Helena Rico Pereira

Licenciada em Ciências de Engenharia Biomédica

Classification of patients with parkinsonian syndromes using medical imaging and artificial intelligence algorithms

Dissertação para obtenção do Grau de Mestre em
Engenharia Biomédica

Orientador: Doutor Hugo Alexandre Teixeira Duarte Ferreira, Professor Auxiliar, Faculdade de Ciências da Universidade de Lisboa

Co-Orientador: Doutora Carla Maria Quintão Pereira, Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa [elo de ligação]

Júri:
Presidente: Doutora Carla Maria Quintão Pereira,
Vogais: Doutor José Manuel Matos Ribeiro da Fonseca
Doutor Hugo Alexandre Teixeira Duarte Ferreira

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

hele janeiro 2019 ra

Licenciada em Ciências de Engenharia Biomédica

Classification of patients with parkinsonian syndromes using medical imaging and artificial intelligence algorithms

Dissertação para obtenção do Grau de Mestre em
Engenharia Biomédica

Orientador: Doutor Hugo Alexandre Teixeira Duarte Ferreira, Professor Auxiliar, Faculdade de Ciências da Universidade de Lisboa
Co-Orientador: Doutora Carla Maria Quintão Pereira, Professor Auxiliar, Faculdade de Ciências e Tecnologia da Universidade NOVA de Lisboa [elo de ligação]

Júri:

Presidente: Doutora Carla Maria Quintão Pereira,
Vogais: Doutor José Manuel Matos Ribeiro da Fonseca
Doutor Hugo Alexandre Teixeira Duarte Ferreira

janeiro 2019

Classification of patients with parkinsonian syndromes using medical imaging and artificial intelligence algorithms

For those who suffer from Parkinsonian Syndromes,
patients and their families.

Acknowledgments

Interestingly, this page is the one that truly makes me think about the study that I had made together with a mix of feelings such as nostalgia and self-achievement. It was, for sure, a long period of learning and limit testing. This work would not be possible without the support of several people.

Firstly, I want to thank to IBEB, Instituto de Biofísica e Engenharia Biomédica for the warm reception and for giving me a nice space to work in. To my adviser, Hugo Ferreira, I left a special thank you, for the proposal of such innovative and challenging study and, of course, for the motivation and tips given. Thanks to this piece of work, I really feel that I have learned a lot in these few months. To Professor Carla Quintão which supported me in these five years of Biomedical Engineering increasing in me the interest about the human brain even more.

To all the people that I met at IBEB, especially to Professor Nuno Matela for the help in the installation process and to Professor Alexandre Andrade. To Dona Beatriz for the all cheerful "Bons dias" and for the help in the PC transportation to be fixed. To Daniela Godinho for the explanation about the framework Caffe and MedCNN. To my collegues Tiago, Rita, Magdalena and Sevinç I shared my PC room with, for all our agreeable and cheerful lunches, for the experiences and culture sharing and for the ice-cream moments.

To Sérgio Cláudio of the Computer Services Direction of Faculty of Sciences at University of Lisbon, for the precious and well-disposed help in the installation process.

To all my friends that I made in this last five years at FCT UNL, especially Olivia Pinto, the funniest girl I ever met, Vanessa Ramos for the company at IBEB too, João Borrego for sharing your thesis life with me and Diogo Tecelão, dinofriend, for all the tips. Maria Gonçalves, Cláudia Espalha, Claudia Alves, Sara Lucas a huge thank you too!

To my academic family, Vera Moutinho, Marta Ferreira and Miguel Barbosa for all the support, science videos, true friendship and for understating why I was not so present as I wish to.

To Big, my thai friend, for being the first person to ask me if he could read my thesis after I have finished.

To Pedro Cerqueira for all the by chance meetings at every corner in FCUL.

To Rita Gonçalves, my neighbour and oldest friend, for the support and for sending me all the nature photographs which allow to clear my mind.

The following people are those that when I wrote those lines made me feel that there are no words or sentences to describe how thankful I am. Thus, besides those words, I promise you all a sweat hug.

To my mom for all the wise mother advices, for your smile, for having inspiring me, for the support and specially for making me see the positive side of everything. For listening me about my thesis doubts even though it is not your area. Thanks for being my mom with all letters!

To my dad, the first genius that I met, for inspiring me and motivating me to learn more and more, to help me to stay focused in my objectives, for all the science talks after dinner or in the car trips, and especially for make me curious since I was a child.

To my little sister. For all your patience, advices, your concerning about me, for your funny jokes and our laughs and for almost force me to watch TV with you to take a rest.

To my uncle Zé, for your cheerful mood and for the clever and shortly speech about maths and informatics and for telling me once: "there are no genius, but hard workers."

To grandma Lina, for your delicious and stunning food, for your concerning about my eyes in front of PC and to make me realize how lucky I am to have gone so far.

To grandma Antónia, for your care and your concerning about the hours I spent working.

To grandad Lino for having been an example of never giving up and grandad Chainho, for telling me indirectly that knowledge is one of the most important things in the world.

Lastly, but not the least, to Duarte. Thanks for your strength, kind support and to make me see that there are no impossible things. For your time, even though you were also writing your PhD thesis. For your company and your smile that were fundamental during this period. For all our boat trips.

To all the mysterious and/or unknown forces of the universe.

Obrigada.

This work was financially supported under the grant **PAC nº 16428 – MEDPERSYST**, funded by national funds FCT / MCTES (PIDDAC) under the "Programa de atividades conjuntas"/ P2020.

Abstract

The distinction of Parkinsonian Syndromes (PS) is challenging due to similarities of symptoms and signs at early stages of disease. Thus, the need of accurate methods for differential diagnosis at those early stages has emerged. To improve the evaluation of medical images, artificial intelligence turns out to be a useful tool. Parkinson's Disease, the commonest PS, is characterized by the degeneration of dopamine neurons in the *substantia nigra* which is detected by the dopamine transporter scan (DaTscan™), a single photon-emission tomography (SPECT) exam that uses of a radiotracer that binds dopamine receptors. In fact, by using such exam it was possible to identify a sub-group of PD patients known as "Scans without evidence of dopaminergic deficit" (SWEDD) that present a normal exam, unlike PD patients. In this study, an approach based on Convolutional Neural Networks (CNNs) was proposed for classifying PD patients, SWEDD patients and healthy subjects using SPECT and Magnetic Resonance Imaging (MRI) images. Then, these images were divided into subsets of slices in the axial view that contains particular regions of interest since 2D images are the norm in clinical practice. The classifier evaluation was performed with Cohen's Kappa and Receiver Operating Characteristic (ROC) curve. The results obtained allow to conclude that the CNN using imaging information of the Basal Ganglia and the mesencephalon was able to distinguish PD patients from healthy subjects since achieved 97.4% accuracy using MRI and 92.4% accuracy using SPECT, and PD from SWEDD with 97.3% accuracy using MRI and 93.3% accuracy using SPECT. Nonetheless, using the same approach, it was not possible to discriminate SWEDD patients from healthy subjects (60% accuracy) using DaTscan™ and MRI. These results allow to conclude that this approach may be a useful tool to aid in PD diagnosis in the future.

Keywords: Parkinsonian Syndromes Classification - SWEDD - DatScan™ - MRI – Convolutional Neural Networks

Resumo

A distinção de Síndromes Parkinsonianos (SP) tornou-se um desafio dado que os sintomas e os sinais característicos destas doenças são muito semelhantes entre si nas fases iniciais. Assim surge a necessidade da criação de métodos precisos para um diagnóstico que as diferencie. Uma possível solução é o recurso à inteligência artificial que se tem vindo a tornar numa ferramenta útil para a melhoria da interpretação das imagens médicas. A doença de Parkinson (DP), a SP mais comum, é caracterizada pela neurodegeneração dos neurónios dopaminérgicos. Esta neurodegeneração é detetada pelo DaTscan™, um exame de tomografia computorizada por emissão de fotão único (do inglês Single-Photon Emission Computed Tomography, SPECT). A utilização deste exame permitiu a identificação de um subgrupo de pacientes diagnosticado com DP, denominados de "exames sem evidência de défice dopaminérgico" (do inglês "Scans without evidence of dopaminergic deficit", SWEDD). Os SWEDD são pacientes que possuem um DatScan normal ao contrário dos pacientes com PD. Nesta dissertação, é proposta uma abordagem com recurso às redes neurais convolucionais (do inglês: Convolutional Neural Networks, CNNs) usando imagens SPECT e Imagens por Ressonância Magnética (IRM) que foram divididas em cortes axiais contendo regiões de interesse específicas, dado que na prática clínica as imagens convencionais são imagens a 2D, para classificação de doentes com PD, SWEDD e pessoas saudáveis. Para a avaliação do classificador calculou-se o Kappa de Coehn e a Característica de Operação do Receptor. Os resultados obtidos sugerem que CNN foi capaz de diferenciar DP do grupo de controlo (exatidão: 97,4% IRM e 92,4% SPECT) e DP de SWEDD (exatidão: 97.4% IRM e 93.3% SPECT) com recurso a informação de imagem dos gânglios da base e mesencéfalo, mas não foi capaz de discernir controlos de SWEDD. Deste modo, conclui-se que esta abordagem poderá ser útil para auxiliar no diagnóstico de DP no futuro.

Palavras-Chave: Classificação de síndromes parkinsonianos – SWEDD - DaTscan™ – IRM – Redes neuronais convolucionais

Contents

List of Figures

List of Tables

Acronyms

AC	Anterior Commissure
AD	Alzheimer's Disease
ANN	Artificial Neural Networks
AP	Atypical Parkinsonism
CBD	Corticobasal Degeneration
CBGTC	Cortico-Basal Ganglia-Thalamo-Cortical
CNN	Convolutional Neural Networks
CPU	Central Processing Unit
DaT	Dopamine Transporter
DICOM	Digital Imaging and Communication in Medicine
ET	Essential Tremor
FC	Fully Connected
FWHM	Full Width at Half Maximum
GE	Gradient-Echo
GPe	External Globulus Pallidus
GPi	Internal Globulus Pallidus
GPU	Graphics Processing Unit
HY	Hoehn and Yahr
LMDB	Lightning Memory-Mapped Database Manager
LSSVM	Least Squares Support Vector Machine to classification
MAS	Multiple System Atrophy
MDS- UPDRS	Movement Disorder Society - Unified Parkinson's Disease Rating Scale
MedCNN	Convolutional Neural Networks for Medical Applications
MNI	Montreal Neurologic Institute
MNIST	Modified National Institute of Standards and Technology
MRI	Magnetic Resonance Imaging
NIFTI	Neuroimaging Informatics Technology Initiative
OSA	Objective Striatal Analysis
PC	Posterior Commissure
PD	Parkinson's Disease
MCI	Mild Cognitive Impairment
PPMI	Parkinson's Progression Markers Initiative
PS	Parkinsonian Syndromes

PSP	Progressive Supranuclear Palsy
ReLU	Rectified Linear Unit
RF	Radiofrequency
SE	Spin-Echo
SGD	Stochastic Gradient Descent
SNc	Substantia Nigra pars compacta
SNr	Substantia Nigra pars reticulata
SOM	Self-Organized Map
SPECT	Single Photon-Emission Tomography
SPM	Statistical Parametric Mapping
STN	Substantia Nigra
SVM	Support Vector Machine
SWEDD	Scans Without Evidence of Dopaminergic Deficit
T1-W	T1-Weighted
T2-W	T2-Weighted
TE	Echo Time
TR	Repetition Time
UKPSBB	United Kingdom Parkinson's Disease Society Brain Bank
VBM	Voxel-Based Morphometry
VLN	Ventral Lateral Nucleus

Chapter 1 Introduction

1.1 Motivation

Parkinsonian Syndromes (PS) are progressive neurodegenerative disorders that mainly affect elderly people [1], [2]. PS, also known as Parkinsonism, are chronic movement disorders that affect the central nervous system and are characterized by motor symptoms such as slow/impairment movement with rigidity and/or tremor [1], [3]. Inside of the group of PS, Parkinson's Disease (PD) is the second neurodegenerative disorder more frequent in the world and the commonest PS followed by the Atypical Parkinsonism (AP) [1]. AP includes Multiple System Atrophy (MSA), Progressive Supranuclear Palsy (PSP) and Corticobasal Degeneration (CBD) [1], [3]. About of 75% of PS cases are PD, 5% are PSP and 5% are CBD [4].

According to epidemiologic and demographic studies, as world population is increasing as well as aging, it is estimated that PS, specially PD, will drastically increase in the following decades [1]. According to World Health Organization, PD has an incidence rate of 9.7 to 13.8 per 100,000 population per year [5] and approximately a prevalence of 1 million people in United States and in Western Europe [1]. In Portugal, it is estimated that about 18,000 inhabitants are also suffering from this disease [6].

PS are not only characterized by motor symptoms. Patients may also suffer from mood disorders, sleep disturbances, cognitive impairment, urinary incontinence, sexual dysfunction and orthostatic hypotension [1], [5]. However, it is important to notice that these features and the motor symptoms tend only to be manifested at advanced stages of the diseases [1], [7]. At early stages of these diseases, the symptoms are not yet manifested or are very subtle [7]. Besides, when features are manifested even at early-onset, they tend to be very similar between PS and other movement disorders such as Essential Tremor (ET) [8]. Thus, it still is very difficult to differentiate them and have a precise diagnosis [9].

Currently, there are approaches and scales to diagnose PS based on their clinical features, namely the United Kingdom Parkinson's Disease Society Brain Bank (UKPDSBB) clinical diagnostic criteria [10] and the Hoehn and Yahr Scale [11]. This criteria and scale include motor symptoms and a good response to Levodopa, a PD medicine. PD patients have an excellent response to that medication unlike AP who present poor response [1]. This medication acts in the affected and pathological regions of PS [1].

Pathologically, PS are characterized by a deficit of dopamine that originates motor disorders. The dopamine is produced in the *substantia nigra*, a structure in the mesencephalon [1] In PS, this structure is degenerated faster than normal when compared with elderly people [1].

Despite the fact that PS diagnosis remain clinical, medical imaging have been useful for the study of neurodegenerative diseases [12–14]. For instance, Magnetic Resonance Imaging (MRI) has been used to detect structural changes and distinguish PS and Single-Photon Emission Computed Tomography (SPECT) with Dopamine transporter (DaT) imaging has been used to detect the referred dopamine degeneration and distinguish PS from other motor diseases. Both techniques turn out to be useful in providing neuroimaging biomarkers [1], [12]. In fact, these SPECT images led to the discovered of a sub group of patients early diagnosed with PD that presented Scans Without Evidence for Dopaminergic Deficit (SWEDD) [12], [13], [15], [16]. These subjects are now referred to SWEDD patients. Some studies refer that 10-20% of PD patients are SWEDD patients and point out that this may related with misdiagnosis and instead of PD these patients may suffer from another motor disorder in which there is no *substantia nigra* degeneration like dystonic tremor [17], [18].

Usually, DaT SPECT images are interpreted visually by experts but the European Association of Nuclear Medicine Neuroimaging Committee recommends that a quantitative analysis should also be performed to improve the interpretation [19], which includes the use of Artificial Intelligence (AI) techniques [16].

AI has recently become an useful tool in the medical field, specially applied in the classification and analysis of neuroimaging data [16], [20]–[22] which includes Convolution Neural Networks (CNN) which are an useful tool for pattern recognition and visual-classification problems and have been used in the study of neurodegenerative disorders [22]–[25].

Therefore, distinguishing early stages PS and other movement disorders with similar symptoms has become a need and a challenge [26]. An early diagnosis not only reduces the economic health care systems costs but also aid in the treatment adjustment since the therapeutics is different for each PS and for the SWEDD patients [5]. Moreover, this will also increase their quality of life of patients [5].

In this study, it is proposed an approach to aid and improve early diagnosis of Parkinson's Disease and SWEED patients based on a CNN framework using 2D MRI and 2D DaT SPECT images since the conventional medical images, in clinical practice, are 2D images [27] which are faster to acquire than 3D images.

1.2 Goals of the study

The main goal of this dissertation is to classify MRI and SPECT images of PD, SWEDD and healthy subjects using CNN.

To achieve this, the following secondary objectives will have to be reached:

1. Extract MRI and DaT (DaTscan™ in this case) SPECT images of the Parkinson's Progression Markers Initiative (PPMI) database and analyse demographic information about the patients, such as age and sex;
2. Perform MRI and SPECT images pre-processing;
3. Divide MRI and SPECT images into slices of the axial anatomical view;
4. Train, validate and test the CNN model.

With this approach it is intended to aid and to improve the early diagnosis of PD and to understand SWEED and PD imaging differences.

1.3 Dissertation Structure

This dissertation is comprised by this present chapter, the Theoretical Background chapter, the Methodology chapter, the Results chapter, the Discussion chapter and the Conclusions chapter.

The present chapter, Introduction, briefly introduce some important concepts used in this dissertation, contextualized the study and enumerate the objectives. In Chapter 2, Theoretical Background, it is explained theorical concepts about PS, MRI and SPECT images and CNN, and it is described the state of the art of MRI and SPECT images in the study of PS and CNN applications in the medical field as well as other algorithms to classify PS. In Chapter 3, the Methodology, the criteria used for the formation of the data set is described as also MRI and SPECT images pre-processing steps, CNN architecture and how these were trained, validated and tested. In Chapter 4, the results obtained are described in the form of tables and representative plots. In Chapter 5, Discussion, the discussion of the results obtained is presented with a brief comparison with other studies. In the last Chapter, the conclusion and future work are presented.

Chapter 2 Theoretical Background

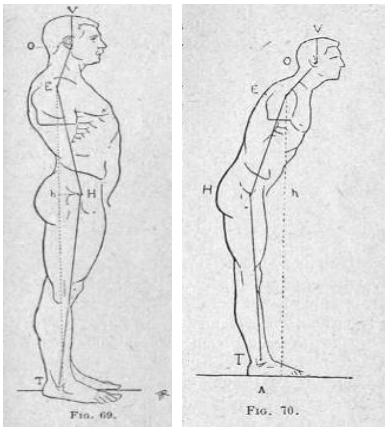
In this Chapter, the theoretical concepts related with PS, MRI and SPECT DaTscan™ images acquisition and CNN architectures is described as the state of the art in the PS field including how MRI and SPECT images are used for the study of these diseases and the studies that used CNN as tool for PS classification.

2.1 Parkinsonian Syndromes

2.1.1 The History of Parkinsonian Syndromes

PS has been described since early ages. The first texts found go back to ancient China and India, between 1000 and 500 B.C.[28], [29]. These documents describe symptoms very similar to those in parkinsonism [28], [29].

In 1817, PD was set as a neurological syndrome by the physician James Parkinson who observed and studied six cases of patients with PD, initially know as Shaking Palsy [30], [31] Parkinson described Shaking Palsy patients as having involuntary tremor, decrease of muscular strength and tendency to lean forward [30], [31]. In the middle-to-late of 19th century, Jean-Martin Charcot and his students could describe clinically this disease and discovered that Shaking Palsy patients do not necessarily present tremor as symptom and that these patients were not so weak as previously thought, which made them change the name of the disease for Parkinson' Disease [30]. In 1895, Richer and Miege gave a very important contribution to the understanding of PD progression through the publication of drawings and statues photos that illustrated the disability stages of this disease [30], [32], as illustrated by Figure 2.1.



a) b)

Figure 2.1 Drawings of the normal and pathologic erect position. a) Normal men have an erect position. Head in the vertical axis, as well as his back b) Parkinsonian patients tend to tilt their heads forward to the foot vertical axis. Adapted from Goetz [30]

A few years later, Brissaud advanced with the possibility that *substantia nigra* damage could be the anatomical cause of PD and Greenfield and Bosanquet reported a full pathological analysis of PD as well as Parkinsonism based on brain injuries [30], [33]. Charcot studies also reported that PD patients can also present bradykinesia, that is, slow movements, as the responsible cause for difficulty in performing ordinary daily activities instead of the tremor or rigidity [30], [32], [34]. Moreover, this author discovered that some PD patients differed in their symptoms [32], [34].

Initially, Charcot aimed to distinguish PD from other neurological diseases such as Multiple Sclerosis (MS) [34]. Charcot found that PD patients differ from MS and other diseases by the fact that they present rest tremor, rigidity, very soft speech, bradykinesia and a hunched posture [30], [32], [34]. He and his students were the first to report the atypical features present by some patients. These patients were characterized by an extended posture, a different facial expression and no tremor as shown in Figure 2.2, which are different features from those manifested by PD patients [34]. Because of that, the term Parkinson-plus-syndromes or PS has coined [29].

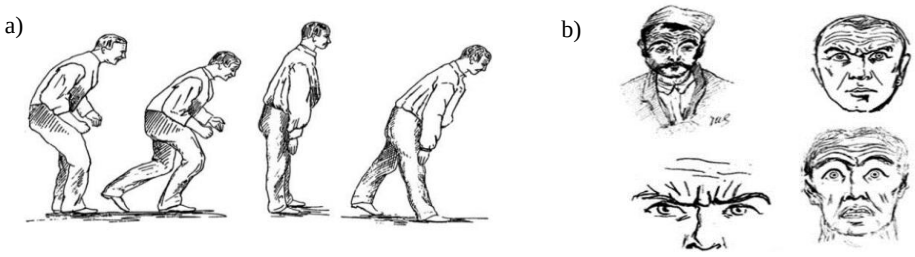


Figure 2.2 Charcot's drawings at his lesson (June 1888) illustrating Parkinsonian Syndromes. a) Men at left has PD and it is characterized by a curved posture. The men on the right has Parkinsonism with an extended posture. b) These four drawings illustrate an atypical Parkinson's Disease case of a Charcot's patient. This patient shows a frightened expression, forehead muscles contracted despite the placid and blank stare of PD patients. Adapted from: [34]

Nowadays, PS include PD, PD with dementia with Lewy Bodies, AP, Secondary Parkinsonism and other neurodegenerative diseases such as Huntington's disease and Alzheimer's disease with parkinsonism [1], [3]. As referred previously AP includes MSA, PSP and CBD and, in its turn, Secondary Parkinsonism covers drug-induced, infections and toxins [1].

2.1.2 Clinical and pathophysiology of the Parkinsonian Syndromes

The following lines describes the clinical status and pathology of the PD, MSA, PSP and CBD.

Patients with PS usually manifests identical motor symptoms such as bradykinesia, tremor and rigidity [1], [3]. These symptoms are associated with damage in structures responsible for motor control in the brain, namely, basal ganglia., and these injuries depend on the type of PS that the patient has. This damage in the basal ganglia is associated to dopaminergic neurons degeneration and consequently the dopamine deficit in specific brain areas [1], [3], [35].

Since PS are neurodegenerative diseases it is important to briefly explain some concepts related

with brain structures before the explanation of PS pathophysiology itself. The brain is comprised by cells called neurons which are small structures that communicate to each other by transmitting stimulus and sensations from the environment and command the body to respond to this stimulus, [36]–[38], as shown on Figure 2.3 a). These cells communicate by sending the stimulus through a contact zone between neurons called synapse [37], [38]. The synapses are divided into two types, electric synapses and chemical synapses. The last ones are the commonest in the brain. In the electric synapses, the neurons terminals are physically connected by protein channels that enable the passage of the electric impulse. In the chemical synapse the impulse electric pass from one neuron to another through chemical substances called neurotransmitters. The axon terminal from the transmitter neuron has a terminal button that contains vesicles with neurotransmitters inside. These vesicles merge with the pre-synaptic membrane and the neurotransmitters are released in the synaptic gap. Then, these neurotransmitters bind to specific receptors located in the postsynaptic membrane in dendrites' terminations of the receptor neuron. Thus, the electric impulse is propagated through the neuron [37], [38], as illustrated on Figure_2.3 b).

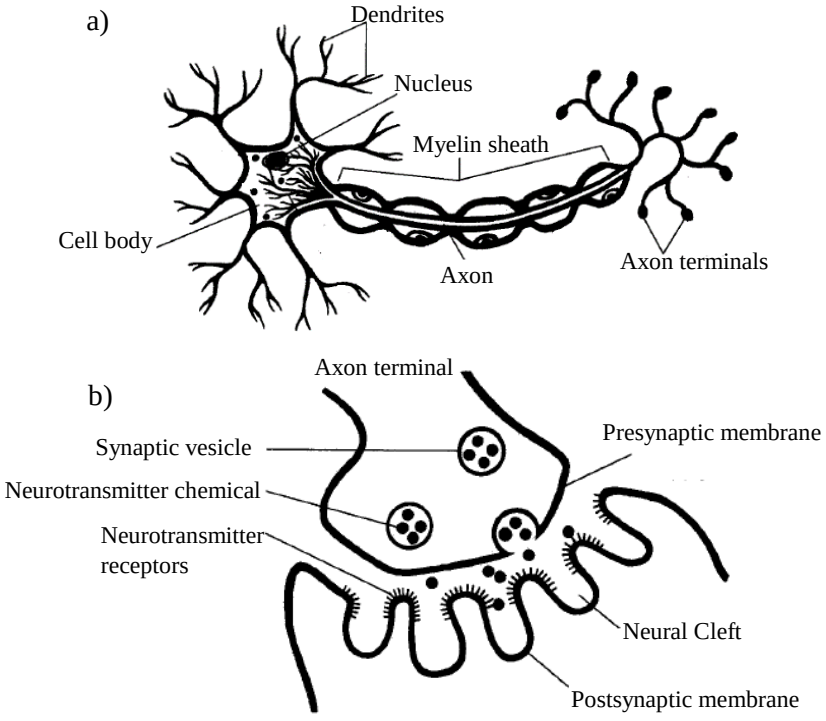


Figure 2.3 Neuron structure and chemical synaptic. a) Normal structure of a neuron b) Chemical synaptic. Neurotransmitters are messengers that send the information from one neuron to other. Adapted from [38]

Regarding PD, this disease is strictly related with neurons degeneration and deficit of dopamine, an important neurotransmitter related with motor control among other functions. [1]

The pathological hallmarks of PD are degeneration of the dopaminergic neurons located at the *substantia nigra,* as shown in Figure 2.4, reduction of the striatal dopamine, and abnormal

accumulation of protein aggregates, known as Lewy bodies, in injured cells of the *substantia nigra*, [1], [39]. These pathological features are responsible for the motor symptoms and signs that PD patients present. For instance, the neurodegeneration that occurs in *substantia nigra* affects all the basal ganglia cycle which, together with thalamus, are responsible for motor control functions. Once the cycle is affected, the information that reaches the motor cortex is not regulated, which leads to changes in movement [1], [40].

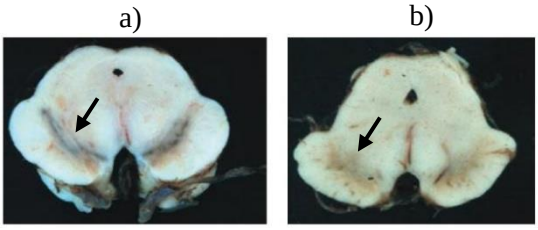


Figure 2.4 *Substantia nigra* a) *Substantia nigra* of a healthy subject, the presence of the neuromelanin dopaminergic neurons are visible in the image due to its characteristic black color. b) *Substantia nigra* from a PD patient, wherein the absence of dopaminergic neurons is visible. Adapted from: [12]

In healthy people, the basal ganglia system controls the motor system and avoid involuntary movements. In addition to motor control, basal ganglia are also responsible for cognitive functions and behaviour control [1]. The process of motor control is possible due to connections between basal ganglia structures along with the thalamus and the cortex. This junction responsible for motor control is called the Basal Ganglia model or Cortico-Basal Ganglia-Thalamo-Cortical (CBGTC) loop. The basal ganglia are comprised by two primary inputs, striatum (caudate nucleus and putamen) and subthalamic nucleus (STN), by two primary outputs, internal segment globulus pallidus (GPi) and *substantia nigra pars reticulata* (SNr) and by two intrinsic structures, external segment globulus pallidus (GPe) and *substantia nigra pars compacta* (SNc), as shown in Figure 2.5. [1], [41], [42]

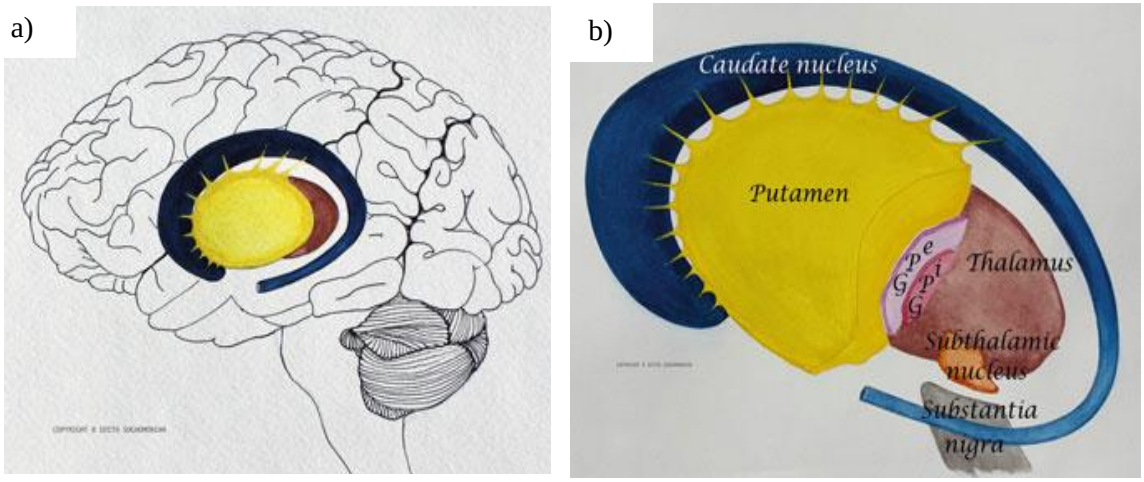


Figure 2.5 Basal ganglia structures. a) The color structures indicate the anatomical position in the brain of the basal ganglia, putamen and caudate nucleus b) Names of the structures of the basal ganglia with striatum (putamen and caudate nucleus) and thalamus. Adapted from: [41]

The basal ganglia model is organized into two pathways: the direct pathway and indirect pathway that act together to control the body movements. The direct pathway occurs when the motor cortex is stimulated, for instance, to initiate or perform a movement. In its turn, the indirect pathway stops a movement or prevent the body from doing involuntary movements [1], [42].

In the direct pathway, the motor cortex sends an excitatory stimulus to the striatum and the STN. In its turn, the striatum sends an inhibitory stimulus to GPi and SNr, whereas STN sends an excitatory stimulus to GPi, SNc, and SNr. SNc modulates neuronal firing and equilibrates the basal ganglia cycle by transmitting and increasing dopamine neurotransmitters to the striatum receptors called D1, which will raise the activity in the striatum cells. Thus, the striatum increases the inhibitory stimulus sent to the GPi. In its turn, GPi reduce its activity and does not send any signal to the Ventral Lateral nucleus (VLN) of thalamus which allows it to send excitatory stimulus to the motor cortex and spinal cord to allow motor functions [40], [42].

In the indirect pathway, the striatum and STN also receive the excitatory stimulus from the cortex. Then, the striatum sends an inhibitory stimulus to the GPe and the STN to the SNc. The SNc will increase the dopamine in the striatum in the D2 receptors which lead to a decrease of an inhibitory impulse to the GPi. Consequently, GPi and SNr can send inhibitory signals to VLN. Then, the activity of the motor cortex is decreased, and no movement is executed since VLN cells get the stimulus of inhibition. All this process is illustrated in Figure 2.6 [40], [42].

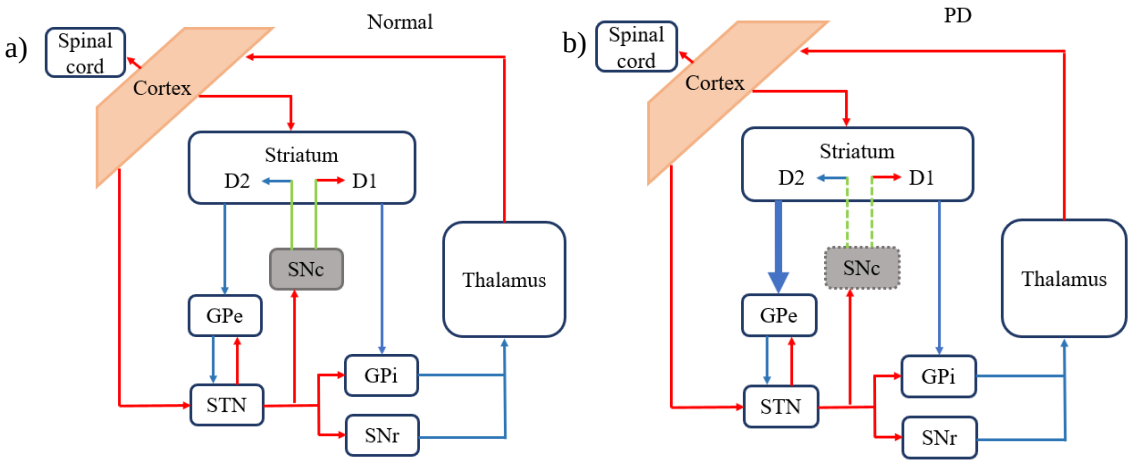


Figure 2.6 Comparison of the CBGTC from a heathy person and PD patient. The red arrows represent the excitatory stimulus, the blue arrows represent the inhibitory stimulus and the green illustrate dopaminergic neurons. These neurons originate an excitatory stimulus if they are bounded to the D1 receptor or inhibitory if they are bounded to the D2 receptor. In a) there is represented the normal CBGTC. b) illustrates the CBGTC of a subject with PD. In this case, the dashed green lines represent the dopamine deficit which causes an increase of the inhibitory stimulus in the GPe (blue arrow in bold) which lead to motor symptoms. Adapted from:[1]

In PD, the basal ganglia suffer dopaminergic neurons degeneration in the SNc which become pathologically degenerated and leads to the dopamine deficit in the striatum. Consequently, there is an increase of neurons fire in SNr and GPi. These two structures give rises to excessive inhibition of the thalamus and consequently an extreme reduction of the activity in the motor

cortex that originates the parkinsonian features [1], [40]. These motor symptoms arise when approximately 50% of the *substantia nigra* has already degenerated [43].

Besides the degeneration of the dopaminergic neurons and the reduction of dopamine in the system, post-mortem studies show an evidence of an abnormal aggregated of specific proteins in the injured neurons, known as Lewy Bodies [1]. These structures have a spherical form and are filamentous inclusions formed by presynaptic protein alpha-synuclein that combines with other components present in the neurons [1], [44], [45]. The reason why this agglomerated forms and why the protein α-synuclein pull out of its binding sites in the presynaptic axon terminal still unknown [44], [45].

To summarize, clinical PD is based on several features that include motor and non-motor symptoms [1]. Non-motor symptoms such as depression and dementia tend to manifest in the last stages of the disease, as shown in Table 2.1[1].

Table 2.1 Clinical features of Parkinson's disease. This table introduces the clinical features that PD patients manifest through the course of the disease. Anosmia is usually the onset of PD. Motor signals tend to manifest at advanced stages since they are related with degeneration of the SNc. Adapted from [1]

Clinical features of PD		
Principal Motor Features	**Other Motor Features**	**Non-Motor Features**
Bradykinesia **(2)**	Reduced eye blinking	Anosmia **(1)**
Rest tremor **(2)**	Soft voice	Sensory disturbances (e.g.,
Rigidity	Dysphagia	pain)
Gait impairment	Masked facies (hypomimia)	Mood disorders (e.g.,
		depression) **(3)**
		Sleep disturbances **(1)**
		Autonomic disturbances
		Cognitive impairment **(3)**

(1) E.g. of symptoms/signals that manifests are early stages of PD

(2) E.g. of symptoms/signals that manifests at middle stages of PD

(3) E.g. of Symptoms/signals that manifests at advanced stages of PD

Concerning PD treatment, since the 1960s that this therapy includes a medicine called levodopa [1], [46]. Since artificial dopamine does not cross the blood-brain barrier, a barrier that protects the brain from strange particles and molecules, scientists create levodopa (commonly known as L-DOPA) which is a dopamine precursor. Levodopa is capable of penetrating in the brain and then is converted into dopamine [1], [46]. However, Levodopa has some side effects, namely, nausea, vomiting, and orthostatic hypotension. Moreover, at advanced stages of this illness and with a long-term of taking this medicine, PD patients may manifest excessive involuntary muscle

contractions, which may originate abnormal postures and diphasic dyskinesias [1], [47], [48]. These symptoms appear due to chronic administration of the Levadopa, which causes fluctuations as a decrease of its effect in the organism [1], [49].

