameter grid across all possible parameter value combinations. Then, we applied the cross-validation method to avoid overfitting.

Applying this method to the SVM model, the best combination was retained, assessing its performance according to the precision, recall, F1-score and accuracy.

3.3.2 Random Forest

Random Forest, proposed in 1995 by Ho et al [24], is a ensemble method that fits decision trees into subsamples of the data set and averages them to improve predictive accuracy and control overfitting. This averaging method was also used in [23] as PD patients classifier, although with a different database type, and selected as the best performing classifier versus logistic regression and KNN. In this work, we will compare their performance based on the precision, recall, F1-score and accuracy of each model.

As in SVM, this classifier was also submitted into GridSearchCV to find the best parameter combination.

3.3.3 Logistic Regression

Logistic regression model makes possible binary classification transforming a linear regression value into a value between 0 and 1 with a logistic function. GridSearchCV was also applied to found out the hyper-parameters to use in the model. The resulting combination was C=10, penalty='l1', random_state=0, solver='liblinear'. Since we are using 'liblinear' solver, which favor us for having a small dataset, the training model is limited to one-vs-rest scheme.

The model performance will be measured according to the precision, recall, F1-score and accuracy.

3.3.4 k-Nearest Neighbors (k-NN)

K-Nearest Neighbor (KNN), a classical classifier, chooses the most similar "K" records, that is, the closest K neighborhood data points. Although it's commonly used in comparative research such as in previously mentioned [23], it will mainly help us to compare the performance of a simple algorithm versus others more complex applied here such as ANN.

3.3.5 Artificial Neural Networks (ANN)

Unlike previous ML algorithms, artificial neural networks (NN) are a more complex set of algorithms, that honoring his name, aims to imitate human brain neural networks. Known for their performance on image processing, ANN arrange layers in a way that captures the complexity of a pattern.

This algorithm feature extraction structure is composed of three convolutional layers, each one with a 3x3 kernel layer, a ReLU activation function, followed by one max pooling layer size of (2, 2) and one flatten with activation 'relu'. The classification layer included a dense layer, activation function 'sigmoid', that allows full connectivity between neurons in preceding and succeeding layers. Finally, compile the model using a 'binary_crossentropy' loss, 'adam' optimizer and measure the performance of the model with metric 'accuracy'. We validate each model with a cross-validation method, which also avoid overfitting due to the limited data sample. The structure is illustrated in Figure 3.2.

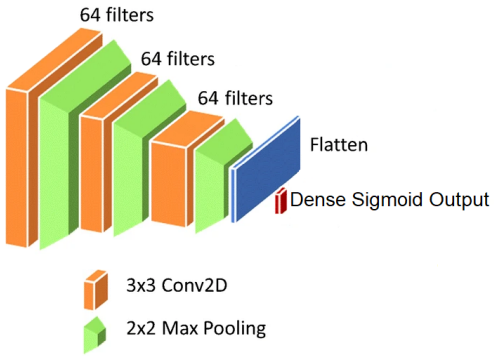


Figure 3.2: Detailed structure used in ANN model.

Beyond comparison indicators between the different algorithms, to select the best ANN model, we tested multiple layer combinations until the one that best behaved was selected. Accuracy and loss value metrics were used in this task. Graphing both indicators allowed us to compare how the model performed with the training and validation data after each iteration. On one hand, the graphical interpretation of the accuracy allow us to visualize the performance of the model, that is, how precisely it is to classify the pattern of an image. On the other hand, the loss value indicates the error that adds each training epoch. The convergence behavior of both graphs makes it easier to select the best model for later classification of the test set. After 20 epochs, the model converged as shown in Figure 3.3.

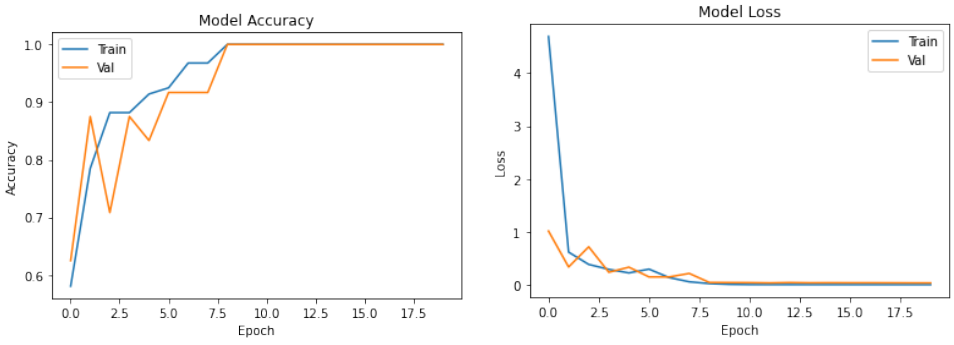


Figure 3.3: A. Accuracy value graph B. Loss value graph

3.4 Validation dataset results

As said before, the dataset was separated into 37% for test (75 subjects) and 63% (129 subjects) for train. From this 63%, we split into 80%-20% for validation. This second split, plus cross-validation application, provide an unbiased evaluation of a model fit on the training dataset, while adjusting the model hyper-parameters. The preliminary classification results are shown in Table 3.2.

$Classifier$	$Precision$	$Recall$	$F1-score$	$Accuracy$
SVM	0.97	0.95	0.96	96.15%
Random Forest	0.96	0.96	0.96	96.15%
Logreg	0.94	0.97	0.96	96.15%
k-NN	**1.00**	**1.00**	**1.00**	**100.0%**
ANN	0.95	0.97	0.96	96.15%

Table 3.2: Classification performance of five hand-craft classifiers measuring precision, recall, F1-score and accuracy with 2D validation dataset.

Still k-Nearest Neighbor (k-NN) is the best rated model, we will verify in Chapter 5 if it doesn't describes a overfitted model. The other models also present promising results, but as well as k-NN, all of them will be tested with an unseen dataset in Chapter 5 for statistical analysis.

Chapter 4

Second ML and DL approach

4.1 [18F]PR04.MZ PET 3D normalized scans

Three-dimensional [18F]PR04.MZ-PET/CT scans were acquired within 60-90 min after intravenously injected 5.16±1.03 mCi (191±38 MBq, range: 88.8-322 MBq), using a WB PET/CT scanner (Biograph mCT Flow, Siemens Healthineers, Erlangen, Germany). Raw data was corrected on TrueX software, reconstructed with an ordered subset expectation maximization (OSEM) algorithm, followed by post-reconstruction smoothing (Gaussian, 4 mm FWHM). PMOD (PMOD Technologies, Zurich, Switzerland) was used for spatial normalization to Montreal Neurological Institute (MNI) space. Manual inspection allows identification of motion-related and intensity technicalities. Final preprocessed images were 91 slices of 109x91 voxels[1] 1.59x1.59x1.5 mm3, exported on DICOM format. Among the many different languages that allow reading and analyzing DICOM data, we decided to use Python, due to it's functionality with neuroimaging, open source and free access [25].

Despite efforts to complete the 3D database, for different reasons only 144 were obtained out of a total of 204 full-head scans. We used an augmentation method, which was respectively added to the data preprocessing pipeline in Figure 2.2, to compensate the disadvantageous sample. For this end, Python Imaging Library (PIL) was essentially useful, whose parameter transpose(method=Image.FLIP_LEFT_RIGHT) returns a flipped copy of the im-

[1] Volumetric pixel.

age, keeping everything else the same. This step doubled the size of the dataset.

The augmentation step was followed by a procedure that aims to play the role of a radiologist by using an unsupervised learning algorithm to select from the DICOM data the best 10 on-a-row slices, from a scan with 91 x 109 x 91 dimension. The results of this process will submit the input for the supervised algorithm classifier, intending the role of a neurologist to discriminate PD subjects versus NC.

4.2 Unsupervised learning and preprocessing

Even though the basic pipeline is the same for both approaches (Figure 2.1), in this section, we are going to see how feature selection and extraction gets complicated in parallel with the input dataset.

The process suggested in Figure 2.2 aim the automated selection of 10 in-a-row slices from PD and NC subjects for the subsequent analysis. The selection of ”in-a-row slices” submits small ”brain boxes” for image classification. The image augmentation process was followed by five steps detailed below:

1. Select volume of interest (VOI). From the full head scan, 45% and 30% were removed from the upper and lower slices respectively. This standardized inner volume was used as a first filter to select the VOI. Figure 4.1 illustrates the process applied for all scans.

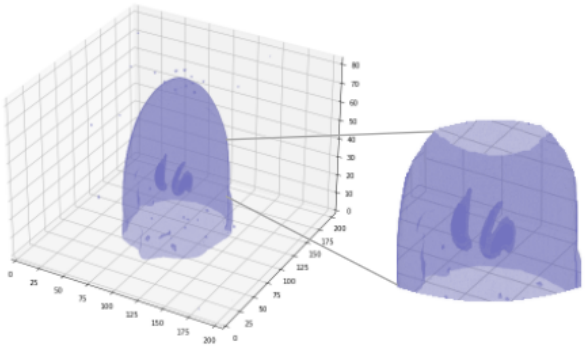


Figure 4.1: 3D superior-inferior margin removal.

2. Apply DBSCAN to select slices. Density-based spatial clustering of applications with noise (DBSCAN), is a technique of cluster building around an initial arbitrary point.

The parameter 'eps' marks the neighborhood radius of the point to analyze if it has 'min_samples', the minimum number of points that the cluster can contain. If not, it's considered as noise. We used sklearn.cluster.DBSCAN function (http://scikit-learn.org) to select, for each scan, the slices with the highest density. In other words, each cluster reflects the dopamine uptake in the target region of interest (ROI), i.e. anterior putamen, posterior putamen, anterior caudate and posterior caudate of both left and right brain hemispheres.

3. Crop selected slices. From the selected high-density slices, we reduced the target ROI for a second DBSCAN application. Since the initial point for cluster building is arbitrary, reducing the area into 20 x 14 pixels decrease computational cost.