Concerning AP and Secondary Parkinsonism, these syndromes are also neurodegenerative diseases but differ from PD since the damage of the neurons can also occur in the striatum and globus pallidus [1]. Besides, these types of parkinsonism do not present Lewy bodies in the neurons of the SNc. [1]. In the beginning, AP and Secondary Parkinsonism patients present a good response to levodopa, but along the disease progression of the disease, the subjects starts to present a poor response [1].

More specifically, MSA can be parkinsonism (MSA-p) or cerebellar (MSA-c) predominant according to their main symptoms [1]. Clinically, MSA-patients can also present autonomic features because the autonomic nervous system is also affected [1], [50], [51]. The autonomic nervous system innervates all the organs of the body and is responsible for maintaining the homeostasis of the body by controlling blood pressure and heart rate [52], [53]. One example of autonomic dysfunctions that these patients manifest is orthostatic hypotension which can cause symptoms such as erectile dysfunction in males or urinary problems [50]. According to Harrison's Principles of Internal Medicine, "Pathologically, MSA is characterized by degeneration of the SNc, striatum, cerebellum, and inferior olivary nuclei coupled with characteristic glial cytoplasmic inclusions (GCIs) that stain for α-synuclein*"* [1].

Regarding PSP, also known as Steele–Richardson–Olszewski syndrome, is pathologically characterized by loss of neurons in the striatum, SNc, subthalamic nucleus, midline thalamic nuclei and abnormal inclusions in the tau protein [1], [54]. PSP symptoms are restricted eye movements, gait impairment that leads to falls, and cognitive dysfunctions. These symptoms occur at different stages since it affects different nuclei [54]–[56]. Unfortunately, PSP develops at a very fast pace. The first symptoms usually manifest at the mid-60s with a faster progression that leads to death in 7 years on average [57]. Until now, the advanced age it the only risk factor of this disease [57].

CBD is a rare condition that affects the cerebral cortex and the basal ganglia by loss and atrophy of the neurons [1], [58]. The progression of the disease is slow and occurs during 6 to 8 years [58] and causes dystonic contractions and sensory disturbances. [1]

Overall, there are specific symptoms for each PS. For instance, the absence of tremor, early speech, and gait impairment suggests a possible Atypical Parkinsonism; manifestations of hallucinations and dementia as the first signal as well as Parkinsonian features indicate a Dementia with Lewy bodies; diplopia, impaired down gaze may suggest PSP; predominant orthostatic hypotension may indicate an MSA [1].

It is important to notice that the causes of Parkinsonian Syndromes are still unknown. [1], [3], [57] Early onset of PD before the age of 40 may indicate a genetic form of PD caused by genetic factors including gene mutations, especially mutation of the alpha-synuclein gene. However, most of the cases, approximately 85–90%, the factors that induce PD is not known yet. Some studies correlate environmental and lifestyle factors with idiopathic PD, i.e., Parkinson with an unknown cause [1], [59]. For instance, Hwang studied the possibility of the oxidative stress due to neuroinflammation being related to the SNc degeneration [60]. In 2003, H. Braak et al. advanced with the hypothesis that the protein alpha-synuclein arises at the gastrointestinal system and then spreads to the brain through the vagus nerve [39]. However, this lacks more studies and evidence to prove that this process may happen [61].

2.2 De novo PD patients

De novo PD patients is a clinical expression to define a group of PD-patients that either does not receive any Levodopa treatment or were recently diagnosed as having Parkinson's Disease [62].

2.3 Non-Parkinson's disease tremor: Essential Tremor

Essential Tremor (ET) is a common movement disorder, monosymptomatic, of unknown cause and sometimes misdiagnosed with neurodegenerative diseases such as PS [1], [63], [64]. ET patients manifest symmetric high-frequency tremor (6–10 Hz) and voice or head tremor but show no evidence of changes in the dopamine in the dopamine system [1], [7], [63].

2.4 State-of-art: diagnosis of Parkinsonian Syndromes

Over the years, several authors have proposed methods for PS diagnosis by elaborating scales that enumerate characteristics and signals for each illness [65]. In this study, only the most common were described.

In 1967, Hohen and Yhar studied the progression and morbidity of patients with parkinsonism and developed a scale of PD progression, the Hoehn and Yhar (HY) Scale, which are nowadays accepted and internationally used [11]. HY Scale enumerates PD stages by describing symptoms that occur at each level. This scale was created to be used in follow-up progression of PD before, during and after treatment and consequently to promote a better evaluation of the therapy [11]. The scale comprises five stages that are based on the symptoms and the clinical degree of disability of the patients. In Table 2.2, the HY Scale is entirely reproduced without alterations. In the 1990s, HY Scale was slightly modified and it was added two intermediary levels: Stage 1.5: "Unilateral and axial involvement " and Stage 2.5: "Mild bilateral disease with recovery on pull test " [66]. This modified HY Scale emerged due to HY scale limitations, namely the fact that some patients may manifest the same symptoms at different stages [66]. Despite this, HY scale still be used for medical evaluation worldwide for PD symptom stage evaluation. [66]

Table 2.2 - Hoehn and Yahr Scale. This scale created in 1967 represents 5 stages of Parkinson´s Disease. Stages I, II and III represent the less levels of disability whereas Stage IV and V represent severe levels of disability. Entirely obtained from [11]

Hoehn and Yahr Scale	
Stage I	"Unilateral involvement only, usually with minimal or no functional impairment."[11]
Stage II	"Bilateral or midline involvement, without impairment of balance." [11]
Stage III	"First sign of impaired righting reflexes. This is evident by unsteadiness as the patient turns or is demonstrated when he is pushed from standing equilibrium with the feet together and eyes closed. Functionally the patient is somewhat restricted in his activities but may have some work potential depending upon the type of employment. Patients are physically capable of leading independent lives, and their disability is mild to moderate."[11]
Stage IV	"Fully developed, severely disabling disease; the patient is still able to walk and stand unassisted but is markedly incapacitated."[11]
Stage V	"Confinement to bed or wheelchair unless aided."[11]

In 2003, H. Braak *et al.* studied 41 cases of subjects clinically diagnoses as PD. Based on the clinical and pathological analysis of the patients, Braak et al. present a study that describes six stages of PD progression and the distinction of the initial, intermediate and final phases of the disease [39]. These stages were formed taking into account the presence of Lewy bodies as well all pathologies and symptoms manifested by patients and related studies. The Braak stages are following described.

- **Stages 1 and 2:**

  In this phase of the illness, PD patients present non-motor symptoms that precede motor symptoms [67]. According to the literature, PD-patients tend to suffer from an impaired sense of smell that arises before the motor symptoms [68], [69]. Thus, one study points out that an early exam of the olfactory performance when this symptom rise could evidence a possible early PD diagnosis [70]. Besides smell impairment, PD patients may also manifest autonomic dysfunctions and Idiopathic Rapid Eye Movement Sleep Behaviour Disorder, during the early stage of the disease [39], [71].

- **Stages 3 and 4:**

  According to Braak et al., the pathology of stage 3 and 4 given by the injuries, i.e., the dopaminergic neurons start to get injured or degenerated which leads to the appearance of motor symptoms[39]. In stage 4, the damage in the olfactory nucleus is getting more severe.

It also in these two stages, 3 and 4, that Lowy bodies get formed and present in the lesioned neurons [39].

- **Stages 5 and 6:**

These two stages are the most severe in this disease [39]. In stage five, the pathology reaches areas in the neocortex and prefrontal neocortex. In stage six, the neurodegeneration in the SNc is evident, and the disease develops affecting deeper regions such as the sensory and motor areas [39].

In 2007, Goetz et al. presents a revision of the Unified Parkinson's Disease Rating Scale (UPDRS), one of the most used scales for the clinical study of PD [72]. The UPDRS scale is a questionnaire with several questions divided into four parts concerning disabilities and impairments to classify PD severity [73]. This revision, also known as Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS), adds a new section that incorporates nonmotor features of PD [72] and allows the identification of the PD problems that a PD-patient may develop [72], [73]. Part I - Nonmotor Aspects of Experiences of Daily Living includes features of dopamine dysregulation syndrome, urinary problems, and cognitive impairment.; Part II - Motor Experiences of Daily Living contains walking and balance, tremor impact on activities and eating tasks; Part III - Motor Examination includes speech, Hand movements, gait, Postural tremor of hands, rest tremor amplitude and Part IV: Motor Complications covers dyskinesias (involuntary muscle movements) and motor fluctuations [1], [72]-

Nowadays, the most used criteria are the UKPSBC [74]. With 80% of accuracy, UKPSBC were created by the analyses of 100 patients with Idiopathic Parkinson's Disease [10], [65]. This study found out that some of these patients were misdiagnosed as PD, instead of having this disease, they suffer from some other PS or Alzheimer's Disease. UKPSBC is formed by three steps. The first step enumerates and characterizes motor symptoms of PS such as bradykinesia or rest tremor, as shown in Figure_2.7. The second step lists pathologies that are not present at PS namely severe dementia at early ages and negative response to PD medication, levodopa. The last step describes signs and characteristics that the patient must have to be diagnosed with PD and includes good response to PD medication Levodopa and rest tremor presence. In this step, the diagnosis is valid when the patient presents at least three attributes [10].

To study mild cognitive impairment in neurologic diseases it is was develop a 10-minute cognitive screening tool called The Montreal Cognitive Assessment (MoCA) [75]. This test analyses cognitive domains such as attention and conceptual thinking [76].

UK Parkinson's Disease Society Brain Bank clinical diagnostic criteria

Step 1 Diagnosis of Parkinsonian syndrome

- Bradykinesia (slowness of initiation of voluntary movement with progressive reduction in speed and amplitude of repetitive actions)
- And at least one of the following:
  - muscular rigidity
  - 4–6 Hz rest tremor
  - postural instability not caused by primary visual, vestibular, cerebellar, or proprioceptive dysfunction.

Step 2 Exclusion criteria for Parkinson's disease

- History of repeated strokes with stepwise progression of parkinsonian features
- History of repeated head injury
- History of definite encephalitis
- Oculogyric crises
- Neuroleptic treatment at onset of symptoms
- More than one affected relative
- Sustained remission
- Strictly unilateral features after 3 years
- Supranuclear gaze palsy
- Cerebellar signs
- Early severe autonomic involvement
- Early severe dementia with disturbances of memory, language, and praxis
- Babinski sign
- Presence of cerebral tumour or communicating hydrocephalus on CT scan
- Negative response to large doses of levodopa (if malabsorption excluded)
- MPTP exposure

Step 3 Supportive prospective positive criteria for Parkinson's disease

(Three or more required for diagnosis of definite Parkinson's disease)

- Unilateral onset
- Rest tremor present
- Progressive disorder
- Persistent asymmetry affecting side of onset most
- Excellent response (70–100%) to levodopa
- Severe levodopa-induced chorea
- Levodopa response for 5 years or more
- Clinical course of 10 years or more

Figure 2.7 United Kingdom Parkinson's Disease Society Brain criteria. Elaborated by [77]

For AP such as MSA, PSP and CBD several authors proposed clinical diagnostic criteria based on post-mortem studies and specific features [65]. In 1994, Quinn proposed a diagnostic criteria based on specific features of MSA: Possible diagnostic of MSA includes sporadic adult-onset and poor response to levodopa, and Probable diagnostic of MSA includes Possible diagnostic of MSA and severe autonomic failure [78]. For PSP, Lees describes clinical features that differ PSP from PD, namely the type of response for Levodopa (poor or excellent) and differences between balance, speech and facial appearance in the two illness[56]. Regarding CBD, there are no diagnostic criteria that were validated so far [65], [79].

The first diagnoses of Parkinsonian syndromes were merely based on visual signs and symptoms [1], [30] However, post-mortem studies found a 24% of misdiagnosis [1]. Thus, several clinically criteria such as those mention above have been created. Besides clinical diagnosis, it is also essential to study the pathology of the disease, through methods that may aid in the diagnosis. [80]. Due to this need, medical imaging in Parkinsonian syndromes has emerged to show anatomical and functional changes related to PD pathologies [1].

2.5 Medical Imaging in Parkinsonian Syndromes

Through the years, medical imaging techniques have been aiding physicians and researchers in the study and understanding of PS [1], [25], [26]. Nowadays, the scientific community use several techniques such as MRI including Diffusion Tensor Imaging (DTI) SPECT Positron-emission tomography (PET). This study only focuses on MRI and SPECT [81].

2.5.1 Magnetic Resonance Imaging

In the following lines, the physics behind MRI will be briefly explained. MRI is an anatomical image that represents the structure and the tissues, in this case, of the brain.

MRI is an imaging modality based on physical principles of the atomic nuclei [82]. These nuclei have magnetic properties that can be translated into in a signal that is captured by coils. This signal allows a formation of an image [83]. Usually, hydrogen is the most nuclei used in MRI because it is the most abundant element in the human body, being present in water and lipids [82], [83] The physics behind the formation of this signal relies on the magnetic moment and angular momentum of the nuclei, the hydrogen proton [82]. The magnetic moment gives the magnetic strength and the orientation of the moving electric charge when this charge is interacting with an external magnetic field. When no external influences are applied such as magnetic and electric fields, the proton spins about itself, i.e., rotate around itself [82]. The hydrogen proton is positively charged and has that spinning movement which allow the production of an electric current that creates magnetic movement. The angular moment is then created due to the mass of the rotating proton multiplied by its angular velocity. [82]

In MRI a constant magnetic field is applied to the body, in this case, the head, to gives a precession movement and a specific orientation to protons that were previous randomly orientated as illustrated in Figure 2.8 [82], [83]. These protons can be orientated parallel or antiparallel to the constant magnetic field applied. The parallel protons have lower energy while antiparallel protons have higher energy level.

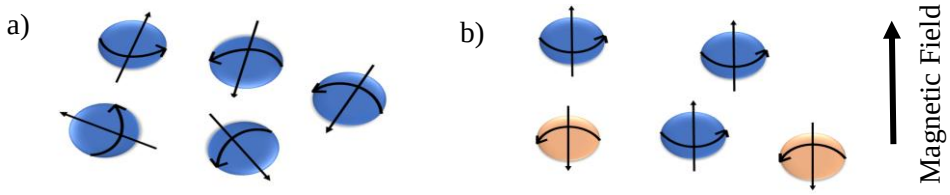


Figure 2.8 Proton orientation in MRI. a) Protons randomly oriented in free space b) Protons oriented according to the magnetic field. The protons which are aligned with the magnetic field are represent in blue.

The process movement is the same as the gyroscopic motion but in this case the proton process due to the constant magnetic field and its orientation, as shown in Figure 2.9.

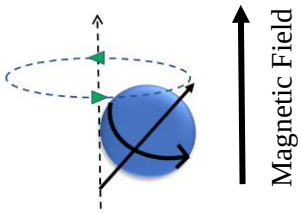


Figure 2.9 Precession Movement. The proton rotates around the direction of constant magnetic field.

The MRI techniques measure the magnetization of all spins in the free space. The magnetization can be represented by one vector with two components, the longitudinal and the transverse [82]. The longitudinal component has the same direction as the constant magnetic field while the transverse component is perpendicular and has the zero value. When all the protons are aligned with the magnetic field, parallel or antiparallel, one pulse sequence perpendicular to the magnetic field is applied to the system. The pulse sequence is characterized by radiofrequency (RF) pulses and gradient pulses, which control timings and other parameters that allow the acquisition of the image [82], [83]. This pulse changes the amplitude of the longitudinal and transversal magnetization, by changing the magnetization of the longitudinal axis to the transverse plane. During the pulse, all protons are in the transverse plane, rotating with the same phase [82].

After pulse application, the magnetization tends to return to the equilibrium position and occurs transverse relaxation and longitudinal relaxation. The transverse relaxation also known as spin-lattice relaxation is the decrease of net magnetization because of the loss of spins coherence of the protons [82]. In its turn, the longitudinal relaxation or spin-spin relaxation is the recuperation of the net magnetization along the horizontal axis because the protons tend to return to the equilibrium state, parallel to the magnetic field [82]. In MRI signal, these events are traduced by two relaxation times T2 and T1. In MRI signal, these events are traduced by two relaxation times T2 and T1. T2 is a time constant that represents the exponential decay of the net magnetization in the transverse wherein protons spin at different phases, as shown in Figure 2.10 [82].

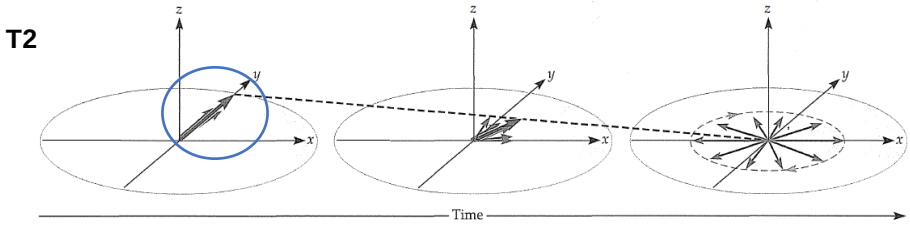


Figure 2.10 Illustration of the T2 decay. During the pulse sequence the spins are perfect aligned. Once the pulse sequence is turn off, protons lose their phase coherence and gain different orientations. Adapted from [85]

The recovery phase of the longitudinal magnetization is given by the time constant T1 which indicates the time that protons took to return to the equilibrium state before the excitation caused by the radiofrequency pulse, as illustrated in Figure 2.11 [82].

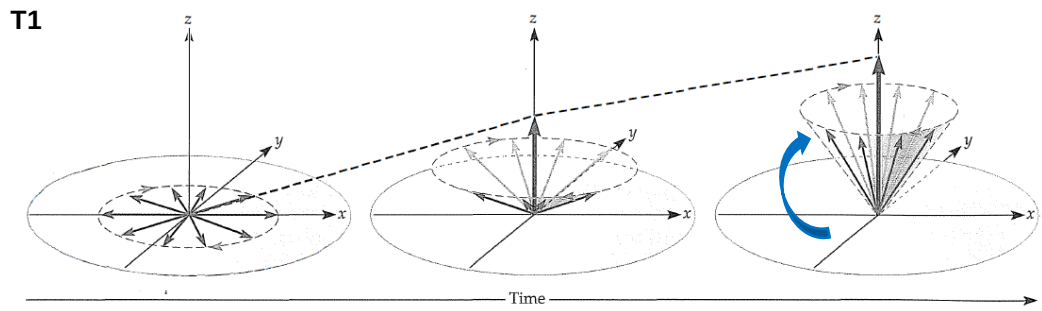


Figure 2.11 Illustration of T1 recovery. The spins, initially in the transverse plan, return to the initial position and orientation in the longitudinal axis, represented as z in the image. Adapted from [82]

During the image acquisition, the constant magnetic field is always turn on and may have slight differences in its intensity, which originate spatial inhomogeneities since the spins will be processing at different speeds. [82]. This event is translated by the time constant T2*. T2* reflects the quickly lost of coherence between the spins due to the spatial inhomogeneities [82].

T1 and T2 values depend on the proton density of the tissues. T1 is always bigger than T2. Tissues with a high number of hydrogen protons have long t1, e.g., fluids have longer T1 than T2 [83].

The contrast of the MRI images is related to T1 and T2 parameters and Proton density [82], [83]. Each constant time gives different structure contrast to each image. Thus, the images can be T1-Weighted (T1-W), T2-Weighted (T2-W) and Proton Density weighted. These images are produced using Spin-Echo (SE) or Gradient-Echo (GE) which are pulse sequences. These pulse sequences are related with two concepts, Repetition Time (TR) and Echo Time (TE), respectively. The contrast in MRI is possible due to variations in TR and TE [83]. The SE sequence comprises an echo with two RF pulses. In its turn, the GE is formed by one RF pulse followed by one gradient that creates the echo. In both sequences, the echo measures the intensity of the signal. The function of the gradient is used to accelerates spins dephasing [82].

T1-W images give information about the RT of each tissue and are characterized by an intermediate TR and a short TE. These images are brighter in voxels containing tissues with short T1 because the spins in these tissues are the first to recovery to longitudinal magnetization and darker in voxels comprising tissues with long T1. For instance, the white matter is very bright, the water-based tissues such as grey matter are mid-grey and the fluids are dark. Thus, T1-W images shows the boundaries between the tissues and fat-based tissues [83].

T2-W have long T1 and intermediate TE. In this cases, the grey matter and fluids get bright and the fat-based tissues get dark since the spins had totally recovered [82]. PD Weighted images allows to distinguish fluid from cartilage [83].

2.5.2 State-of-the-Art: T1-Weighted and T2-Weighted images in PS

MRI is usually used as routine to differentiate symptomatic parkinsonism from controls and to detect if there are any cerebrovascular damage causing the PS features [14].

On the one hand, according to the literature, anatomical changes in early-PD are very subtle in structural MRI [81], [84]. For instance, T1-W images present poor contrast in regions of interest areas such as STN and GP which difficult the detection of pathological anatomic findings [81], [84], [85].

On the other hand, the literature also describes MRI findings in structural alterations in PD brains. Due to the high quantity of iron present in these structures, the T1 value is short and consequently T2-W and T2*-Weighted images allow the detection of changes in structures like STN and GP[85]. Moreover, a study with 27 individuals with PD that filled Unified Parkinson's Disease Rating Scale and HY scale, pointed out for a reduction of 11% putamen in volume and 8% in caudate volume in PD when used a high-resolution 3T structural T1-W images when compared with control group. Nevertheless, this study also indicated that it is still unclear the relationship between the volume reduction and the dopamine deficit. [86]

Concerning SNc degeneration in PD, some studies report no evidence of volumetric changes in SNc of patients [87], [88]. Minate and colleagues found alterations in the SN of PD patients and in T1 value in comparison with the control group. However, the authors conclude that these differences in SN of PD patients were also influenced by the midbrain size variability for each subject [89].

According to the literature, increasing the magnetic field in MRI image acquisition may improve structural image quality. Kwon *et al.* use a 7T magnetic field to acquire T2*-weighted MRI images of the SN from 10 PD-patients and 10 control subjects. The authors report that the shape of the *substantia nigra* of PD patients differs from Control individuals. Moreover, they also identified an increase of the hypointensity of the MRI signal due to a rise in the volume of the SN in PD-Patients. The increase of the SN shape is related to high concentrations of iron in PD-patients, unlike Control subjects which are translated in variations in T2*-weighted MRI signal [90].

To aid in the detection of structural changes in MRI some techniques have been used, such as Voxel-Based Morphometry (VBM). VBM is an MRI technique that allows the analysis of focal brain anatomic differences between healthy and patients subjects. VBM uses image processing methods such as spatial normalization and segmentation into grey matter, white matter, and Cerebrospinal fluid[84], [91]. Using VBM, Burton *et al.* analyzed T1-W MRI scans from 26 PD-patients and 26 PD-patients with dementia with the aim to investigate brain volume loss. The

study reported that PD-patients presented a reduction of grey matter volume in the frontal lobe comparatively to the control group. Regarding PD-patients with dementia, a loss of volume in the grey matter of the temporal lobe was found [92]. Other studies with PD patients identified reduced grey matter in the right side of the hippocampus and in the left superior temporal gyri, loss of volume in the olfactory bulb and volume reduction in grey matter in brain areas related with the olfactory system such as inner olfactory cortex and amygdala [93], [94]. The authors conclude the volume changes in the olfactory system agrees with the stages hypothesis of Braak. [39], [93].

Moreover, one study that used volumetric analysis and segmentation applied to T1-W images of PD patients reported volume reduction of the bilateral putamen in comparison with the control group [95] and Tinaz et all. applied an automated reconstruction method to T1-W images of PD-patients and found a reduction volume of the striatum [96]. These findings are related with PD pathology and motor symptoms.[1] However, one study that also used T1-W images as well, reported mean striatal, cerebellar and brainstem volumes remain normal in comparison with control [97].

Even with these findings, MRI is mostly used to distinguishing atypical syndromes from Parkinson's disease due to limited and subtle changes in early-PD brain structures. Nevertheless, new approaches to MRI pointed that this medical imaging technique may help in the discrimination of nigral degeneration and its progression. [98]

In Atypical Parkinsonism, MRI has been used to detect changes in PSP, MSA and CBD [99]. According to the literature, the most affected areas in these syndromes are basal ganglia, brainstem, cortex and cerebellum. The atrophy and changes in these regions originate alterations in MRI signal relatively to healthy subjects and PD-patients. [100], [101] The Table 2.3 summarize the some of the findings achieved so far and present the in literature. [13], [99]

To sum up, Meijer et al. analysed several studies and proposed that MRI should be a tool to aid in PS diagnosis [14]. According to them, T1-Weighted may be used to detect brain atrophy and tissue loss but have some limitations, namely, poor contrast in some areas due to low sensitivity to signal intensities changes. Regarding T2-Weighted images, those are useful to detect changes in the basal ganglia which are detected by signal intensity changes in MRI [14].

.

Table 2.3 Findings in structural MRI in PS

AP	MRI findings	Reference
MSA	- No alteration in the Olfactory system	Chen *et al.* [93]
	- Atrophy in putamen (also known as putaminal rim sign)	Yekhlef *et al.* [102]
	- Hypointensities in T2-weighted images	
	- Atrophy in putamen of MSA subjects differ from PD-patients	Sako *et al.* [103]
PSP	- Volumetric reductions in brain-stem, midbrain and frontal grey matter	Gröschel *et al.* [104]
	- T2-weighted images hyperintensity in the midbrain	Massey *et al.* [105]
	- Midbrain atrophy, known as Hummingbird sign	
	- Midbrain tegmentum atrophy known as morning glory sign	
CBD	- Cortical atrophy, lateral ventricle dilatation and mindbrain atrophy	Yekhlef *et al.* [102]
	- Atrophy of parietal cortex and corpus callosum	Gröschel *et al.* [104]
PD	- Volume reduction in caudate	Pitcher *et al.* [86]
	- Volume reduction in putamen	Pitcher *et al.*. [86], Geng *et al.* [95]
	- Differences in *substantia nigra* shape	Kwon *et al.* [90]
	- Volume reduction in the grey matter of the frontal lobe	Burton *et al.* [92]
	- Reduced grey matter in Hippocampus, temporal lobe and olfactory bolb	Chen *et al.*[93], Summerfield *et al.* [94]
	- Loss of volume in areas of the olfactory system	

2.5.3 Single-Photon Emission Computed Tomography in Parkinsonian Syndromes

In the literature, some studies have shown that SPECT imaging can be useful to improve PS diagnosis criteria, especially when there is the hypothesis of PS diagnosis [15], [80]. In the following subchapter, it is described SPECT concepts and how related they are with Parkinsonian Syndromes.

SPECT is a functional nuclear imaging technique. Functional images allow the study and the understanding of physiological activities, such as metabolism and blood flow, that occur within organs and tissues [13], [106].

SPECT images are obtained using radiopharmaceuticals comprise by radioligands that are injected in the patient. Once injected, these substances spread in different internal tissues or organs and the radioligands bind to specific molecules presented in those tissues and organs. This binding originates a reaction that causes gamma rays emission, photons. These photons are

capture by radiation detectors, called collimators [106]. The imaged is formed by the acquisition of these gamma rays from different views to guarantee that the distribution of radiation in the three-dimensional perspective of the body is detected. Then, using imaging reconstructions techniques the imaged is formed [106].

In this work, SPECT images are obtained from the brain to detect dopaminergic changes. In this case, the radiopharmaceuticals usually used are cocaine analogues [84]. These nuclear medicine tracers are used for the analysis of a specific brain function or brain pathology, namely the dopaminergic function, in which this injected substance binds to dopamine transporters or dopamine D2 receptors [84]. The radioactivity dose given to patients is stipulated taking into account safety rules [106].

2.5.4 Dopamine transporter SPECT

One type of SPECT imaging is DaT SPECT imaging, usually used to differentiate PS and other motor diseases such as ET [13]. DaT is a transporter protein located in the presynaptic membrane of the striatum that mediates the reuptake of the free dopamine within synaptic gap [107]. DaT SPECT exam allows to find out dysfunctions in the dopamine transporter [13], [80]. One of the radiopharmaceuticals used in the DaT SPECT imaging for PS is an active substance called ioflupane (^{123}I). When injected in the body through the bloodstream, the ioflupane accumulates in the dopamine transporters of the striatum. In patients with parkinsonism, there are dopaminergic neurons degenerated and consequently less DaT. Thus, the concentration of the radioligand is less in the striatum in comparison with a healthy person. This reduction is visible on the scan [108]. Factors such as sex and age usually influence DaT density in healthy and pathologically subjects. For instance, aging is related with a reduction of DaT in the striatum of healthy subjects. However, this reduction is not so strong and evident as the loss in PD-patients [109]. Female subjects, healthy or PD-patients, differ slightly from male healthy or PD-patients by presenting a higher concentration of DaT [110], [111].

To obtain SPECT images it also used radiotracers like ^{123}I-iodobenzamide (123I-IBZM) and 123I-iodolisuride (^{123}I-IBF) for the detection of dopamine D2 receptor to obtain information about the density of postsynaptic neuronal loss [13].

2.5.5 PD patients with Scans Without Evidence of Dopaminergic Deficit, SWEDD

In the literature, drug trails and studies that had used DaT SPECT images to analyse PD-patients reported that a few early-PD patients presented normal DaTscan [112]–[114]. For instance, in a study with 62 patients with clinical features of parkinsonism and tremor disorders, it was acquired DaT SPECT images from all the individuals. Within the participants, 38 fulfilled UK Brain Bank step 1 PD criteria and 24 failed standard Parkinson's disease and essential tremor criteria. The

results obtained showed that 14 of 24 subjects that do not fill in specific criteria had normal scans and 5 of 28 patients that were classified as PD had a normal scan [113]. These PD-patients that presented a normal scan which indicates that there is no dopaminergic deficit are now called SWEDD patients [15].

Moreover, follow-up studies using DaT SPECT images show that approximately 10% of the cases clinically diagnosed as early PD had SWEDD [18] and that those patients tend to have a normal DaT SPECT along disease progression [18], [115].

However, some findings corroborate the fact that patients with SWEDD may have been misdiagnosis as PD-patients [116]. On the one hand, there are strong evidence that at early-PD stages, when the symptoms first appear, 80 % of striatal dopamine and 50 % of nigral dopamine cells are already lost. This reduction can be detect by a DaT SPECT exam which results in an abnormal scan. [43], [84], [117] On the other hand, recent studies found that SWEDDs patients had normal olfaction score unlike PD patients [17] and present different structural connectivity [118].

Clinically, SWEDD's patients can be divided into two categories: tremor dominant or non-tremor dominant. Within the causes of tremor dominant SWEDD type are dystonic tremors [15], [18], [119] and the causes behind non-tremor dominant SWEDD type are related with Vascular Parkinsonism or Huntington's disease [15]. Besides the evidence of the dopamine neurons without degeneration, SWEDD patients are also characterized by the fact that the glucose metabolic patterns are also normal [120]. Regarding the response of SWEDD patients to L-Dopa, in the literature this is dubious. [112]

2.5.6 Diagnosis of Parkinsonian Syndromes with DAT-SPECT

The literature suggests that DaT SPECT scans can be an asset to improve diagnostic accuracy as well as to aid in the identification of individuals with dopaminergic deficit before the onset symptoms [80], [121]–[123]. For instance, in PD, 60-70% of the dopamine neurons are degenerated and consequently DaT ligand uptake is substantially reduced even at early-stages comparatively to aged-matched healthy individuals. This loss in PD is substantially detected in DaT SPECT [80], [121]–[123]. Unfortunately, the sensitivity of the DaT SPECT is known to be less than 100%, which may be related with SWEDD cases [80].

Besides, in the comparison of PS, Dat SPECT is limited and does not aid to distinguish PD, MSA, PSP and CBD [80] because DaT SPECT scan only detects loss of DaT which is typical in all PS. Consequently, DaT SPECT imaging is not used to differentiate PS. For instance, all of PS have abnormal DaT SPECT with asymmetry in the striatum [124] in comparison with healthy subjects

but with different progression and pathology. However, even with those changes DaT SPECT is only able to differentiate PS and Control groups, as illustrated in the Figure 2.12. [125], [126]

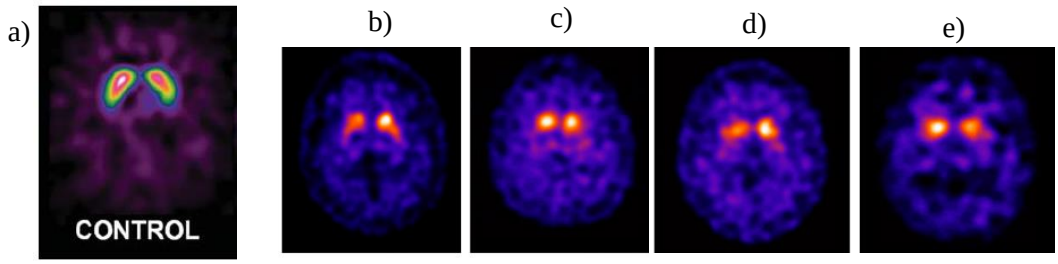


Figure 2.12 DaT SPECT images from Control subjects and Parkinson's Diseases. DaTscan SPECT of a) Control subjects, b) PD-patients, c) MSA- patients, d) PSP-patients and e) dementia with DLB. The differences between the PS and the healthy group are evidenced in the images but the changes among PD and AP may not be comparable because all present evident asymmetries in the striatum. Adapted from [80], [123]

Despite this difficulty of PS distinction, DaT SPECT is very useful to differentiate PS from other diseases that have similar symptoms like those manifested in those syndromes [124]. For instance, studies reported that the visually analysis of [123I]-FP-CIT SPECT by experts can allow the diagnosis of essential tremor and parkinsonism because ET patients present normal a scan which indicates that there is no pathological change in the dopaminergic system [127], [128].

In the Figure 2.13, it is enumerated the cases and the groups of subjects that are nowadays considered as having normal or abnormal DaT SPECT.

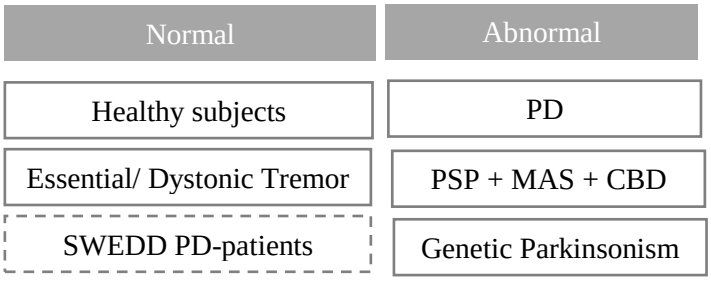


Figure 2.13 - Normal vs Abnormal DaT SPECT. According to the literature Essential tremor and Dystonic tremor present normal DaT scans since there is no degeneration in the basal ganglia. SWEDD PD-patients present a normal DaT SPECT but are early diagnosed as having PD. Parkinsonian syndromes are inserted into abnormal DaT SPECT since these patients have loss of dopamine.

2.6 Artificial Intelligence

Artificial intelligence is a concept that emerged after World War II, and it is used nowadays in diverse areas, namely, robotics, bioinformatics, speech and imaging processing, finance and chemistry [129].

Several authors have been trying to define Artificial Intelligence. Haugeland sees AI as the area that "*makes computers* think . . . machines with minds, in the full and literal sense." [130] and Rich and Knight define that concept as "The study of how to make computers do things at which,

at the moment, people are better"[131]. So far, all the definitions are inserted in two pillars: Thinking/Action and Humanly/Rationally, because a computer to be considered intelligent has to have human and/or animal characteristics [129]. Those pillars can be divided into four approaches: Thinking Humanity, Thinking Rationally, Acting Humanly and Acting Rationally [129]. Within of the topic Acting Humanly is Machine Learning (ML).

ML as Goodfellow *et al.* said "is the ability [of computers] to acquire their own knowledge, by extracting patterns from raw data", that is, computers with AI must have the ability to learn from data of any situation in the world with an automatic learning process and, thus, predict or make decisions based on what have learned [132]. Inside of Machine Learning field is Deep Learning. In this subchapter, Deep Learning is explained with especially focus on Convolutional Neural Network algorithm and its applications in the medical imaging field.

2.6.1 Deep Learning

Deep Learning is a subfield of ML that nowadays have been more used due to the improvement of hardware and software [133]. In deep learning, the process of learning is based on getting the characteristics of the object. For instance, considering an airplane image, the algorithm extracts the colour, the shape of the wings and other characteristics on the image to learn what defines an airplane. In AI, these characteristics are called features.

It is also important to mention that these images are constituted by a bunch of pixels that contain a lot of information and that deep learning algorithms are able to extract features at a high level, namely, abstract features. For instance, to detect a cat, an airplane, or a face of a human, these algorithms first extract small and simple concepts such corners or edges and then extract complex concepts, called abstract features, such as objects parts.

All these features are extracted by which is called hidden layers [132]. A usually deep learning model is constituted by the visible layer, which contains the input image and several hidden layers that capture features of the image, first the simple ones then the complex ones as exemplified in Figure 2.14. The information travels from one hidden layer to another, that is, the 2nd hidden layers extract features from the resultant pixels of the 1st hidden layer and so on. The hidden layers are layers that extract indirectly information about the input.

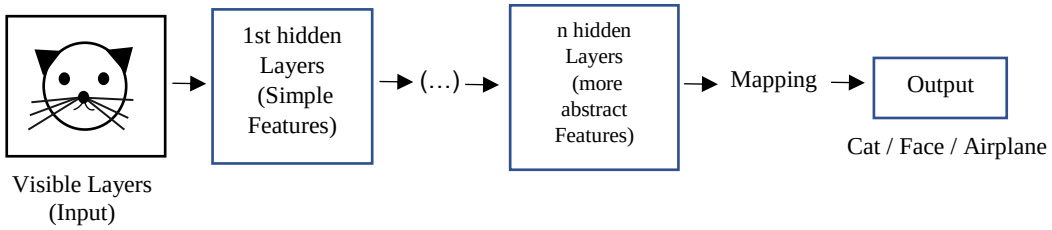


Figure 2.14 Schematic of a deep learning model. Adapted from [132]

In AI and consequently, in deep learning, there are two learning process:

- **Supervised learning**: in the learning process, also called training, the neural network receives the input with the respective label, i.e., the class of the input [132]. Thus, the network learns input features by knowing that they belong to a certain class.
- **Unsupervised learning:** in this case, the input of the network does not have any additional information such as the class that belongs to, that is, the learning process is based on the network ability to learn and recognized features from non-label data [132].

2.6.2 History Behind Convolutional Neural Networks

In this subchapter, it is briefly described the history of CNN and the progress present in literature that originates this algorithm.

Deep learning can be implemented using Artificial Neural Networks (ANN). ANN architectures are based on animal biological neuronal systems [134] and they are considered deep if it have more than one hidden layer. All the artificial neural networks are made of computational units called neurons, also known as nodes, that mimics neurons in the brain [135] and are connected with each other as biological neurons.

In 1943, McCulloch and Pitts proposed a mathematical approach for an artificial neuron based on the biological neurons [136]. This model mimics the synaptic strength by expressing it as the weighted sum of n input signals in the neuron minus a threshold. This modulation is represented in the Equation 2.1, where positive weights reflect excitatory stimulus and negative weights reflect inhibitory stimulus [137]. The θ is a non-linear function called activation function, w_i is weight of the i^{th} input where $i = 0,1,2,3 \dots, n$, x_i is input signal in the neuron and u is the threshold value, usually $w_0 = -u$, and $x_0 = 1$ [137].

$$y = \theta\left(\sum_{i=1}^{n} w_i x_i - u\right) \tag{2.1}$$

The Equation 2.1 can also be written as in Equation 2.2 [135].

$$y(x) = G(wx^T + b) \tag{2.2}$$

In Equation 2.2, G is the activation function, wx^T is the array of weights and the input array, where T indicates a matrixial operation called transpose and b is the threshold called bias.

The activation function acts like biological neurons, where when the input overcame a specificity threshold, the neuron fire the information across the axon [135]. In this case, the information flows to the next neuron and so on in a process called feedforward. A feedforward neural network, also known as multilayer perceptron, is a process that occurs from the first layer to the last layer without returning to the beginning point or to an intermediary stage [135].

The *b* threshold is unique for each neuron and it is responsible to define if the activation of the neuron is significant to pass forward in the network. Both weight and bias are trainable parameters, that is, the CNN learn and adjusts them in the feedforward process [135]. An example of a feedforward ANN architecture and their parameters is represented in the Figure 2.15.

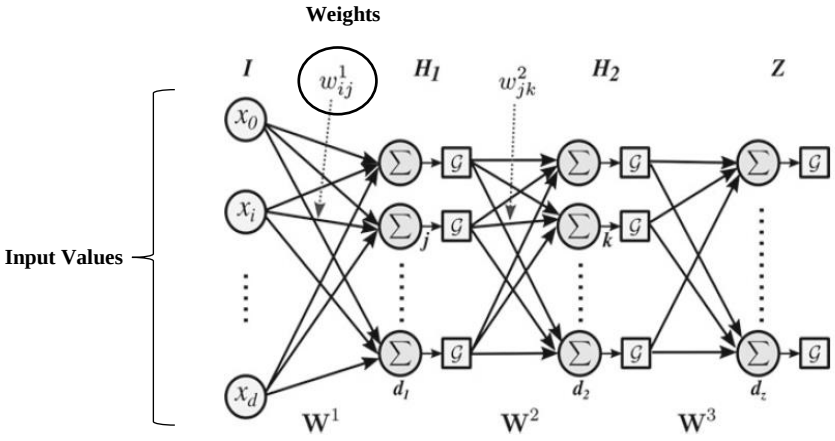


Figure 2.15 Feedforward neural network architecture. I is the input layer where x_i are the input values, H denoted the hidden layers and Z the last layer. W is the weighted array and w_{ij} is two-dimensional weighed and d_1, d_2 and d_z are neurons. Adapted from [134]

2.6.3 State-of-the-Art of CNN

Convolutional Neural Networks are one type of ANN usually used for pattern recognition [138], [139]. The basis of this network is based on David Hubel and Torsten Wiesel work [132]. These two neurophysiologists spend several years studying the visual cortex of cats with specially focus on the receptive field cells of the primary visual cortex located in the retina, the area responsible for first processing the visual input [132], [140].

The receptive fields cells are sensory neurons that detect light brightness, control the firing of visual neurons firing and are divided in simple or complex cells [140], [141]. Simple receptive fields cells have two types of subdivisions, the excitatory and the inhibitory and these regions are activated only if they receive the specific stimuli, light on or off. Through observation of cells mapping, Hubel and Wiesel realized that these cells were sensible to simple features of the input such as bars and edges and that and this cells were also able to detect position variations [132], [140], [142]. Unlike simple cells, complex cells are insensitive to edges, bars and position variations but are able to detect more complex features, such if cells have the same orientation [140].

These findings inspired Fukushima that in 1980 created the first neural network visual cortex-inspired. This network called Neocognitron was based on the cells of the cat visual cortex [143] previously mentioned and it was develop to have the ability to recognized patterns as a human

been learns using unsupervised learning. The Neocognitron are constituted by two layers: S-Layer and C-layer, that are applied one after the other. S-Layers behavior like filters to mimic simple cells proprieties of edges and bars detection. In its turn, C-layers receive as input S-layer values, and act like complex cells by merging values. This structure was elaborated to be able to recognize shapes of the input image based on their geometrical features at any position and without being sensitive to small distortions of that shapes. The Neocognitron which means that the network learn by itself by receiving as input repetitive stimulus patterns without any additional information such as the class that the input belongs to [143]. Fukushima realized that complex cells in the last layers started to respond only to a specific stimulus pattern and conclude that Neocognitron was able to recognize patterns correctly [143].

A few years later, other studies used an Artificial Convolutional Neural Network based on Neocognitron to apply to medical imaging for pattern recognition [139], [144]. The structure of this Artificial CNN is simpler than the Neocognitron and differ in the fact that it is not constituted by C-layers and instead of using unsurprised learning, the automatic machine learning method used was back-propagation algorithm, a supervised learning approach [139], [144].

In 1998, LeCun and colleagues proposed a CNN called LeNet that uses a Gradient-Based Learning Approach and the backpropagation algorithm [138]. Even with some similiters with the previous CNN algorithms, LeCun proposed an improved CNN structure organized in several layers based also in the cat visual cortex. This CNN is constituted by the layers that mimics simple cells, the convolutional layers, and by pooling layers that have the same proprieties as complex cells. The CNN was training and tested with Modified National Institute of Standards and Technology (MNIST) database of handwritten digits [138], [145]. LeCun et al. conclude that Gradient-Based Learning improve CNN performance and suggests that CNN is a useful tool to feature extraction [138].

2.6.4 Convolutional Neural Networks Architecture

Convolutional Neural Networks are a trainable feedforward neural network constituted by multiple stages called layers [138]. CNN are constituted by the input layer, hidden layers and the output layer and all of these layers are characterized by having three dimensions: width, height and depth [135]. Each hidden layer receives a two-dimension input from the previous layer and originates a two-dimensional output. These two-dimensional arrays are called feature maps and contain the features extracted [138].

Concerning CNN architecture, this network is organized in two parts: the feature extraction and classification [134]. This network is mainly based on two mathematical operations: cross-correlation and pooling, that occur in the feature extraction stage.

The feature extraction stage is constituted by a convolutional layer, an activation layer and a pooling layer.

The classification part is constituted by a Fully Connected (FC) layer. After FC layer, usually there is final layer with a loss function that originates a output traduced by single vector with probability scores [135].

Therefore, CNN has the following architecture: an input layer receives normalized images with the same size [138], convolutional layers, activation function layers, pooling layers, fully connected layer and loss function layer. The main layers are illustrated in the Figure 2.16.

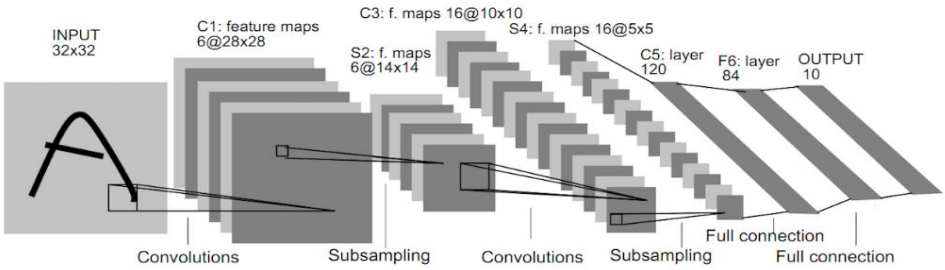


Figure 2.16 Convolutional Neural Network Architecture. This architecture was drawn by LeCun et al. [145] to classify handwritten digits of the MNIST database. Between the convolution layer and the subsampling layer also called pooling layer is usually an activation layer. After the fully connected layer is also a loss function layer.

The following subtopics are a deep description of each stage and of the components of CNN.

- **Convolutional layer CNN**

Firstly, it is important to notice that CNN has more than one convolutional layer to extract several features. Moreover, together with activations functions, it mimics simple cells of the visual cortex of the cat [132], [135] by acting like a detector or filter.

Convolution approach is a mathematical operation that englobes two functions, in which one affect the other, to create a third that represents the modifications that that the second function caused in the first one [132], [146].

$$y(t) = \int_{-\infty}^{+\infty} f(a)g(t-a)da \tag{2.3}$$

The *g* function in the Equation 2.3 represent a weight that it is applied to function *f*. This operation is usually represented with and asterisk, Equation 2.4.

$$y(t) = (f * g)(t) \tag{2.4}$$

In convolutional neural networks the f function can be represented as I and represents the input, an initial image or sets of arrays with visual information. The g function is a kernel function

usually represented as K. The kernel is also an array with parameters that are changed by the learning algorithm to improve CNN. [132] Both I and K functions are multi-dimensional arrays, usually two-dimensional, i.e. matrix, and can be expressed as in Equation 2.5 [132].

$$S(i,j) = (I * K)(i,j) = \sum_{\mathrm{m}=0}^{P-1} \sum_{n=0}^{Q-1} I(i - \mathrm{m}, \mathrm{j} - \mathrm{n}) K(\mathrm{m}, \mathrm{n}) \tag{2.5}$$

In the eq.2.5, the convolution is represented by summation over a finite number of the input dimensions, $S(i,j)$ represents the feature map wherein i are the rows and j are the columns of the image [132]. The input image is a two-dimensional array of size $W \times H$ where W is width and H is height which values are usually pixels. The kernel size is $P \times Q$ [135]. Each feature map is an image with size $W - P + 1 \times H - Q + 1$. For instance, an input image of size 8×8 and a kernel size of 2×2 originates an output image with 7×7. In CNN, it is common to use the Equation 2.5 with some slight differences to guarantee that the kernel function is not inverted when applied in the network [135]. Thus, the equation usually applied, also known as cross-correlation, is expressed in Equation 2.6 [132].

$$S(i,j) = (I * K)(i,j) = \sum_{\mathrm{m}=0}^{P-1} \sum_{n=0}^{Q-1} I(i + \mathrm{m}, \mathrm{j} + \mathrm{n}) K(\mathrm{m}, \mathrm{n}) \tag{2.6}$$

The kernel function is used to extract features from the input image. The kernel is also known as filter because its acts like a learning filter where only specific characteristics of the image are extracted to proceed to the next layer.[135] Each kernel has weights which are values that indicate the importance that the algorithm needs to give to a specific feature. [135]

The cross-correlation function is applied in the convolutional layer, and basically, the kernel convolves across the input image along width and height. The kernel is constituted by local receptive fields that are activated when a feature is detected [135] and allow neurons to extract all visual features such as corners or edges at any spatial position of the input image or from a feature map if the convolution layer is after a pooling layer.

After that, CNN learns which weight must give to each feature to allow the activation of the receptive fields [132], that is, the weights of each units in kernel are constantly being adjusted in the training process to improve feature detection [135]. Besides, this network it is also able to detected these features in any position [138].

Then, all these features pass again through other layers to extract high-order features. The convolution output contains several features maps in which each map contains information about a specific feature. Thus, the convolutional layer output has width, height and depth [132], [135].

- **Activation function Layer – Rectified Linear Unit**

The activation layer is usually used immediately after the convolutional layer. This layer contains a non-linear function applied element-wise to determine if the values of the convolutional layer output traduces important features. Basically, this function defines a threshold that indicates which neurons are activated [132], [135]. The advantage of the application of a non-linear function it to modulate the linear output as a non-linear function [147]. Thus, CNN can learn and modulate any input-output relationship.

The most common activation function applied to neural networks is Rectified Linear Unit (ReLU) [135]. This function is expressed in Equation 2.7 and graphically in Figure 2.17 [132], [135].

$$G_{reLU}(x) = \max\{0, \mathrm{x}\} \tag{2.7}$$

When applied to CNN, Equation 2.2 previously mentioned is given as follows, Equation 2.8:

$$y(x) = G(wx^T + b) = \max\{0, wx^T + b\} \tag{2.8}$$

In Equation 2.8, $wx^T + b$ is the output from linear operations such as cross-correlation and pooling and, in this case, the input of the activation layer [135].

Figure 2.17 illustrates that the activation function sets all the negative values in the feature map to zero and all the positive values remain unaltered. In CNN, the negative values that were set to zero will not pass forward in the network, that is, that neurons will not fire. Concerning the positive values, those will pass through to the following layer because they traduce a meaningful feature [132].

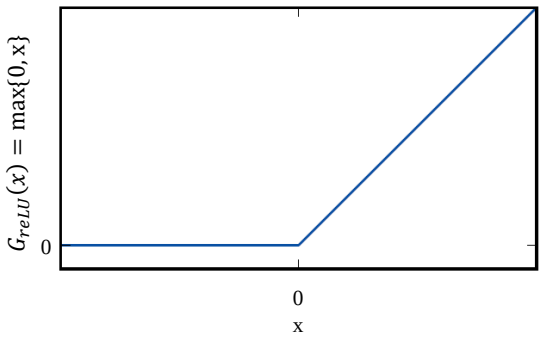


Figure 2.17 Graphic representation of Rectified Linear Unit. Negative values x, usually the weights in neural networks, are set to zero. Positive values have different values according to their weight. Adapted from: [132]

The main advantage of the activation function such as ReLU is the fact that accelerates the process of training and makes it more "effective" as Habibi et al suggest [133]. ReLU has a disadvantage called dead neurons, which are neurons that have a negative weight during training and that always conducted to zero [135].

Besides ReLU there are other activation function, such as Sigmoid Function and Hyperbolic Tangent Function Tanh, that are represented in Figure 2.18 [135]. However, the literature suggests that ReLU is the function that conducts to better results [147], [148].

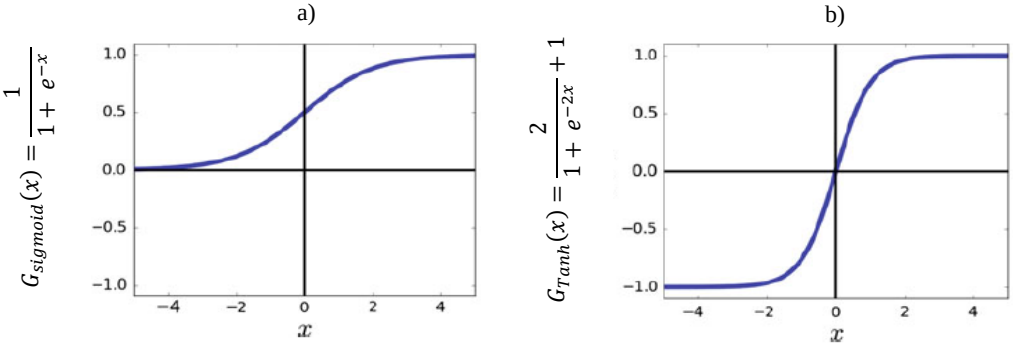


Figure 2.18 Activation Functions. a) Sigmoid Function and b) Hyperbolic Tangent Function.

- **Pooling Layers**

The main objective of adding a pooling layer, also called down sampling, to a neural network is to reduce spatial resolution of the output of the previous layer, that is, reduce feature maps size [135], [138] only along their with and height. This layer is important to prevent CNN to detect features by its position in the receptive field and to reduce their sensibility to ships and distortions in the features and consequently reduces overfitting [138]. Besides, once the output is simplified and the CNN has to learn less parameters, this layer also reduces computational costs. [133]

The pooling layer contains a pooling function that changes the output of the previous layer based on the information in the nearby outputs [132]. For instance, the pooling operation changes locally the feature map which is the output usually from the activation layer.

In CNN, the most common pooling functions are max pooling, average and weighted average [132] and these are functions convolves along the image as a filter. The max pooling operation along the input selects a rectangular neighbourhood and within these pixels select the maximum value. The output is the maximum values of the rectangular areas considered. In the average operation the output value is the average of the rectangular neighbourhood pixels and in the weighted average the output is calculated based on distance of the other pixels to the central one [132]. According to Sherer et al. the max-pooling presents better results than average pooling [135]. In this layer, it is also common to use the stride operation, which it will be explained further, to avoid possible overlapping [135].

- **Sparse Connectivity, Shared weights, Zero-padding, Stride in Feature Extraction**

According to LeCun et al, the convolution operation has two important ideas: sparse connectivity and shared weights [135] and pooling operation has shared weights. Zero-padding and Stride are common both to convolutional and pooling layers.

Sparse connectivity: In feature extraction stage, all the layers are organized in the following manner: a small portion of one-layer is connected with a single neuron of the next layer. Sparse connectivity refers to that connection between layers and it also is related with the fact that the kernel size is smaller than the input image which has several pixels [132]. In fact, input image size makes almost impossible to connect all the neurons to each pixel, because it would be necessary millions of neurons and computational power. Thus, CNN by having sparse connectivity, the local receptive fields of the kernel pass through the input image. Then, for each small area of the input with the size of the kernel, the output of that area with the kernel is only connected with one neuron, as illustrated in Figure 2.19. Even considering that this originates a size reduction, it does not affect feature extraction. Besides, even small features can be detected. Sparse connectivity does not only increases the statistical efficiency but also reduce the parameters needed for learning. [132], [138]

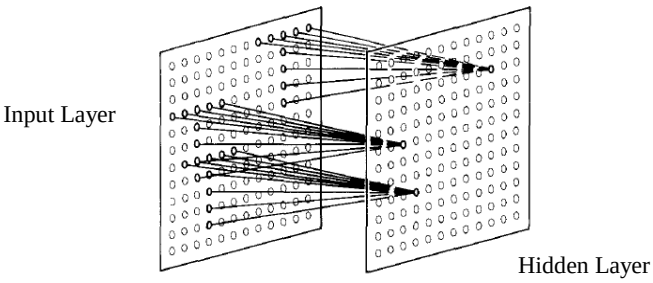


Figure 2.19 Connection of input layer or one hidden layer with the following layer, in CNN. In feature extraction state, each small parts of a layer are connected to only one unit of the following layer. Adapted from: [142]

Shared weights: The shared weights also known as parameter sharing emerged to reduce parameters in the CNN and are used in convolutional layers [132], [138]. This means that the receptive field used in convolutional layers are always the same for all the image, that is, it does not change along the image. Thus, CNN by using the same parameters is able to detect several features with just one kernel which values and regardless their position [132].

Zero-padding: Zero-padding refers to the addiction of pixels with zero value to the border of the resultant feature map. During feature extraction, the size of the features maps decreases due to the application of the cross-correlation and pooling operations [132], [135]. Thus, zero-padding guarantees that spatial dimensions of the resultant feature map remain the same or with a specific dimension [132], [135]. This process is illustrated in Figure 2.20.

Stride: The stride function is used in situations wherein it is wished to skip some pixels. This technique reduces substantially computational costs and possible loss of features information is not significant [132]. An example of stride application in a feature map it is showed in Figure 2.21.

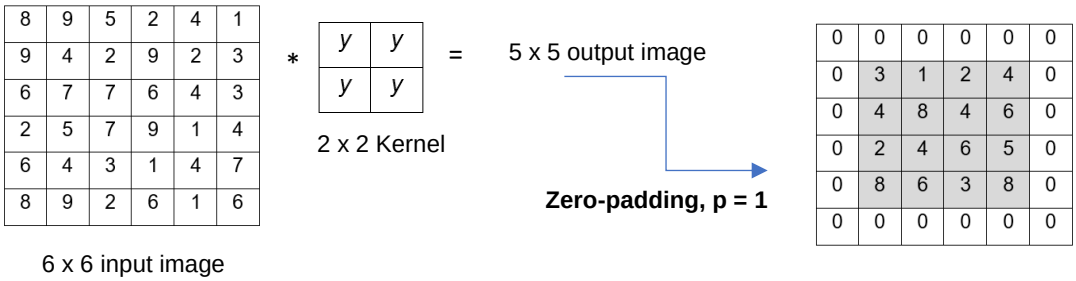


Figure 2.20 Zero-padding. This figure represents an example of a 2x2 kernel applied to a 6x6 image. The output is a 5x5 matrix. The application of the zero-padding equals to 1, in this case, makes the dimensions remain the same.

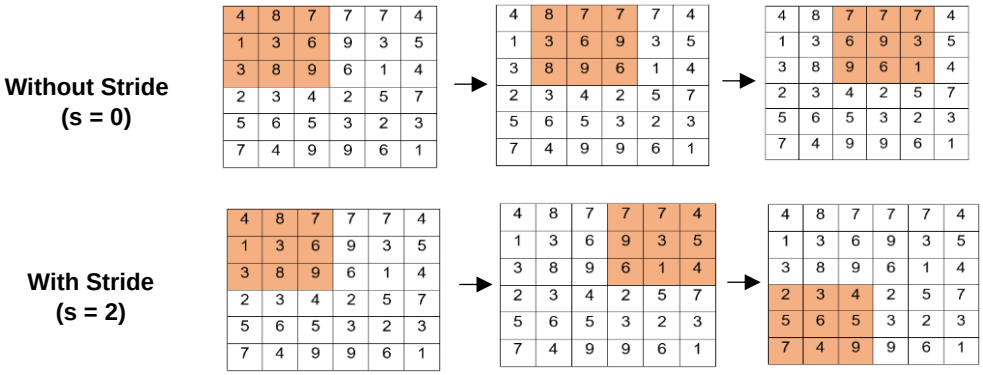


Figure 2.21 Example of application of a kernel with and without stride. Without stride there is no pixel skipping. With stride equals 2, the kernel skips two pixels at once. The s number represents the stride value.

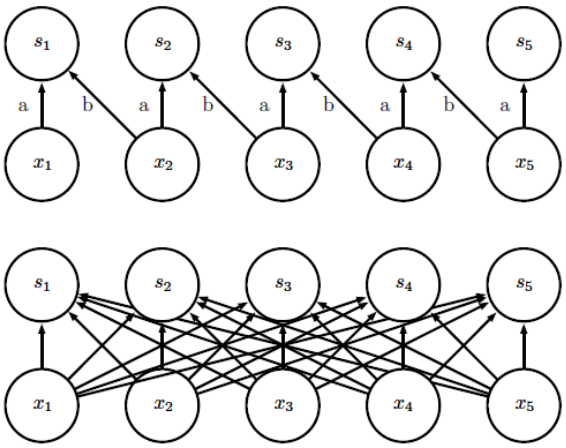


Figure 2.22 Layers at feature extraction vs fully connected layer. At the top, it is represented layers with sparse connectivity wherein *a* and *b* are the weights, x_i are the input and s_i the output. Each neuron couple are connected with only one neuron of the *s* layer. At the bottom, it is represented a FC layer wherein each neuron x_i is connected with all neurons of the following layer.

- **Fully Connected Layer**

The fully connected layer is usually connected with the last pooling layer of the network or with the last layer of the feature extraction part. Unlike convolutional and pooling layers, in FC layer, each unit is connected with each neuron of the previous layer [135] as shown in Figure 2.22. For

each input value of FC layer, it is multiplied a weight and it is added a bias [133] which traduces on the relation of high-features and classes. In CNN, this layer originates an array with linear unnormalized log probability scores as output with *n* classes, wherein *n* is the number of classes that the input image may belong too [135].

- **Loss layer**

The loss layer is the last layer of the CNN and consists in the application of a loss function to the output obtained in the previous layer, usually a FC layer [132], [135]. This function also known as cost function is used to measure the difference between the prediction and the correct value, that is the loss [132].

In deep learning, the most common loss functions are Softmax and Multiclass Support Vector Machine (SVM) [149].

Softmax function:

Softmax function also known as multinomial logistic regression basically computes scores from the FC layer into a meaningful probabilities scores and calculates the loss based on those scores [132], [149]. These values obtained in the FC layer can be represented as a vector *k* and, basically, this vector pass through the softmax layer and originates an output which is a normalized probability distribution (between 0 and 1) [132]. Each predictive value in *k* can be expressed as z, wherein $z = wx_i^T + b$ .The softmax function is expressed in Equation 2.10 [132], [149] wherein x_i is the input image, z_k is score obtained and z_j are the scores along *k* array in the j^{th} position. $\sum_j \exp(z_j)$ is responsible for normalize the scores. Thus, the output of a CNN with a softmax function is meaningful normalize probability distribution [132].

$$Softmax\ (x)_i = P(Y = k|X = x_i) = \frac{e^{z_k}}{\sum_j e^{z_j}} \tag{2.10}$$

Regarding the loss value, this is calculated through the log of the probability of the score obtained in the FC layer matching to the correct class, as expressed in Equation 2.11 [149] wherein y_i is the real class of the input image and it is an integer and z_{yi} is the score that the FC layer gives to the correct class.

$$L_i = -\log(P(Y = y_i|X = x_i)) = \frac{e^{z_{yi}}}{\sum_j e^{z_j}} \tag{2.11}$$

Multiclass SVM loss function:

Multiclass SVM loss function also known as Hinge loss has a different approach to loss calculation. Unlike Softmax, this function does not originate an output with meaningful scores. Instead, it quantifies how bad was the classifier performance based on loss calculations [149]. This function is usually applied when there are multiple classes.

This loss function analyses if the score of the correct class is higher than the score of the incorrect classes by a margin Δ, a previous stipulated constant. Basically, this function subtracts the scores of the incorrect classes with the score of the correct layer. Then, the sum of that difference between scores is made and the sum result traduces the loss [149]. The Multiclass SVM loss function is expressed in the Equation 2.12, wherein z_{yi} corresponds to the score of the correct class, z_j the scores of the incorrect classes and *j* the j^{th} position in the *k* array and each *j* correspond to one class.

$$Loss = L_i = \sum_{j \neq y_i} \max(0,\ z_j - z_{yi} + \Delta) \tag{2.12}$$

This function can be represented graphically as illustrated in Figure 2.23 a). In Figure 2.23 b) is illustrated an example of $Loss = 0$ because $z_{yi} \geq z_j + \Delta$.

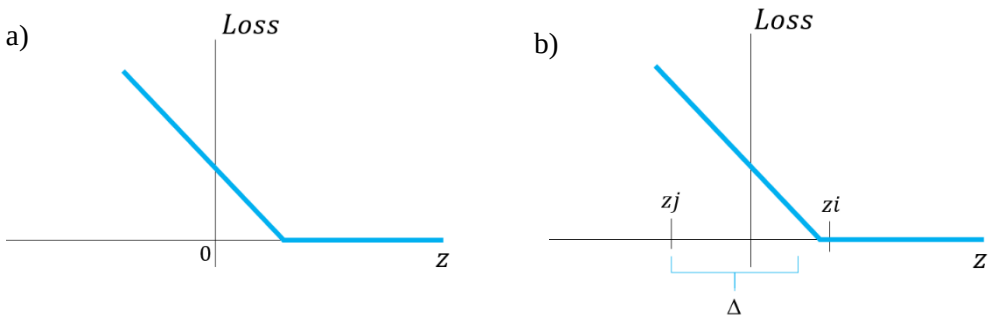


Figure 2.23 Hinge Loss function. a) Hinge Loss b) Hinge Loss wherein the score of correct class (z_i) is bigger than the incorrect class (z_j).

The total and final loss is given by the sum of the losses in the network, as expressed in Equation 2.13, wherein N is the number of images in the dataset [132], [149]. This Equation is valid for any loss function.

$$Loss = \frac{1}{N}\sum_{i} L_i(z, y_i) \tag{2.13}$$

- **CNN Regularization**

In deep learning, during the learning process the network creates a complex function to modulate and describe the outputs. However, sometimes the model fits too well the training data but does not fit in the test data and tends to occur underfitting or overfitting [132].

Underfitting emerge when the low error values is not good enough, and the network cannot obtain good results in the train test and consequently cannot replicate results to new data. Overfitting occurs when the CNN model fits to well in the data and the error between the training error and the test is to big [132].

To prevent the network model to overfit is common to add a term in the loss function to allow the algorithm to choose the best generalized function that is able to modulate unseen data, the test set.

This term penalizes the weights of the neurons by making them small and near zero or by setting them to zero which are associated with simple functions [132], [149]. This process is called regularization and prevent the network to create overly complex functions. The regularization term is given by $\lambda R(W)$ and thus, the final loss function is given by the Equation 2.14, wherein R(W) is the regularization function, λ a hypermeter called regularization strength and W the weight vector. The increase of the λ parameter leads to small weight [132].

$$L\ (W) = \frac{1}{N}\sum_{i=1}^{N} L_i(z, y_i) + \ \lambda R(W) \tag{2.14}$$

The most common $R(W)$ functions used are L_1 regularization, L2 regularization and dropout.

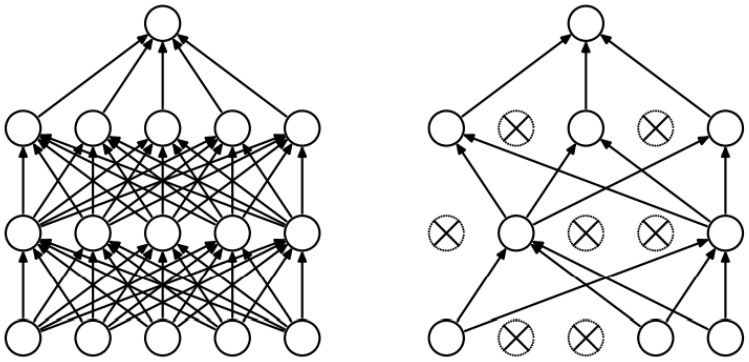
L_1 regularization: This regularization, also known as weighted decay, set the weights to smaller values, and it is expressed as the squared norm of the sum of all the weights which leads to small weights, Equation 2.15 [132], [149].

$$L_1: \quad R\ (W) = \sum_{k}\sum_{l} W_{k,l}^2 \tag{2.15}$$

L_2 regularization: This regularization, set some weights to zero, and it is expressed as the norm of the sum of all the weights, Equation 2.16 [132], [149].

$$L_2: \quad R\ (W) = \sum_{k}\sum_{l} |W_{k,l}| \tag{2.16}$$

Dropout: This regularization is usually placed after FC. Dropout consists in temporally remove random neurons during training process in forward pass to increase the ability of the network to learn the weights more precisely [132], [150]. An example is illustrated in Figure 2.24. Thus, the network act like if it was training several networks because it is training several subnetworks with different neurons removed. Even though it increases the time of training this does not lead to computational costs [150].



a) **b)**

Figure 2.24 Dropout layer. a) CNN structure without dropout b) CNN structure with a dropout layer. Adapted from: [150]

- **CNN Training**

In AI, the intent of training is to find the weights and biases that minimizes the loss function, that is, which values minimizes the discrepancy between the real class and the output [132]. One of the most successful approaches to train a neural network is the Gradient-Based Learning [145] because those parameters are updated very faster. In this case, the loss function minimization is achieved by the calculation of its gradient with respect of weights [138].

In CNN, it is usually used the Stochastic Gradient Descent (SGD) to adjust the weights. SGD is obtained using the backpropagation algorithm. Basically, this algorithm computes the negative gradients that faster minimize the loss function and send them backwards to beginning of the network to adjust the weights. Then, these update weights are propagated forward in the network and the process repeats until the error/loss is minimized.

The SGD is a linear combination of the previous weight update, V_t, and the negative gradient $\nabla L(W_t)$. The weight update is expressed in Equation 2.17 and Equation 2.18, wherein t is the iteration number, μ momentum, α is the learning rate, $V_{t+1}$the update value, $W_{t+1}$the update weight and W_t the previous rate [151].

$$V_{t+1} = \mu V_t + \alpha \nabla L(W_t) \tag{2.17}$$

$$W_{t+1} = W_t + V_{t+1} \tag{2.18}$$

The parameters μ and α are hyperparameters, where the momentum, μ, is the weight that indicates how much the previous weight influence the weight of the following layer, and the learning rate, α, is the weight of the negative gradient [151]

It is important to notice that the initialization of the weight, bias and hyperparameters is made according to specific methods that will not be explored in this work.

2.6.5 State-of-the-art: CNN applications to Medical field

CNN have many applications in several fields, such as, traffic signs identification [135], video, speaking and image recognition [152]. The usage of this AI algorithm has been increasing through the years due to the improvement of the computer power and because it has been proved that CNN are useful in those areas [132], [135], especially in the medical field [24], [25], [153], [154].

For instance, in 1995, a study used CNN for microcalcifications detection in digital mammograms and lung nodules on radiographs [139]. Later, in 1996, Sahiner et al. performed abnormal mass and normal breast tissue classification with recourse to CNN [144].

In 2017, Zou et al. used a 3D-CNN to classify Attention Deficit Hyperactivity Disorder with structural and functional MRI and conclude that a multi-modality approach is useful to find neuroimaging biomarkers in brain related diseases [155].

Regarding neurodegenerative diseases, Pereira et al. used CNN to classify electric signals of PD patients and healthy subjects [156]. These electric signals translate the tremors of the hand when subjects are drawing specific figures such as circles and spirals. Then, the electronic signal was converted into a time series-based image and tested with several CNN types [156].

Liu et al. proposed a 3D cascaded CNN to study Alzheimer's Disease (AD) in multi-modality approaches [24]. The study was conducted with MRI and PET images of 93 AD-patients, 204 mild cognitive impairment (MCI)patients and 100 control subjects. The MRI and PET images where divided into 3D local patches. Then to each patch is was assigned a 3D CNN to extract the high-level features. After that, all the high-level features were compacted, and then they were the input of a 2D CNN. The 3D CNN was constituted by convolutional and pooling layers with three dimensional kernels and 2D CNN was formed by two consecutive convolutional layers, two fully connected layers and one softmax layer. The 3D CNN and 2D CNN architectures are illustrated in Figure 2.25.