4. Apply DBSCAN to select single slice. To improve the clusters quality, we increased both 'eps' and 'min_samples' in 300 and 10 times respectively for the cropped selected slices from the previous step. This allowed us to select the highest density slice, which will be the central slice of the output optimized "brain box".

5. Select 10 slice margin. Finally, we crop a "margin" around the previous step highest density slice, getting an output of 10 x 109 x 91 dimension scan.

The success of this process allowed reducing the input of the next step from 18564 to 2040 images, to split them into the training/validation set and the test set for the supervised learning classification.

4.3 Supervised learning classification models development

As in the first approach, we select four ML algorithms and one ANN for image exploration. The biggest difference in this second approach is the input dimensions. The output dataset from the unsupervised learning algorithm submitted a three-dimensional array. For each subject, we reduced the 10 x 109 x 91 scan down into a vector of pixels.

4.3.1 ML models

Machine learning models used the grid search technique for hyper-parameter optimization, and applied the cross-validation to avoid overfitting. Table 4.2 summarizes the results from the train/validation dataset. Comparison between the different models will be made thought the indicators used up to now: Precision, Recall, F1-score and Accuracy.

4.3.2 DL model

To select the best combination of hyperparameters, we used Ax (Adaptive Experimentation Platform), https://ax.dev, that simplifies the search an optimal neural network configuration. The metric that we defined in the network compile seeks a 'binary_accuracy' optimization. AX library .get_next_trial () allows to iteratively create a neural network with a combination of parameters for 25 experimental trials. Finally, ax_client.get_best_parameters() gets the best set of parameters, the structure of which is summarized in Table 4.1.

Hyperparameter	*Best combination*
'learning_rate'	0.000402981024536774
'dropout_rate'	0.013573833052722957
'num_hidden_layers'	1
'neurons_per_layer'	17
'batch_size'	64
'activation'	'tanh'
'optimizer'	'rms'
'keras_cv'	1.0

Table 4.1: Summary of the best set of parameters for neural network binary accuracy optimization.

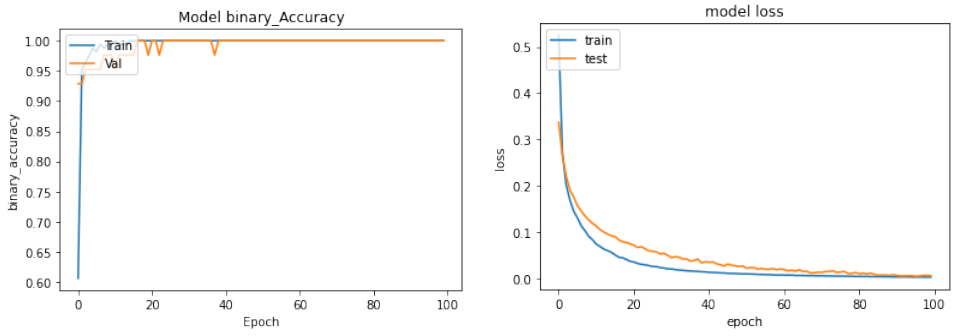


Figure 4.2: Neural Network Train/Validation Model: A. Accuracy value graph. B. Loss value graph.

We carry out 10-fold cross-validation for the best parameter combination model. Cross validation method, a data partitioning strategy, avoids overfitting by creating a generalized model, adaptive to unseen data. After 100 epochs, our model converged into the best val_binary_accuracy between all trials, as shown in Figure 4.2. The monotonously decreasing loss behaviour indicates that the model is not overfitted and therefore, the testing dataset is classifiable.

4.4 Validation dataset results

The results we got from the second approach process are summarized in Table 4.2. As in Chapter 3, all of the models are highly qualified and will be tested for the final classification in the next chapter to verify overfitting.

$Classifier$	$Precision$	$Recall$	$F1-score$	$Accuracy$
SVM	0.94	0.93	0.93	93.48%
Random Forest	1.00	1.00	1.00	100.0%
Logreg	1.00	1.00	1.00	100.0%
k-NN	1.00	1.00	1.00	100.0%
ANN	1.00	1.00	1.00	100.0%

Table 4.2: Classification performance of machine and deep learning classifiers measuring precision, recall, F1-score and accuracy.

Chapter 5

Results of CADs Approaches

In this chapter, we will see the test results obtained for each approach. Section 6.1 summarizes the best results of the first approach, discussed in Chapter 3, while Section 6.2 presents the best results for the second approach, developed in Chapter 4. Finally, Section 6.3 compile the achievements made through this research.

5.1 First approach Test dataset results

After the point where we left it in Chapter 3, the trained models were tested with the remaining dataset, followed by statistical analysis. Results are shown in Table 5.1.

$Classifier$	$Precision$	$Recall$	$F1-score$	$Accuracy$
Google Cloud AI	0.93	0.93	0.93	93.33%
SVM	0.83	0.82	0.81	81.33%
Random Forest	0.74	0.63	0.57	61.33%
Logreg	0.78	0.74	0.73	73.33%
k-NN	0.69	0.56	0.47	54.66%
ANN	**0.99**	**0.99**	**0.99**	**98.67%**

Table 5.1: Classification performance of the Google Cloud algorithm versus five image-based ML classifiers measuring precision, recall, F1-score and accuracy with 2D testing dataset.

Table 5.1 also show the results from the Google Cloud Vision API. Google's efforts to stay

on the top of the IA battle couldn't miss the use of deep learning for pathology predictions through image recognition. Vision API can "assign labels to images and quickly classify them into millions of predefined categories. Detect objects and faces, read printed and handwritten text, and build valuable metadata into your image catalog". As far as the experiment went, it didn't required file preparation or image preprocessing. Still, this row will only be used as a gold-standard CADs to compare the Python classifiers, the replicability of which we can guarantee.

As supposed in Chapter 3, test-dataset results using k-NN showed that the model was overfitted. In this case, all four indicators presented the worst performance versus other models.

Due to it's intrinsic purpose of correlating the probability of class membership, SVM it's widely used to solve binary classification problems. Still, results using test dataset were less accurate that the ones for training; moreover, it doesn't reach the proposed Google Vision API gold-standard .

Finally, we have the ANN applied for this first approach. The last model outperformed all previous ones in terms of metric performance: precision, recall, F1-score and accuracy. This results were better than the Google Vision API, which exceeded the expectations of the PositronMed research team.

5.2 Second approach Test dataset results

Table 5.2 shows the best results achieved by each model. With the exception of the deep learning model, all ML classifiers performed equally, and also very accurately. Even so, they differ in the predicted probability of class membership, which is displayed in the Appendix.

In this second approach, as in the first one, artificial neural network outperformed over the other models in identifying PD compatible patients. Despite its closeness to the 100%, this much accuracy is not desirable due to the overfitting potential and labeling errors. The Ax Platform was very useful to select the best parameter combination, that otherwise would have consumed a lot of time and work.

Even though k-NN showed a poor performance in 2D images, the 3D-input-model man-

Classifier	*Precision*	*Recall*	*F1 – score*	*Accuracy*
SVM	**1.00**	**0.97**	**0.98**	**98.28%**
Random Forest	**1.00**	**0.97**	**0.98**	**98.28%**
Logreg	0.97	0.97	0.97	96.55%
k-NN	1.00	0.94	0.97	96.55%
ANN	**1.00**	**0.97**	**0.98**	**98.28%**

Table 5.2: Classification performance of four ML classifiers and a ANN measuring precision, recall, F1-score and accuracy with 3D testing dataset.

aged to improve its results. In general terms, we found out models performed much better with more 3D dataset than 2D dataset, but this topic is going to be discussed in Chapter 6.

5.3 Achievements

Computational development in healthcare systems for decision-making is gaining importance. Computer-aided diagnosis is iteratively overcoming medical imaging analysis. In Chapter 2, we took a view on the gold-standard in PD detection. Studies vary in terms of radiotracers such as [123I]FP-CIT, 18F-(FDG), 99mTc-TRODAT-1, and image techniques as MRI, SPECT and PET; and even more, computational approaches open to branches of an infinite set of model structures and performance metrics.

This investigation studied the ability of Machine and Deep Learning to discriminate Normal Controls (NC) versus Parkinson's Disease compatible patients using [18F]PR04.MZ PET tracer, and deliver the probability of diagnosis. There's still no literature for a computational approach with the radiotracer we used, so the results here presented mark the state-of-art for dopamine transporter imaging with these specifications.

PET/CT [18F]PR04.MZ is a non-invasive imaging technique used in PositronMed molecular imaging lab, for the study of movement disorders associated with neurodegenerative diseases such as Parkinson's. Previous work [15, 16, 17] found out [18F]PR04.MZ radiotracer is highly selective for imaging dopamine transporters. Now, we validate this knowledge with several image-based algorithms designed for the detection of patterns.

Chapter 6

Conclusions and future work lines

Almost at the end of this research, we can intuit artificial intelligence can facilitate the diagnosis for patients in the earliest stage of Parkinson's desease. In this chapter, we compile this work conclusions, limitations and future work lines.

6.1 Conclusions

Neurodegenerative disorders are making presence in an increasingly aging world population, and Parkinson's disease (PD) patients have globally doubled to over 6 million in 26 years(1990-2016), a greater increase compared to other neurological disorders [47]. Studies attribute it to specific factors, such as the increasing life expectancy, diagnostic and therapeutic advances, and better health education of the population [45, 46, 47].

For a disease that still has no cure, early diagnosis is an important factor in curbing symptoms, which include motor and non-motor deficits. PD is characterized by a dopaminergic degeneration in the ventral area of the substantia nigra pars compacta (SNpc) of the midbrain, where nigro-striatal neurons originate. This behavior is detectable with the quantification of dopamine transporters (DAT) through nuclear medicine imaging. Positron emission tomography (PET) is a nuclear medicine technique to extract and project images, enabling the measure of metabolic activity in specific cells of the body. The close relationship between the presence of DAT in the SNpc and the diagnosis of PD motivates the study of radiophar-

maceuticals clinically validated and with high-affinity [15].