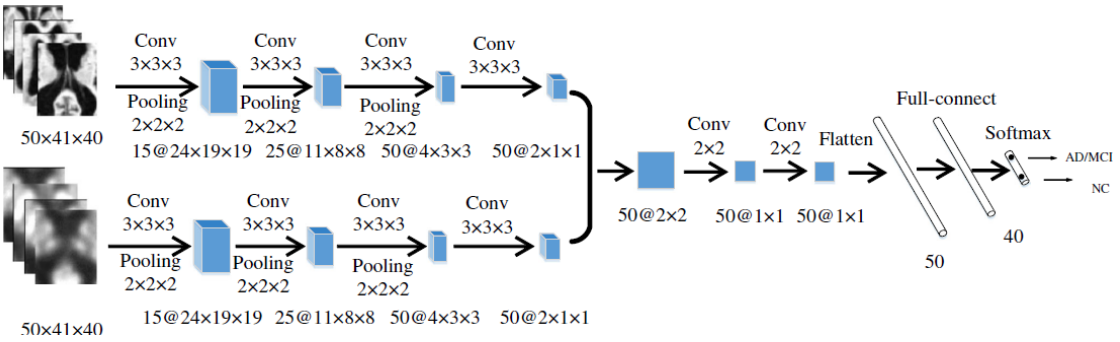


Figure 2.25 Multi-modality cascaded CNN architecture. This CNN was drawn by Lio et al. [24]

The authors conclude for the approach proposed there is no need for segmentation and rigid registration steps in the image pre-processing steps and that comparatively with other classifications models, this new method present higher accuracy. The results obtained in this study was 93.26% accuracy for AD vs Control classification and 82.95% accuracy for MCI vs controls [24].

CNN were also applied to Multiple Sclerosis identification. A study conducted by Zhang et al., obtained 98.23% accuracy, 98.22% sensitivity and 98.24% specificity [154]. In this study participated 38 patients and 26 healthy controls. Each image was divided into slices and those which contained the regions associated to Multiple Sclerosis pathology were collected to the classification. To increase the data training, the authors performed data augmentation by applying

the following methods: rotation, gamma correction, noise injection, random translation and scaling. The CNN was constituted by 7 convolutional layers, 3 fully connected layers, dropout layers and the parametric ReLU layer. The authors conclude that the introduction of dropout and the parametric ReLU increases CNN accuracy [154].

CNN have also been used for medical imaging segmentation. For instance, Karimi et al. study showed that a CNN-based segmentation framework was able to predict the coordinates of the prostate surface based on prostate MRI images [157] and Chen et al. used CNN to brain tumour MRI image segmentation [158].

Moreover, CNN was also applied to the classification of DaTScan SPECT and MRI images. The following subchapter describes studies that used CNN to classify PS as well as other machine learning techniques used for the same propose.

2.7 State-of-the-art: Classification methods for PS

Through the years, several approaches have been used to attempt to classify PS. In 2017, Choi et al. designed a 3D-CNN to classify DaTScan images of early PD-patients, SWEDD patients, control (healthy) subjects and patients with non-PS [25]. Besides classification with CNN, the authors also analysed the performance of professionals. The images were extracted from the PPMI database and from the Seoul University Hospital. The 3D CNN, PD Net, is illustrated in the Figure 2.26. This 3D-CNN is constituted by 3D-convolutional layers interleaved with pooling layers with the max pooling function, followed by a ReLU function, a FC layer and the softmax layer. The input images were normalized by the range 0 to 255 and zero-padding was also applied. The training process was performed with stochastic gradient descent. The training and validation sets were executed using images from PD and control subjects. The classification of SWEDD subjects were performed using the pre-trained PD Net with PD and Control images to obtain a result of “normal” or “abnormal” image. Then, the prediction was compared with the diagnosis of the PPMI database. Concerning PD patients, in the test set it was obtained 96.0% accuracy, 100% specificity and the 94.2% Sensitivity. The authors also verified that these results are higher than those achieved by the visual and conventional quantification [25].

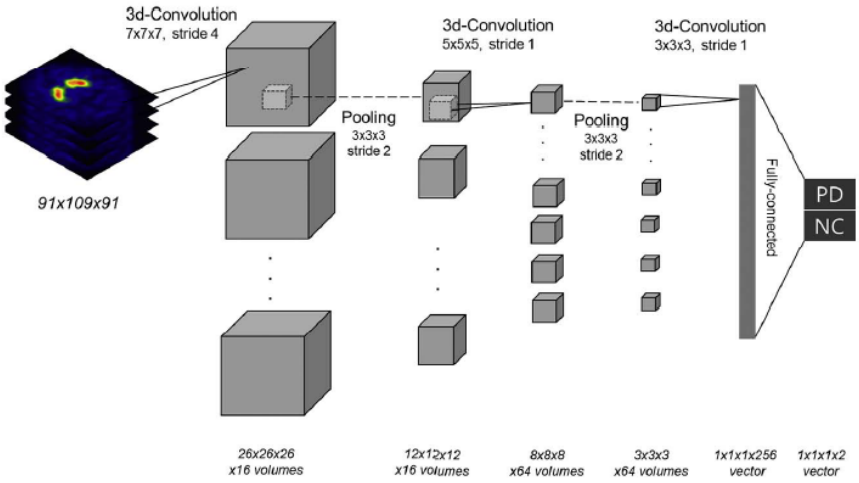


Figure 2.26 Deep convolutional neural network for DaTscan SPECT images classification designed by Choi et al. [25].

Another study conducted by Martinez-Murcia also used CNN to distinguish DaTscan SPECT images of PD-patients, SWEDD patients and Healthy subjects. Firstly, using a traditional CNN, similar to the one designed by LeCun et al., it was classified PD vs Control which result in 95.5% accuracy, 96.1% sensitivity and 94.5% specificity. Then, it was classified PD vs Control vs SWEDD which leads to 82.0% accuracy [153].

Moreover, Esmaeilzadeh et al. developed a deep learning framework for PD classification with 3D-CNN using MRI images. The architecture of the network is constituted by 3 blocks of 2 convolutional layer and 1 pooling layer wherein the activation function for each layer is the Leaky-ReLU, two fully connected layers and one softmax layer. Besides MRI images, the authors also tested an approach of introducing two features in the last FC layer: age and sex. The introduction of these two variables increases the accuracy of the classifier. Furthermore, to a better understanding and visualization the output of the CNN model they create a Parkinson Heat-Map of Brain of the axial, sagittal and coronal views. These maps are images with a colour scale that translates the most significant and critical areas in the image that the CNN considered as an important feature as illustrated in Figure 2.27 [22].. In this case, it coincides with the pathological regions in PD: *Substantia Nigra*, Basal Ganglia and the Superior Parietal [22].

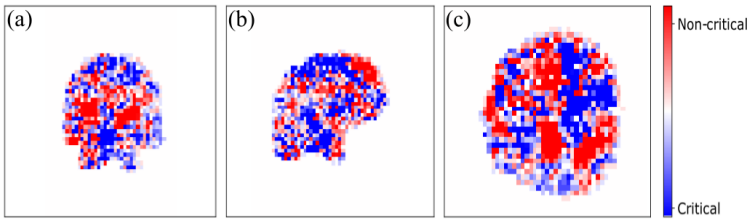


Figure 2.27 Parkinson Heat-Map of Brain in the Esmaeilzadeh et al study. a) Coronal view b) Axial view c) Sagittal view. At blue colors indicate the critical regions for PD diagnosis. Adapted from: [24]

Besides CNN, other approaches have been used to analyze and/or classify SPECT images. Some studies used a fully automated Objective Striatal Analysis (OSA) to calculate the striatal binding ratio [159]–[161]. The striatal binding ratio is a measure used by specialists to analyse the SPECT images, since this ratio is different in PD-patients in comparison with healthy subjects. This ratio is given by the ratio of striatal region and the occipital minus 1 [162].

Regarding OSA, in a succinctly explanation, in this technique the DaT SPECT images are re-oriented with a standard average template and the algorithm OSA automatically detects the striatum and the occipital structures and calculate the binding ratio of these structures [159]. Once this ratio is lower in PS patients [116], this algorithm is able to detect that difference in comparison with healthy subjects and that is why it is considered an useful tool [160]. A study that made usage of this technique achieved similar ratios to those manually obtained by a specialist [159]. Interestingly, OSA allows the detection of the laterally asymmetry in the dopamine of the striatum [161] which causes unilateral symptoms [163].

Besides OSA, Machine Learning methods such as SVM were also used for PS diagnosis [164]–[166]. SVM is a supervised learning algorithm used for pattern recognition [167].

Briefly, SVM is an algorithm that transforms the input vectors into a feature map space that is filled with points originated from the training process. Then this feature map has a hyperplane, the optimal mathematical model, that works like a decision boundary to separate training data points from each class. The best hyperplane is the one that optimizes the algorithm, that is, the hyperplane that it is maximally distant from both classes and that best divide them [167]. In complex cases, wherein, the training data points from the both classes are mixed and the linear hyperplane is not able to divide the points of each class, it is usually used kernel techniques [168].

Prashanth et al. used SVM to classify SPECT images of early PD-patients and controls and performed a logistic regression to estimate the risk of PD. The SVM classification was performed using the striatal binding ratio as a feature, and it was obtained an 96% accuracy using a kernel and 92% accuracy when used linear SVM [164]. Other studies using SVM and SPECT images classified early-PD and controls based on the side of the brain that was affected [166].

In 2015, Oliveira et al. obtained a high accuracy result with the SVM using a voxels-as-features approach [21]. The classifier achieved 97.86% accuracy using specific voxels of the volume of interest as features for the SVM. These voxels were from the cortex and occipital area [21]. Other study elaborated by Tagare *et al.* that also used a voxel-based logistic analysis reveal that voxels in the caudate, the putamen and globus pallidus are useful for PD and Control classification. Moreover, this approach could distinguish SWEDD-patients from PD-patients, but not SWEDD-patients from control group. They also develop a new method called logistic component that

revealed interesting findings related with age and sex of the subjects. In fact, this study showed that exists a normal increase of dopaminergic degeneration due to aging but that it is smaller and less significant in comparison with PD-Patients. Moreover, the authors also showed that there is no evidence of difference in the dopamine degeneration between female and male PD-patients [20].

Regarding MRI images studies that also used other algorithms than CNN, Singh et al. also used two combined machine learning algorithms to classify T1-weighted images of PD, SWEDD and control subjects of the PPMI database. Those methods were an unsupervised Self-Organized Map (SOM) used for feature extraction and a supervised learning based Least Squares Support Vector Machine to classification (LSSVM). The (SOM) algorithm was able to detect the biomarkers that differ SWEDD from PD and the test was performed by dividing the data set in two groups: Age-unrelated groups and age-related groups. The case that they considered that had the best performance was when it was used age-related groups and LSSVM classifier obtained 99% accuracy for PD vs SWEDD [169].

Other study performed by Duchesne using also T1-weighted MRI images to classify PD, PSP and MSA achieved higher accuracy results (90.6%) and 93.3% sensitivity and 88.2% specificity. They used features from the characteristics of the hindbrain in the MRI images, namely shape and intensity, as input for the SVM algorithm [170]. Salvatore et al. also applied SVM algorithm to distinguish PD from PSP. The authors performed a voxel-based analysis and conclude that best areas to discriminate this two PS were: midbrain, corpus callosum and thalamus [171]. Zhang *et al.* discriminate PD and ET using machine learning framework based on SVM and statistical analyses for classification. The significant features were extracted with the Principal Component Analysis technique from the most significant regions of interest: caudate, calcarine and right medial frontal gyrus. The resulting accuracy was 84.54%, 84.48% and 87.22% for caudate, calcarine and right medial frontal gyrus, respectively [172].

Chapter 3 Methodology

This section is divided in four main topics: study population, MRI and DaTscan™ pre-processing steps, train, validation and test set formation and CNN architecture. Firstly, it is presented the data population selection, the criteria for data extraction and data set formation will be explained. Secondly, the pre-processing of the MRI and SPECT images will be described as well as the formation of the slices that will be used for the classification. Then, the CNN framework used in this work will be described together with the parameters and images used for train, validation and test sets. Lastly, the methods used for the evaluation of the model performance will be presented and described as well.

3.1 Study population

The images used in this study were extracted from the PPMI database. PPMI is an international, multi-center and longitudinal study constituted by data that came from several clinical sites in the world, such as, United States, Europe, Israel and Australia [173].

This database is a landmark for the clinical study of PD progression and its goal is to improve the understanding of the disease and boost the research to refine the PD therapeutics and adequate it for each patient [174]. PPMI is comprised of biological samples, clinical behavioural assessments and medical images obtained from different techniques based on standardized protocols. These biomarkers are available for the scientific community to aid the investigation, study and identification of biomarkers related with PD progression [174].

In the present moment of this study, the database was constituted by a total of 26295 brain scans and 1184 participants. These participants are divided into six cohorts according to their characteristics, diagnoses and exams results. The six cohorts are:

- **Control Subjects:** This group is comprised by health participants who are 30 years or older and do not present first degree blood relatives with PD. This exclude the risk of the disease being inherited from relatives [175].
- ***de novo* PD subjects:** This group contains participants early-diagnosed with PD (up two years) and that never were medicated with PD medication [175].
- **Prodromal subjects:** These participants do not have PD but were diagnosed with hyposmia or sleep behaviour disorder, which are features that may suggest the beginning of PD [175].
- **SWEDD subjects:** This group contains people early diagnosed with PD that present a normal DaTscan™ [175].

- **Genetic Cohort Subjects:** This group is comprised by subjects who have a genetic mutation in a specific gene, and that were diagnose with or without PD [175].
- **Genetic Registry Subjects:** This group is comprised by subjects diagnosed with or without PD who have a genetic mutation or a first-degree relative [175].

All subjects, except controls, had been submitted to the cognitive tests MoCA, and two motor assessments MDS-UPDRS and Hoehn and Yahr scale performed in PPMI clinical sites.

PPMI is constituted by several brain scans obtained from different medical imaging techniques: Computed Tomography (CT), Diffusion Tensor Imaging (DTI), SPECT, Positron Emission Tomography (PET), MRI and Functional Magnetic Resonance (fMRI) [174]. Although, all of these techniques are important for PD study [84], in this work it only will be used images acquired with MRI and SPECT. Moreover, it was also needed that the subjects in the study had some characteristics in common. Therefore, exclusion criteria based on 7 criteria were applied as described in the following section and summarized in Table 3.1.

3.1.1 Exclusion and inclusion criteria

1. **Group Exclusion:**

   This study is focused on PD, SWEDD and Healthy subjects due to the need to distinguish them. Therefore, all groups except *de novo* PD, SWEDD and Control were excluded. After this selection, the data sample was comprised by 23163 images and 749 participants.

2. **Images Exclusion:**

   As referred previously, only MRI and DaTscan™ SPECT images were used. On one hand, MRI images give anatomic information and allow to find changes in pathological brains when compared to controls. On the other hand, DaTscan™ provides information about striatum function. After this selection, the data sample was comprised by 7831 images and 749 participants.

3. **Exclude pre-processing MRI data:**

   In PPMI, some MRI are already processed. Since in this study it is needed that they must undergo the same processing, only the original MRI images will be considered. SPECT images do not need this criterion because PPMI only have pre-processed DaTscan images. After this selection, the data sample was comprised by 6425 images and 749 participants.

4. **T2 and PD MRI images exclusion:**

- The PPMI database consists of several MRI images, namely Proton Density Weighed (PD-W), T1-W image an T2-W images. In this study only the T1-W images will be used due to the following reasons:

- The conventional T1-W images have been considered important for the study of PS because they can translate biomarkers that indicates structural changes that occur during the progression of these diseases [84].
- In fact, T1-Weighted images are usually used for distinguishing PS such as PSP from PD and normal subjects [176]. This happens because PD patients present subtle changes in T1-Weighted images unlike other PS that have visible changes in MRI images [84].
- Besides, T1-Weighted can also detect changes in putamen volume in both early and advanced stage of PD [95].

After this selection, the data sample was comprised by 3130 images and 749 participants.

5. **MRI exclusion according to their acquisition characteristics:**

   PPMI database is comprised by several T1-weighted MRI images which have different acquisition parameters from each other such as the magnetic field, sequence acquisition and their dimension (2D or 3D). According to the literature, 3D images have higher spatial resolution than 2D images [177]. This is due to the fact that 3D images have a smaller section thickness that ensures that there is no loss of spatial information, which does not happen with 2D images [178]. Regarding the intensity of the magnetic field, it is possible to obtain good results with either 1.5T or 3T [179]. Images acquired with 3T allow that early stages of PD can be detected, but whether 3T or 1.5T, good results can be obtained to detect PS using both intensity scans [95], [180]. Moreover, according to the Canadian Agency for Drugs and Technologies in Health, it cannot yet be clinically concluded that 3T is more effective than 1.5T since it cannot be ascertained whether or not patients receive a personalized treatment and whether or not it translates into a different clinical outcome [181]. Thus, according to this, all the 3D T1-W images which acquired whether with 1.5T or 3T were included in this study. After this selection, the data sample was comprised by 3111 images and 749 participants.

6. **Subjects must have both exams - MRI-SPECT pair:**

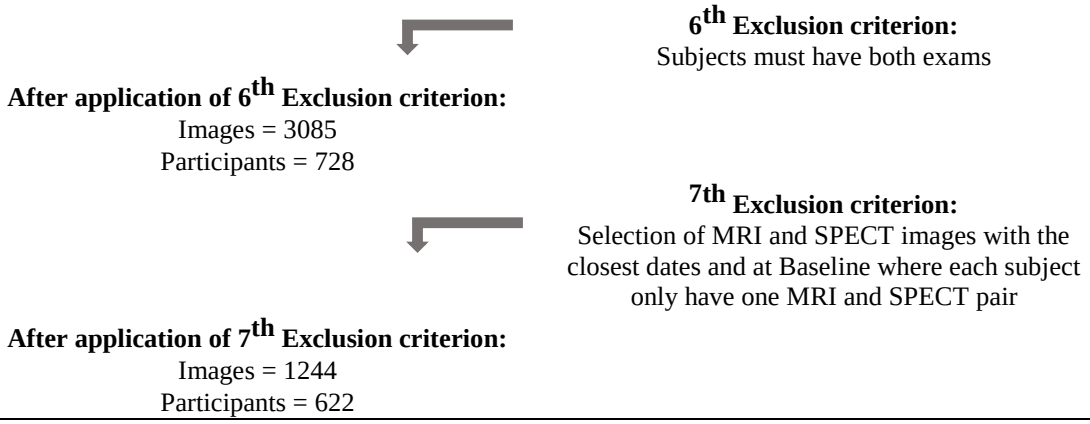
   This criterion is one of the most important. To compare images from different imaging techniques and to guarantee that SPECT images are adjusted to the respective subject, it is mandatory that each subject has at least one T1-weighted image and one DaTscan™ SPECT. Thus, all the participants that only have one MRI image or one DaTscan™ were excluded. After this selection, the data sample was comprised by 3085 images and 728 participants.

7. **Selection of MRI and SPECT images with the closest dates:** This criterion joins SPECT and MRI images that have the closest acquisition dates with the aim to form

image pairs. According to National Institute for Health and Care Excellence of the United Kingdom, patients with PD should be seen regularly to review their diagnosis, with intervals of 6-12 months [182]. In PPMI, there are participants that have more than one DaTscan™ or more than one T1-W image. In these cases, it is necessary to select which exam will be used to form MRI-SPECT pair. Thus, the MRI and SPECT exams that are selected to form a pair are the ones with closest dates, that is, the ones that only differ in a maximum of 6-12 months. Besides, since PPMI is a longitudinal study, several subjects have more than one SPECT and MRI pair. In this work it was considered one MRI-SPECT pair for each subject at baseline. After this selection, the data sample was comprised by 1244 images and 622 participants.

Table 3.1 Exclusion Criteria for Population of the study

Data	Criteria
PPMI Images = 26295 Participants = 1184	
	1st Exclusion criterion: Exclude all the groups except: Control, SWEDD and *de novo* PD
After application of 1st Exclusion criterion: Images = 23163 Participants = 749	
	2nd Exclusion criterion.: Removal of all the exams that will not be used (DTI, CT, fMRI and PET)
After application of 2nd Exclusion criterion: Images = 7831 Participants = 749	
	3th Exclusion criterion: Exclude all the processed MRI data
After application of 3nd Exclusion criterion: Images = 6425 Participants = 749	
	4th Exclusion criterion: T2-W and PD-W MRI images exclusion
After application of 4th Exclusion criterion: Images = 3130 Participants = 749	
	5th Exclusion criterion: MRI exclusion according to their acquisition characteristics
After application of 5th Exclusion criterion: Images = 3111 Participants = 749	



3.1.2 MRI and SPECT Acquisition

MRI and SPECT images used in this work were acquired in clinical sites of PPMI. This acquisition is based on a standard protocol developed by PPMI [174]. Although acquired with the same protocol, it is also important to notice that the images were acquired in MRI machines from different companies. As previously referred, those images were acquired at Baseline [183]. The 3D T1-Weighted images used in this study were acquired from Sagittal or Axial planes [183]. Regarding DaTscan™ SPECT images, those are acquired after 4 ± 0.5 hours after the radiotracer Ioflupane I123 injection [184].

The acquisition parameters of the MRI and SPECT images used in this work are described in the Appendix A. This information was extracted of the header of the DICOM MRI and DaTscan brain scans and from the PPMI database website (https://ida.loni.usc.edu/login.jsp).

3.1.3 Statistical data sample analysis

In this dissertation, it is intended groups with age and sex matched. Thus, the PD patients, SWEDD patients and Control subjects were considered categorical independent groups, wherein the age of each subject is a numerical variable and sex is a categorical variable. Then, it is was performed the Mann-Whitney U Test to test if there are age differences between the groups and the Chi-Square Test to test if there are associations between the sex of the groups since for the subjects of those groups must share similar characteristics to allow a reliable comparison.

Mann-Whitney U Test is a non-parametric test used to compare differences between two independent groups, when the dependent variable is numerical or ordinal [185]. In its turn, Chi-Square Test is used to detect if there are any relationship between two categorical variables [185]. Both methods test the Null hypothesis (H_0). This hypothesis postulates that there is no difference two phenomena using Mann-Whitney U Test and no relationship using Chi-Square Test [186]. The Null hypothesis is rejected when a value, called p-value, is less than a significant level. In this work, it was considered that the significant level is 0.05 because it usually the suggested by

the literature. Thus, if the p-value is less than 0.05, there is a significant difference for Mann Whitney U test or association between two groups using Chi-Squared Test. Otherwise, if the p-value is higher than 0.05, then there is no significant difference or association [185] depending on the test used. Both Mann Whitney U and Chi-squared tests were performed using the SPSS software.

3.2 MRI and SPECT pre-processing and registration

The principal aim of this procedure is to align brain scans even coming from different subjects into a common space using a standard template. All the pre-processing steps performed in this study followed typical procedures described and used in the literature [24], [25], [123], [187].

Firstly, both type of images, MRI and SPECT, were downloaded in Digital Imaging and Communications in Medicine (DICOM) format. After extracted, T1-Weighted and DaTscan images were converted from DICOM to Neuroimaging Informatics Technology Initiative (NIfTI) format with Statistical Parametric Mapping 12 (SPM) and dcm2nii – Neuroimaging tools & Resources Collaboratory, respectively. This format is the most used format in imaging pre-processing programmes because it facilitates data analysis [188].

Concerning the MRI pre-processing steps, they were performed using SPM12. SPM12 is an academic toolkit develop by Wellcome Trust Centre for Neuroimaging at University College London for functional data analysis and image processing based on statistical and mathematical concepts [189]. The pre-processing steps were: Anterior Commissure (AC) – Posterior Commissure (PC) alignment correction, bias regularisation, affine registration to Montreal Neurological Institute - International Consortium for Brain Mapping (MNI-ICBM) brain template, spatial normalisation, and smoothing. These steps were performed as follows:

1. The AC – PC alignment correction was manually performed to ensure that the T1-Weighted images have the closest orientation to the template and to facilitate the following steps [190]. This was executed using the “Display” option of the SPM12 interface;
2. To to reduce artefacts in MRI images and correct non-uniform intensity inhomogeneities caused by high-fields, it was applied bias regularisation = light regularisation (0.001) and Bias Full Width at Half Maximum (FWHM) = 60 mm cutoff (default values). This step was performed using the “Segment” option.
3. Affine registration was performed to register the MR images to the MNI- ICBM template. This process was performed with the “Segment” option. This step allows the formation of Deformation Fields images which are images that express information about the shapes of the subject brain and traduces the amount of displacement of these shapes in the location of 3D space [191]. Then, the Deformation Fields are used in the following step to complete the registration process.

4. Spatial normalization was used to handle scans into standard stereotactic space [191]. This step was performed using "Normalize: Write" with the deformation fields obtained in the previous step.
5. Smoothing is used to improve signal-to-noise ratio, reduce artefacts in MRI images and correct non-uniform tissue intensities 2. This step was performed using "Smooth" option with FWHM =2 [191], [192].

Regarding SPECT images, the DaTscan™ were already pre-processed when extracted from PPMI. According with the information provided by PPMI [193], SPECT raw projection data were acquired at PPMI Imaging centres and were sent to the HERMES (Hermes Medical Solutions, Stockholm, Sweden) to be reconstructed with the system for iterative reconstruction based on hybrid ordered-subset expectation maximisation, an algorithm for image reconstruction. In this step, no filters were applied [193]. Then the subsequent processing was applying a standard 3D Gaussian 6.0 mm filter to the resultant image. [18] Following that, all the scans were normalized to Montreal Neurologic Institute (MNI) space, a template that was created based on healthy young subjects [193].

Therefore, in this study, the only processing step that was performed to the SPECT images were the co-registration with the MRI images resultant from the smoothing step. The co-registration allows that SPECT and MRI images have the same spatial normalization. All DaTscan™ were co-registered with the respective T1-weighted image of the same subject. This was performed using "Coregister: Estimate & Reslice" option [192].

3.2.1 Input images

In clinical practice, the conventional medical images are 2D images [27], since 3D images present higher time of acquisition [194]. Thus, the 3D pre-processed MRI and SPECT images were divided into two-dimensional slices in PNG format which is compatible with software application used [195]. The slices were obtained from the Axial anatomical view illustrated in Figure 3.1.

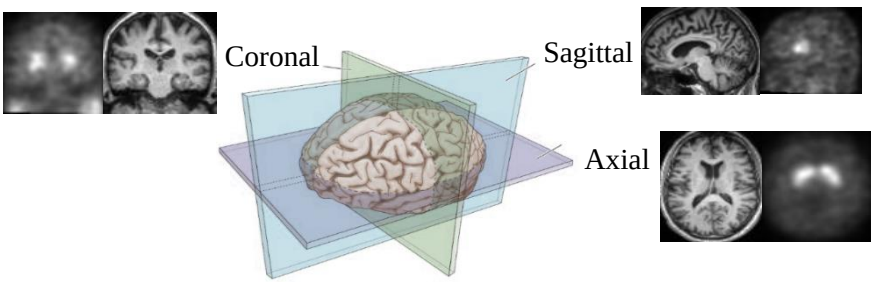


Figure 3.1 Anatomical planes of the brain. The Coronal view separate the front from the back of the brain. The Sagittal separates the left side from the right side. The Axial view is from the bottom of the brain to the top and it is the one used in this study. Adapted from [196].

Both MRI and SPECT images have the same dimensions due to the registration process, 79x95x79 voxels. Thus, each 3D Axial and each DaTscan™ image was divided in 79 slices. One

example of the slices obtained of one healthy subject is represented in the Appendix B. Each Axial slice has a size of 79x95 mm^2. The division in slices was performed using MATLAB.

After the division, it was applied a Contrast Limited Adaptive Histogram Equalization (CLAHE) to the resultant slice that were used in the classification. The CLAHE is an approach based on the Histogram Equalization (HE). HE is an image processing method used to increase the quality of the images based on contrast enhancement [197]. Basically, the HE uniforms the greys levels of an image using a histogram that reflects the uniform distribution of that grey levels. In its turn, unlike HE, CLAHE has a local approach in which it is performed an histogram for each region of the image [197]. The CLAHE was performed with the function "adapthisteq" of the MATLAB.

3.3 Classification with Convolutional Neural Networks

3.3.1 MedCNN and Caffe

In this study, it was used an application based on the deep learning framework Caffe. This application called *MedCNN: convolutional neural networks for medical applications* was developed by Godinho at Institute of Biophysics and Biomedical Engineering (IBEB) to be applied to the classification of biomedical signals and medical images [195].

Caffe is an open source framework of deep learning algorithms namely convolutional neural networks, developed by Berkeley AI Research. Caffe makes use of Graphics Processing Units, known as GPU, to fasten computation than using only the CPU [198]. GPU are powerful electronic circuits that are able to increase the parallel computing with a good energy efficiency [199], Figure 3.2.

The Caffe framework is comprised by C++ libraries, Compute Unified Device (CUDA®) and Blobs. All the implementation of this framework is C++ based [198] and the acceleration of Caffe computation [135] is performed by CUDA and cuDNN library , develop by NVIDIA [200]. The Blobs act like Protocol Buffers. Protocol Buffers are methods that serialize data, that is, convert data to formats for storing and communication [201] and they are compatible with several programming languages.

In its turn, Blobs are 4-dimensional arrays to store and communicate data within Caffe libraries [198]. Usually, Blobs stores batches of images and parameters of the network [198]. In this case that large-data is used, the data is store in Lightning Memory-Mapped Database Manager (LMDB) format. The LMDB library stores the data as byte arrays and it is characterized by its memory efficiency and elevated performance [202].

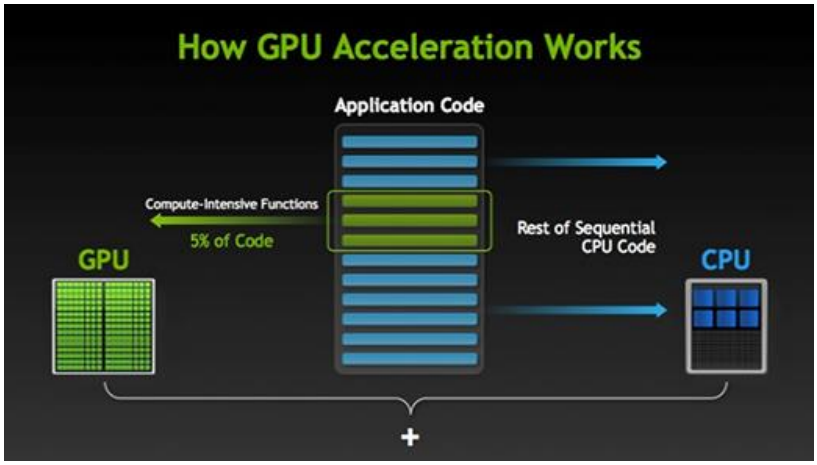


Figure 3.2 GPU versus CPU. During GPUusage, several of operations are performed by the GPU while CPU just runs some code. GPU has thousands of cores to process work in parallel and intensive functions, like shown in the image, e.g., three lines are processed at the same time unlike CPU which only have multiple cores and process in a serial sequence and one line at once. Adapted from: [200]

Caffe also provides interfaces for MATLAB and Python [198]. For instance, the MedCNN interface was developed using Python code [195]. See Figure 3.3

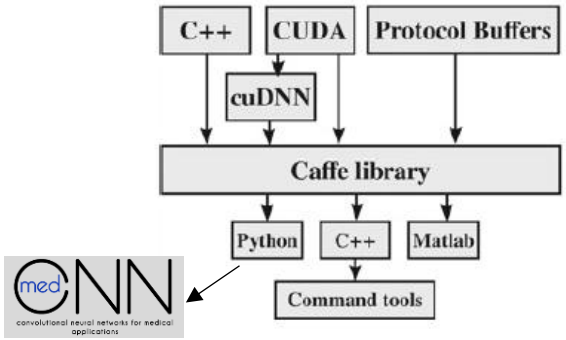


Figure 3.3 Framework Caffe and MedCNN application. Protocols Buffers, C++ and CUDA allow the criation of the Caffe Library that contains files that can be used by the Caffe user. The MedCNN application was develeoped using the python interface. Adapted from:[135]

Concerning the installation process of both Caffe and application of MedCNN, firstly is installed CUDA 8.0 and respectively cuDNN, followed by some Caffe dependencies required and described at Caffe website. The Anaconda Python 2.7 is also installed to allow the use of Caffe python code. Then, the framework Caffe is installed and MedCNN can be lunched trough the Spider, a python development environment, of Anaconda.

In this work, it was used a CentOS Linux 7, with the processor Intel® Xeon(R) CPU E5530 @ 2.40GHz × 4 and the graphic card GeForce GTX TITAN X/PCIe/SSE2. This GPU is characterized by its high performance due to its features: 3072 CUDA cores, 7.0 Gbps of memory clock and 12 GB of random access [200] which make it one of best NIVIDIA graphic cards [203].

3.3.2 CNN architecture

The MedCNN application works with the following network architecture. Firstly, the input data of the untrained network (the PNG slices) are converted to the LMDB format and scaled to 1/255 to set the intensity of the pixels of the image into a range of 0 to 1 to guarantee an intensity normalization of the images [204]. This scaling also enables faster computing since set all the inputs with the same covariance which helps in calculations in the training process [205]. Then, the data pass through the CNN architecture designed in the Figure 3.4.

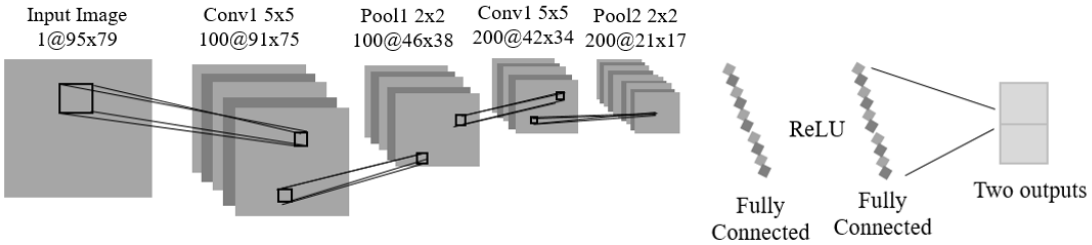


Figure 3.4 CNN designed in the MedCNN application and adapted to this study. The input and features maps have size denoted as Height x Width (Caffe notation). Adapted from: [195]

This CNN is inspired in the LeNet develop by LeCun et al. [138] and used in Godinho work [195]. As illustrated in Figure 3.4. The CNN is comprised by the following order and layers:

- **Convolutional layer** with a kernel size of 5x5, stride equal 1 and an output of 100 features maps in the first layer and 200 features maps in the second layer;
- **Pooling layer** with max-pooling function, kernel size of 2x2 and stride equal 2 and an output of 100 features maps;
- **Convolutional layer** with a kernel size of 5x5, stride equal 1 and an output of 200 features maps;
- **Pooling layer** with max-pooling function, kernel size of 2x2 and stride equal 2 and an output of 200 features maps;
- **Inner Product Layer**, the Fully Connected layer originates 100 output values
- **Activation layer** with the ReLU function
- **Inner Product Layer**: from the Fully Connected layer results 2 outputs which are the number of classes that the image may belong to.
- **Softmax layer** with the Softmax loss function.

A weight filler was applied to all layers except ReLU and Softmax layers. A weight filler is an algorithm that initializes the weights. In this work, it was applied the weight filler Xavier which is based on Glorot and Bengio study [206], and basically set the initial values of the weights based on the number of inputs and output nodes.

For the learning process it was used the SGD with the Backpropagation algorithm. In this procedure, it was performed a learning rate decay policy that defines how the algorithm learns the weights.

In this approach, it was used a learning rate decay policy defined as $\alpha_{in} \times (1 + gamma \times iter)^{-power}$, wherein α_{in} is the initial learning rate, $gamma$ is a value that drops the learning rate, $iter$ is the current iteration, and $power$ is related with the decay function. The parameters were initialized as follows: α_{in} has the value 0.001, gamma has the value 0.0001, power has the value 0.75 and momentum has the value 0.9. In this case, the regularization is given by the weight decay that was initialized with the value 0.0005. The maximum iterations was 5000 and for each 500 iterations occurs a validation test [195]. These backpropagation hyperparameters are stored in a file created by Caffe, called solver. These parameters are based on Godinho work [195] and caffe documentation.

3.3.3 Train, Validation and Test Sets

To train, validate and test the algorithm it was used the Monte Carlo cross-validation method present in the MedCNN application. This method randomly splits data into two sets: train set and test set Then, this process is repeated several times and several batches with one train set and one test set are formed. The resultant accuracy is given by the average of the accuracy of the batches [207]. In this case, the data is divided into three parts: the train set, the validation set and test set. Moreover, this method also prevents biased estimations.