Latest approaches study CADs performance through DAT PET image processing. Chapter 2 summarized the research done for various tracers, also showing image recognition can improve diagnosis certainty and reinforce clinicians diagnostic decisions especially when there's lack of expertise. Up to date, no research has studied the performance of a CADs using [18F]PR04.MZ PET radiotracer, a DAT-selective, high-affinity [15, 16, 17], monoamine reuptake inhibitor [18, 19].

Throughout this thesis, we achieved our main objective by presenting the first computer-aided diagnosis performance using [18F]PR04.MZ PET radiotracer that discriminates NC vrs early PD compatible patients with above 98% of accuracy, providing an assisting resource for early PD detection. The resulting image-based algorithms can preprocess, train and classify 2D and 3D datasets, and allows comparison between five different ML techniques.

6.1.1 Comparison between datasets

The approaches performed in five different models used two and three-dimensional datasets. The results highlighted in both cases the performance of artificial neural networks, but differ in the other models when looking at both databases separately. Figure 6.1 illustrates a comparison between 2D/3D dataset performance. With the exception of the ANN, models with an three-dimensional input performed much better than those that used 2D images, although all images were processed in the algorithms as a vector of pixels.

Computational complexity required for running every algorithm differs from one approach to another, since it varies on the input size. Let's remember that first approach images were resized into 40 x 80 resolution, while 3D images used in the second approach were preselected with the help of an unsupervised learning algorithm, which delivered an output of 10 on-a-row slices of 109 x 91 resolution. However, studies have shown image resolution does not necessarily reduce model performance [48, 49]. This lead us to conclude that amount of slices provided by the second approach allowed the algorithm to be trained with a higher sensibility and specificity of [18F]PR04.MZ uptake in DAT midbrain regions.

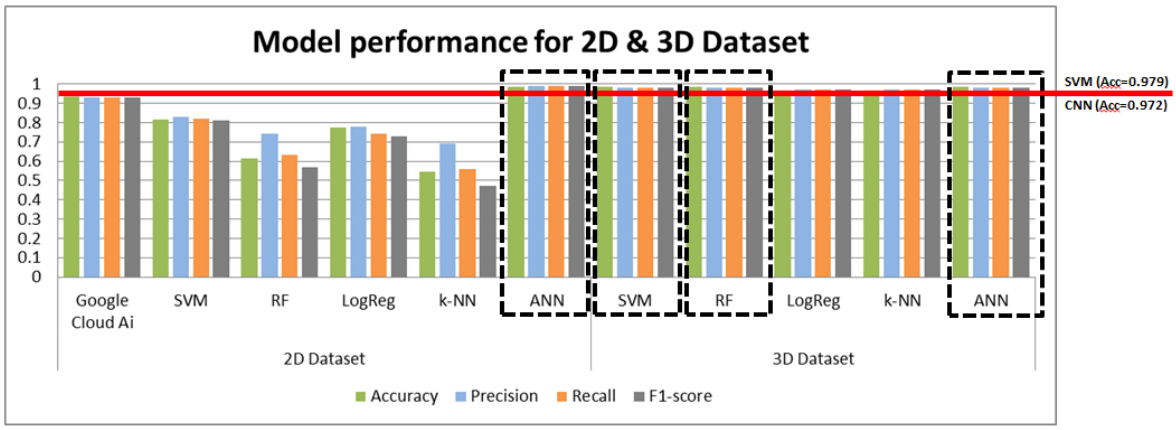


Figure 6.1: State-of-the-art versus performance based on accuracy, precision, recall and F1-score metrics of the five models used in this research: Support Vector Machine, Random Forest, Logistic Regression, k-Nearest Neighbor and Neural Network.

6.1.2 Interpretable and Explainable

The increasingly use of computational development to make highly risk decisions, as medical diagnosis, has brought ethical dilemmas to debate: are we going to trust in artificial intelligence rather than human intelligence? Do more accurate performance make more trustful the use of artificial neural networks than human brain neural networks?

In Machine Learning, interpretable models correlate with the level of trust, given the traceability of its prediction or decision steps. Models with low interpletability but high accuracy are less explainable, either called black-box models. In the higher degree of explainable models we found logistic regression, while neural networks tend to be more accurate but lowly explainable. Figure 6.2 illustrates a graphic representation of the degrees of accuracy and explainability for the models used in this research work.

The computer-aided diagnosis approach we worked on suggest the use of artificial neural networks as an assisting resource for early PD compatible detection due to it's outperformance over the other models. However, it's by definition a poorly interpretable model. Some studies use techniques to incorporate interpretability in neural networks such as Local Interpretable Model-Agnostic Explanations(LIME) [28], but this approach is beyond the scope of this investigation. Still, other models developed in this work performed very well.

Artificial intelligence, which hasn't even reached its mature age, still distantly sees the goal to become a multi-tasking tool comparable to human behavior. The models here used

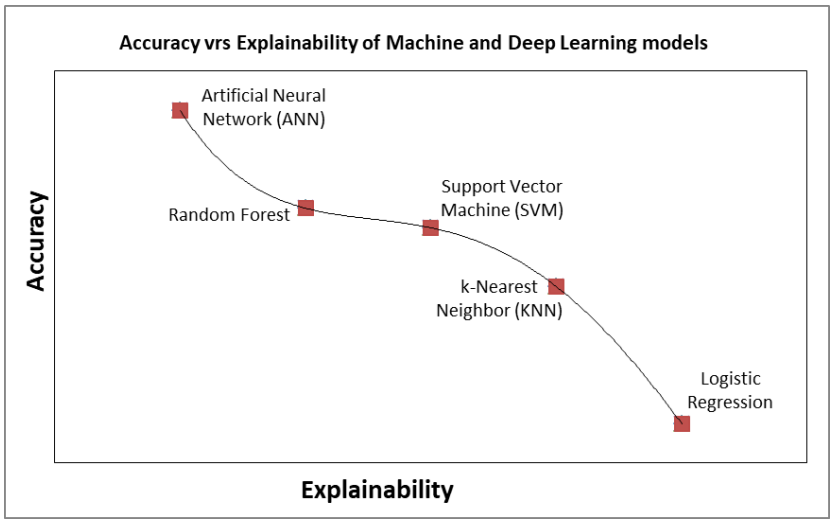


Figure 6.2: Accuracy and Explainable representation of the models used in this research: SVM, RF, LogReg, k-NN and ANN.

analyze the problem within a research framework, call it the same Parkinsonian syndrome. Unprecedentedly in medicine, ML and DL techniques certainly have opened up possibilities to identify neurodegenerative disorders. However, it does not manage to receive the infinite variables that could come before a final diagnosis. So even though the expectations for the CADs we created is to employ it as a clinical tool for a quantifiable diagnostic examination, it remains on the side of the physician and the patient to take the step of trusting it.

6.2 Limitations and Future work lines

The gold standard on which the accuracy and other performance metrics reported are based was the criteria issued by three specialists. Thus, the CAD system that resulted from this investigation intended to reproduce this assessment, which may even not be exempt of misdiagnosis. The definite diagnosis is only obtained at an autopsy [4], and to date, all PD compatible patients participating in this study are still alive.

Future work lines should improve 2D model results combining images with clinical data, including a feature selection and extraction approach. With this projection, PositronMed team have started the compilation of qualitative information, but it's still work in progress.

Another investigation field of deepening is the differentiation of atypical Parkinsonian syn-

dromes. A dataset with variation in medical diagnosis, i.e. patients suffering other disorders, could train an algorithm that classifies subjects with PD, MSA (multiple system atrophy), corticobasal degeneration (CBD), dementia with Lewy bodies (DLB), etc. Studies that apply this multiclass classification, although they work on a limited database [26], showed future work lines for different image and computational techniques.

Within the framework of [18F]PR04.MZ PET tracer validation, a dataset that not only tag early PD compatible subjects, but also follow the deterioration of patients in mid and advanced stages of PD, could allow the classification of multiple stages of Parkinson's disease.

Bibliography

[1] Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol. 2017 Sep;10(3):257-273. doi: 10.1007/s12194-017-0406-5. Epub 2017 Jul 8. PMID: 28689314.

[2] Surmeier DJ. Determinants of dopaminergic neuron loss in Parkinson's disease. FEBS J. 2018 Oct;285(19):3657-3668. doi: 10.1111/febs.14607. Epub 2018 Aug 14. PMID: 30028088; PMCID: PMC6546423.

[3] Vaughan RA, Foster JD. Mechanisms of dopamine transporter regulation in normal and disease states. Trends Pharmacol Sci. 2013 Sep;34(9):489-96. doi: 10.1016/j.tips.2013.07.005. Epub 2013 Aug 20. PMID: 23968642; PMCID: PMC3831354.

[4] Marsili L, Rizzo G, Colosimo C. Diagnostic Criteria for Parkinson's Disease: From James Parkinson to the Concept of Prodromal Disease. Front Neurol. 2018 Mar 23;9:156. doi: 10.3389/fneur.2018.00156. PMID: 29628907; PMCID: PMC5877503.

[5] Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, Obeso J, Marek K, Litvan I, Lang AE, Halliday G, Goetz CG, Gasser T, Dubois B, Chan P, Bloem BR, Adler CH, Deuschl G. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015 Oct;30(12):1591-601. doi: 10.1002/mds.26424. PMID: 26474316.

[6] Postuma RB, Poewe W, Litvan I, Lewis S, Lang AE, Halliday G, Goetz CG, Chan P, Slow E, Seppi K, Schaffer E, Rios-Romenets S, Mi T, Maetzler C, Li Y, Heim B, Bledsoe IO, Berg D. Validation of the MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2018 Oct;33(10):1601-1608. doi: 10.1002/mds.27362. Epub 2018 Aug 25. PMID: 30145797.