In the train set, the train data is used to optimize the function that best fits the data, for instance, the weight model. In the validation set, the data it is used to optimize the hypermeters such as number of iterations and learning rate, to estimate the error between the value predicted and the real value and, thus, to verify if the network does not overfits. The validation occurs during network training. In its turn, the test set are non-labelled and unseen data by the trained network that test and evaluates the network [208].

In this study, Monte Carlo 5-fold cross-validation was performed to obtain 5 batches for each classification group. It was applied an approach of splitting the data in 50% for the train set, 25% for the validation set and 25% for the test set. Thus, for each batch it was used a split relation of 2:1:1 (train, validation and test).

In this work, the group of subjects that had the less subjects was considered the limiting group and *n* corresponded to number of individuals of that group. Thus, for each classification *2n* subjects was used which corresponds to the sum of the *n* subjects of the limiting group with *n* subjects random extracted from the group with more subjects.

In this study, it was performed the following classification: Control vs PD, Control vs SWEDD and PD vs SWEDD. The Table 3.2 resumes each classification group.

Table 3.2 Characteristics of each classification group

Groups	Nº batches	Sets*		
		Train	Validation	Test
Control vs PD	5	50% ($\frac{1}{2}n$ Control + $\frac{1}{2}n$ PD)	25% ($\frac{1}{4}n$ Control + $\frac{1}{4}n$ PD)	25% ($\frac{1}{4}n$ Control + $\frac{1}{4}n$ PD)
Control vs SWEDD	5	50% ($\frac{1}{2}n$ Control + $\frac{1}{2}n$ SWEDD)	25% ($\frac{1}{4}n$ Control + $\frac{1}{4}n$ SWEDD)	25% ($\frac{1}{4}n$ Control + $\frac{1}{4}n$ SWEDD)
PD vs SWEDD	5	50% ($\frac{1}{2}n$ PD + $\frac{1}{2}n$ SWEDD)	25% ($\frac{1}{4}n$ PD + $\frac{1}{4}n$ SWEDD)	25% ($\frac{1}{4}n$ PD + $\frac{1}{4}n$ SWEDD)

***** n is the number of subjects of the limiting group

The cases where this division into data sets gives non-integers numbers, the size number of the train and validation sets is round half down to the nearest integer and in the test set the size number is round half up to the nearest integer.

To increase the information given to the network regarding regions of interest the slices were assembled into groups of three, as illustrated in Figure 3.5, and the average of the three slices was performed to produce a slice without loss of texture information as described in Savio *et al.* study. [209]. Thus, the resulting voxels are the average of the intensities of the matching pixels. The resulting slice is, then, the input of the CNN. The generation of slices was for all subjects.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	57	57	58	59	60
61	62	63	64	65	67	67	68	69	70
71	72	73	74	75	77	77	78	79	---

Figure 3.5 Illustration of the groups of three slices that formed the 3-channel slices.

In this case, it was only considered slices from 21 to 47 since they contain specific regions of interest. The slice 21 contains the beginning of the mesencephalon and the slice 47 contains the end of the septum pellucidum.

To summarize, the classification was performed according to the following order:

1. Classification of MRI slices:
   1.1 Control vs PD
   1.2 Control vs SWEDD
   1.3 PD vs SWEDD
2. Classification of the DaTscan slices:
   2.1 Control vs PD
   2.2 Control vs SWEDD
   2.3 PD vs SWEDD

3.3.4 Test/ Evaluation of the Classifier CNN

The test results are stored in a comma-separated value (.cvs) file. Thus, it is possible to analyse the results using statistical methods. The analysis of the test set results is based on the calculation of the Accuracy, the Sensitivity, the Specificity, the Positive Predictive Value (PPV) and the Negative Predictive Value (NPV).

The accuracy defines the proximity of the results obtained to the real value, i.e., discrimination between patients from healthy subjects. The sensitivity evaluates the ability of the test to correctly predict if the subject has the disease and the specificity translates the capacity of the test to rightly predict healthy subjects.

In its turn, the PPV defines the probability of having the disease if the test is positive and the NPV defines the probability of do not have the disease if the test is negative [210]. The calculation of those parameters was performed using the confusion matrix in Table 3.3. In this study, the probability of having PD or to not having PD (control group), the probability of being SWEDD to not having SWEDD (control group) and the probability of having PD to having SWEDD was compared.

To evaluate the performance of the classifier, the Cohen's Kappa coefficient and the Receiver operating characteristic (ROC) curve was used. The Cohen's Kappa (κ) method measures the interrater reliability, that is, the agreement between two or more classifiers by taking into account that may occur agreement by chance. The agreement happens when two or more classifiers set the same score to the variables. Thus, if $\kappa=1$, the agreement is perfect [211]. The calculation of the Cohen's Kappa is expressed in Table 3.4.

In its turn, the ROC curve is a method that evaluates the classifier performance [212]. This curve is given by the plot of sensitivity vs false positive rate (FPR) that it is calculated by 1-specificity, as illustrated in Figure 3.6.

Table 3.3- Confusion matrix for the calculation of Sensitivity, Specificity and PPV and NPV. The ‘a’, ’b’, ‘c’ and ‘d’ defines the number of subjects for each situation

	Disease	**Not Disease**	
Positive Test	a	b	PPV = $\frac{a}{a+b}$
Negative Test	c	d	NPV = $\frac{d}{c+d}$
	Sensitivity = $\frac{a}{a+c}$	Specificity = $\frac{d}{d+b}$	Accuracy = $\frac{a+d}{a+b+c+d}$

Table 3.4 Calculation of Cohen’s Kappa. The a,b,c and d are the same as in the Table3.3

Cohen’s Kappa	
	N = a + b + c + d
$k = \frac{p_o - p_e}{1 - p_e}$	$p_o = \frac{(a+b)}{N}$
	$p_e = \frac{(a+b)}{N} \times \frac{(a+c)}{N} + \frac{(c+d)}{N} \times \frac{(b+d)}{N}$

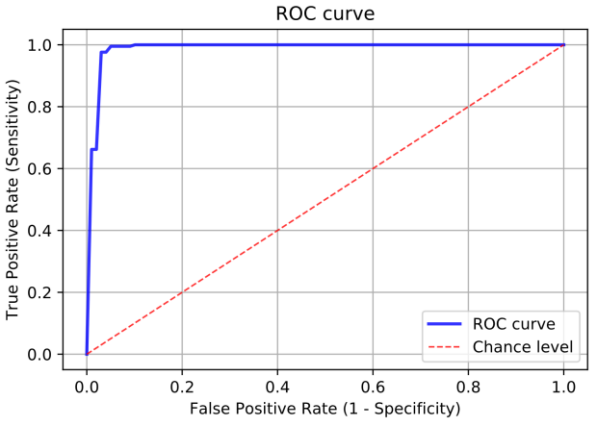


Figure 3.6 Example of a ROC curve

The closer is the curve to the sensitivity axes, the greater is the power of discrimination of the classifier. When the curve is near the chance line, it suggests that the classifier is not good at distinguish the two classes under testing (e.g. disease from healthy) [212]. The measure that translates the meaning of ROC curve is the area under the curve (AUC). AUC that analysis effectiveness of the classifier to distinguish disease from healthy and it based on the sensitivity and FPR. If AUC = 1, the classifier is perfect at performing that distinction. If AUC = 0.5, then the classifier is not better than change for balanced groups (i.e.groups with the same number of

subjects). If AUC = 0 the classifier set all the healthy subjects as unhealthy and unhealthy subjects as healthy [212], for example.

In this work, the ROC curve was calculated with 95% confidence interval and it is given by the average of the ROC curve and AUC of the 5 batches. The standard deviation represents the difference between the resulting average and the batches of each imaging technique. The ROC curve was only performed for the slices that showed the highest accuracies. The ROC curve and the AUC was performed using python.

Chapter 4 Results

4.1 Study Population

The resulting data sample obtained after imaging pre-processing steps and respective demographic information of the subjects are summarized in Table 4.1. Due to noise and image artefacts, 18 MRI brain scans and the matching SPECT images were excluded.

Thus, in this work, 1,208 images and 604 subjects were used. Table 4.1 shows that from the 604 subjects, 168 were Control subjects, 338 PD patients and 58 SWEDD patients, in which each subject has one MRI and one SPECT images. The average age for Controls is 60 ± 11 years old, 62 ± 8 years old for PD and for SWEDD is 62 ± 10 years old. Regarding sex, the gender ratio in terms of F/M was 0.54 in the Control group, 0.56 in the PD group and 0.53 in the SWEDD group.

Table 4.1 Subjects demographic information

Group	**s** (subjects)	**N** images	**Age (yo)** [31,50[	[50,70[	[70,86[	Avg±std	**Sex (F/M)**
Control	168	336 (168 MRI + 168 SPECT)	27	106	35	60 ±11	59/109
PD	378	756 (378 MRI + 378 SPECT)	40	252	86	62±10	136/242
SWEDD	58	116 (58 MRI + 58 SPECT)	7	38	13	62 ±10	20/38

yo.: years old; Avg: Average; std: Standard Deviation, F: Female; M: Male.

Regarding the statistics tests performed to analyse the data sample, the results of the Mann-Whitney U Test and Chi-square test are expressed in Table 4.2.

Table 4.2 Test statistics results of Mann-Whitney U Test and Chi-square Test for Control, PD and SWEDD groups

Groups	**Mann-Whitney U Test**		**Chi-Square**	
	U	p-value	χ^2	p-value
Control vs PD	29,884.50	0.272	0.037	0.847
Control vs SWEDD	4,354.50	0.228	0.008	0.930
PD vs SWEDD	10,523.00	0.623	0.049	0.825

The p-value, regarding age comparison, was 0.272 for Control and PD groups, 0.228 for Control vs SWEDD, and 0.623 for PD vs SWEDD. About sex comparison the p-value was 0.847 for Control and PD groups, 0.930 for Control and SWEDD and 0.825 for PD and SWEDD.

Appendix C illustrates an example of the slices MRI on Figure C.1 and DaTscan slices on Figure C.2 from control (healthy) subject obtained after the pre-processing steps and the images splitting, not including the CLAHE processing since it was just applied to slices that were used for training models.

4.2 Classification

Table 4.3 shows the number of slices that were used in the data sets of each classification of MRI and SPECT slices, separately taking into account order mentioned in Chapter 3.3.3.

In Control vs PD comparison, the limiting group was the Control group that comprises 168 subjects against the 378 PD patients. In both comparison (Control vs SWEDD and PD vs SWEDD), the limiting group was the SWEDD group which comprises 58 slices.

As can be seen on Table 4.3, the train set comprises 168 slices in Control vs PD and 58 slices in the other comparisons, which is 50% of the total of slices used in each case. Each validation and tests set comprise 84 slices in Control vs PD, 14 slices and 15 slices, respectively, in Control vs SWEDD and PD vs SWEDD.

Table 4.3 Number of slices used for each classification group in Train, Validation and Test

Classification Group	Total slices used	Train	Validation	Test
Control vs PD	336 (168 Control/ 168 PD)	84 Control 84 PD	42 Control 42 PD	42 Control 42 PD
Control vs SWEDD	116 (58 Control/ 58 SWEDD)	29 Control 29 SWEDD	14 Control 14 SWEDD	15 Control 15 SWEDD
PD vs SWEDD	116 (58 PD/ 58 SWEDD)	29 PD 29 SWEDD	14 PD 14 SWEDD	15 PD 15 SWEDD

The slices that encompass regions of interest, namely, mesencephalon and basal ganglia, are illustrated in Figure 4.1.

4.1.1 Comparison of Control and Parkinson's Disease

The accuracy results obtained in the comparison of Control and PD subjects using MRI images for each batch are expressed in Table B.1 and using SPECT images are in Table B.2 of

Appendix B. These tables show that the slices that lead to the highest accuracy were MRI slices 27_28_29 and SPECT slices 39_40_41.

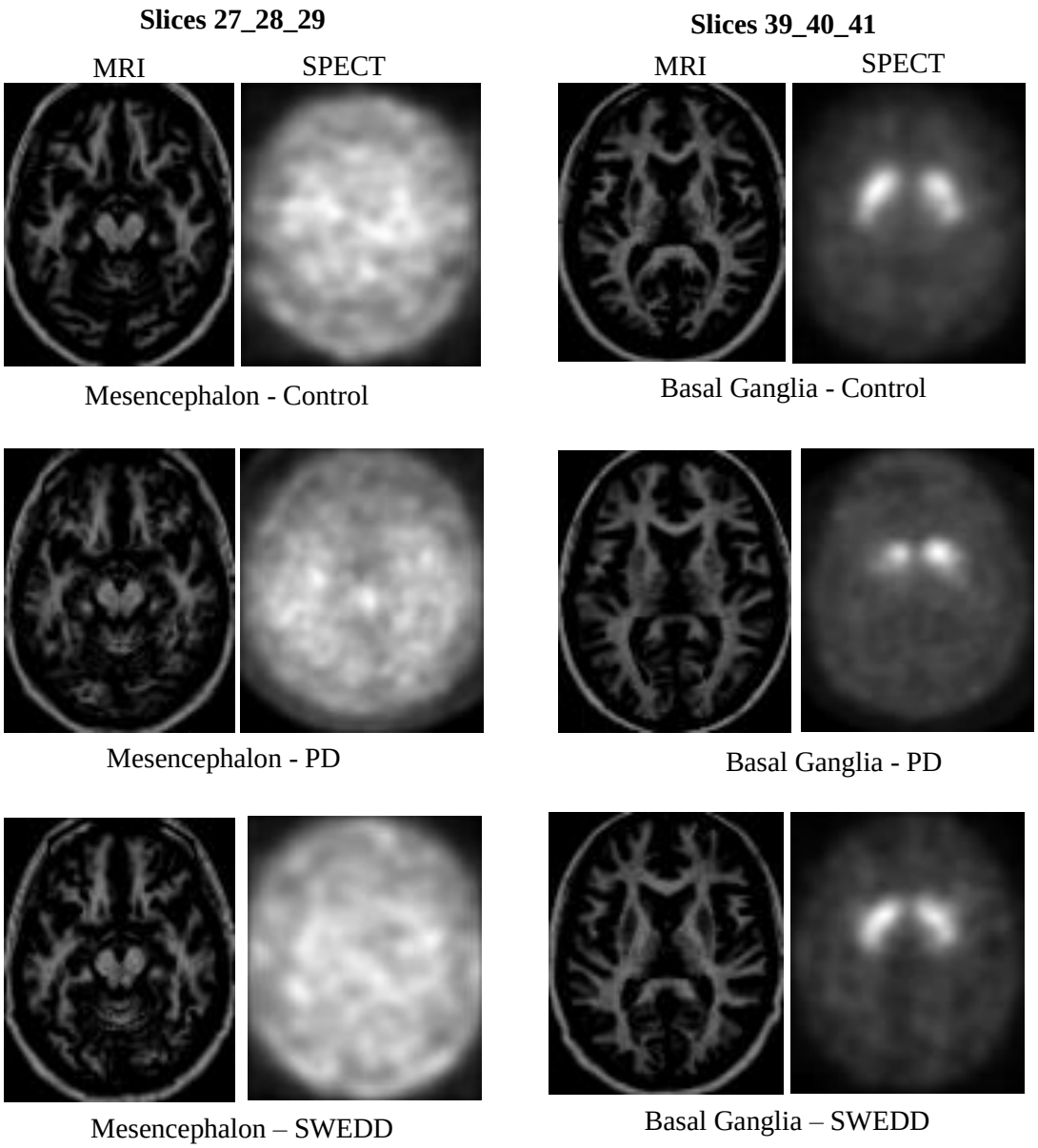


Figure 4.1 Slices that comprises the mesencephalon and the basal ganglia. These slices are from a Control subject, PD patient and a SWEDD patient.

Figure 4.2 shows the average accuracy of the five batches achieved with MRI and SPECT slices. As is shown in that figure, the CNN was able to distinguish Control from PD with 97.4±1.2% average accuracy using MRI slices 27_28_29 that encompassing the mesencephalon. This result was the highest in the distinction of PD from Control within the MRI and SPECT results, followed by SPECT slices 39_40_41 comprising the basal ganglia that had 92.4±3.6% average accuracy. Unlike MRI slices 27_28_29, the matching SPECT slices had a lower accuracy 62.0±2.5%. MRI slices 39_40_41 had also a small average accuracy 45.7±2.9 %. The remain MRI slices showed an accuracy approximately to 50% as SPECT slices 21_22_23 and SPECT slices 24_25_26. From SPECT slices 30_31_32 to slices 45_46_47 the average accuracy was always above 73%.

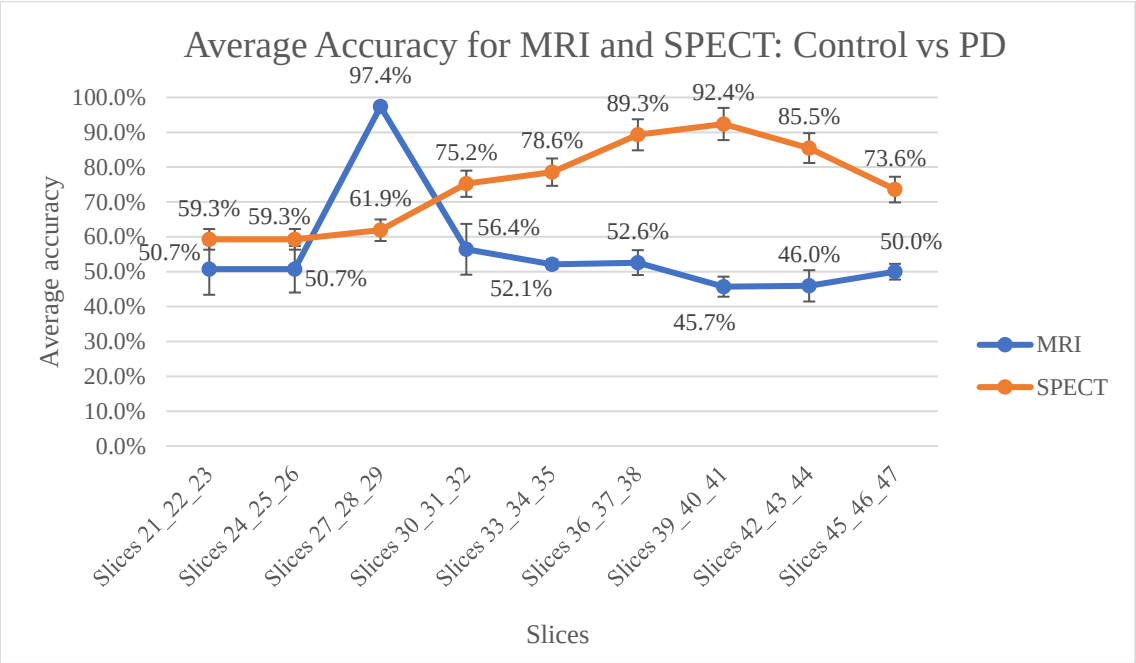


Figure 4.2 Average accuracy using MRI and SPECT slices in Control vs PD

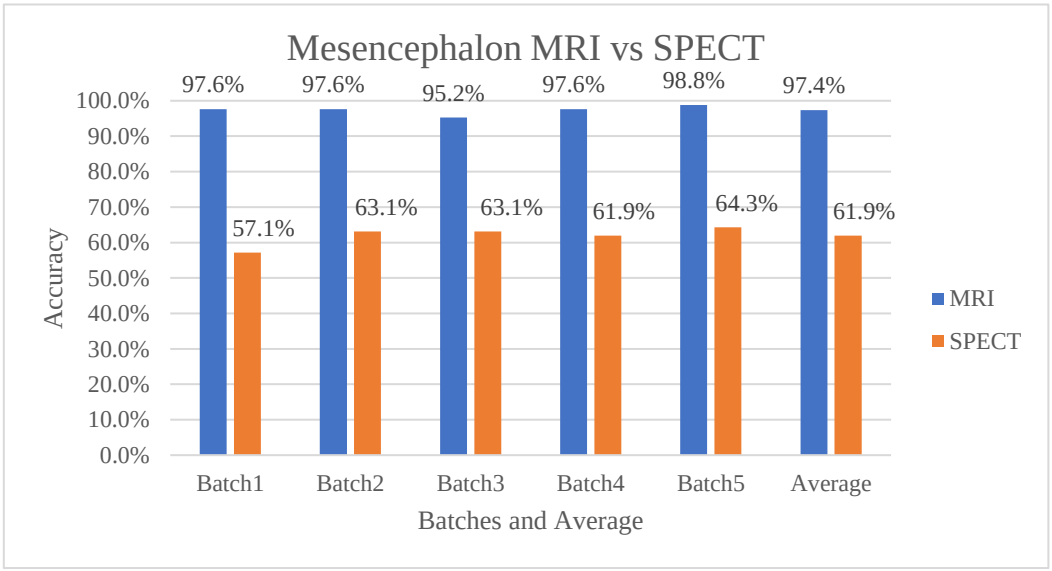


Figure 4.3 Accuracy results for Mesencephalon MRI vs SPECT in Control vs PD

Figure 4.3 illustrates the accuracy of the five batches resulting from the cross-validation of MRI and SPECT slices 27_28_29 that include the mesencephalon. This figure shows that the highest accuracy result achieved was 98.8% in batch 5 using MRI slices. The accuracy concerning SPECT slices 27_28_29 was close to 60% in all batches.

Regarding the accuracy obtained using slices 39_40_41 encompassing the basal ganglia, those results are illustrated in Figure 4.4. The highest accuracy in was 97.6% in Batch 2 and the lower was 88.1% using Batch 5 using SPECT slices. The other SPECT batches had similar results, unlike MRI batches that presented an accuracy close to 45%.

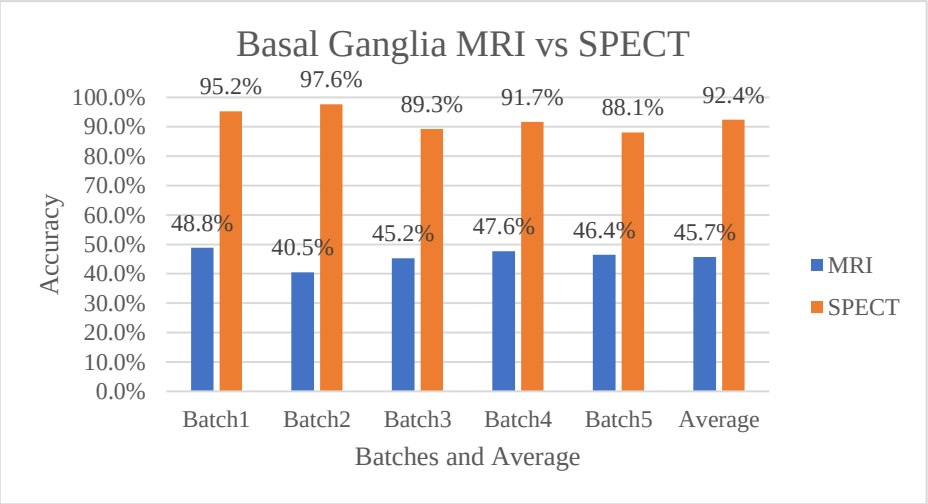


Figure 4.4 Accuracy results for Basal Ganglia - MRI vs SPECT in Control vs PD

Table 4.4 Sensitivity, Specificity, PPV, NPV and Cohen's Kappa for Control vs PD using MRI and SPECT images

Slices		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Kappa + std
Slices 21_22_23	MRI	51.43	55.24	53.55	53.14	0.07±0.11
	SPECT	59.05	59.52	59.61	59.09	0.19±0.11
Slices 24_25_26	MRI	50.95	50.48	50.02	51.56	0.01±0.11
	SPECT	55.71	62.86	60.04	58.96	0.19±0.11
Slices 27_28_29	MRI	100.00	94.76	95.07	100.00	0.95±0.03
	SPECT	62.38	61.43	61.82	62.23	0.24±0.11
Slices 30_31_32	MRI	54.29	58.57	56.05	57.01	0.13±0.11
	SPECT	74.29	76.19	76.82	74.76	0.50±0.09
Slices 33_34_35	MRI	53.81	50.48	52.15	52.16	0.04±0.11
	SPECT	82.38	74.76	77.03	80.81	0.57±0.09
Slices 36_37_38	MRI	49.05	56.19	52.49	52.76	0.05±0.11
	SPECT	89.05	89.52	89.81	89.34	0.79±0.07
Slices 39_40_41	MRI	48.10	43.33	45.84	45.45	-0.09±0.11
	SPECT	92.86	91.90	92.11	92.81	0.85±0.06
Slices 42_43_44	MRI	44.76	47.14	45.80	46.09	-0.08±0.11
	SPECT	81.90	89.05	88.42	83.35	0.71±0.08
Slices 45_46_47	MRI	48.10	51.90	50.14	49.88	0.00±0.11
	SPECT	75.24	71.90	73.03	74.64	0.47±0.10

std: Standard Deviation

The results of Sensitivity, Specificity, PPV, NPV, and Kappa of MRI and SPECT slices of the classification, are summarized in Table 4.4. The higher specificity and sensitivity results were 100.0% and 94.76% respectively, obtained with MRI slices 27_28_29.

Unlike MRI, the matching SPECT slices had 62.38% sensitivity and 61.43% specificity. SPECT slices 39_40_41 achieved 92.86% sensitivity and 91.90% specificity, and the matching MRI slices had 48.10% sensitivity and 43.33% specificity.

Figure 4.4 also shows that the PPV and NPV was 95.07% and 100.0%, respectively, in MRI slices 27_28_29, and 92.11% and 92.81% in SPECT slices 39_40_41. Regarding the Cohen's Kappa, MRI slices 27_28_29 achieved $\kappa = 0.95 \pm 0.03$ which is higher than the result of SPECT slices 39_40_41 that was $\kappa = 0.85 \pm 0.06$.

The ROC curve and the respective AUC of the mesencephalon and basal ganglia are illustrated in Figure 4.5 and Figure 4.6, respectively. Figure 4.5 shows that the AUC was higher using MRI slices since these slices presented AUC = 0.99, CI 95%: [0.96 - 1.00] using MRI slices of the mesencephalon.

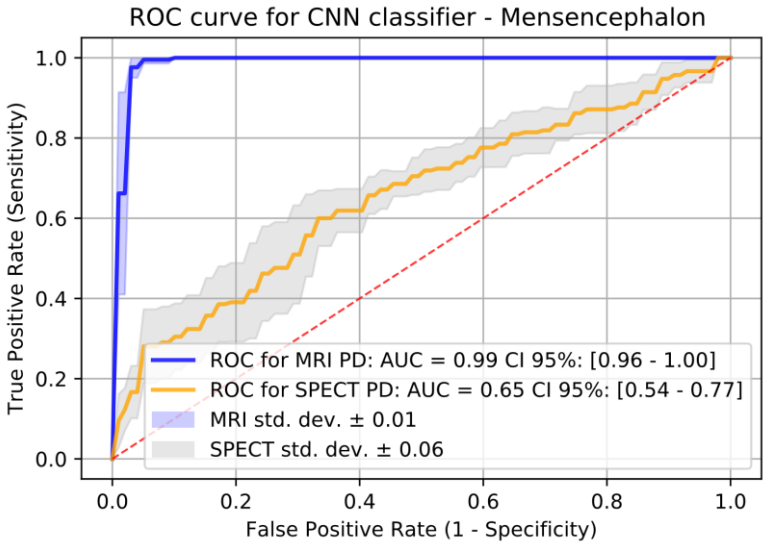


Figure 4.5 ROC curve and AUC for PD in mesencephalon classification of Control vs PD.

Regarding slices encompassing the basal ganglia, Figure 4.6 shows that the AUC was 0.98, CI 95%: [0.94 - 1.00] using SPECT slices which is higher than the results obtained using the matching MRI slices.

Figure 4.7 expresses the confusion matrixes of the batches with higher accuracy of the classification of the mesencephalon. The confusion matrix of the batch 5 of the classification using MRI slices shows that the CNN was able to correctly predict 100% of the PD patients and 97.62% of the Controls subjects. The CNN using SPECT slices of the batch 5 was able to correctly classify 61.90% of the PD-patients and 66.67% of the Control Subjects.

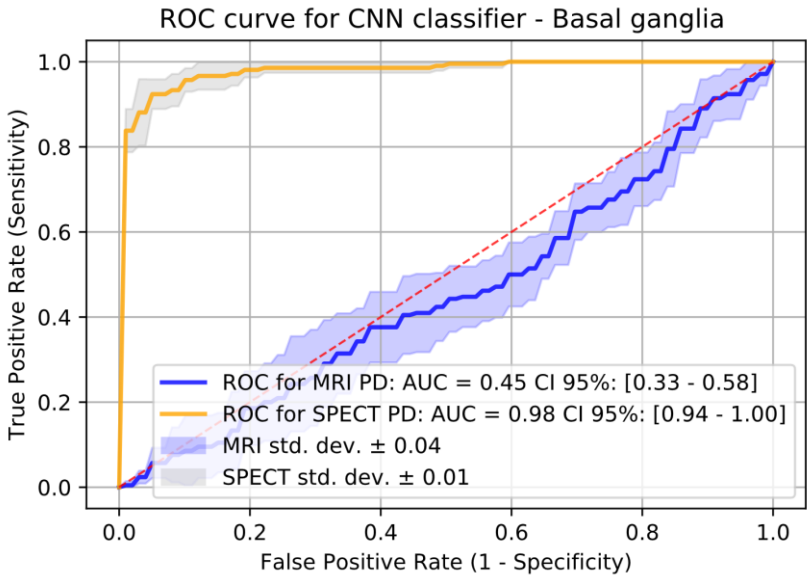


Figure 4.6 ROC curve and AUC for PD in basal ganglia classification of Control vs PD.

Confusion matrix of MRI slices - Mesencephalon

Batch 5	**Real Class**		
Predicted Class	**PD**	**Control**	
PD	42	1	43
Control	0	41	41
	42	42	**N = 84**

Confusion matrix of SPECT slices - Mesencephalon

Batch 5	**Real Class**		
Predicted Class	**PD**	**Control**	
PD	26	14	40
Control	16	28	44
	42	42	**N = 84**

Figure 4.7 Confusion matrix of MRI slices of the batch 5 and SPECT slices of the batch 5. These slices comprise the mesencephalon (Slices 27_28_29).

Regarding the classification of the basal ganglia, Figure 4.8 shows the predictions of the classification using MRI and SPECT slices 39_40_41.

The CNN, using MRI slices of the Batch 2, was able to rightly predict 45.24% of the PD patients and 35.72% of the Controls. These results were lower than those achieved by the classification using SPECT slices. With SPECT slices, the CNN was able to predict 95.24% of the PD patients and 95.24% of the Control subjects.

Confusion matrix of MRI slices – Basal Ganglia

Batch 2	Real Class		
Predicted Class	**PD**	**Control**	
PD	19	27	46
Control	23	15	38
	42	42	**N = 84**

Confusion matrix of SPECT slices – Basal Ganglia

Batch 1	Real Class		
Predicted Class	**PD**	**Control**	
PD	40	2	42
Control	2	40	42
	42	42	**N = 84**

Figure 4.8 Confusion matrix of MRI sliced of the batch 2 and SPECT slices of the batch 1. These slices correspond to the basal ganglia (slices 39_40_41).

4.2.2 Comparison of Control and SWEDD

The accuracy results of the classification of Control vs SWEDD are expressed in Table B.3 and Table B.4 for MRI slices and in Table 4.8 for SPECT slices, respectively, in Appendix B. Those tables show that the accuracy results of MRI and SPECT slices are similar.

As can be seen in Figure 4.9, the results of the average accuracy of the MRI slices were slightly higher than those achieved using SPECT slices.

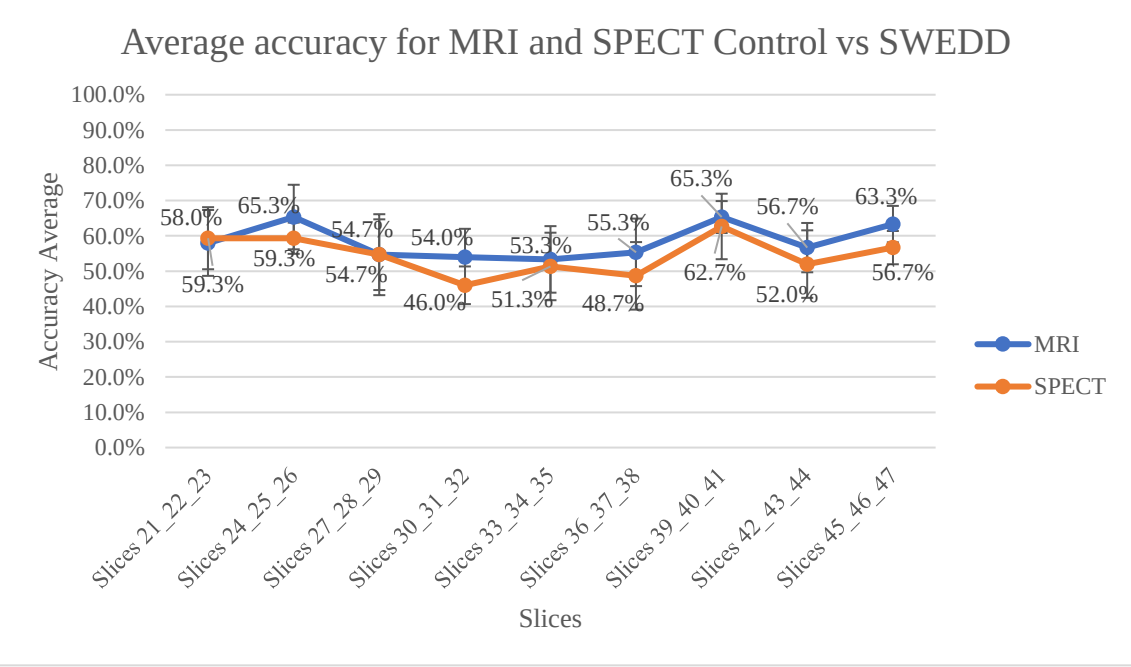


Figure 4.9 Average Accuracy for MRI and SPECT Control vs SWEDD

For instance, both MRI and SPECT slices that encompass the mesencephalon had 54.7% average accuracy. Regarding the slices that comprise the basal ganglia, SPECT slices 39_40_41 had 62.7% average accuracy and the matching MRI slices had 52.0%.

Figure 4.10 and Figure 4.11 show the resulting accuracy of the classification of Control vs SWEDD using MRI and SPECT slices comprising regions of interest.

Figure 4.10 illustrates that the batches have distinct results in both medical imaging techniques. For instance, MRI slices had 50.0% accuracy in Batch 1 and 73.33% in Batch 2. The same occurred with SPECT slides, e.g., Batch 1 presented 40.0% accuracy and Batch 5 had 70.00%.

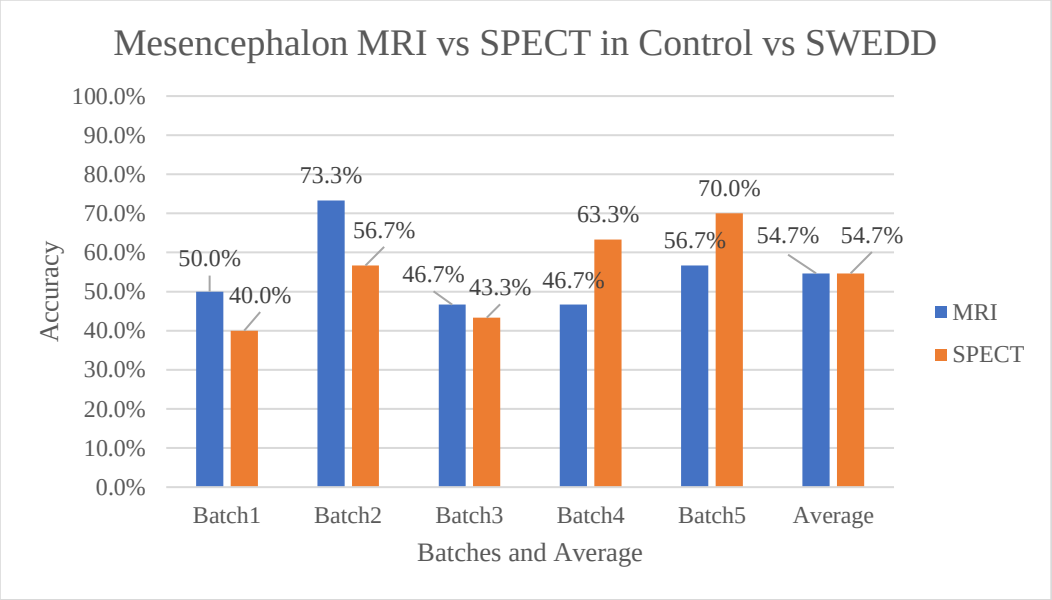


Figure 4.10 Accuracy results of Mesencephalon classification - MRI vs SPECT in Control vs SWEDD

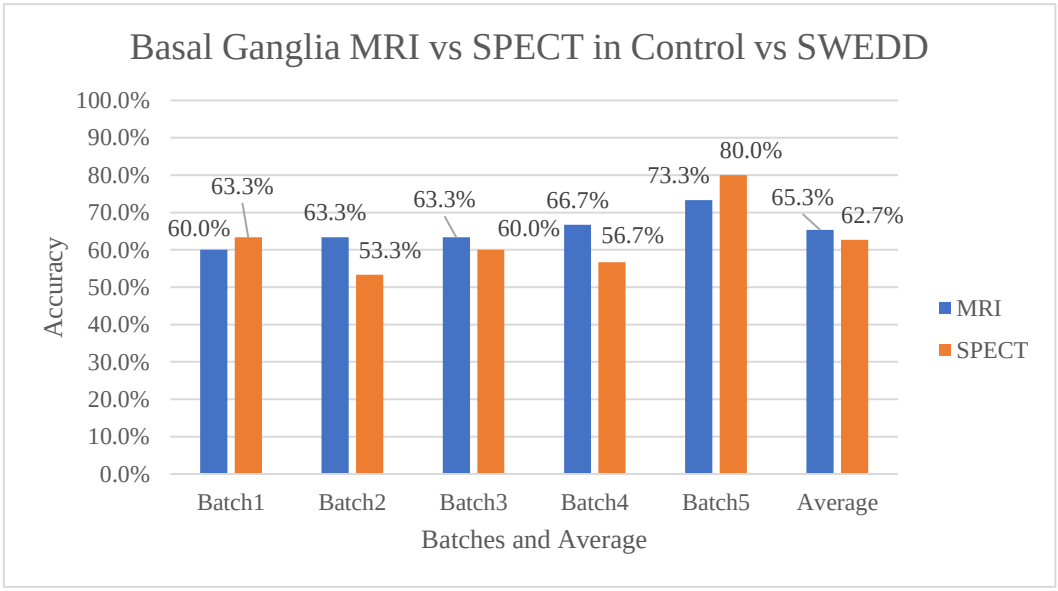


Figure 4.11 Accuracy results of Mesencephalon classification - MRI vs SPECT in Control vs SWEDD

Regarding the basal ganglia slices, Figure 4.13 shows that the batches in the classification with MRI slices have an accuracy between 60.00% and 73.33%. Concerning SPECT batches, those have an accuracy between 53.33% and 80.00%. As in the previous case of the mesencephalon classification, the results are similar.

The Table 4.9 expresses the Sensitivity, Specificity, PPV, NPV and Cohen's Kappa obtained in Control vs SWEDD classification.

Table 4.5 Sensitivity, Specificity, PPV, NPV and Cohen's Kappa for Control vs SWEDD using MRI and SPECT images

Slices		**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Kappa ± Std**
Slices 21_22_23	MRI	44.00	72.00	62.65	55.94	0.16±0.18
	SPECT	58.67	60.00	59.31	60.04	0.19±0.18
Slices 24_25_26	MRI	66.67	64.00	66.01	66.29	0.31±0.17
	SPECT	54.67	64.00	60.08	58.80	0.19±0.18
Slices 27_28_29	MRI	58.67	50.67	54.66	54.66	0.09±0.18
	SPECT	49.33	60.00	56.13	53.92	0.09±0.18
Slices 30_31_32	MRI	49.33	58.67	55.25	53.44	0.08±0.18
	SPECT	48.00	44.00	46.01	45.83	-0.08±0.18
Slices 33_34_35	MRI	58.67	53.33	55.10	57.20	0.12±0.18
	SPECT	53.33	49.33	52.67	47.54	0.03±0.18
Slices 36_37_38	MRI	53.33	57.33	57.39	54.75	0.11±0.18
	SPECT	53.33	44.00	50.20	46.55	-0.03±0.18
Slices 39_40_41	MRI	65.33	65.33	65.34	65.81	0.31±0.17
	SPECT	68.00	57.33	62.81	63.33	0.25±0.17
Slices 42_43_44	MRI	49.33	64.00	58.87	56.07	0.13±0.18
	SPECT	50.67	49.33	52.62	48.77	0.00±0.18
Slices 45_46_47	MRI	66.67	60.00	62.86	64.92	0.27±0.17
	SPECT	54.67	49.33	51.19	52.27	0.04±0.18

Std: standard deviation

As show in Table 4.5, MRI slices 27_28_29 had 58.67% sensitivity, 50.67% specificity, 54.66% PPV and 54.66 NPV and SPECT slices 39_40_41 had 68.00% sensitivity, 57.33% specificity, 62.81% PPV and 63.33 NPV. Regarding Cohen's Kappa, MRI slices encompassing the mesencephalon and the matching SPECT slices had both $\kappa = 0.09 \pm 0.18$. SPECT slices 39_40_41 presented $\kappa = 0.25 \pm 0.17$ and the matching MRI slices had $\kappa = 0.31 \pm 0.17$.

Concerning the ROC curve and the respective AUC, the results of AUC achieved using slices comprising the mesencephalon and the basal ganglia are illustrated in Figure 4.12 and Figure 4.13, respectively. In the classification of the Control vs SWEDD, Figure 4.12 shows that MRI slices encompassing the mesencephalon had AUC = 0.57 with CI 95% [0.36-0.78] and SPECT slices had AUC = 0.54 with CI 95% [0.33-0.75] and these results are close to the chance level.

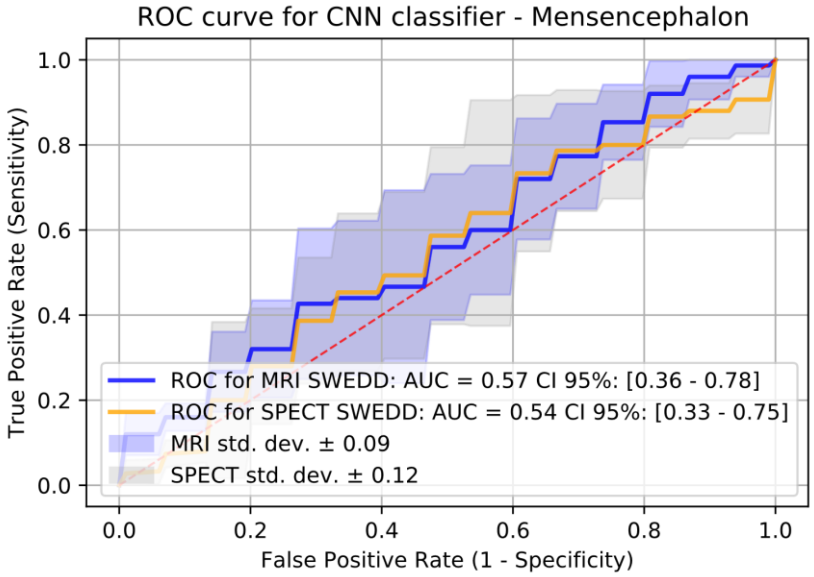


Figure 4.12 ROC curve and AUC obtained in the classification of mesencephalon slices, in Control vs SWEDD

As it is shown in Figure 4.13, the AUC was 0.67 with CI 95% [0.48-0.87] using MRI slices and 0.66 with CI 95% [0.46-0.85] using SPECT slices.

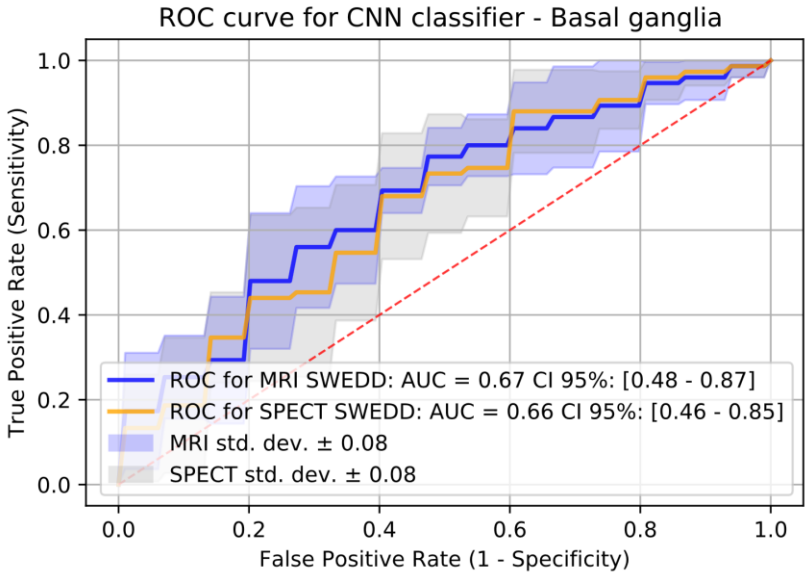


Figure 4.13 ROC curve and AUC obtained in the classification of basal ganglia slices, in Control vs SWEDD

The confusion matrixes of the classification of Control and SWEDD using MRI and SPECT slices comprising regions of interest are expressed in Figure 4.14 and Figure 4.15.

Figure 4.14 shows that CNN using MRI slices 27_28_29 of Batch 2 was able to predict 66.67% of the SWEDD patients and 60% of the Controls and SPECT slices of Batch 5 correctly predict 80% of the SWEDD cases and 80% of the healthy subjects. It is important to notice that these were the batches that presented the highest accuracy.

In its turn, Figure 4.15 shows that using MRI slices comprising the basal ganglia, the CNN was able to correctly predict 73.33% of the SWEDD cases and 73.33% of the Controls.

Regarding SPECT slices, the CNN correctly classified 80% of the SWEDD patients and 80% of the healthy subjects.

Confusion matrix of MRI slices - Mesencephalon

Batch 2	**Real Class**		
Predicted Class	**SWEDD**	**Control**	
SWEDD	10	6	16
Control	5	9	14
	15	15	**N = 30**

Confusion matrix of SPECT slices - Mesencephalon

Batch 5	**Real Class**		
Predicted Class	**SWEDD**	**Control**	
SWEDD	12	3	15
Control	3	12	15
	15	15	**N = 30**

Figure 4.14 Confusion matrix of MRI slices of the batch 2 and SPECT slices of the batch 5. These slices correspond to the mesencephalon (Slices 27_28_29).

Confusion matrix of MRI slices – Basal Ganglia

Batch 5	**Real Class**		
Predicted Class	**SWEDD**	**Control**	
SWEDD	11	4	15
Control	4	11	15
	15	15	**N = 30**

Confusion matrix of SPECT slices – Basal Ganglia

Batch 5	**Real Class**		
Predicted Class	**SWEDD**	**Control**	
SWEDD	12	3	15
Control	3	12	15
	15	15	**N = 30**

Figure 4.15 Confusion matrix of MRI slices of the batch 5 and SPECT slices of the batch 5. These slices correspond to basal ganglia (Slices 39_40_41).

4.2.3 Comparison of Parkinson's Disease and SWEDD

The accuracy results achieved in the classification of PD vs SWEDD using MRI slices are in Table B.5 and using SPECT slices are in Table B.6 in the Appendix B. These results show that the highest accuracy was obtained using MRI slices 27_28_29, SPECT slices 36_37_38 and SPECT slices 39_40_41. The accuracy of these slices is expressed in the following figures.

For instance, Figure 4.16 shows the average accuracy obtained with all the slices considered in this work. As can be seen in Figure 4.16, the highest average accuracy was achieved using SPECT slices 36_37_38 which presented 93.3±3.7%, followed by Slices 39_40_41 comprising the basal ganglia that achieved 86.0±4.9%. Concerning MRI slices including the mesencephalon, these obtained 73.3±3.7% average accuracy. The remaining MRI slices had average accuracy results in a range of 51.3% to 69.3% and SPECT slices between 48.7% and 80.0%.

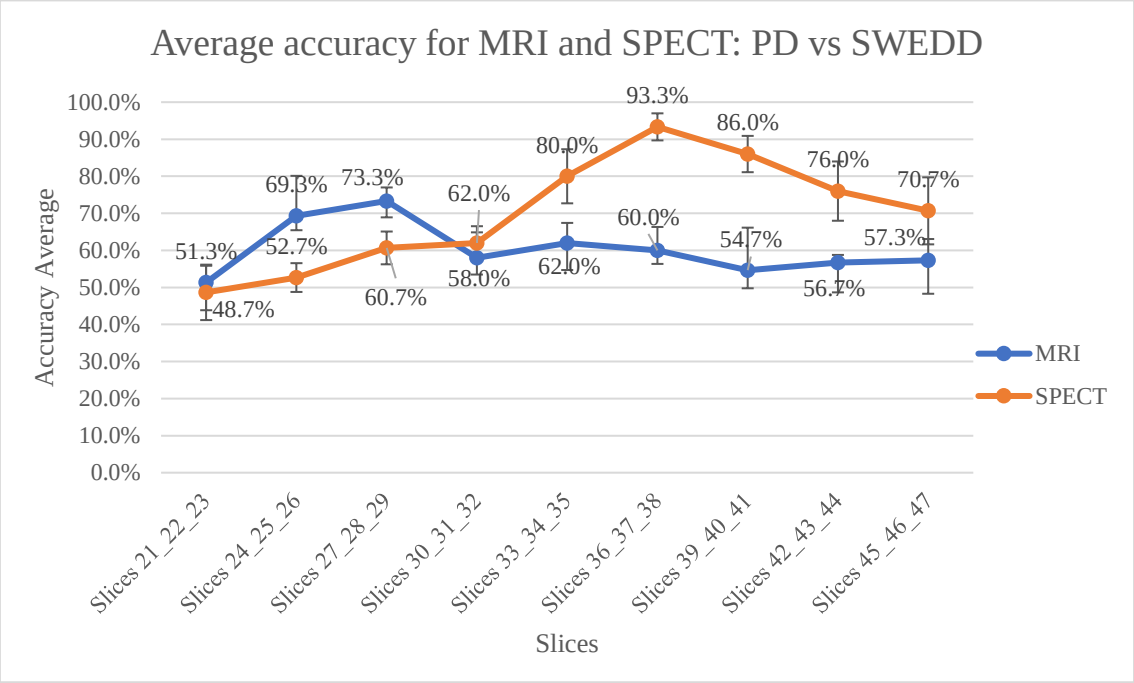


Figure 4.16 Average Accuracy for MRI and SPECT PD vs SWEDD

In the classification of PD vs SWEDD, using MRI slices 27_28_29, the CNN achieved 76.7% accuracy in batches 2 and 4, and 73.3% accuracy in batches 3 and 5, as shown in Figure 4.17. Unlike MRI, Figure 4.17 also shows that the accuracy of the matching SPECT slices was 66.67% in batch 4, 63.3% in batch 3 and 60.0% in batches 2 and 5.

Regarding the classification using slices that encompasses the basal ganglia, Figure 4.18 shows that using SPECT slices, the batch with the highest accuracy was the batch 1 with 93.3% accuracy, followed by the batch 3 that presented 90.0% accuracy. Regarding MRI slices, these presented approximately 56% accuracy.

The slices that presented higher accuracy, the results of the batches are displayed on Figure 4.19. SPECT slices 36_37_38 achieved 96.7% accuracy in batches 1 and 3 and 86.7% accuracy in batch 4 which is the lower result of these slices. Unlike SPECT, the accuracy of the matching slices was close to 60%.

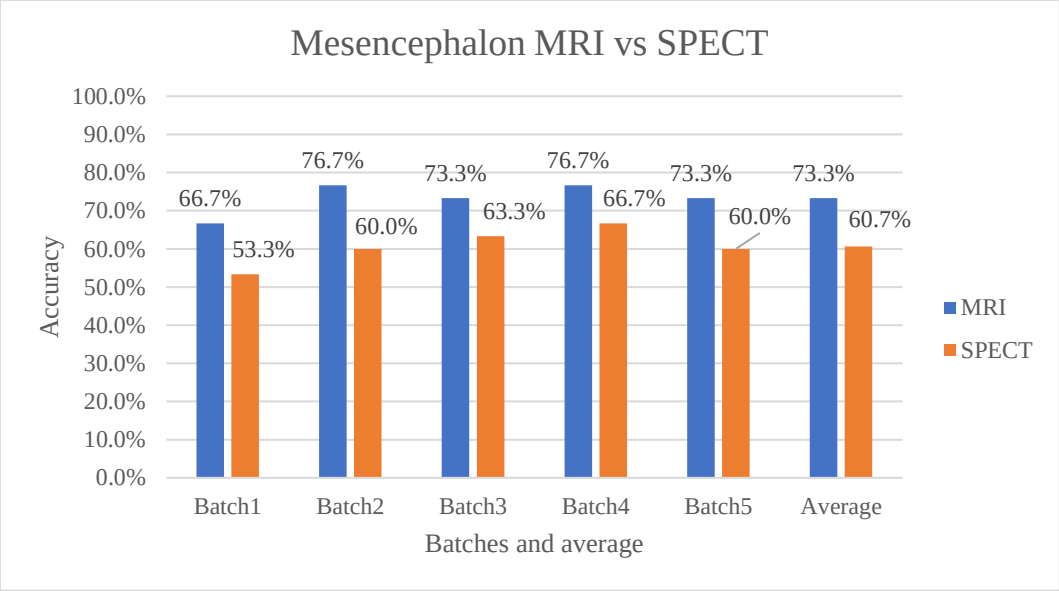


Figure 4.17 Accuracy results of Mesencephalon classification - MRI vs SPECT in PD vs SWEDD

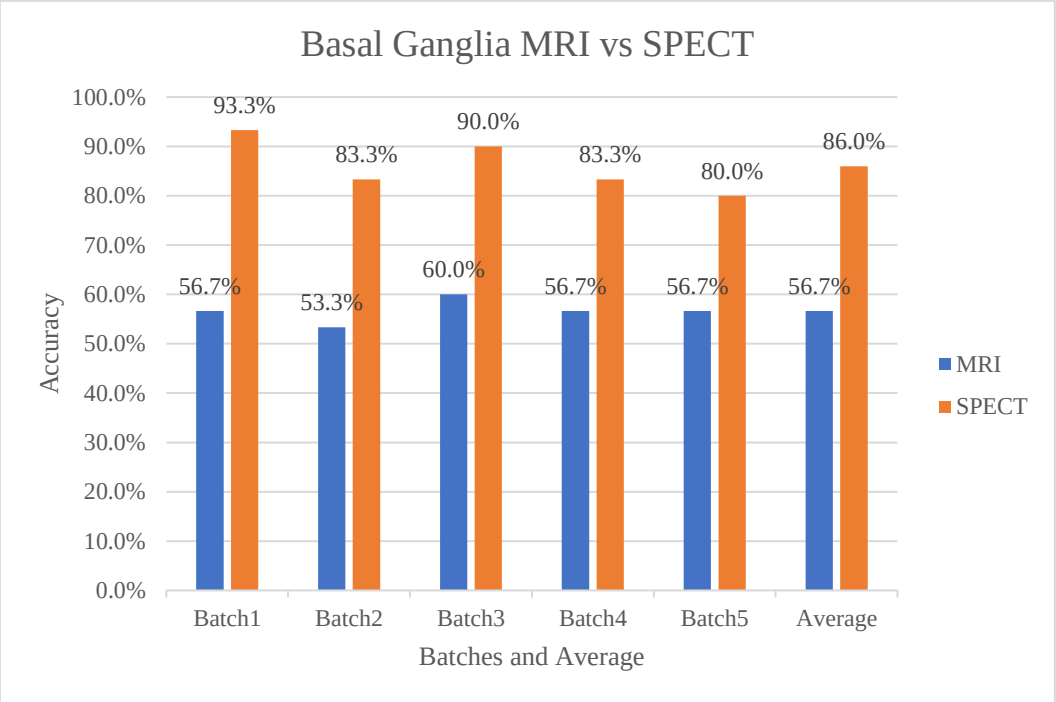


Figure 4.18 Accuracy results of Basal Ganglia classification - MRI vs SPECT in PD vs SWEDD

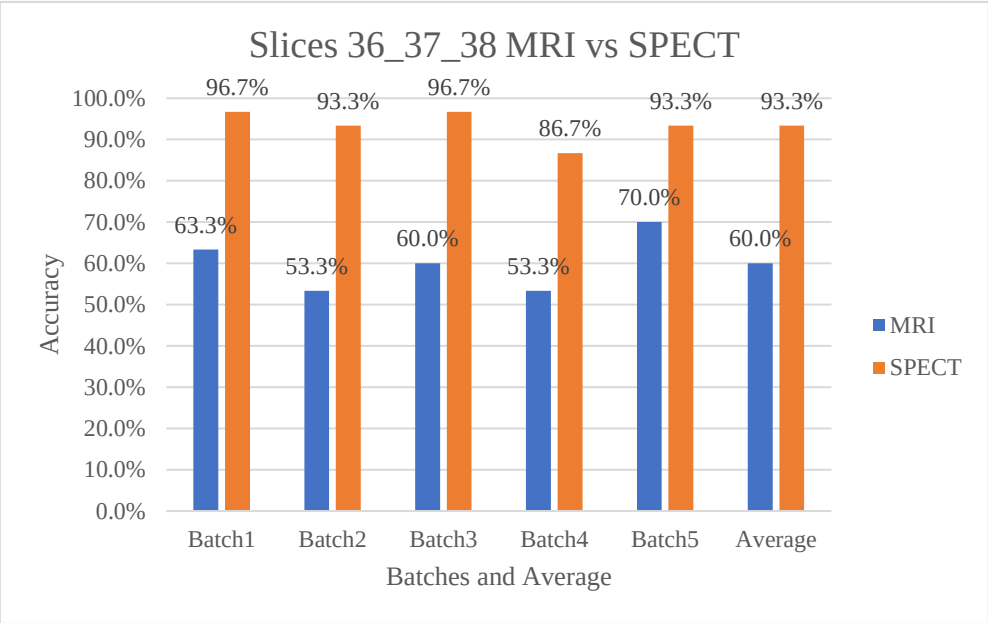


Figure 4.19 Accuracy results of slices 36_37_38 classification - MRI vs SPECT in PD vs SWEDD

The results of Sensitivity, Specificity, PPV, NPV and Cohen's Kappa are summarized in Table 4.12. The distinction of PD vs SWEDD using MRI slices encompassing the mesencephalon (Slices 27_28_29) showed 65.33% sensitivity, 81.33% specificity, 79.01% PPV, 70.88% NPV.

Table 4.6 Sensitivity, Specificity, PPV, NPV and Cohen's Kappa for PD vs SWEDD using MRI and SPECT slices

Slices		**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**	**Kappa ± Standard error**
Slices 21_22_23	MRI	50.67	52.00	51.36	51.20	0.03±0.18
	SPECT	50.67	46.67	47.96	50.28	-0.03±0.18
Slices 24_25_26	MRI	73.33	65.33	68.71	70.36	0.39±0.16
	SPECT	57.33	48.00	52.92	52.54	0.05±0.18
Slices 27_28_29	MRI	65.33	81.33	79.01	70.88	0.47±0.16
	SPECT	58.67	50.67	54.66	54.66	0.09±0.18
Slices 30_31_32	MRI	60.00	56.00	58.10	57.89	0.16±0.18
	SPECT	61.33	62.67	62.30	62.09	0.24±0.18
Slices 33_34_35	MRI	65.33	58.67	60.90	63.65	0.24±0.18
	SPECT	74.67	85.33	84.59	78.33	0.60±0.14
Slices 36_37_38	MRI	69.33	50.67	59.44	63.24	0.20±0.18
	SPECT	96.00	90.67	91.53	95.88	0.87±0.09
Slices 39_40_41	MRI	56.00	53.33	54.83	55.45	0.09±0.18
	SPECT	86.67	85.33	85.74	86.45	0.72±0.12
Slices 42_43_44	MRI	65.33	48.00	55.76	58.43	0.13±0.18
	SPECT	74.67	77.33	79.60	75.10	0.52±0.15
Slices 45_46_47	MRI	62.67	52.00	56.44	58.67	0.15±0.18
	SPECT	76.00	65.33	67.80	75.91	0.41±0.16

SPECT slices 27_28_29 had 58.67% sensitivity, 50.67% specificity. Regarding the slices that comprise the basal ganglia, the CNN achieved 86.87% sensitivity and 85.33% specificity using SPECT slices 38_40_41. The slices with higher accuracy, slices 36_37_38, presented 96% sensitivity, 90.67% specificity, 91.53% PPV and 95.88% NPV.

Concerning the Cohen's Kappa obtained in the comparison of PD and SWEDD, MRI slices 27_28_29 had $\kappa = 0.47 \pm 0.16$, SPECT slices 39_40_41 had $\kappa = 0.72 \pm 0.12$. SPECT slices 36_37_38 showed presented the highest κ (0.87 ± 0.09).

The results of the ROC curve to the classification PD vs SWEDD of the slices 27_28_29, slices 36_37_38 and slices 39_40_41 are in Figure 4.20, Figure 4.21 and Figure 4.22, respectively.

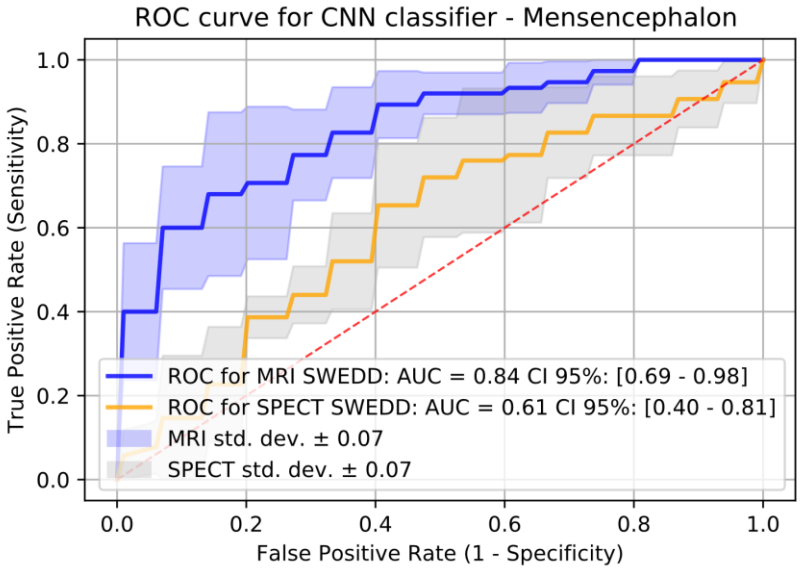


Figure 4.20 ROC curve and AUC obtained in the classification of mesencephalon slices, in PD vs SWEDD

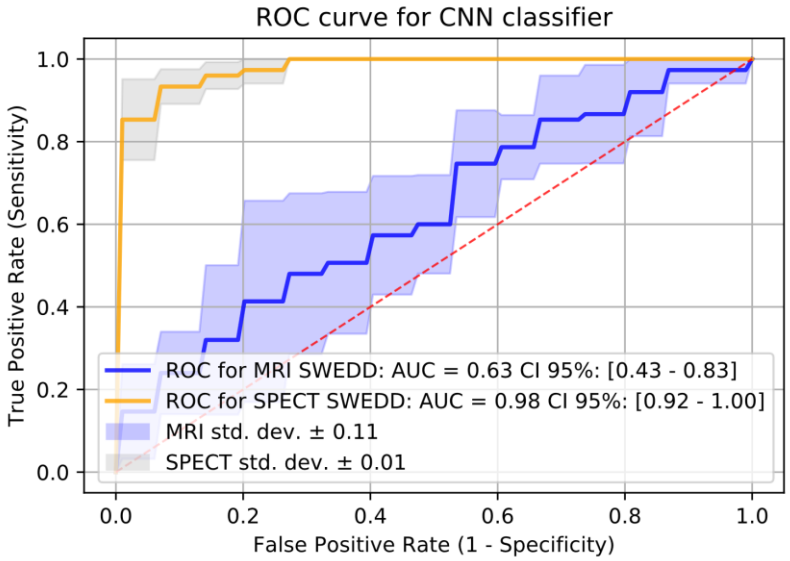


Figure 4.21 ROC curve and AUC obtained in the classification of slices 36_37_38, in PD vs SWEDD.

The classification of PD vs SWEDD using MRI slices including the mesencephalon had AUC = 0.84 [0.69 – 0.98] with CI of 95%. In its turn, SPECT slices presented AUC = 0.61 [0.40 – 0.81] with CI of 95%, as illustrated in Figure 4.22.

Regarding the slices of the basal ganglia, the ROC curves of MRI and SPECT slices are represented in Figure 4.21. In its turn, SPECT slices presented AUC = 0.90 [0.79 – 1.00] with CI of 95%, and a standard deviation of 0.05.

Figure 4.22 shows the ROC curves and respective AUC of MRI and SPECT slices 36_37_38. In this case, MRI slices showed AUC = 0.63 [0.43 – 0.83] with CI of 95%, SPECT slices presented AUC = 0.98 [0.92 – 1.00] with CI of 95%, and a standard deviation of 0.01.

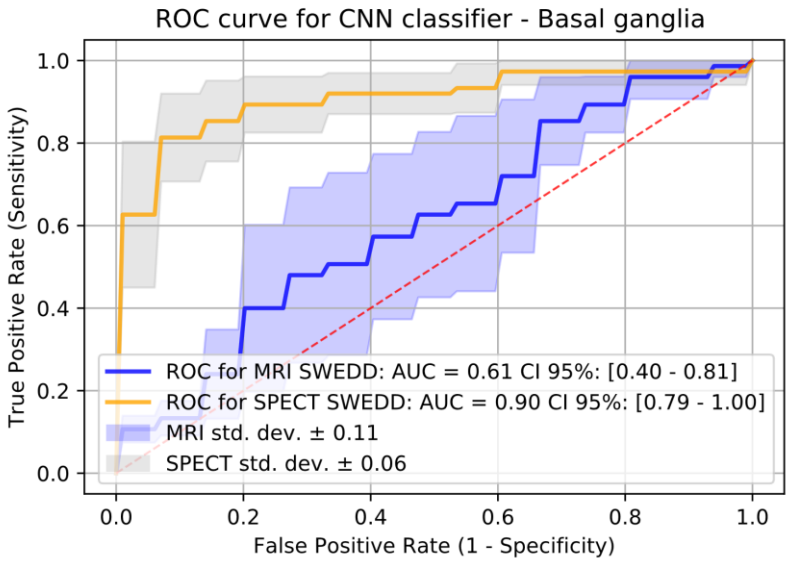


Figure 4.22 ROC curve and AUC obtained in the classification of basal ganglia slices, in PD vs SWEDD.

Confusion matrix of MRI slices - Mesencephalon

Batch 2	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	9	1	10
PD	6	14	14
	15	15	**N = 30**

Confusion matrix of SPECT slices - Mesencephalon

Batch 2	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	8	5	13
PD	7	10	17
	15	15	**N = 30**

Figure 4.23 Confusion matrix of the MRI and SPECT slices of the batch 2. These slices correspond to the mesencephalon (Slices 27_28_29).

Figure 4.23 contains the confusion matrixes of MRI and SPECT slices 27_28 _29 of the batch 2. The CNN classifier using MRI slices was able to correctly predict 60% of the PD cases and 93.33% of the Control subjects.

Using MRI slices 39_40_41 encompassing the basal ganglia, the CNN was able to correctly predict, in batch 5, 80% of the PD patients and 40% of the Control subjects, as illustrated in Figure 2.24. Unlike MRI, using the matching SPECT slices of PD, the algorithm was able to predict 100% of the PD cases and 93.33% of the Controls.

Confusion matrix of MRI slices – Basal Ganglia

Batch 5	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	13	6	19
PD	2	9	11
	15	15	**N = 30**

Confusion matrix of SPECT slices – Basal Ganglia

Batch 5	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	12	3	15
PD	3	12	15
	15	15	**N = 30**

Figure 4.24 Confusion matrix of the MRI and SPECT slices of the batch 5. These slices correspond to basal ganglia (slices 39_40_41).

Confusion matrix of MRI slices – Slices 36_37_38

Batch 3	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	12	9	21
PD	3	6	6
	15	15	**N = 30**

Confusion matrix of SPECT slices – Slices 36_37_38

Batch 3	**Real Class**		
Predicted Class	**SWEDD**	**PD**	
SWEDD	15	1	16
PD	0	14	14
	15	15	**N = 30**

Figure 4.25 Confusion matrix of the MRI and SPECT slices of the batch 3. These slices correspond to slices 36_37_38.

As it is shown it Figure 4.25, using SPECT slices 36_37_38, the CNN was able to correctly classify 80% of the PD cases and 80% of the Controls subjects unlike using MRI slices that rightly predict 86.67% of the PD patients and 86.67% of the Controls.

Chapter 5 Discussion

5.1 Main results

The main and most significant results obtained in this dissertation are following enumerated:

a) The discrimination of Control from PD:
   - Using MRI slices 28_29_30 encompassing the mesencephalon, the CNN achieved 97.4% average accuracy, which were the highest accuracy result in this classification, and AUC = 0.99
   - Using SPECT slices 39_40_41 encompassing the basal ganglia, the CNN obtained 92.4% average accuracy and AUC = 0.98

b) The discrimination of Control from SWEDD:
   - MRI and SPECT slices had similar results, for instance, approximately 50-60% accuracy

c) The discrimination of PD from SWEDD:
   - Using MRI slices 28_29_30 encompassing the mesencephalon, the CNN achieved 73.3% average accuracy and AUC = 0.84
   - Using SPECT slices 39_40_41 encompassing the basal ganglia, the CNN obtained 86.0% and AUC = 0.90
   - Using SPECT slices 36_37_38, the CNN achieved 93.3% average accuracy and AUC = 0.98

5.2 Classification

Prior to the classification, the Mann-Whitney test and Chi-Square were performed. The Mann-Whitney results showed that there are no had no significant differences in terms of age between the subjects of the following groups: Control and PD ($p = 0.272$), Control and SWEDD ($p = 0.228$), and PD and SWEDD ($p = 0.623$). In its turn, the Chi-Squared showed that there is no significant association between the sex of the subjects that comprises these groups: Control and PD ($p = 0.847$), Control and SWEDD ($p = 0.930$), and PD and SWEDD ($p = 0.825$).

Taking into account these results, the data sample used in Control vs PD, in Control vs SWEDD, and in PD vs SWEDD was considered age and sex-matched as well as the groups of subjects present in each test set used in the classification.

Concerning the classification of Control vs PD, the CNN classifier was able to distinguish PD from healthy subjects using MRI slices that encompassed the mesencephalon and using SPECT slices comprising the basal ganglia. Moreover, these MRI slices lead to higher accuracy in the

classification of PD and Control. For instance, MRI slices embracing the mesencephalon presented higher average accuracy (97.4%) than the matching SPECT slices (61.70%) and SPECT slices including the basal ganglia presented also higher accuracy (92.4%) than the same MRI slices (41.75%). The results also showed that in the classification using SPECT slices, the CNN tend to present higher accuracy when slices containing the basal ganglia or structures near it are used. Unlike SPECT, the accuracy of the classifier using the matching MRI slices tends to decrease to 45-50% in slices including the basal ganglia or other structures than mesencephalon.

Concerning Cohen's Kappa results, the kappa of the MRI slices encompassing the mesencephalon was 0.95±0.03 which, according to McHugh, is an almost perfect level of agreement [213] with the real PPMI label. In its turn, SPECT slices including the basal ganglia showed a strong level of accordance ($\kappa = 0.85 \pm 0.06$) [213] with PPMI database. The SPECT slices, containing structures anatomically located close to the basal ganglia, presented a moderate level of agreement with the PPMI labels ($0.60 < \kappa < 0.80$). The classification with the remaining slices presented a weak level of agreement ($\kappa < 0.50$) [213] with the medical diagnosis of PPMI.

Moreover, the analysis of the ROC curve shows that using MRI slices 27_28_29 and SPECT slices 39_40_41 the classifier had an outstanding ability to discriminate PD from Controls [214], according to Hosmer and Lemeshow, presenting AUC = 0.99 and AUC = 0.98, respectively. Unlike these slices, the AUC of the SPECT slices 27_28_29 and MRI slices 39_40_41 shows that the CNN was not able to discriminate PD from Control.