[7] Saeed U, Lang AE, Masellis M. Neuroimaging Advances in Parkinson's Disease and Atypical Parkinsonian Syndromes. Front Neurol. 2020 Oct 15;11:572976. doi: 10.3389/fneur.2020.572976. PMID: 33178113; PMCID: PMC7593544.

[8] Latif J, Xiao C, Imran A, Tu S. Medical Imaging using Machine Learning and Deep Learning Algorithms: A Review. 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), 2019, pp. 1-5, doi: 10.1109/I-COMET.2019.8673502.

[9] Mei J, Desrosiers C, Frasnelli J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. Front Aging Neurosci. 2021 May 6;13:633752. doi: 10.3389/fnagi.2021.633752. PMID: 34025389; PMCID: PMC8134676.

[10] Khachnaoui H, Mabrouk R, Khlifa N. Machine learning and deep learning for clinical data and PET/SPECT imaging in Parkinson's disease: a review. IET Image Process., 2020; 14: 4013-4026. doi: 10.1049/iet-ipr.2020.1048

[11] Prashanth R, Dutta Roy S. Early detection of Parkinson's disease through patient questionnaire and predictive modelling. Int J Med Inform. 2018 Nov;119:75-87. doi: 10.1016/j.ijmedinf.2018.09.008. Epub 2018 Sep 9. PMID: 30342689.

[12] Vasquez-Correa JC, Arias-Vergara T, Orozco-Arroyave JR, Eskofier B, Klucken J, Noth E. Multimodal Assessment of Parkinson's Disease: A Deep Learning Approach. IEEE J Biomed Health Inform. 2019 Jul;23(4):1618-1630. doi: 10.1109/JBHI.2018.2866873. Epub 2018 Aug 23. PMID: 30137018.

[13] Tăuţan AM, Ionescu B, Santarnecchi E. Artificial intelligence in neurodegenerative diseases: A review of available tools with a focus on machine learning techniques. Artif

Intell Med. 2021 Jul;117:102081. doi: 10.1016/j.artmed.2021.102081. Epub 2021 Apr 30. PMID: 34127244.

[14] Saeed U, Compagnone J, Aviv RI, Strafella AP, Black SE, Lang AE, Masellis M. Imaging biomarkers in Parkinson's disease and Parkinsonian syndromes: current and emerging concepts. Transl Neurodegener. 2017 Mar 28;6:8. doi: 10.1186/s40035-017-0076-6. PMID: 28360997; PMCID: PMC5370489.

[15] Kramer V, Juri C, Riss PJ, Pruzzo R, Soza-Ried C, Flores J, Hurtado A, Rösch F, Chana-Cuevas P, Amaral H. Pharmacokinetic evaluation of [18F]PR04.MZ for PET/CT imaging and quantification of dopamine transporters in the human brain. Eur J Nucl Med Mol Imaging. 2020 Jul;47(8):1927-1937. doi: 10.1007/s00259-019-04594-z. Epub 2019 Dec 1. PMID: 31788709.

[16] Juri C, Kramer V, Riss PJ, Soza-Ried C, Haeger A, Pruzzo R, Rösch F, Amaral H, Chana-Cuevas P. [18F]PR04.MZ PET/CT Imaging for Evaluation of Nigrostriatal Neuron Integrity in Patients With Parkinson Disease. Clin Nucl Med. 2021 Feb 1;46(2):119-124. doi: 10.1097/RLU.0000000000003430. PMID: 33323728; PMCID: PMC7774816.

[17] Juri C, Chana P, Kramer V, Pruzzo R, Amaral H, Riss PJ, Rösch F. Imaging Nigrostriatal Dopaminergic Deficit in Holmes Tremor with 18F-PR04.MZ-PET/CT. Clin Nucl Med. 2015 Sep;40(9):740-1. doi: 10.1097/RLU.0000000000000868. PMID: 26204203.

[18] Riss PJ, Debus F, Hummerich R, Schmidt U, Schloss P, Lueddens H, Roesch F. Ex vivo and in vivo evaluation of [18F]PR04.MZ in rodents: a selective dopamine transporter imaging agent. ChemMedChem. 2009 Sep;4(9):1480-7. doi: 10.1002/cmdc.200900177. PMID: 19588472.

[19] Riss PJ, Roesch F. Efficient microwave-assisted direct radiosynthesis of [(18)F]PR04.MZ and [(18)F]LBT999: selective dopamine transporter ligands for quantitative molecular imaging by means of PET. Bioorg Med Chem. 2009 Nov 15;17(22):7630-4. doi: 10.1016/j.bmc.2009.09.054. Epub 2009 Oct 4. PMID: 19846314.

[20] Illan IA, Gorrz JM, Ramirez J, Segovia F, Jimenez-Hoyuela JM, Ortega Lozano SJ. Automatic assistance to Parkinson's disease diagnosis in DaTSCAN SPECT imaging. Med Phys. 2012 Oct;39(10):5971-80. doi: 10.1118/1.4742055. PMID: 23039635.

[21] Oliveira FPM, Faria DB, Costa DC, Castelo-Branco M, Tavares JMRS. Extraction, selection and comparison of features for an effective automated computer-aided diagnosis of Parkinson's disease based on [123I]FP-CIT SPECT images. Eur J Nucl Med Mol Imaging. 2018 Jun;45(6):1052-1062. doi: 10.1007/s00259-017-3918-7. Epub 2017 Dec 23. PMID: 29275487.

[22] Vapnik VN. The nature of statistical learning theory. Springer science business media, 2013.

[23] Patra AK, Ray R, Abdullah AA, Dash SR. Prediction of Parkinson's disease using Ensemble Machine Learning classification from acoustic analysis. Journal of Physics: Conference Series. 2019;1372:012041–012041.

[24] Ho TK. Random decision forests. Proceedings of 3rd International Conference on Document Analysis and Recognition. 1995;1:278-282. doi: 10.1109/ICDAR.1995.598994.

[25] Muschelli J. Recommendations for Processing Head CT Data. Front Neuroinform. 2019 Sep 4;13:61. doi: 10.3389/fninf.2019.00061. PMID: 31551745; PMCID: PMC6738271.

[26] Martins R, Oliveira F, Moreira F, Moreira AP, Abrunhosa A, Januário C, Castelo-Branco M. Automatic classification of idiopathic Parkinson's disease and atypical Parkinsonian syndromes combining [11C]raclopride PET uptake and MRI grey matter morphometry. J Neural Eng. 2021 Apr 29;18(4). doi: 10.1088/1741-2552/abf772. PMID: 33848996.

[27] Hsu SY, Yeh LR, Chen TB, Du WC, Huang YH, Twan WH, Lin MC, Hsu YH, Wu YC, Chen HY. Classification of the Multiple Stages of Parkinson's Disease by a Deep Convolution Neural Network Based on 99mTc-TRODAT-1 SPECT Images. Molecules. 2020 Oct 19;25(20):4792. doi: 10.3390/molecules25204792. PMID: 33086589; PMCID: PMC7587595.

[28] Magesh P, Myloth R, Tom R. An Explainable Machine Learning Model for Early Detection of Parkinson's Disease using LIME on DaTSCAN Imagery. Computers in Biology and Medicine. 2020;126:104041. doi: 10.1016/j.compbiomed.2020.104041.

[29] Chakraborty S, Aich S, Kim HC. Detection of Parkinson's Disease from 3T T1 Weighted MRI Scans Using 3D Convolutional Neural Network. Diagnostics (Basel). 2020 Jun 12;10(6):402. doi: 10.3390/diagnostics10060402. PMID: 32545609; PMCID: PMC7345307.

[30] Segovia F, Górriz JM, Ramírez J, Martínez-Murcia FJ, Castillo-Barnes D. Assisted Diagnosis of Parkinsonism Based on the Striatal Morphology. Int J Neural Syst. 2019 Nov;29(9):1950011. doi: 10.1142/S0129065719500114. Epub 2019 Mar 4. PMID: 31084232.

[31] Mabrouk R, Chikhaoui B, Bentabet L. Machine Learning Based Classification Using Clinical and DaTSCAN SPECT Imaging Features: A Study on Parkinson's Disease and SWEDD. IEEE Transactions on Radiation and Plasma Medical Sciences, 2019 March; 3(2). doi: 10.1109/TRPMS.2018.2877754.

[32] Nicastro N, Wegrzyk J, Preti MG, Fleury V, Van de Ville D, Garibotto V, Burkhard PR. Classification of degenerative parkinsonism subtypes by support-vector-machine analysis and striatal 123I-FP-CIT indices. J Neurol. 2019 Jul;266(7):1771-1781. doi: 10.1007/s00415-019-09330-z. Epub 2019 Apr 29. PMID: 31037416; PMCID: PMC6586917.

[33] Shen T, Jiang J, Lin W, Ge J, Wu P, Zhou Y, Zuo C, Wang J, Yan Z, Shi K. Use of Overlapping Group LASSO Sparse Deep Belief Network to Discriminate Parkinson's Disease and Normal Control. Front Neurosci. 2019 Apr 29;13:396. doi: 10.3389/fnins.2019.00396. PMID: 31110472; PMCID: PMC6501727.

[34] Wenzel M, Milletari F, Krüger J, Lange C, Schenk M, Apostolova I, Klutmann S, Ehrenburg M, Buchert R. Automatic classification of dopamine transporter SPECT: deep convolutional neural networks can be trained to be robust with respect to variable im-

age characteristics. Eur J Nucl Med Mol Imaging. 2019 Dec;46(13):2800-2811. doi: 10.1007/s00259-019-04502-5. Epub 2019 Aug 31. PMID: 31473800.

[35] Wu Y, Jiang JH, Chen L, Lu JY, Ge JJ, Liu FT, Yu JT, Lin W, Zuo CT, Wang J. Use of radiomic features and support vector machine to distinguish Parkinson's disease cases from normal controls. Ann Transl Med. 2019 Dec;7(23):773. doi: 10.21037/atm.2019.11.26. PMID: 32042789; PMCID: PMC6990013.

[36] Yagis E, De Herrera AGS, Citi L. Generalization Performance of Deep Learning Models in Neurodegenerative Disease Classification. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2019; 1692-1698. doi: 10.1109/BIBM47256.2019.8983088.

[37] Li S, Lei H, Zhou F, Gardezi J, Lei B. Longitudinal and Multi-modal Data Learning for Parkinson's Disease Diagnosis via Stacked Sparse Auto-encoder. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019; 384-387. doi: 10.1109/ISBI.2019.8759385.

[38] Martinez-Murcia FJ, Górriz JM, Ramírez J, Ortiz A. Convolutional Neural Networks for Neuroimaging in Parkinson's Disease: Is Preprocessing Needed? Int J Neural Syst. 2018 Dec;28(10):1850035. doi: 10.1142/S0129065718500351. Epub 2018 Jul 26. PMID: 30215285.

[39] Oliveira FPM, Faria DB, Costa DC, Castelo-Branco M, Tavares JMRS. Extraction, selection and comparison of features for an effective automated computer-aided diagnosis of Parkinson's disease based on [123I]FP-CIT SPECT images. Eur J Nucl Med Mol Imaging. 2018 Jun;45(6):1052-1062. doi: 10.1007/s00259-017-3918-7. Epub 2017 Dec 23. PMID: 29275487.

[40] Choi H, Ha S, Im HJ, Paek SH, Lee DS. Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. Neuroimage Clin. 2017 Sep 10;16:586-594. doi: 10.1016/j.nicl.2017.09.010. PMID: 28971009; PMCID: PMC5610036.

[41] Prashanth R, Roy SD, Mandal PK, Ghosh S. High-Accuracy Classification of Parkinson's Disease Through Shape Analysis and Surface Fitting in 123I-Ioflupane SPECT Imaging. IEEE J Biomed Health Inform. 2017 May;21(3):794-802. doi: 10.1109/JBHI.2016.2547901. Epub 2016 Mar 29. PMID: 28113827.

[42] Taylor JC, Fenner JW. Comparison of machine learning and semi-quantification algorithms for (I123)FP-CIT classification: the beginning of the end for semi-quantification? EJNMMI Phys. 2017 Nov 29;4(1):29. doi: 10.1186/s40658-017-0196-1. PMID: 29188397; PMCID: PMC5707214.

[43] Peng B, Wang S, Zhou Z, Liu Y, Tong B, Zhang T, Dai Y. A multilevel-ROI-features-based machine learning method for detection of morphometric biomarkers in Parkinson's disease. Neurosci Lett. 2017 Jun 9;651:88-94. doi: 10.1016/j.neulet.2017.04.034. Epub 2017 Apr 21. PMID: 28435046.

[44] Hirschauer TJ, Adeli H, Buford JA. Computer-Aided Diagnosis of Parkinson's Disease Using Enhanced Probabilistic Neural Network. J Med Syst. 2015 Nov;39(11):179. doi: 10.1007/s10916-015-0353-9. Epub 2015 Sep 29. PMID: 26420585.

[45] Leiva AM, Martínez-Sanguinetti MA, Troncoso-Pantoja C, Nazar G, Petermann-Rocha F, Celis-Morales C. Chile lidera el ranking latinoamericano de prevalencia de enfermedad de Parkinson. Revista médica de Chile, 2019;147(4). doi: 10.4067/S0034-98872019000400535.

[46] Benito-León J. Epidemiología de la enfermedad de Parkinson en España y su contextualización mundial. Rev Neurol 2018;66 (04):125-134. doi: 10.33588/rn.6604.2017440

[47] GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. 2018 Nov;17(11):939-953. doi: 10.1016/S1474-4422(18)30295-3. Epub 2018 Oct 1. PMID: 30287051; PMCID: PMC6191528.

[48] Albukhanajer WA, Jin Y, Briffa JA. Trade-off between computational complexity and accuracy in evolutionary image feature extraction. 2015 IEEE Congress on Evolutionary Computation (CEC), 2015; 2412-2419. doi: 10.1109/CEC.2015.7257184.

[49] Correa G, Amado Assunção P, Agostini L, da Silva Cruz L. Performance and Computational Complexity Assessment of High Efficiency Video Encoders. IEEE Transactions on Circuits and Systems for Video Technology. 2012 Dec;22(12):1899-1909. doi: 10.1109/TCSVT.2012.2223411.

Appendix

Appendix A - PD burden in Latinamerican countries (1990-2016)

Countries	Deaths		Prevalence		DALYs	
	2016 counts	Percentage change in age-standardised rates, 1990–2016	2016 counts	Percentage change in age-standardised rates, 1990–2016	2016 counts	Percentage change in age-standardised rates, 1990–2016
Argentina	2798 (2099 to 3623)	0.1% (−13.4 to 13.8)	68 048 (52 574 to 85 157)	2.5% (−8.5 to 13.5)	39 297 (29 708 to 49 971)	1.8% (−10.2 to 14.6)
Chile	1064 (756 to 1437)	16.5% (−7.1 to 43.0)	25 845 (20 232 to 32 298)	19.9% (12.8 to 27.5)	14 860 (10 643 to 19 579)	17.1% (−2.7 to 40.4)
Uruguay	287 (216 to 375)	7.1% (−2.9 to 18.9)	6289 (4860 to 7952)	10.9% (2.8 to 18.7)	3775 (2870 to 4861)	8.2% (−1.2 to 19.2)
Colombia	796 (607 to 1022)	12.2% (1.4 to 25.2)	25 930 (20 527 to 32 111)	15.5% (9.8 to 20.8)	13 140 (10 059 to 16 860)	13.4% (3.5 to 25.1)
Costa Rica	110 (83 to 143)	12.8% (2.4 to 24.4)	3230 (2532 to 4003)	15.3% (10.0 to 21.2)	1700 (1300 to 2192)	14.0% (4.6 to 24.3)
El Salvador	128 (97 to 168)	13.8% (0.9 to 28.3)	3436 (2680 to 4268)	18.6% (12.2 to 25.1)	1902 (1440 to 2462)	13.7% (2.9 to 26.5)
Guatemala	182 (130 to 244)	15.7% (−8.2 to 41.6)	5194 (4068 to 6433)	18.0% (12.3 to 23.8)	2825 (2078 to 3721)	15.7% (−5.8 to 37.6)
Honduras	107 (76 to 146)	19.8% (−3.2 to 47.4)	2741 (2166 to 3383)	18.4% (12.8 to 24.4)	1626 (1168 to 2140)	18.7% (−2.0 to 43.7)
Mexico	2299 (1752 to 2952)	14.4% (10.3 to 18.6)	68 715 (54 711 to 83 874)	17.7% (15.9 to 19.6)	35 633 (27 612 to 44 856)	16.4% (12.8 to 20.0)
Nicaragua	77 (58 to 102)	13.8% (−1.7 to 32.1)	2273 (1799 to 2788)	15.9% (10.6 to 22.0)	1185 (908 to 1519)	14.8% (1.6 to 30.4)
Panama	83 (62 to 108)	11.4% (−3.3 to 29.4)	2369 (1878 to 2906)	14.3% (8.5 to 19.4)	1255 (949 to 1629)	12.2% (−0.6 to 26.9)
Venezuela	464 (335 to 613)	9.3% (−6.2 to 28.6)	15 235 (12 126 to 18 584)	14.0% (6.9 to 19.9)	7758 (5734 to 10 079)	11.0% (−2.9 to 29.0)
Bolivia	202 (147 to 269)	21.1% (3.1 to 43.7)	5114 (4011 to 6330)	14.1% (8.8 to 19.8)	3003 (2237 to 3871)	19.2% (3.0 to 38.5)
Ecuador	295 (224 to 383)	10.8% (0.7 to 22.5)	8688 (6872 to 10 735)	13.6% (8.5 to 19.3)	4589 (3519 to 5833)	11.2% (2.3 to 21.8)
Peru	596 (436 to 806)	15.9% (−4.6 to 37.7)	16 915 (13 398 to 20 887)	12.3% (6.8 to 18.0)	9106 (6795 to 11 952)	15.2% (−2.4 to 34.8)
Brazil	4033 (3074 to 5199)	15.0% (11.1 to 19.6)	128 836 (102 469 to 159 395)	16.4% (14.2 to 18.7)	66 204 (50 914 to 84 027)	15.0% (11.6 to 19.1)
Paraguay	98 (74 to 129)	23.4% (8.0 to 39.4)	2912 (2300 to 3602)	19.4% (13.7 to 25.2)	1573 (1189 to 2015)	23.1% (10.7 to 37.0)
Cuba	507 (381 to 662)	7.9% (−4.4 to 20.2)	12 678 (9985 to 15 908)	9.0% (2.0 to 16.5)	7203 (5459 to 9294)	9.0% (−2.0 to 19.9)
Dominican Rep	202 (149 to 264)	13.8% (−3.2 to 30.9)	5456 (4317 to 6690)	17.5% (11.5 to 23.6)	2999 (2282 to 3831)	14.6% (−0.3 to 30.6)
Puerto Rico	181 (136 to 235)	14.1% (2.4 to 26.9)	4300 (3392 to 5366)	12.2% (7.1 to 18.6)	2478 (1872 to 3177)	13.5% (3.0 to 24.7)

Figure 8.1: Deaths, prevalence, and DALYs for Parkinson's disease and percentage change between 1990 and 2016 in age-standardised rates by Latinamerican location. Data extracted from [47].