Concerning the anatomically and physiologically point-of-view, the classification of PD and Control show that the slices that achieved higher accuracy and AUC were those that comprise the mesencephalon or the basal ganglia, which are, according to the literature, regions of interest in PD [1]. Moreover, the results of the MRI slices 27_28_29 classification are in accordance with the studies conducted by Minati *et al.* and Kwon *et al.* which suggest that *substantia nigra* of PD suffer changes that are detectable by MRI [89], [90]. Regarding the results of the SPECT slices 39_40_41 classification, these are in accordance with the literature that suggests that PD patients had alterations in the basal ganglia due to the dopaminergic deficit [1], [3] that it is detected using DaTscan SPECT scans [123].

In Table 5.1 it is shown the results obtained in this work in comparison with those achieved by other studies. For instance, the study conducted by Esmaeilzadeh *et al.* presented 100% accuracy in the classification of PD and Control using 3D MRI images [22], which is a higher value than the accuracy obtained using the MRI slices encompassing the mesencephalon (97.4%). However, it is important to note that in the study of Esmaeilzadeh *et al.*, there is no explicit information either if the authors performed more than one test or if they used cross-validation methods as used in this work.

Table 5.1 Comparison of studies that performed classification of Control vs PD

Authors	Control vs PD (test sets results*)	Acc (%)	Sen (%)	Spe (%)	AUC (%)
Esmaeilzadeh et al [22]	- MRI 3D images - Datasets: (85%,10%,5%) Train set: 1040 images; Validation set: 120 images; Test set: 56 images - Subjects: 452 PD: 292M Age: 63.3±9.8 years old 160F, Age: 61.9±9.9 years old 204 Control: 134M, Age: 61.7±10.9 years old 70F, Age: 59.2±11.6 years old - Data Augmentation - Algorithm: 3D CNN + Dropout + group normalization + Demographic information about the subjects (age and sex)	100	NA	NA	1.00
Singh et al. [169]	MRI images - Subjects: 518 PD: 346M/172F, Age: 61.8±9.6 years old 245 Control: 155M, 90F, Age: 60.1±11.4 years old - Algorithms: SOM for Feature Extraction (features: intensity) LSSVM for training and classification (10 folds) It was performed two cases: 1. Age Unrelated Groups 2. Age Related Groups	 87.4 97.2	 71.9 93.4	 94.7 98.9	 NA NA
Amoroso et al. [215]	MRI images - Datasets: NA - Subjects: 374 PD: 243M/131F, Age: 61.6±9.8 years old 169 Control:	93.0	93.0	92.0	0.92

	107M, 62F, Age: 60±11.5 years old - Algorithms: Random Forest for Feature Extraction (features: connectivity measures + clinical data) SVM for classification (10 fold-cross validation)				
Choi et al. [25]	SPECT DaTscan 3D images - Subjects: 374 PD: 245M/134F, Age: 61.5±9.9 years old 170 Control: 112M, 58F, Age: 60.9±11.5 years old - Data Sets: Train/Validation sets: 379 PD/ 170 Control Test set: 52 PD/ 23 NC - Algorithms:				
	1. 3D CNN	96.0	94.2	100	0.988
	2. Visual interpretation	87.4	97.2	65.2	0.81
	3. Visual interpretation + conventional quantification (regional DaT binding ratio of putamen/ caudate and occipital cortex)	92.0	96.2	82.6	0.92
Oliveira et al.[21]	SPECT DaTscan 3D images - Subjects: 445 PD: 287M/158F, Age: 61.6±9.8 years old 209 Control: 136M, 73F, Age: 61.8±1135 years old - Algorithms: SVM Input of SVM: Features that contains the Binding potential ratio of Striatal VOIs with cortex and occipital reference VOIs (20-fold cross-validation)	97.86	97.75	98.09	NA
Martinez-Murcia et al. [153]	SPECT DaTscan 3D images - Subjects: 158 PD	95.5	96.1	94.5	NA

	111 Control - Data Augmentation - Algorithms: 3D CNN CNN input: portions of the images containing the striatum (10-fold cross validation)				
The present study	MRI and SPECT slices - Datasets: 2:1:1 - train, validation and test Train set: 84 PD/ 84 Control Validation set: 42 PD/ 42 Control Test set: 42 PD/ 42 Control - Subjects: 378 PD 242M, 136F, Age: 62.0±10.0 168 Control 109M, 59F, Age: 60.0±11.0 years old - Algorithms: 2D CNN (5 batches: Monte Carlo cross-validation) 2 cases:				
	MRI mesencephalon slices	97.4	100.0	94.8	0.99
	SPECT basal ganglia slices	92.4	92.8	91.9	0.98

Acc: Accuracy, Sen: Sensitivity, Spe: Specificity; AUC: Area Under the Curve; VOIs: Voxels of interest, Age is mean ±standard deviation. M: Male; F: Female.

Furthermore, the present work used balanced datasets to prevent biased results unlike the study mentioned above that did not specified if the datasets contains the same number of PD and Controls subjects. A drawback of this work in comparison with the study of Esmaeilzadeh *et al.* is the fact that the size of the train set is smaller than the 1040 images used in the 3D CNN. These images were obtained using data augmentation - flip technique [22] - and in this work these techniques were not applied.

Although the CNN was capable of classifying PD and Controls using MRI slices 27_28_29, this classifier was not able to discriminate those groups using the MRI slices 39_40_41 comprise the basal ganglia. This fact does not agree with the classifier in the study of Esmaeilzadeh *et al.* and the SOM algorithm in Singh *et al.* study that considered this region as an important feature for the classification [22] when MRI images are used. However, these supports the evidence of changes in PD SPECT slices 39_40_41.

As it is shown in Table 5.1, unlike this work, the study conducted by Singh *et al.* used a Feature Extraction algorithm. In this case, the SOM algorithm extract features based on the variations in the intensity of the PD images in comparison with Controls as well as differences between voxels [187] to create a vector with regions of interest.

In the comparison of Control vs PD, the present work achieved higher accuracy than the classification using the group of Age Unrelated Groups (AUG) (87.4%) in the Singh *et al.* study. In the classification of AUG, the authors extracted as features the variations in the intensity of the grey matter images, white matter images and original images [169]. Moreover, the PPV (95.07%) and NPV (100.0%) were also higher in this study in comparison with Singh *et al.* (PPV = 87.5 and NPV = 86.5) [169]. In the classification of the Age-Related Groups (ARG), the features were extracted from the difference in the average intensity in each age group. In comparison with this case, the CNN using MRI slices 27_28_29 presented the same accuracy result as the SVM classifier. However, the specificity is lower and differ in approximately 4% but has higher sensitivity. In this study, the NPV was higher but the PPV was lower comparatively with the that study (NPV = 97.3% and PPV = 97.2%) [169]. It is also important to notice that the Singh *et al.* used an unbalance sample (245 Controls and 518 PD patients), unlike this work, which may lead to biased results.

In comparison with Amoroso *et al.* study, the CNN using MRI slices 27_28_29 presented higher accuracy (97.4%) and AUC (0.99), as shown in Table 5.1. Although the number of batches used in that study is higher (10 fold-cross validation), they used an unbalanced data set (169 Controls and 374 PD patients) that may influence the results.

The present study uses CNN to extract characteristics such as the textures of the input image in contrast to the Amoroso *et al.* study that uses Random Forests to extract features based on connectivity measures [215]. These connectivity measures using the Pearson's correlation were obtained from a network model of the brain regions based on T1-W 3D MRI images [215]. The present work considered the mesencephalon and the basal ganglia as important regions of interest unlike the Random Forests that considered the frontal, occipital and temporal [215] as signification regions and gave less importance to the mesencephalon.

The classification using the SPECT slices of the basal ganglia obtained a slightly lower accuracy test (92.4%) than the accuracy achieved by Choi *et al.* (96.0%) using 3D CNN [25], Martinez-Murcia *et al.* (95.5%) [151] and Oliveira et al. (97.86%) [21]. However, the same SPECT slices present higher accuracy and Cohen's Kappa ($\kappa = 0.85 \pm 0.06$) than the accuracy of the visual interpretation by experts and by the conventional techniques calculated in Choi *et al.* study ($\kappa = 0.65 \pm 0.11$) [21].

In comparison with the results of the MRI slices contains the mesencephalon, these achieved higher accuracy and sensitivity than the studies mentioned in Table 5.1 concerning classification if Control and PD using SPECT but lower specificity (94.8%) in comparison with Oliveira *et al.* study (98.09%) [21]. However, it is also important to notice that although Oliveira *et al.* and Martinez-Murcia *et al.* studies used cross-validation methods, the data sample used in the classification is unbalance, since the number of PD patients is higher than the number of Controls subjects.

In this study, the features were extracted from the entire SPECT slices. However, instead of all image, Martinez-Murcia *et al.* used only portions of the image that contained volumes of interest (VOIs), namely the striatum, which obliges the use of another pre-processing step for segmentation. Despite this difference, the accuracy results in comparison with this study just had approximately 2% of difference.

In the study of Oliveira *et al.*, the features were not extracted from a supervised algorithm as in this work. Instead, the authors calculated the binding potential ratio of the striatum by taking the occipital and the cortex (excluding the basal ganglia) as a reference. Besides, in comparison with the present work and the other studies in Table 5.1, the Oliveira *et al.* study present the highest accuracy achieved.

In Control vs SWEDD, although the classification using MRI slices presented higher results than SPECT slices, the CNN was not able to discriminate Control from SWEDD. For instance, the accuracy was 54.7% for MRI and SPECT slices 27_28_29, 65.3% and 62.8% with for MRI and SPECT slices encompassing the basal ganglia, respectively.

Moreover, Cohen's Kappa results were less than 0.2 in all the cases which, according to McHugh, evidence that the CNN has a minimal level of agreement [213] with the PPMI database. Regarding the ROC curve, both MRI and SPECT slices 27_28_29 had AUC close to 0.50 that shows, according to Hosmer et al. [214], Control and SWEDD were not discriminated.

The results obtained with SPECT DaTscan slices are accordance with the literature that indicates that SWEDD patients presented a normal DaTscan [12], [13], [15]. In comparison with the literature, the MRI classification showed lower accuracy (63.33%), sensitivity and specificity. For instance, the study of Singh *et al.* previously mentioned also classified Control vs SWEDD and obtained 99.4% accuracy in the age related group classification and 96.4% in the age unrelated group [169], as shown in Table 5.2. However, in that article the proportion of Control and SWEDD subjects are not the same (245 Control and 68), in contrast to this work that used a balanced train set (29 Control and 29 SWEDD) and test set (15 Control and 15 SWEDD). Besides,

it is important to note that the data sample used in this classification was lower than in Control vs PD since more subjects were considered.

Concerning the classification of PD versus SWEDD, the CNN was able to discriminate those groups with 73.3% accuracy using the MRI slices of the mesencephalon and 76.00% accuracy using the SPECT slices of the basal ganglia. The highest accuracy result was obtained with SPECT slices 36_37_38 (93.3%). This result was not expected since in Control vs PD the highest accuracy result was obtained using SPECT slices of the basal ganglia. This may be related with the fact that each test set of the batches and between classification groups is different from the others because they are all result from a random selection of the data sample or because there may be something else in those slices than enable to discriminate PD vs SWEDD. Regarding the classification with the remaining slices, the CNN was not capable to discriminate PD versus SWEDD.

The results of the Cohen's Kappa indicate that the classification using slices 36_37_38 had perfect agreement level and it is the higher result obtained in this case.

Table 5.2 Comparison of studies that performed classification of Control vs SWEDD

Authors	**Control vs SWEDD (test sets results)**	**Acc (%)**	**Sen (%)**	**Spe (%)**	**AUC (%)**
Singh et al. [169]	MRI images - Subjects: 245 Control: 155M, 90F, Age: 60.09±11.35 68 SWEDD: 48M/20F, Age: 61.53±9.59 - Algorithms: SOM for Feature Extraction (features: intensity) LSSVM for training and classification (10 folds) It was performed two cases:				
	1. Age Unrelated Groups	96.4	98.8	88.1	NA
	2. Age Related Groups	99.4	1	97.2	NA
The present study	MRI and SPECT slices - Datasets: 2:1:1 - train, validation and test Train set: 29 PD/ 29 Control Validation set: 14 PD/ 14 Control Test set: 15 PD/ 15 Control				

	- Subjects: 168 Control 109M, 59F, Age: 60.0±11.0 years old 58 SWEDD 38M, 20F, Age: 62±10.0 years old - Algorithms: 3D CNN (5 batches: Monte Carlo cross-validation)				
	MRI mesencephalon slices	54.7	58.7	50.67	0.57
	SPECT basal ganglia slices	65.33	68.0	57.3	0.46

Acc: Accuracy, Sen: Sensitivity, Spe: Specificity; AUC: Area Under the Curve; VOIs: Voxels of interest

The results of the Cohen's Kappa indicate that the classification using SPECT slices 36_37_38 had, according to McHugh, perfect agreement level ($\kappa = 0.87$)[213] with the PPMI database. Using the SPECT slices of the basal ganglia, the results had a moderate level of agreement. However, the MRI slices of the mesencephalon lead to a weak level of agreement.

The results of the ROC curve indicate that the classifier using SPECT DaTscan slices of the basal ganglia present higher AUC ($\kappa = 0.72$) than MRI basal ganglia slices ($\kappa = 0.47$) and consequently are greater discriminator.

Physiologically, the SPECT results are in accordance with Marek et al. study that suggests that SWEDD patients do not have PD [216] since these patients did not present dopaminergic deficit [1], [12]. The results of the MRI slices suggest that the PD slices containing the mesencephalon are different from the SWEED slices.

In comparison with Singh *et al.* study that also classified PD and SWEDD, the accuracy of the MRI slices was lower (73.33%) [169] as shown in Table 5.3. However, the authors did not use a balanced sample in the training set, which may have led to biased estimations (518 PD and 68 SWEDD) [169].

To summarize, the higher accuracy results as well as higher AUC common in the three classification groups was obtained using MRI slices 27_28_29 that correspond to mesencephalon and SPECT slices 39_40_41 that correspond to basal ganglia. Differences between batches or between classification slices results may be related with the fact that the test set was always created randomly for each batch.

Moreover, unlike some studies expressed in Table 5.1 that use 3D images, the approach with 2D slices used in this work has the advantage that can be reproduced in clinical practice since the physicians make usage of 2D MRI images to aid in the diagnosis of brain diseases [27]. Besides, another benefit of using 2D slices, is the fact these images need smaller scanning (acquisition)

times than 3D images [194]. However, the disadvantage of the 2D slices usage is the fact these slices may contain less information than the volume of interest in 3D images.

Table 5.3 Comparison of studies that performed classification of PD vs SWEDD

Authors	**PD vs SWEDD (test sets results)**	**Acc (%)**	**Sen (%)**	**Spe (%)**	**AUC (%)**
Singh et al. [169]	MRI images - Subjects: 518 PD: 346M/172F, Age: 61.79±9.58 68 SWEDD: 48M/20F, Age: 61.53±9.59 - Algorithms: SOM for Feature Extraction (features: intensity) LSSVM for training and classification (10 folds) It was performed two cases:				
	1. Age Unrelated Groups	94.6	99.2	60.7	NA
	2. Age Related Groups	98.9	99.5	95.7	NA
The present study	MRI and SPECT slices - Datasets: 2:1:1 - train, validation and test Train set: 29 PD/ 29 Control Validation set: 14 PD/ 14 Control Test set: 15 PD/ 15 Control - Subjects: 378 PD 242M, 136F, Age: 62.0±10.0 58 SWEDD 38M, 20F, Age: 62±10.0 years old - Algorithms: 2D CNN (5 batches: Monte Carlo cross-validation) 2 cases:				
	MRI mesecenphalon slices	73.3	58.7	50.7	0.84
	SPECT basal ganglia slices	86.0	68.0	57.3	0.90

Acc: Accuracy, Sen: Sensitivity, Spe: Specificity; AUC: Area Under the Curve; VOIs: Voxels of interest

Concerning the CNN performance, unlike 3D images, the usage of 2D slices has the advantages of requiring less computing power and memory usage and training the classification models is less time-consuming.

Singh *et al.* [167], Oliveira *et al.* [21] and Amoroso *et al.* [215] performed a classification based on SVM algorithm which achieved similar results to those obtained in this work with CNN, which indicates that both these approaches leads to higher accuracy results. Moreover, the feature extraction using other algorithm may also be time consuming. The drawback of the CNN is the fact that to training the studies typically are used several images. For instance, Lecun *et al.* [138] used 60,000 images to train the CNN, but it is also important to notice that the number of classes was 9 since it was intended to classify nine digits in contrast to this study that only was classified two classes.

Thus, the main limitation of this dissertation was related with data sample, namely, the number of SWEDD subjects. PPMI database only have 58 SWEDD patients with both MRI and SPECT images. These 58 SWEDD subjects are significantly lower than the 169 Control subjects and 378 PD-patients used. As described in the methodology chapter, the number of subjects used in training the classification models is based on the number of subjects of the limiting group, that is, the group with less subjects in order to guarantee that in the classification step the same number of individuals of each group is analyzed. Because of that, only 58 Control and 58 PD-patients were considered in the Control vs SWEDD and PD vs SWEDD classification.

Chapter 6 Conclusion and Future Work

The main goal of this dissertation is to classify PD patients, SWEDD patients and Control (healthy) subjects using an approach based on MRI and SPECT slices and CNN. The results obtained allow to conclude that this objective was partially achieved since the CNN was able to distinguish PD from Controls and PD from SWEDD, but not Control from SWEDD.

The discrimination of PD from Controls and PD from SWEDD was reached using specific slices that encompassed regions of interest associated with PD according to the literature. In particular, were MRI slices included the mesencephalon and SPECT slices included the basal ganglia, which are regions of interest known to be related with the dopamine deficit that PD-patients present. Moreover, according to the ROC curve results, these regions also lead to higher accuracy results in comparison with slices including other brain structures. However, unlike the expected, the MRI slices of the basal ganglia presented lower accuracy which may be related to the dataset and the number of subjects used.

Regarding Control vs SWEDD, the CNN was not able to discriminate those subjects using SPECT slices which are in accordance with the literature wherein studies indicates that SWEDD patients have normal DaTscan. Furthermore, the findings using MRI slices of the mesencephalon shows there are no differences in Control slices and SWEDD slices.

The classification of PD vs SWEDD suggests that SWEDD patients differ from PD patients in the fact that they had different MRI mesencephalon and different SPECT basal ganglia slices imaging patterns.

Thus, the results obtained in this work suggest the fact that CNN is a useful tool for medical imaging classification based on imaging patterns.

In comparison with other studies that make use of 3D images, this work has the advantage of using 2D slices which are the norm in clinical use and which are faster to acquire. Thus, this approach enables the usage of 2D images for diagnosis. Moreover, this also makes it possible to review the diagnosis using previously acquired data. The classification using the combination of 2D images and CNN should in principle be simpler and faster in terms of computational effort enabling faster outputs obtained at the imaging console or Picture Archiving and Communication System (PACS).

However, this methodology still needs improvement to be applied in the clinical environment. For instance, a lot of images were left out in the classification since it was used a balanced data

sample in the data sets. Additionally, data augmentation techniques can also be applied, such as mirroring, translations, and rotations, to increase the train set size.

Furthermore, the classification with the axial anatomical view opens the way to the classification of the sagittal and coronal views as well.

Moreover, the CNN performance can also be tested with the implementation of other loss and regularization functions, that were mentioned in subchapter 2.6. Regarding CNN architecture, this can be modified to accept more than three slices as input to perform, for instance, a multi-modality classification.

To conclude, the approach proposed in this dissertation may be considered a promising and an innovative method for PD and SWEDD classification as well as other diseases based since it not only gives local information but also it indicates which medical imaging technique has higher accuracy.

Bibliography

[1] Kasper DL, Fauci AS, Longo DL, et al. Harrison's Principles of Internal Medicine, 19th ed. New York: McGraw Hill Education; 2015.

[2] Ferirnho P, Bugalho M, Migeul JP. For better health in Europe. Report with the support of the European Commission. Fundação Merck Sharp & Dohme. 2004.

[3] Williams DR, Litvan I. Parkinsonian syndromes. Continuum (Minneap Minn). 2013; 19(5): 1189-212.

[4] Pahwa R, Lyons K. Handbook of Parkinson's disease.4th ed. CRC Press. 2007.

[5] World Health Organization. Neurological disorders: public health challenges. Geneva: World Health Organization. 2006.

[6] Ferreira JJ, Gonçalves N, Valadas A, et al. Prevalence of Parkinson's disease: a population-based study in Portugal. Eur J Neurol. 2017; 24(5):748-50.

[7] Cummings JL, Henchcliffe C, Schaier S et al.. The role of dopaminergic imaging in patients with symptoms of dopaminergic system neurodegeneration. Brain. 2011; 134(Pt 11):3146-66.

[8] Bajaj N, Hauser RA, Grachev ID. Clinical utility of dopamine transporter single photon emission CT (DaT-SPECT) with (123I) ioflupane in diagnosis of parkinsonian syndromes. J. Neurol. Neurosurg. Psychiatry. 2013; 84(11). 1288–95.

[9] Yu F, Barron DS, Tantiwongkosi, P. Fox P, Patterns of gray matter atrophy in atypical parkinsonism syndromes: A VBM meta-analysis. Brain Behav. 2015; 5(6):e00329.

[10] Hughes AJ, Daniel SE, Kilford L. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry. 1992;55(3):181-4.

[11] Hoehn MM, Yahr MD. Parkinsonism: onset, progression and mortality. Neurology. 1967;17(5):427-42.

[12] Brooks DJ. Imaging Approaches to Parkinson Disease. J Nucl Med. 2010;51(4):596-609.

[13] Dierckx RAJO, Otte, A, de Vries EFJ, van Waarde A, Leenders KL. PET and SPECT in Neurology. (Eds.) Berlin, Heidelberg: Springer; 2014.

[14] Meijer FJA, Goraj B, Bloem BR, et al. Clinical Application of Brain MRI in the Diagnostic

Work-up of Parkinsonism. J Parkinsons Dis. 2017; 7(2): 211–17.

[15] Bajaj N. SWEDD for the General Neurologist. 2010; Acnr.2010;10(4):30-1

[16] Oh YS, Choi JH, Kwon DY. Classification of Scans Without Evidence of Dopamine Deficit (SWEDD) According to the Olfactory Function. J Parkinsons Dis. 2016; 19;6(4):771-8.

[17] Silveira-Moriyama L, Schwingenschuh P, O'Donnell A, et al. Olfaction in patients with suspected parkinsonism and scans without evidence of dopaminergic deficit (SWEDDs). J Neurol Neurosurg Psychiatry. 2009;80(7):744-8.

[18] Schneider SA, Edwards MJ, Mir P, et al. Patients with adult-onset dystonic tremor resembling Parkinsonian tremor have scans without evidence of dopaminergic deficit (SWEDDs). Mov Disord. 2007;22(15):2210-5.

[19] Darcourt J, Booij J, Tatsch K et al.. EANM procedure guidelines for brain neurotransmission SPECT using123I-labelled dopamine transporter ligands, version 2. Eur J Nucl Med Mol Imaging. 2010;37(2):443-50.

[20] Tagare HD, DeLorenzo C, Chelikani S et al.. Voxel-based logistic analysis of PPMI control and Parkinson's disease DaTscans. Neuroimage Neuroimage. 2017;152:299-311.

[21] Oliveira FP, Castelo-Branco M. Computer-aided diagnosis of Parkinson's disease based on [123 I]FP-CIT SPECT binding potential images, using the voxels-as-features approach and support vector machines. J Neural Eng. 2015;12(2):026008.

[22] Esmaeilzadeh S, Yang Y, Adeli E. End-to-End Parkinson Disease Diagnosis using Brain MR-Images by 3D-CNN. 2018. (eds) Machine Learning in Medical Imaging. Lecture Notes in Computer Science. Springer, Cham. 2018; 11046.

[23] Chollet F. Deep Learning with Python. 1st ed. United Stated of America: Manning Publications Co. 2017.

[24] Liu M, Cheng D, Wang K, et al. Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. Neuroinformatics. 2018 ;16(3-4):295-308.

[25] Choi H, Ha S, Im HJ, et al. Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. Neuroimage Clin. 2017;16:586-594.

[26] Heim B, Krismer F, De Marzi R et al.. Magnetic resonance imaging for the diagnosis of

Parkinson's disease. J Neural Transm (Vienna). 2017;124(8):915-964.

[27] Loizou CP, Papacharalambous C, Samaras G. Brain Image and Lesions Registration and 3D Reconstruction in Dicom MRI Images. Proc. - IEEE Symp. Comput. Med. Syst. 2017; 1:419–22.

[28] Manyam BV. Paralysis agitans and levodopa in 'Ayurveda': Ancient Indian medical treatise. Mov Disord. 1990;5(1):47-8.

[29] Goetz CG. The History of Parkinson's Disease: Early Clinical Descriptions and Neurological Therapies. Cold Spring Harb Perspect Med. 2011; 1(1):1–16.

[30] C. G. Goetz, "The History of Parkinson's Disease: Early Clinical Descriptions and Neurological Therapies," Cold Spring Harb. Perspect. Med., vol. 1, no. 1, pp. 1–16, 2011.

[31] Parkinson J. An Essay on the Shaking Palsy. 1817. J Neuropsychiatry Clin Neurosci. 2002;14(2):223-36.

[32] Walusinski O. Jean-Martin Charcot's lessons on Parkinson's disease. Revue Neurologique 2017; 173(10):604-5.

[33] Greenfield JG, Bosanquet FD. The brain-stem lesions in Parkinsonism. J Neurol Neurosurg Psychiatry. 1953;16(4):213-26.

[34] Charcot JM. Lectures on the diseases of the nervous system: delivered at la Salpêtrière. London: The New Sydenham Society.1879.

[35] Simuni T. Is It Essential Tremor or a Parkinsonian Syndrome? Diagnostic Considerations in Primary Care. Medscape, LLC, [Internet]. 2014. [cited 2018 Aug15] Available from: https://www.medscape.org/viewarticle/805072_transcript.

[36] Bear M, Connors B, Paradiso M. Neuroscience: Exploring the Brain, 3rd ed. Philadelphia, PA, US: Lippincott Williams & Wilkins Publishers; 2007.

[37] Caldas AC. A herança de Franz Joseph Gall : o cérebro ao serviço do comportamento humano. Amadora: McGraw-Hill; 2000.

[38] Marieb EN, Hoehn K. Human Anatomy & Physiology. 9th ed. Pearson College DivPublication; 2012.

[39] Braak H, Del Tredici K, Rüb U, et al. Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging. 2003;24(2):197-211.

[40] Miller R. A theory of the basal ganglia and their disorders. Boca Raton: CRC Press, 2008.

[41] Soghomonian JJ. The basal ganglia : novel perspectives on motor and cognitive functions. 1st ed. Springer; 2016.

[42] Mink JW. The basal ganglia : focused selection and inhibition of competing motor programs Prog Neurobiol. 1996;50(4):381-425.

[43] Fearnley MJ, Lees AJ. Ageing and Parkinson's disease: substantia nigra regional selectivity. Brain. 1991;114(Pt 5):2283-301.

[44] Forno LS. Neuropathology of Parkinson's disease. J Neuropathol Exp Neurol. 1996;55(3):259-72.

[45] Petrucelli L, Dickson DW. Neuropathology of Parkinson's Disease. Parkinsons. Cold Spring Harb Perspect Med. 2012; 2(8): a009258.

[46] Fahn S, Oakes D, Shoulson I. Levodopa and the Progression of Parkinson ' s Disease. JN Engl J Med. 2004;351(24):2498-508.

[47] A. Q. Rana and J. Depradine, "Abdominal pain: a symptom of levodopa end of dose wearing off in Parkinson's disease.," West Indian Med. J., vol. 60, no. 2, pp. 223–4, Mar. 2011.

[48] Jankovic J, Tintner R. Dystonia and parkinsonism. Parkinsonism Relat Disord. 2001;8(2):109-21.

[49] Jinnah HA, Factor SA. Diagnosis & treatment of dystonia. Neurol Clin. 2015; 33(1): 77–100.

[50] Gilman S, Wenning GK, Low AP, et al. Second consensus statement on the diagnosis of multiple system atrophy. Neurology. 2008; 71(9): 670–76.

[51] McKay JH, Cheshire WP. First symptoms in multiple system atrophy. Clin Auton Res. 2018;28(2):215-221.

[52] Buijs RM. The autonomic nervous system: a balancing act. Handb Clin Neurol. 2013;117:1-11.

[53] Hall WD. An Overview of the Autonomic Nervous System. In: Walker HK, Hall WD, Hurst JW, editors. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths; 1990. Available from: https://www.ncbi.nlm.nih.gov/books/NBK399/

[54] Ling H. Clinical Approach to Progressive Supranuclear Palsy. J Mov Disord. 2016; 9(1): 3–13.

[55] Litvan I, Hauw JJ, Bartko JJ, et al. Validity and reliability of the preliminary NINDS neuropathologic criteria for progressive supranuclear palsy and related disorders. J Neuropathol Exp Neurol. 1996 Jan;55(1):97-105.

[56] Morris HR, Wood NW, Lees AJ. Progressive supranuclear palsy (Steele-Richardson-Olszewski disease). Postgraduate Medical Journal 1999;75:579-84.

[57] Williams DR, de Silva R, Paviour DC. Characteristics of two distinct clinical phenotypes in pathologically proven progressive supranuclear palsy: Richardson's syndrome and PSP-parkinsonism. Brain. 2005;128(Pt 6):1247-58.

[58] National Institute of Neurological Disorders and Stroke. Corticobasal Degeneration Information Page [Internet]. National Institute of Neurological Disorders and Stroke. 2018. [Cited: 2018 Aug 12]. Available from: https://www.ninds.nih.gov/Disorders/All-Disorders/Corticobasal-Degeneration-Information-Page

[59] Lill CM, Klein C. Epidemiologie und Ursachen der Parkinson-Erkrankung. Nervenarzt. 2017; 88(4): 345–55.

[60] Hwang O. Role of oxidative stress in Parkinson's disease. Exp Neurobiol. 2013; 22(1): 11–17.

[61] Lionnet A, Leclair-Visonneau L, Neunlist, et al. Does Parkinson's disease start in the gut?. Acta Neuropathol. 2018;135(1):1-12.

[62] European Medicines Agency. Questions and answers on additional clarification for inclusion criteria in the "Guideline on clinical investigation of medicinal products in the treatment of Parkinson's disease" rev. 2. [Internet] European Medicines Agency. 2012. [Cited: 2018 Aug 12]. Available from: https://www.ema.europa.eu/documents/scientific-guideline/questions-answers-additional-clarification-inclusion-criteria-guideline-clinical-investigation_en.pdf

[63] Elble RJ. What is essential tremor?. Curr. Neurol. Neurosci. Rep. 2013; 13(6):353.

[64] Kieburtz K. Movement Disorders: Neurologic Principles and Practice. 1998;50 (4):1199;

[65] Litvan I, Bhatia KP, Burn DJ, et al. Movement Disorders Society Scientific Issues Committee Report SIC Task Force Appraisal of Clinical Diagnostic Criteria for Parkinsonian Disorders. Mov Disord. 2003;18(5):467-86.

[66] Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations. Mov Disord. 2004; 19(9):1020-8.

[67] Poewe W. Non-motor symptoms in Parkinson's disease. Eur J Neurol. 2008;15 Suppl 1:14-20.

[68] Hawkes CH, Shephard BC, Daniel SE. Olfactory dysfunction in Parkinson's disease. JJ Neurol Neurosurg Psychiatry. 1997; 62(5): 436–46.

[69] C. H. Hawkes, B. C. Shephard, and S. E. Daniel, "Is Parkinson's disease a primary olfactory disorder?," QJM, vol. 92, no. 8, pp. 473–80, Aug. 1999.

[70] M. W. Geisler and C. Murphy, "Event-related brain potentials to attended and ignored olfactory and trigeminal stimuli.," Int. J. Psychophysiol., vol. 37, no. 3, pp. 309–15, Sep. 2000.

[71] Goetz CG, Lutge W, Tanner CM. Autonomic dysfunction in Parkinson's disease. Neurology. 1986;36 (1):73.

[72] Goetz CG, Fahn S, Martinez-Martin P. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. Mov Disord. 2007;22(1):41-7.

[73] Goetz CG. Movement Disorder Society-Unified Parkinson's Disease Rating Scale (MDS-UPDRS) : une nouvelle échelle pour l'évaluation de la maladie de Parkinson," Rev. Neurol. Rev Neurol (Paris). 2010 Jan;166(1):1-4.

[74] National Collaborating Centre for Chronic Conditions (UK). Parkinson's Disease: National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care. [Internet] London: Royal College of Physicians (UK); 2006. [Cited: 12-Aug-2018] (NICE Clinical Guidelines, No. 35.) Available from: https://www.ncbi.nlm.nih.gov/books/NBK48513/

[75] Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. JJ Am Geriatr Soc. 2005;53(4):695-9.

[76] Federico A, Tinazzi M, Tamburin S. MoCA for cognitive screening in Parkinson's disease: Beware of floor effect. Mov Disord. 2018;33(3):499.

[77] Hoehn MM, Yahr MD. Parkinsonism : onset, progression, and mortality. Neurology. 1967;17(5):427-42.

[78] Quinn N. Multiple system atrophy--the nature of the beast. J Neurol Neurosurg Psychiatry. 1989;Suppl:78-89.

[79] Riley DE, Lang AE, Lewis A, et al. Cortical-basal ganglionic degeneration., Neurology.

1990;40(8):1203-12.

[80] Kägi G, Bhatia KP, Tolosa E. The role of DAT-SPECT in movement disorders. J Neurol Neurosurg Psychiatry. 2010;81(1):5-12.

[81] Saeed U, Compagnone J, Aviv RI, et al. Imaging biomarkers in Parkinson's disease and Parkinsonian syndromes: current and emerging concepts. Transl Neurodegener. 2017; 28;6:8.

[82] Huettel SA, Song AW, McCarthy G. Functional Magnetic Resonance Imaging. 2nd ed. Mass: Sinauer Associates; 2008.

[83] McRobbie D, Moore E, Graves M , et al. MRI from picture to proton. Cambridge: Cambridge University Press; 2006.

[84] Eidelberg D. Imaging in Parkinson's disease. 1st ed. New York: Oxford University Press; 2011.

[85] Pyatigorskaya N, Gallea C, Garcia-Lorenzo D, et al. A review of the use of magnetic resonance imaging in Parkinson's disease. Ther. Adv. Neurol. Disord. 2014; 7(4):206–20.

[86] Pitcher TL, Melzer TR, Macaskill MR, et al. Reduced striatal volumes in Parkinson's disease: a magnetic resonance imaging study. Transl Neurodegener. 2012;1(1):17.

[87] Oikawa H, Sasaki M, Tamakawa Y, et al. The substantia nigra in Parkinson disease: proton density-weighted spin-echo and fast short inversion time inversion-recovery MR findings. AJNR Am J Neuroradiol. 2002;23(10):1747-56.

[88] Martin WR, Wieler M, Gee M. Midbrain iron content in early Parkinson disease: A potential biomarker of disease status. Neurology. 2008;70(16 Pt 2):1411-7.

[89] Minati L, Grisoli M, Carella F, et al. Imaging degeneration of the substantia nigra in Parkinson disease with inversion-recovery MR imaging. AJNR Am J Neuroradiol. 2007;28(2):309-13.

[90] Kwon DH, Kim JM, Oh SH, et al., even-Tesla magnetic resonance images of the substantia nigra in Parkinson disease. Ann Neurol. 2012 Feb;71(2):267-77.

[91] Ashburner J, Friston KJ. Voxel-Based Morphometry—The Methods. Neuroimage. 2000;11(6 Pt 1):805-21.

[92] Burton EJ, McKeith IG, Burn DJ, et al. Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with Lewy bodies and controls. Eur J Radiol. 2014;83(3):564-70.

[93] Chen S, Tan HY, Wu ZH, et al. Imaging of olfactory bulb and gray matter volumes in brain areas associated with olfactory function in patients with Parkinson's disease and multiple system atrophy. Eur J Radiol. 2014;83(3):564-70.

[94] Summerfield C, Junqué C, Tolosa E. Structural Brain Changes in Parkinson Disease With Dementia. Arch Neurol. 2005;62(2):281-5.