Appendix B - Results for ML and DL models applied to 2D dataset

B.1 Logistic Regression

Probability of class membership (2D test dataset)		
Code	Label	LogReg model prediction
12	0	96.46 % Normal
13	0	99.89 % Normal
121	0	71.69 % Normal
122	0	98.45 % Normal
124	0	98.23 % Normal
129	0	99.98 % Normal
135	0	59.1 % Parkinsonismo
138	0	99.69 % Normal
140	0	99.96 % Normal
141	0	99.76 % Normal
142	0	100.0 % Normal
143	0	99.67 % Normal
147	0	79.01 % Normal
148	0	65.51 % Parkinsonismo
149	0	99.94 % Normal
150	0	92.73 % Normal
151	0	90.67 % Parkinsonismo
152	0	90.69 % Normal
154	0	92.33 % Parkinsonismo
156	0	99.98 % Normal
158	0	91.82 % Normal
160	0	69.78 % Normal
164	0	99.98 % Normal
165	0	99.98 % Normal
168	0	98.24 % Normal
172	0	99.61 % Normal
174	0	98.91 % Normal
176	0	97.51 % Normal
177	0	69.72 % Parkinsonismo
181	0	68.8 % Normal
182	0	99.98 % Normal
183	0	99.97 % Normal
184	0	99.68 % Normal
185	0	97.38 % Normal
186	0	99.98 % Normal
190	0	88.05 % Normal
192	0	95.69 % Normal
197	0	98.22 % Normal
203	0	99.96 % Normal

Probability of class membership (2D test dataset)		
Code	Label	LogReg model prediction
3	1	70.61 % Normal
4	1	93.92 % Parkinsonismo
157	1	99.77 % Parkinsonismo
159	1	99.78 % Parkinsonismo
161	1	99.99 % Parkinsonismo
162	1	99.95 % Normal
163	1	98.41 % Parkinsonismo
166	1	99.98 % Parkinsonismo
167	1	99.6 % Parkinsonismo
169	1	99.98 % Parkinsonismo
170	1	98.8 % Parkinsonismo
171	1	99.99 % Parkinsonismo
173	1	80.34 % Parkinsonismo
175	1	99.97 % Parkinsonismo
178	1	86.77 % Parkinsonismo
179	1	90.22 % Parkinsonismo
180	1	100.0 % Parkinsonismo
187	1	99.95 % Parkinsonismo
188	1	98.43 % Parkinsonismo
189	1	99.99 % Parkinsonismo
191	1	85.87 % Parkinsonismo
193	1	97.37 % Parkinsonismo
194	1	100.0 % Parkinsonismo
195	1	100.0 % Parkinsonismo
196	1	62.86 % Normal
198	1	83.84 % Parkinsonismo
199	1	99.97 % Parkinsonismo
200	1	99.99 % Parkinsonismo
201	1	99.98 % Parkinsonismo
202	1	99.94 % Parkinsonismo
204	1	99.99 % Parkinsonismo
205	1	99.83 % Parkinsonismo
206	1	99.96 % Parkinsonismo
207	1	99.97 % Parkinsonismo
208	1	99.83 % Parkinsonismo
209	1	99.69 % Parkinsonismo

Figure 8.2: Predicted probability of class membership for Logistic Regression model using 2D dataset.

B.2 Artificial Neural Network (ANN)

Probability of class membership (2D test dataset)		
Code	Label	ANN model prediction
12	0	99.83 % Normal
13	0	100.0 % Normal
121	0	100.0 % Normal
122	0	100.0 % Normal
124	0	100.0 % Normal
129	0	100.0 % Normal
135	0	99.99 % Normal
138	0	100.0 % Normal
140	0	100.0 % Normal
141	0	100.0 % Normal
142	0	100.0 % Normal
143	0	100.0 % Normal
147	0	99.81 % Normal
148	0	96.0 % Normal
149	0	99.27 % Normal
150	0	100.0 % Normal
151	0	100.0 % Normal
152	0	99.93 % Normal
154	0	79.71 % Normal
156	0	100.0 % Normal
158	0	100.0 % Normal
160	0	100.0 % Normal
164	0	100.0 % Normal
165	0	100.0 % Normal
168	0	100.0 % Normal
172	0	100.0 % Normal
174	0	99.87 % Normal
176	0	100.0 % Normal
177	0	99.95 % Normal
181	0	99.07 % Normal
182	0	100.0 % Normal
183	0	100.0 % Normal
184	0	100.0 % Normal
185	0	100.0 % Normal
186	0	100.0 % Normal
190	0	100.0 % Normal
192	0	100.0 % Normal
197	0	100.0 % Normal
203	0	100.0 % Normal

Probability of class membership (2D test dataset)		
Code	Label	ANN model prediction
3	1	63.83 % Parkinsonismo
4	1	100.0 % Parkinsonismo
157	1	99.99 % Parkinsonismo
159	1	99.91 % Parkinsonismo
161	1	99.97 % Parkinsonismo
162	1	100.0 % Normal
163	1	100.0 % Parkinsonismo
166	1	100.0 % Parkinsonismo
167	1	100.0 % Parkinsonismo
169	1	100.0 % Parkinsonismo
170	1	73.24 % Parkinsonismo
171	1	100.0 % Parkinsonismo
173	1	99.99 % Parkinsonismo
175	1	100.0 % Parkinsonismo
178	1	99.21 % Parkinsonismo
179	1	98.59 % Parkinsonismo
180	1	100.0 % Parkinsonismo
187	1	100.0 % Parkinsonismo
188	1	100.0 % Parkinsonismo
189	1	100.0 % Parkinsonismo
191	1	99.98 % Parkinsonismo
193	1	99.58 % Parkinsonismo
194	1	100.0 % Parkinsonismo
195	1	99.98 % Parkinsonismo
196	1	95.07 % Parkinsonismo
198	1	97.61 % Parkinsonismo
199	1	100.0 % Parkinsonismo
200	1	99.91 % Parkinsonismo
201	1	100.0 % Parkinsonismo
202	1	100.0 % Parkinsonismo
204	1	100.0 % Parkinsonismo
205	1	67.97 % Parkinsonismo
206	1	98.27 % Parkinsonismo
207	1	100.0 % Parkinsonismo
208	1	100.0 % Parkinsonismo
209	1	99.99 % Parkinsonismo

Figure 8.3: Predicted probability of class membership for ANN model using 2D dataset.

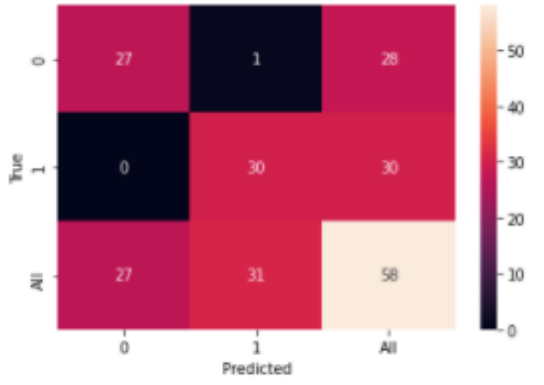
Appendix C - Results for ML and DL models applied to 3D dataset

C.1 Support Vector Machine (SVM)

Probability of class membership (3D test dataset)		
Code	Label	SVM model prediction
016_P180_(N)	0	99.25 % Normal
020_P187_(N)	0	72.59 % Normal
027_P194_(N)-flip	0	99.9 % Normal
028_P195_(N)	0	99.79 % Normal
035_C022_(N)	0	94.91 % Normal
044_P219_(N)-flip	0	98.41 % Normal
060_P240_(N)-flip	0	99.96 % Normal
067_P247_(N)-flip	0	99.41 % Normal
068_P248_(N)-flip	0	99.3 % Normal
076_P256_(N)	0	98.13 % Normal
076_P256_(N)-flip	0	98.97 % Normal
096_P289_(N)-flip	0	99.39 % Normal
107_P300_(N)	0	52.37 % Normal
113_P314_(N)-flip	0	99.92 % Normal
114_P315_(N)	0	99.22 % Normal
120_P324_(N)	0	99.61 % Normal
120_P324_(N)-flip	0	99.65 % Normal
148_P418_(N)-flip	0	99.96 % Normal
149_P423_(N)-flip	0	97.92 % Normal
151_P427_(N)-flip	0	99.88 % Normal
156_P460_(N)-flip	0	96.98 % Normal
160_P467_(N)	0	95.05 % Normal
162_P469_(N)-flip	0	93.75 % Normal
168_P486_(N)	0	98.98 % Normal
176_P510_(N)-flip	0	98.36 % Normal
181_P526_(N)-flip	0	98.97 % Normal
190_P545_(N)	0	94.62 % Normal
003_P154_(P)-flip	1	99.44 % Normal

Probability of class membership (3D test dataset)		
Code	Label	SVM model prediction
003_P154_(P)-flip	1	99.44 % Normal
004_P155_(P)	1	97.71 % Parkinsonismo
005_P156_(P)-flip	1	99.29 % Parkinsonismo
015_P179_(P)	1	98.7 % Parkinsonismo
021_P188_(P)-flip	1	98.15 % Parkinsonismo
026_P193_(P)	1	98.08 % Parkinsonismo
033_P200_(P)	1	98.34 % Parkinsonismo
043_P218_(P)-flip	1	99.07 % Parkinsonismo
045_P220_(P)-flip	1	96.93 % Parkinsonismo
050_P225_(P)	1	92.11 % Parkinsonismo
058_P238_(P)-flip	1	96.45 % Parkinsonismo
088_P272_(P)	1	98.92 % Parkinsonismo
095_P288_(P)	1	97.52 % Parkinsonismo
098_P291_(P)	1	96.27 % Parkinsonismo
100_P293_(P)-flip	1	99.04 % Parkinsonismo
104_P297_(P)	1	99.44 % Parkinsonismo
104_P297_(P)-flip	1	99.35 % Parkinsonismo
106_P299_(P)-flip	1	97.57 % Parkinsonismo
111_P304_(P)-flip	1	88.65 % Parkinsonismo
145_P392_(P)-flip	1	94.28 % Parkinsonismo
171_P490_(P)	1	97.92 % Parkinsonismo
173_P500_(P)	1	98.03 % Parkinsonismo
180_P523_(P)	1	97.81 % Parkinsonismo
180_P523_(P)-flip	1	97.74 % Parkinsonismo
187_P534_(P)	1	98.19 % Parkinsonismo
188_P535_(P)-flip	1	94.94 % Parkinsonismo
189_P540_(P)	1	97.53 % Parkinsonismo
195_P561_(P)-flip	1	98.2 % Parkinsonismo
200_P585_(P)-flip	1	96.9 % Parkinsonismo
201_P586_(P)-flip	1	98.19 % Parkinsonismo
205_P594_(P)	1	98.3 % Parkinsonismo

Figure 8.4: Predicted probability of class membership for SVM model using 3D dataset.




```
Donde 0 es Normal y 1 Parkinsonismo
Accuracy: 98.28 %
Precision: 100.0 % Normal 96.77 % Parkinsonismo
Recall: 96.43 % Normal 100.0 % Parkinsonismo
```

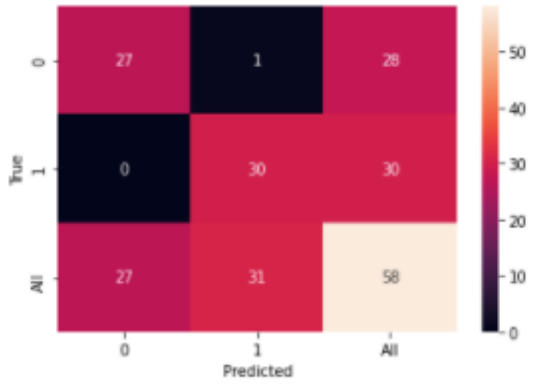
Figure 8.5: Confusion matrix of SVM model using 3D dataset.

C.2 Random Forest

Probability of class membership (3D test dataset)		
Code	Label	Random Forest model prediction
016_P180_(N)	0	95.0 % Normal
020_P187_(N)	0	65.0 % Normal
027_P194_(N)-flip	0	100.0 % Normal
028_P195_(N)	0	95.0 % Normal
035_C022_(N)	0	95.0 % Normal
044_P219_(N)-flip	0	95.0 % Normal
060_P240_(N)-flip	0	90.0 % Normal
067_P247_(N)-flip	0	95.0 % Normal
068_P248_(N)-flip	0	90.0 % Normal
076_P256_(N)	0	90.0 % Normal
076_P256_(N)-flip	0	95.0 % Normal
096_P289_(N)-flip	0	95.0 % Normal
107_P300_(N)	0	65.0 % Normal
113_P314_(N)-flip	0	100.0 % Normal
114_P315_(N)	0	95.0 % Normal
120_P324_(N)	0	100.0 % Normal
120_P324_(N)-flip	0	100.0 % Normal
148_P418_(N)-flip	0	95.0 % Normal
149_P423_(N)-flip	0	90.0 % Normal
151_P427_(N)-flip	0	85.0 % Normal
156_P460_(N)-flip	0	85.0 % Normal
160_P467_(N)	0	80.0 % Normal
162_P469_(N)-flip	0	80.0 % Normal
168_P486_(N)	0	55.0 % Normal
176_P510_(N)-flip	0	95.0 % Normal
181_P526_(N)-flip	0	95.0 % Normal
190_P545_(N)	0	85.0 % Normal

Probability of class membership (3D test dataset)		
Code	Label	Random Forest model prediction
003_P154_(P)-flip	1	100.0 % Normal
004_P155_(P)	1	95.0 % Parkinsonismo
005_P156_(P)-flip	1	95.0 % Parkinsonismo
015_P179_(P)	1	100.0 % Parkinsonismo
021_P188_(P)-flip	1	100.0 % Parkinsonismo
026_P193_(P)	1	100.0 % Parkinsonismo
033_P200_(P)	1	95.0 % Parkinsonismo
043_P218_(P)-flip	1	100.0 % Parkinsonismo
045_P220_(P)-flip	1	95.0 % Parkinsonismo
050_P225_(P)	1	80.0 % Parkinsonismo
058_P238_(P)-flip	1	90.0 % Parkinsonismo
088_P272_(P)	1	100.0 % Parkinsonismo
095_P288_(P)	1	90.0 % Parkinsonismo
098_P291_(P)	1	90.0 % Parkinsonismo
100_P293_(P)-flip	1	95.0 % Parkinsonismo
104_P297_(P)	1	100.0 % Parkinsonismo
104_P297_(P)-flip	1	95.0 % Parkinsonismo
106_P299_(P)-flip	1	90.0 % Parkinsonismo
111_P304_(P)-flip	1	80.0 % Parkinsonismo
145_P392_(P)-flip	1	85.0 % Parkinsonismo
171_P490_(P)	1	95.0 % Parkinsonismo
173_P500_(P)	1	90.0 % Parkinsonismo
180_P523_(P)	1	95.0 % Parkinsonismo
180_P523_(P)-flip	1	100.0 % Parkinsonismo
187_P534_(P)	1	100.0 % Parkinsonismo
188_P535_(P)-flip	1	75.0 % Parkinsonismo
189_P540_(P)	1	85.0 % Parkinsonismo
195_P561_(P)-flip	1	100.0 % Parkinsonismo
200_P585_(P)-flip	1	90.0 % Parkinsonismo
201_P586_(P)-flip	1	100.0 % Parkinsonismo
205_P594_(P)	1	100.0 % Parkinsonismo

Figure 8.6: Predicted probability of class membership for Random Forest model using 3D dataset.




```
Donde 0 es Normal y 1 Parkinsonismo
Accuracy: 98.28 %
Precision: 100.0 % Normal 96.77 % Parkinsonismo
Recall: 96.43 % Normal 100.0 % Parkinsonismo
```

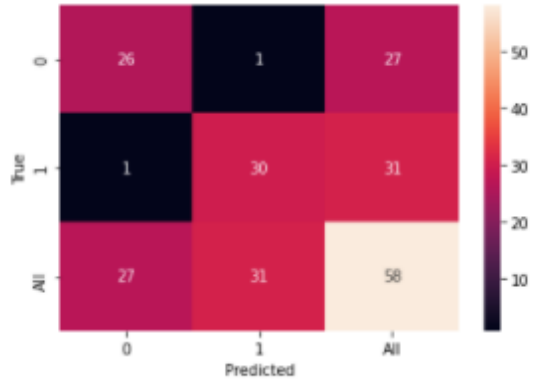
Figure 8.7: Confusion matrix of Ranfom Forest model using 3D dataset.

C.3 Logistic Regression

Probability of class membership (3D test dataset)		
Code	Label	LogReg model prediction
016_P180_(N)	0	99.99 % Normal
020_P187_(N)	0	65.19 % Parkinsonimo
027_P194_(N)-flip	0	100.0 % Normal
028_P195_(N)	0	100.0 % Normal
035_C022_(N)	0	99.96 % Normal
044_P219_(N)-flip	0	100.0 % Normal
060_P240_(N)-flip	0	99.99 % Normal
067_P247_(N)-flip	0	100.0 % Normal
068_P248_(N)-flip	0	98.96 % Normal
076_P256_(N)	0	88.96 % Normal
076_P256_(N)-flip	0	84.32 % Normal
096_P289_(N)-flip	0	99.97 % Normal
107_P300_(N)	0	86.13 % Normal
113_P314_(N)-flip	0	100.0 % Normal
114_P315_(N)	0	99.07 % Normal
120_P324_(N)	0	100.0 % Normal
120_P324_(N)-flip	0	100.0 % Normal
148_P418_(N)-flip	0	100.0 % Normal
149_P423_(N)-flip	0	99.94 % Normal
151_P427_(N)-flip	0	100.0 % Normal
156_P460_(N)-flip	0	98.91 % Normal
160_P467_(N)	0	99.83 % Normal
162_P469_(N)-flip	0	95.7 % Normal
168_P486_(N)	0	99.99 % Normal
176_P510_(N)-flip	0	99.79 % Normal
181_P526_(N)-flip	0	99.76 % Normal
190_P545_(N)	0	91.93 % Normal

Probability of class membership (3D test dataset)		
Code	Label	LogReg model prediction
003_P154_(P)-flip	1	99.25 % Normal
004_P155_(P)	1	99.93 % Parkinsonimo
005_P156_(P)-flip	1	99.99 % Parkinsonimo
015_P179_(P)	1	99.91 % Parkinsonimo
021_P188_(P)-flip	1	99.99 % Parkinsonimo
026_P193_(P)	1	99.7 % Parkinsonimo
033_P200_(P)	1	99.8 % Parkinsonimo
043_P218_(P)-flip	1	99.98 % Parkinsonimo
045_P220_(P)-flip	1	98.25 % Parkinsonimo
050_P225_(P)	1	97.28 % Parkinsonimo
058_P238_(P)-flip	1	95.54 % Parkinsonimo
088_P272_(P)	1	99.99 % Parkinsonimo
095_P288_(P)	1	99.97 % Parkinsonimo
098_P291_(P)	1	99.22 % Parkinsonimo
100_P293_(P)-flip	1	99.97 % Parkinsonimo
104_P297_(P)	1	99.99 % Parkinsonimo
104_P297_(P)-flip	1	99.99 % Parkinsonimo
106_P299_(P)-flip	1	99.71 % Parkinsonimo
111_P304_(P)-flip	1	98.81 % Parkinsonimo
145_P392_(P)-flip	1	99.97 % Parkinsonimo
171_P490_(P)	1	99.85 % Parkinsonimo
173_P500_(P)	1	99.97 % Parkinsonimo
180_P523_(P)	1	63.65 % Parkinsonimo
180_P523_(P)-flip	1	70.79 % Parkinsonimo
187_P534_(P)	1	99.98 % Parkinsonimo
188_P535_(P)-flip	1	99.69 % Parkinsonimo
189_P540_(P)	1	99.81 % Parkinsonimo
195_P561_(P)-flip	1	98.19 % Parkinsonimo
200_P585_(P)-flip	1	98.06 % Parkinsonimo
201_P586_(P)-flip	1	100.0 % Parkinsonimo
205_P594_(P)	1	99.01 % Parkinsonimo

Figure 8.8: Predicted probability of class membership for Logistic Regression model using 3D dataset.




```
Donde 0 es Normal y 1 Parkinsonismo
Accuracy: 96.55 %
Precision: 96.3 % Normal 96.77 % Parkinsonismo
Recall: 96.3 % Normal 96.77 % Parkinsonismo
```

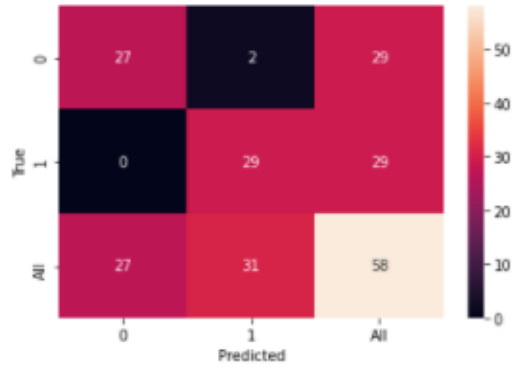
Figure 8.9: Confusion matrix of Logistic Regression model using 3D dataset.

C.4 K-Nearest Neighbor (KNN)

Probability of class membership (3D test dataset)		
Code	Label	KNN model prediction
016_P180_(N)	0	100.0 % Normal
020_P187_(N)	0	100.0 % Normal
027_P194_(N)-flip	0	100.0 % Normal
028_P195_(N)	0	100.0 % Normal
035_C022_(N)	0	100.0 % Normal
044_P219_(N)-flip	0	100.0 % Normal
060_P240_(N)-flip	0	100.0 % Normal
067_P247_(N)-flip	0	100.0 % Normal
068_P248_(N)-flip	0	100.0 % Normal
076_P256_(N)	0	100.0 % Normal
076_P256_(N)-flip	0	100.0 % Normal
096_P289_(N)-flip	0	100.0 % Normal
107_P300_(N)	0	100.0 % Normal
113_P314_(N)-flip	0	100.0 % Normal
114_P315_(N)	0	100.0 % Normal
120_P324_(N)	0	100.0 % Normal
120_P324_(N)-flip	0	100.0 % Normal
148_P418_(N)-flip	0	100.0 % Normal
149_P423_(N)-flip	0	100.0 % Normal
151_P427_(N)-flip	0	100.0 % Normal
156_P460_(N)-flip	0	100.0 % Normal
160_P467_(N)	0	100.0 % Normal
162_P469_(N)-flip	0	100.0 % Normal
168_P486_(N)	0	100.0 % Normal
176_P510_(N)-flip	0	100.0 % Normal
181_P526_(N)-flip	0	100.0 % Normal
190_P545_(N)	0	100.0 % Normal

Probability of class membership (3D test dataset)		
Code	Label	KNN model prediction
003_P154_(P)-flip	1	100.0 % Normal
004_P155_(P)	1	100.0 % Parkinsonismo
005_P156_(P)-flip	1	100.0 % Parkinsonismo
015_P179_(P)	1	100.0 % Parkinsonismo
021_P188_(P)-flip	1	100.0 % Parkinsonismo
026_P193_(P)	1	100.0 % Parkinsonismo
033_P200_(P)	1	100.0 % Parkinsonismo
043_P218_(P)-flip	1	100.0 % Parkinsonismo
045_P220_(P)-flip	1	100.0 % Parkinsonismo
050_P225_(P)	1	100.0 % Normal
058_P238_(P)-flip	1	100.0 % Parkinsonismo
088_P272_(P)	1	100.0 % Parkinsonismo
095_P288_(P)	1	100.0 % Parkinsonismo
098_P291_(P)	1	100.0 % Parkinsonismo
100_P293_(P)-flip	1	100.0 % Parkinsonismo
104_P297_(P)	1	100.0 % Parkinsonismo
104_P297_(P)-flip	1	100.0 % Parkinsonismo
106_P299_(P)-flip	1	100.0 % Parkinsonismo
111_P304_(P)-flip	1	100.0 % Parkinsonismo
145_P392_(P)-flip	1	100.0 % Parkinsonismo
171_P490_(P)	1	100.0 % Parkinsonismo
173_P500_(P)	1	100.0 % Parkinsonismo
180_P523_(P)	1	100.0 % Parkinsonismo
180_P523_(P)-flip	1	100.0 % Parkinsonismo
187_P534_(P)	1	100.0 % Parkinsonismo
188_P535_(P)-flip	1	100.0 % Parkinsonismo
189_P540_(P)	1	100.0 % Parkinsonismo
195_P561_(P)-flip	1	100.0 % Parkinsonismo
200_P585_(P)-flip	1	100.0 % Parkinsonismo
201_P586_(P)-flip	1	100.0 % Parkinsonismo
205_P594_(P)	1	100.0 % Parkinsonismo

Figure 8.10: Predicted probability of class membership for KNN model using 3D dataset.




```
Donde 0 es Normal y 1 Parkinsonismo
Accuracy: 96.55 %
Precision: 100.0 % Normal 93.55 % Parkinsonismo
Recall: 93.1 % Normal 100.0 % Parkinsonismo
```

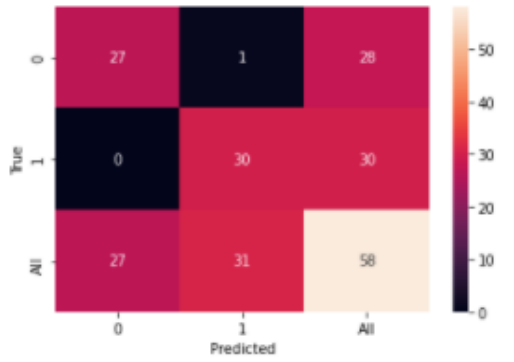
Figure 8.11: Confusion matrix of k-NN model using 3D dataset.

C.5 Artificial Neural Network (ANN)

Probability of class membership (3D test dataset)		
Code	Label	ANN model prediction
016_P180_(N)	0	100.0 % Normal
020_P187_(N)	0	100.0 % Normal
027_P194_(N)-flip	0	100.0 % Normal
028_P195_(N)	0	100.0 % Normal
035_C022_(N)	0	100.0 % Normal
044_P219_(N)-flip	0	100.0 % Normal
060_P240_(N)-flip	0	100.0 % Normal
067_P247_(N)-flip	0	100.0 % Normal
068_P248_(N)-flip	0	100.0 % Normal
076_P256_(N)	0	100.0 % Normal
076_P256_(N)-flip	0	100.0 % Normal
096_P289_(N)-flip	0	100.0 % Normal
107_P300_(N)	0	66.93 % Normal
113_P314_(N)-flip	0	100.0 % Normal
114_P315_(N)	0	100.0 % Normal
120_P324_(N)	0	100.0 % Normal
120_P324_(N)-flip	0	100.0 % Normal
148_P418_(N)-flip	0	100.0 % Normal
149_P423_(N)-flip	0	100.0 % Normal
151_P427_(N)-flip	0	100.0 % Normal
156_P460_(N)-flip	0	100.0 % Normal
160_P467_(N)	0	100.0 % Normal
162_P469_(N)-flip	0	100.0 % Normal
168_P486_(N)	0	100.0 % Normal
176_P510_(N)-flip	0	100.0 % Normal
181_P526_(N)-flip	0	100.0 % Normal
190_P545_(N)	0	100.0 % Normal

Probability of class membership (3D test dataset)		
Code	Label	ANN model prediction
003_P154_(P)-flip	1	100.0 % Normal
004_P155_(P)	1	100.0 % Parkinsonimo
005_P156_(P)-flip	1	100.0 % Parkinsonimo
015_P179_(P)	1	100.0 % Parkinsonimo
021_P188_(P)-flip	1	100.0 % Parkinsonimo
026_P193_(P)	1	100.0 % Parkinsonimo
033_P200_(P)	1	100.0 % Parkinsonimo
043_P218_(P)-flip	1	100.0 % Parkinsonimo
045_P220_(P)-flip	1	100.0 % Parkinsonimo
050_P225_(P)	1	100.0 % Parkinsonimo
058_P238_(P)-flip	1	100.0 % Parkinsonimo
088_P272_(P)	1	100.0 % Parkinsonimo
095_P288_(P)	1	100.0 % Parkinsonimo
098_P291_(P)	1	100.0 % Parkinsonimo
100_P293_(P)-flip	1	100.0 % Parkinsonimo
104_P297_(P)	1	100.0 % Parkinsonimo
104_P297_(P)-flip	1	100.0 % Parkinsonimo
106_P299_(P)-flip	1	100.0 % Parkinsonimo
111_P304_(P)-flip	1	100.0 % Parkinsonimo
145_P392_(P)-flip	1	100.0 % Parkinsonimo
171_P490_(P)	1	100.0 % Parkinsonimo
173_P500_(P)	1	100.0 % Parkinsonimo
180_P523_(P)	1	99.77 % Parkinsonimo
180_P523_(P)-flip	1	99.98 % Parkinsonimo
187_P534_(P)	1	100.0 % Parkinsonimo
188_P535_(P)-flip	1	100.0 % Parkinsonimo
189_P540_(P)	1	100.0 % Parkinsonimo
195_P561_(P)-flip	1	100.0 % Parkinsonimo
200_P585_(P)-flip	1	100.0 % Parkinsonimo
201_P586_(P)-flip	1	100.0 % Parkinsonimo
205_P594_(P)	1	100.0 % Parkinsonimo

Figure 8.12: Predicted probability of class membership for ANN model using 3D dataset.




```
Donde 0 es Normal y 1 Parkinsonismo
Accuracy: 98.28 %
Precision: 100.0 % Normal 96.77 % Parkinsonismo
Recall: 96.43 % Normal 100.0 % Parkinsonismo
```

Figure 8.13: Confusion matrix of ANN model using 3D dataset.