[95] Geng DY, Li YX, Zee CS. Magnetic Resonance Imaging-Based Volumetric Analysis of Basal Ganglia Nuclei and Substantia Nigra in Patients with Parkinson's Disease. Neurosurgery. 2006;58(2):256-62.

[96] Tinaz S1, Courtney MG, Stern CE. Focal cortical and subcortical atrophy in early Parkinson's disease. Mov Disord. 2011;26(3):436-41.

[97] Schulz JB, Skalej M, Wedekind D, et al. Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. Ann Neurol. 1999;45(1):65-74.

[98] Brooks DJ. Morphological and functional imaging studies on the diagnosis and progression of Parkinson's disease. J Neurol. 2000;247 Suppl 2:II11-8.

[99] Pagano G, Niccolini F, Politis M. Imaging in Parkinson's diseaseClin Med (Lond). 2016;16(4):371-5.

[100] Seppi K, Poewe W. Brain Magnetic Resonance Imaging Techniques in the Diagnosis of Parkinsonian Syndromes. Neuroimaging Clin N Am. 2010 ;20(1):29-55.

[101] Schrag A, Good CD, Miszkiel K. Differentiation of atypical parkinsonian syndromes with routine MRI. Neurology. 2000;54(3):697-702.

[102] Yekhlef F, Ballan G, Macia F. Routine MRI for the differential diagnosis of Parkinson's disease, MSA, PSP, and CBD. J Neural Transm (Vienna). 2003;110(2):151-69.

[103] Sako W, Murakami N, Izumi Y. The difference in putamen volume between MSA and PD: Evidence from a meta-analysis. Parkinsonism Relat Disord. 2014;20(8):873-7.

[104] Gröschel K, Hauser TK, Luft A, et al. Magnetic resonance imaging-based volumetry differentiates progressive supranuclear palsy from corticobasal degeneration. Neuroimage. 2004;21(2):714-24.

[105] Massey LA, Micallef C, Paviour DC, et al. Conventional magnetic resonance imaging in confirmed progressive supranuclear palsy and multiple system atrophy. Mov Disord. 2012;27(14):1754-62.

[106] Bronzin JD. The biomedical engineering handbook. 3rd ed. Volume 3. Electrical Engineering Handbook. Taylor & Francis; 2006.

[107] Beaulieu JM, Gainetdinov RR. The physiology, signaling, and pharmacology of dopamine receptors. Pharmacol Rev. 2011;63(1):182-217.

[108] European Medicines Agency. Resumo do EPAR destinado ao público - DaTSCAN ioflupano (123I). [Internet] European Medicines Agency. 2010. [cited 2018 Agu 12]. Available from: https://www.ema.europa.eu/documents/overview/datscan-epar-summary-public_pt.pdf

[109] van Dyck CH, Seiby JP, Robert T, et al. Age-related decline in dopamine transporters: Analysis of striatal subregions, nonlinear effects, and hemispheric asymmetries. Am. J. Geriatr. Psychiatry. 2002;10(1):36–43.

[110] Varrone A, Dickson JC, Tossici-Bolt L, et al. European multicentre database of healthy controls for [123I]FP- CIT SPECT (ENC-DAT): Age-related effects, gender differences and evaluation of different methods of analysis. Eur J Nucl Med Mol Imaging. 2013;40(2):213-27.

[111] Kaasinen V, Joutsa J, Noponen T et al. Effects of aging and gender on striatal and extrastriatal [123I]FP-CIT binding in Parkinson's disease. Neurobiol. Neurobiol Aging. 2015;36(4):1757-1763.

[112] Marshall VL, Patterson J, Hadley DM, et al. Two-year follow-up in 150 consecutive cases with normal dopamine transporter imaging. Nucl Med Commun. 2006;27(12):933-7.

[113] Benamer HT, Oertel WH, Patterson J, et al. Prospective study of presynaptic dopaminergic imaging in patients with mild parkinsonism and tremor disorders: Part 1. Baseline and 3-month observations. Mov Disord. 2003;18(9):977-84.

[114] Batla A, Erro R, Stamelou M, et al. Patients with scans without evidence of dopaminergic deficit: A long-term follow-up study. Mov Disord. 2014;29(14):1820-5.

[115] Marek K, Seibyl J, Eberly S, et al. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. Neurology. 2014;82(20):1791-7.

[116] Booij J, Knol RJ. SPECT imaging of the dopaminergic system in (premotor) Parkinson's disease. Parkinsonism Relat Disord. 2007;13 Suppl 3:S425-8.

[117] Uhl GR, Walther D, Mash D, et al. Dopamine transporter messenger RNA in Parkinson's disease and control substantia nigra neurons. Ann Neurol.1994;35(4):494-8.

[118] Kim M, Park H. Structural connectivity profile of scans without evidence of dopaminergic

deficit (SWEDD) patients compared to normal controls and Parkinson's disease patients. Springerplus. 2016;5(1):1421.

[119] Schwingenschuh P, Ruge D, Edwards MJ, et al. Distinguishing SWEDDs patients with asymmetric resting tremor from Parkinson's disease: A clinical and electrophysiological study. Mov Disord. 2010;25(5):560-9.

[120] Eckert T, Feigin A, Lewis DE, et al. Regional metabolic changes in Parkinsonian patients with normal dopaminergic imaging. Mov Disord. 2007;22(2):167-73.

[121] Marek KL, Seibyl JP, Zoghbi SS. [123I] beta-CIT/SPECT imaging demonstrates bilateral loss of dopamine transporters in hemi-Parkinson's disease. Neurology. 1996;46(1):231-7.

[122] Fuente-Fernández R. Role of DaTSCAN and clinical diagnosis in Parkinson disease. Neurology. 2012;78(10):696-701.

[123] Rahmim A, Salimpour Y, Jain S et al. Application of texture analysis to DAT SPECT imaging: Relationship to clinical assessments. Neuroimage Clin. 2016;12:e1-e9.

[124] Tatsch K, Poepperl G. Nigrostriatal Dopamine Terminal Imaging with Dopamine Transporter SPECT: an update. J Nucl Med. 2013;54(8):1331-8.

[125] Pirker W, Asenbaum S, Bencsits G, et al. [123I]beta-CIT SPECT in multiple system atrophy, progressive supranuclear palsy, and corticobasal degeneration. Mov Disord. 2000;15(6):1158-67.

[126] Varrone A, Marek KL, Jennings D, et al.[(123)I]beta-CIT SPECT imaging demonstrates reduced density of striatal dopamine transporters in Parkinson's disease and multiple system atrophy. Mov Disord. 2001;16(6):1023-32.

[127] Benamer HTS, Patterson J, Grosset DG, et al. Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group. Mov Disord. 2000;15(3):503-10.

[128] Asenbaum S, Pirker W, Angelberger P, et al. [123 I]β-CIT and SPECT in essential tremor and Parkinson's disease. J Neural Transm (Vienna). 1998;105(10-12):1213-28.

[129] Russell SJ, Norvig P. Artificial intelligence: a modern approach. Upper Saddle River, N.J.: Prentice Hall, 2010.

[130] Haugeland J.Artificial Intelligence: The Very Idea. MIT Press. 1st ed. Cambridge, Mass: 1985; 287p.

[131] Rich E, Knight K, Artificial intelligence. 2nd ed. McGraw-Hill; 1991.

[132] Goodfellow I, Bengio Y, Courville A. Deep learning. [Internet] MIT Press; 2016. [Cited: 2018 Aug 12]. Avaliable from: https://www.deeplearningbook.org/

[133] Mathworks. Introducing Deep Learning with MATLAB.[Internet] Mathworks; 2017.[Cited: 2018 Aug 12] Avaliable from: ttps://www.mathworks.com/content/dam/mathworks/tag-team/Objects/d/80879v00_Deep_Learning_ebook.pdf

[134] van Gerven M, Bohte S. Editorial: Artificial Neural Networks as Models of Neural Information Processing. Front Comput Neurosci. 2017; 11: 114.

[135] Aghdam HH, Heravi EJ. Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification. Cham, Switzerland: Springer International Publishing AG; 2017.

[136] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics. 1943; 5 (4):115-33.

[137] Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: a tutorial. Computer (Long. Beach. Calif). 1996; 29(3):31–44.

[138] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc. IEEE, 1998; 86(11): 2278–324.

[139] Lo SCB,Chan HP, Lin S, et al. Artificial convolution neural network for medical image pattern recognition. Neural Networks. 1995; 8(Issue 7-8):1201–14.

[140] Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol. 1962; 160(1): 106–154.

[141] Huberl DH, Wiesel TN. Receptive fields of single neurones in the cat's striate cortex. J Physiol. 1959; 148(3): 574–591.

[142] Hadji I, Wildes RP. What Do We Understand About Convolutional Networks?.[Internet] CoRR.2018.[cited 2018 Aug 12] Available from: https://arxiv.org/pdf/1803.08834.pdf

[143] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics. Sahiner B, Chan HP, Petrick N. Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images. IEEE Trans. Med. Imaging. 1996; 15(5):598–610.

[145] LeCun Y, Boser B, Denker JS, et al. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Comput. 1989;1(4):541–51.

[146] Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on International Conference on Machine Learning. ICML'10. 2010; 807-14.

[147] G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” no. 3, p. 6421113.

[148] Jarrett K, Kavukcuoglu K, Ranzato M, et al. What is the best multi-stage architecture for object recognition?. Proc. IEEE Int. Conf. Comput. Vis. 2009; 2146–53.

[149] Li FF, Johnson J, Yeung S. Lecture 03: Loss Functions and Optimization. [Internet]. Stanford vision and learning lab. 2017. [Cite: 2018 Aug 20] Avaliable from: http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture3.pdf

[150] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. 2014;15(1):1929–58.

[151] Montavon G, Orr G, Müller KR. Neural Networks: Tricks of the Trade. Heidelberg: Springer-Verlag Berlin. 2012.

[152] Poria S, Cambria E, Gelbukh A. Deep Convolutional Neural Network Textual Features and Multiple Kernel Learning for Utterance-level Multimodal Sentiment Analysis. Proc. 2015 Conf. Empir. Methods Nat. Lang. Process. 1995; 17-21.

[153] Martinez-Murcia FJ, Ortiz A, Górriz JM, et al. A 3D Convolutional Neural Network Approach for the Diagnosis of Parkinson’s Disease. In: Ferrández Vicente J., Álvarez-Sánchez J., de la Paz López F., Toledo Moreo J., Adeli H. (eds) Natural and Artificial Computation for Biomedicine and Neuroscience. IWINAC 2017. Lecture Notes in Computer Science. Springer, Cham. 2017; 10337:324–333.

[154] Zhang YD, Pan C, Sun J, Tang C. Multiple sclerosis identification by convolutional neural network with dropout and parametric ReLU. J. Comput. Sci. 2018; 28:1–10.

[155] Zou L, Zheng J, Miao C, McKeown M J, et al. 3D CNN Based Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Functional and Structural MRI. IEEE Access. 2017; 5:23626–36.

[156] Pereira CR, Weber SAT, Hook C, et al. Deep learning-aided Parkinson’s disease diagnosis from handwritten dynamics. Proc. - 2016 29th SIBGRAPI Conf. Graph. Patterns Images, SIBGRAPI 2016. 2016; 340–6.

[157] Karimi D, Samei G, Kesch C, et al. Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models. Int. J.

Comput. Assist. Radiol. Surg. 2018; 13(8):1211–9.

[158] Souza AM, Chen L, Wu Y, et al. MRI tumor segmentation with densely connected 3D CNN. Medical Imaging: Image Processing; 2018.

[159] Zubal IG, Early M, Yuan O, et al. Optimized, automated striatal uptake analysis applied to SPECT brain scans of Parkinson's disease patients. J. Nucl. Med.2007; 48:857–64.

[160] Kuo PH, Avery R, Krupinski E, et al. Receiver-Operating-Characteristic Analysis of an Automated Program for Analyzing Striatal Uptake of 123I-Ioflupane SPECT Images: Calibration Using Visual Reads. J J Nucl Med Technol. 2013;41(1):26-31.

[161] Kuo PH, Lei HH, Avery R. Evaluation of an Objective Striatal Analysis Program for Determining Laterality in Uptake of 123I-Ioflupane SPECT Images: Comparison to Clinical Symptoms and to Visual Reads. J Nucl Med Technol. 2014;42(2):105-8.

[162] Seibyl MJ. Baseline Neuroimaging Characteristics of the Parkinson ' s Progression Marker Initiative (PPMI) Parkinsons and Healthy Cohorts.[Internet] Institute for Neurodegenerative Disorders. [cited 2018 Aug 12] Available from: http://www.ppmi-info.org/wp-content/uploads/2012/06/Seibyl-PPMI-MDS-2012.pdf

[163] Filippi L, Manni C, Pierantozzi M et al. 123I-FP-CIT semi-quantitative SPECT detects preclinical bilateral dopaminergic deficit in early Parkinson's disease with unilateral symptoms. Nucl Med Commun. 2005;26(5):421-6.

[164] Prashanth R, Roy SD, P. K. Mandal PK, et al. Automatic classification and prediction models for early Parkinson's disease diagnosis from SPECT imaging. Expert Syst. Appl. 2014; 41(7):3333–42.

[165] Illan IA, Gorrz JM, Ramirez J, at al. Automatic assistance to Parkinsons disease diagnosis in DaTSCAN SPECT imaging. Med Phys. 2012;39(10):5971-80.

[166] Segovia F, Gorriz JM, Ramirez J, et al. Improved Parkinsonism diagnosis using a partial least squares based approach. Med Phys. 2012;39(7):4395-403.

[167] Cristianini N, Shawe-Taylor J. Support Vector Machines and other kernel-based learning methods. 1st ed. Cambridge University Press; 2000.

[168] Vapnik VN. Statistical Learning Theory. 1st ed. Wiley-Interscience. 1998.

[169] Singh G, Samavedham L. Algorithm for image-based biomarker detection for differential diagnosis of Parkinson's disease. IFAC-PapersOnLine. 2015; 48(8):918–23.

[170] Duchesne S, Rolland Y, Vérin M. Automated Computer Differential Classification in

Parkinsonian Syndromes via Pattern Analysis on MRI. Acad. Radiol. 2009;16(1):61–70.

[171] Salvatore C, Cerasa 2, Castiglioni I. Machine learning on brain MRI data for differential diagnosis of Parkinson's disease and Progressive Supranuclear Palsy. J Neurosci Methods. 2014;222:230-7.

[172] Zhang L, Liu C, Zhang X, et al. Classification of Parkinson's Disease and Essential Tremor Based on Structural MRI. in 2016 7th International Conference on Cloud Computing and Big Data (CCBD). 2016;353–56.

[173] PPMI. Parkinson's Progression Markers Initiative | Who We Are. [Internet]. Parkinson's Progression Markers Initiative; 2018. [cited 2018 sep 09] Available from: https://www.ppmi-info.org/about-ppmi/who-we-are/.

[174] Parkinson Progression Marker Initiative, et al. The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol. 2011;95(4):629-35.

[175] PPMI. Parkinson's Progression Markers Initiative | Study Cohorts. [Internet]. Parkinson's Progression Markers Initiative; 2018. [cited 2018 sep 09] Available from:http://www.ppmi-info.org/study-design/study-cohorts/

[176] Focke NK, Helms G, Scheewe S, et al. Individual voxel-based subtype prediction can differentiate progressive supranuclear palsy from idiopathic Parkinson syndrome and healthy controls. Hum Brain Mapp. 2011;32(11):1905-15.

[177] Runge VM, Wood ML, Kaufman DM. FLASH: clinical three-dimensional magnetic resonance imaging. Radiographics. 1988;8(5):947-65.

[178] Schwaighofer BW, Yu KK, Mattrey RF. Diagnostic significance of interslice gap and imaging volume in body MR imaging. AJR Am J Roentgenol. 1989;153(3):629-32.

[179] Noh Y, Sung YH, Lee J. Nigrosome 1 detection at 3T MRI for the diagnosis of early-stage idiopathic Parkinson disease: Assessment of diagnostic accuracy and agreement on imaging asymmetry and clinical laterality. AJNR Am J Neuroradiol. 2015;36(11):2010-6.

[180] Price S, Paviour D, Scahill R, et al. Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and Parkinson's disease. Neuroimage. 2004;23(2):663-9.

[181] Wood R, Bassett K, Foerster V. 1.5 Tesla Magnetic Resonance Imaging Scanners Compared with 3.0 Tesla Magnetic Resonance Imaging Scanners: Systematic Review of Clinical Effectiveness CADTH Technol Overv. 2012; 2(2): e2201.

[182] NICE. Parkinson's disease in adults. [Internet] NICE guideline [NG71]. The National Institute for Health and Care Excellence; 2017 [cited 2018 sep 09] Available from: https://www.nice.org.uk/guidance/ng71

[183] PPMI. MRI - Technical Operations Manual. [Internet] Parkinson's Progression Markers Initiative; 2015. [cited 2018 sep 09] Available from: http://www.ppmi-info.org/wp-content/uploads/2017/06/PPMI-MRI-Operations-Manual-V7.pdf

[184] PPMI. Manual. maging Technical Operations Manual - The Parkinson progression Marker Initiative [Internet] The Parkinson progression Marker Initiative;2017. [cited 2018 sep 09] Available from: http://www.ppmi-info.org/wp-content/uploads/2017/06/PPMI-TOM-V8_09-March-2017.pdf

[185] Weaver KF, Morales V, Dunn SL, et al. An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences. Hoboken, NJ: JohnWiley & Sons, Inc; 2017.

[186] Goodier J. The Cambridge Dictionary of Statistics. 4th ed. New York: Cambridge University Press; 2010.

[187] Singh G. Samavedham L. Unsupervised learning based feature extraction for differential diagnosis of neurodegenerative diseases: A case study on early-stage diagnosis of Parkinson disease. JJ Neurosci Methods. 2015;256:30-40.

[188] Cox RW, Ashburner J, Breman, et al. A (sort of) new image data format standard: NiFTI-1. 10th Annual Meeting of Organisation of Human Brain Mapping. 2016; 1: 6–7.

[189] Penny W, Friston K, Ashburner J. Statistical Parametric Mapping: The Analysis of Functional Brain Images. 1st ed. Hardbound: Academic Press; 2006.

[190] Hermans E. SPM12 starters' guide [Internet] Statistical Parametric Mapping; 2016. [cited 2018 sep 09] Available from: https://www.ernohermans.com/wp-content/uploads/2016/09/spm12_startersguide.pdf

[191] Ashburner J. Computational anatomy with the SPM software. Magn Reson Imaging. 2009;27(8):1163-74.

[192] Ashburner J, Barnes G, Chen CC, et al. SPM12 Manual The FIL Methods Group (and honorary members). [Internet] London Functional Imaging Laboratory Wellcome Trust Centre for Neuroimaging; 2013. [cited 2018 sep 09] Available from: https://www.fil.ion.ucl.ac.uk/spm/doc/spm12_manual.pdf

[193] Wisniewski G, Seibyl J, Marek K, et al. DatScan SPECT Image Processing Methods for

Calculation of Striatal Binding Ratio (SBR). [Internet] The Parkinson progression Marker Initiative;2017. [cited 2018 sep 09] Avaliable from: http://www.ppmi-info.org/wp-content/uploads/2013/06/Seibyl-PPMI-MDS-2013-Sydney__sjl.pdf

[194] Johnson G, Wadghiri YZ, Turnbull DH. 2D multislice and 3D MRI sequences are often equally sensitive. Magn Reson Med. 1999 Apr;41(4):824-8.

[195] Godinho DM. Desenvolvimento de uma aplicação com recurso à unidade de processamento gráfico para classificação de sinais de Perturbações do Espectro do Autismo. [dissertation] [Lisbon] Nova University; 2016.

[196] Purves D, Augustine GJ, Fitzpatrick D, et al., editors. Neuroscience. 2nd ed. [Internet] Sunderland (MA): Sinauer Associates; 2001. [cited 2018 sep 09] Available from: https://www.ncbi.nlm.nih.gov/books/NBK10799/

[197] Zuiderveld K, Contrast Limited Adaptive Histogram Equalization. Graphics Gems;1994.8(5):474-485p.

[198] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional Architecture for Fast Feature Embedding. MM '14 Proceedings of the 22nd ACM international conference on Multimedia. 2014;675-8p.

[199] Eklund A, Dufort P, Forsberg D, et al. Medical image processing on the GPU - Past, present and future. Med Image Anal. 2013;17(8):1073-94.

[200] NVIDIA. Placa de vídeo versus CPU? O que é computação de placas de vídeo?|NVIDIA. [Internet] NVIDIA;2018. [cited 2018 sep 10] Avaliable from: https://www.nvidia.com.br/object/what-is-gpu-computing-br.html.

[201] Developers Google. Developer Guide | Protocol Buffers | Google Developers. [Internet]. Developers Google;2018. [cited 2018 sep 11] Avaliable from: https://developers.google.com/protocol-buffers/docs/overview.

[202] LMDB. LMDB: Lightning Memory-Mapped Database Manager (LMDB). [Internet] LMDB; 2015. [cited 2018 sep 12] Avaliable from: http://www.lmdb.tech/doc/

[203] NVIDIA. GeForce GTX TITAN X Graphics Card | GeForce.” [Internet] [cited 2018 sep 12] Avaliable from: https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-titan-x

[204] Jahn A. Keras Image Preprocessing: scaling image pixels for training. [Internet] [cited 2018 sep 10] Avaliable from: Available: https://www.linkedin.com/pulse/keras-image-preprocessing-scaling-pixels-training-adwin-jahn

[205] LeCun YA, Bottou L, Orr G.B, et al. Efficient BackProp. In: Montavon G., Orr G.B.,

Müller KR. (eds) Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg; 2012.

[206] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. Pmlr. 2010;9:249–56.

[207] Berrar D, Granzow M, Dubitzky W. Introduction to Genomic and Proteomic Data Analysis. In: Dubitzky W., Granzow M., Berrar D. (eds) Fundamentals of Data Mining in Genomics and Proteomics. Springer, Boston, MA; 2017.

[208] Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Math. Intell.2001; 27(2):83–5.

[209] Savio SJ, Harrison LC, Luukkaala T. Effect of slice thickness on brain magnetic resonance image texture analysis. Biomed Eng Online. 2010 Oct 18;9:60.

[210] Parikh R, Mathai A, Parikh S, et al. Understanding and using sensitivity, specificity and predictive values. Indian J. Ophthalmol. 2008;56(1):45–50.

[211] Hill MM, Hill A.Investigação por questionário. 2nd ed. Sílabo; 2002.

[212] Hajian-Tilaki K, "Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation," Casp. J. Intern. Med. 2013; 4(2):627–35.

[213] McHugh ML. Interrater reliability: the kappa statistic. Biochem. medica. 2012; 22(3):276–82.

[214] Hosmer DW, Lemeshow S. Applied Logistic Regression. New York: Wiley, 2000.

[215] Amoroso N, La Rocca M, Monaco A. Complex networks reveal early MRI markers of Parkinson's disease. Med Image Anal. 2018 Aug;48:12-24.

[216] Marek K, Seibyl J, Eberly S. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. Neurology. 2014 May 20;82(20):1791-7.

Appendix A T1 – Weighted MRI and DaTscan acquisition

Table A.1 T1-weighted MRI and DaTscan images sequence parameters and other critical characteristics. These characteristics include repetition time, echo time, inversion time (for MRI) and dose (for SPECT)

Technique	Images Name	Characteristics
MRI	Sag 3D FSPGR BRAVO straight	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems: Size = 256 x 256 x 152 mm^3; Thickness = 1.2 mm; Voxels size = 1x1x1.2 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo
MRI	Sag 3D FSPGR BRAVO	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 152 mm^3; Thickness = 1.2 mm; Voxels size = 1x1x1.2 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo
MRI	SAG 3D T1	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: Philips Medical Systems; Size = 256 x 256 x 170 mm^3; Thickness = 1mm; Voxels size = 1x1x1 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo
MRI	MPRAGE GRAPPA	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: SIEMENS; Size = 240 x 256 x 176 mm^3; Thickness = 1mm; Voxels size = 1x1x1 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo /Inversion Recovery
MRI	AX T1	Acquisition Plane: Axial; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 512 x 512 x 84 mm^3; Thickness = 2mm; Voxels size = 1x1x2 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo /Inversion Recovery

MRI	AXIAL T1 3D MPRAGE	Acquisition Plane: Axial; Acquisition Type = 3D; Manufacturer: SIEMENS; Size = 216 x 256 x 208 mm^3; Thickness = 1mm; Voxels size = 1x1x1 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo /Inversion Recovery
MRI	Axial spgr	Acquisition Plane: Axial; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 512 x 512 x 248 mm^3; Thickness = 1mm; Voxels size = 1x1x1 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo
MRI	SAG FSPGR 3D	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 248 mm^3; Thickness = 1.4mm; Voxels size = 1x1x1.4 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo
MRI	MPRAGE T1 SAG	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 176 mm^3; Thickness = 1.0 mm; Voxels size = 1x1x1.4 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo
MRI	MPRAGE SAG	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: SIEMENS; Size = 512 x 512 x 80 mm^3; Thickness = 2 mm; Voxels size = 1x1x2 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo/Inversion Recovery
MRI	SAG T1 3D FSPGR	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: SIEMENS; Size = 256 x 256 x 98 mm^3; Thickness = 1.5 mm; Voxels size = 1x1x1.5 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo
MRI	SAG SPGR	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 98 mm^3; Thickness = 1.5 mm; Voxels size = 1x1x1.5 mm^3; Field Strength = 0.7T; Pulse Sequence = Gradient Echo

MRI	3D SAG	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: Philips Medical Systems; Size = 256 x 256 x 154 mm^3; Thickness = 1.2 mm; Voxels size = 1x1x1.2 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo
MRI	FSPGR 3D SAG	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 152 mm^3; Thickness = 1.2 mm; Voxels size = 1x1x1.2 mm^3; Field Strength = 1.5T; Pulse Sequence = Gradient Echo
MRI	SAG FSPGR BRAVO	Acquisition Plane: Sagittal; Acquisition Type = 3D; Manufacturer: GE Medical Systems; Size = 256 x 256 x 152 mm^3; Thickness = 1.2 mm; Voxels size = 1x1x1.2 mm^3; Field Strength = 3T; Pulse Sequence = Gradient Echo
SPECT	DaTscan	Size = 109 x 91 x 91 mm^3; Voxels size = 2x2x2 mm^3; Target dose for subjects = 185 MBq or 5.0 mCi of DaTSCAN™ Dose range for injection = 111 to 185 MBq; Raw projection data dimension = 128 x 128 matrix stepping each 3 degrees for a total of 120 projections or 4 degrees for a total of 90 projections

Appendix B: Accuracy Results

Table B. 1 Accuracy results obtained in the classification of Control vs PD with MRI slices

Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg ± std (%)
Slices 21_22_23	Total	56.0	60.7	47.6	39.3	50.0	50.7±7.3
	Control	54.8	50.0	47.6	45.2	54.8	50.5±3.8
	PD	57.1	71.4	47.6	33.3	45.2	51.0±12.7
Slices 24_25_26	Total	56.0	60.7	47.6	39.3	50.0	50.7±6.7
	Control	54.8	50.0	47.6	45.2	54.8	50.5±3.5
	PD	57.1	71.4	47.6	33.3	45.2	51.0±11.6
Slices 27_28_29	Total	97.6	97.6	95.2	97.6	98.8	97.4±1.2
	Control	95.2	95.2	90.5	95.2	97.6	94.8±2.3
	PD	100.0	100.0	100.0	100.0	100.0	100.0±0.0
Slices 30_31_32	Total	46.4	69.0	54.8	54.8	57.1	56.4±7.3
	Control	57.1	61.9	54.8	52.4	66.7	58.6±5.1
	PD	35.7	76.2	54.8	57.1	47.6	54.3±13.2
Slices 33_34_35	Total	52.4	51.2	53.6	50.0	53.6	52.1±1.4
	Control	50.0	54.8	57.1	40.5	50.0	50.5±5.7
	PD	54.8	47.6	50.0	59.5	57.1	53.8±4.4
Slices 36_37_38	Total	53.6	53.6	58.3	48.8	48.8	52.6±3.6
	Control	52.4	59.5	54.8	54.8	59.5	56.2±2.9
	PD	54.8	47.6	61.9	42.9	38.1	49.0±8.5
Slices 39_40_41	Total	48.8	40.5	45.2	47.6	46.4	45.7±2.9
	Control	38.1	35.7	50.0	54.8	38.1	43.3±7.6
	PD	59.5	45.2	40.5	40.5	54.8	48.1±7.7
Slices 42_43_44	Total	44.0	42.9	42.9	54.8	45.2	46.0±4.5
	Control	45.2	42.9	45.2	54.8	47.6	47.1±4.1
	PD	42.9	42.9	40.5	54.8	42.9	44.8±5.1
Slices 45_46_47	Total	53.6	50.0	47.6	47.6	51.2	50.0±2.3
	Control	59.5	52.4	47.6	45.2	54.8	51.9±5.1
	PD	47.6	47.6	47.6	50.0	47.6	48.1±1.0

Avg: Average; std: Standard Deviation

Table B. 2 Accuracy results obtained in the classification of Control vs PD with SPECT slices

Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg ±std (%)
Slices 21_22_23	Total	50.0	65.5	59.5	64.3	57.1	59.3±5.6
	PD	47.6	71.4	54.8	64.3	59.5	59.5±8.1
	SWEDD	52.4	59.5	64.3	64.3	54.8	59.0±4.9
Slices 24_25_26	Total	66.7	58.3	59.5	60.7	51.2	59.3±5.0
	Control	61.9	69.0	59.5	71.4	52.4	62.9±6.8
	PD	71.4	47.6	59.5	50.0	50.0	55.7±8.9
Slices 27_28_29	Total	57.1	63.1	63.1	61.9	64.3	61.9±2.5
	Control	61.9	66.7	57.1	54.8	66.7	61.4±4.9
	PD	52.4	59.5	69.0	69.0	61.9	62.4±6.3
Slices 30_31_32	Total	75.0	79.8	67.9	72.6	81.0	75.2±4.8
	Control	78.6	81.0	64.3	64.3	92.9	76.2±10.9
	PD	71.4	78.6	71.4	81.0	69.0	74.3±4.6
Slices 33_34_35	Total	71.4	78.6	78.6	79.8	84.5	78.6±4.2
	Control	61.9	73.8	76.2	73.8	88.1	74.8±8.3
	PD	81.0	83.3	81.0	85.7	81.0	82.4±1.9
Slices 36_37_38	Total	92.9	91.7	88.1	86.9	86.9	89.3±2.5
	Control	95.2	92.9	90.5	78.6	90.5	89.5±5.8
	PD	90.5	90.5	85.7	95.2	83.3	89.0±4.2
Slices 39_40_41	Total	95.2	97.6	89.3	91.7	88.1	92.4±3.6
	Control	95.2	97.6	90.5	92.9	83.3	91.9±4.9
	PD	95.2	97.6	88.1	90.5	92.9	92.9±3.4
Slices 42_43_44	Total	85.7	86.9	83.3	85.7	85.7	85.5±1.2
	Control	92.9	88.1	92.9	85.7	85.7	89.0±3.2
	PD	78.6	85.7	73.8	85.7	85.7	81.9±4.9
Slices 45_46_47	Total	69.0	71.4	82.1	71.4	73.8	73.6±4.5
	Control	69.0	64.3	78.6	66.7	81.0	71.9±6.6
	PD	69.0	78.6	85.7	76.2	66.7	75.2±6.8

Avg: Average; std: Standard Deviation

Table B. 3 Accuracy results obtained in the classification of Control vs SWEDD with MRI slices

Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg±Std (%)
Slice-21_22_23	Total	46.7	50.0	60.0	60.0	73.3	58.0±9.3
	PD	53.3	66.7	66.7	80.0	93.3	72.0±13.6
	SWEDD	40.0	33.3	53.3	40.0	53.3	44.0±8.0
Slice-24-25-26	Total	63.3	83.3	60.0	66.7	53.3	65.3±10.0
	Control	46.7	86.7	80.0	60.0	46.7	64.0±16.7
	SWEDD	80.0	80.0	40.0	73.3	60.0	66.7±15.2
Slice-27-28-29	Total	50.0	73.3	46.7	46.7	56.7	54.7±10.0
	Control	40.0	73.3	40.0	46.7	53.3	50.7±12.4
	SWEDD	60.0	73.3	53.3	46.7	60.0	58.7±8.8
Slice-30-31-32	Total	60.0	40.0	53.3	53.3	63.3	54.0±8.0
	Control	66.7	46.7	46.7	53.3	80.0	58.7±12.9
	SWEDD	53.3	33.3	60.0	53.3	46.7	49.3±9.0
Slices 33_34_35	Total	70.0	53.3	56.7	53.3	46.7	56.0±7.7
	Control	66.7	53.3	40.0	46.7	60.0	53.3±9.4
	SWEDD	73.3	53.3	73.3	60.0	33.3	58.7±14.8
Slices 36_37_38	Total	50.0	63.3	46.7	46.7	70.0	55.3±9.6
	Control	40.0	86.7	40.0	53.3	66.7	57.3±17.7
	SWEDD	60.0	40.0	53.3	40.0	73.3	53.3±12.6
Slices 39_40_41	Total	60.0	63.3	63.3	66.7	73.3	65.3±4.5
	Control	73.3	60.0	60.0	60.0	73.3	65.3±6.5
	SWEDD	46.7	66.7	66.7	73.3	73.3	65.3±9.8
Slice-42-43-44	Total	46.7	50.0	63.3	63.3	60.0	56.7±7.0
	Control	53.3	53.3	66.7	60.0	86.7	64.0±12.4
	SWEDD	40.0	46.7	60.0	66.7	33.3	49.3±12.4
Slice-45_46_47	Total	73.3	60.0	60.0	63.3	60.0	63.3±5.2
	Control	73.3	66.7	46.7	46.7	66.7	60.0±11.2
	SWEDD	73.3	53.3	73.3	80.0	53.3	66.7±11.2

Avg: Average; std: Standard Deviation

Table B. 4 Accuracy results obtained in the classification of Control vs SWEDD with SPECT slices

Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg±std (%)
Slices 21_22_23	Total	63.3	60.0	70.0	43.3	60.0	59.3±8.8
	PD	73.3	66.7	60.0	46.7	53.3	60.0±9.4
	SWEDD	53.3	53.3	80.0	40.0	66.7	58.7±13.6
Slices 24-25-26	Total	53.3	56.7	66.7	60.0	60.0	59.3±4.4
	Control	60.0	66.7	66.7	60.0	66.7	64.0±3.3
	SWEDD	46.7	46.7	66.7	60.0	53.3	54.7±7.8
Slices 27-28-29	Total	40.0	56.7	43.3	63.3	70.0	54.7±11.5
	Control	53.3	53.3	40.0	66.7	86.7	60.0±15.8
	SWEDD	26.7	60.0	46.7	60.0	53.3	49.3±12.4
Slices 30_31_32	Total	50.0	40.0	53.3	40.0	46.7	46.0±5.3
	Control	46.7	46.7	53.3	33.3	40.0	44.0±6.8
	SWEDD	53.3	33.3	53.3	46.7	53.3	48.0±7.8
Slices 33_34_35	Total	60.0	50.0	33.3	56.7	56.7	51.3±9.6
	Control	66.7	66.7	6.7	46.7	60.0	49.3±22.5
	SWEDD	53.3	33.3	60.0	66.7	53.3	53.3±11.2
Slices 36_37_38	Total	40.0	63.3	40.0	56.7	43.3	48.7±9.6
	Control	33.3	60.0	20.0	73.3	33.3	44.0±19.6
	SWEDD	46.7	66.7	60.0	40.0	53.3	53.3±9.4
Slices 39_40_41	Total	63.3	53.3	60.0	56.7	80.0	62.7±9.3
	Control	73.3	33.3	60.0	40.0	80.0	57.3±18.2
	SWEDD	53.3	73.3	60.0	73.3	80.0	68.0±9.8
Slices 42_43_44	Total	40.0	66.7	56.7	53.3	43.3	52.0±9.6
	Control	33.3	66.7	46.7	60.0	26.7	46.7±15.2
	SWEDD	46.7	66.7	66.7	46.7	60.0	57.3±9.0
Slices 45_46_47	Total	53.3	63.3	60.0	56.7	50.0	56.7±4.7
	Control	46.7	60.0	66.7	53.3	46.7	54.7±7.8
	SWEDD	60.0	66.7	53.3	60.0	53.3	58.7±5.0

Avg: Average; std: Standard Deviation

Table B. 5Accuracy results obtained in the classification of PD vs SWEDD with MRI slices

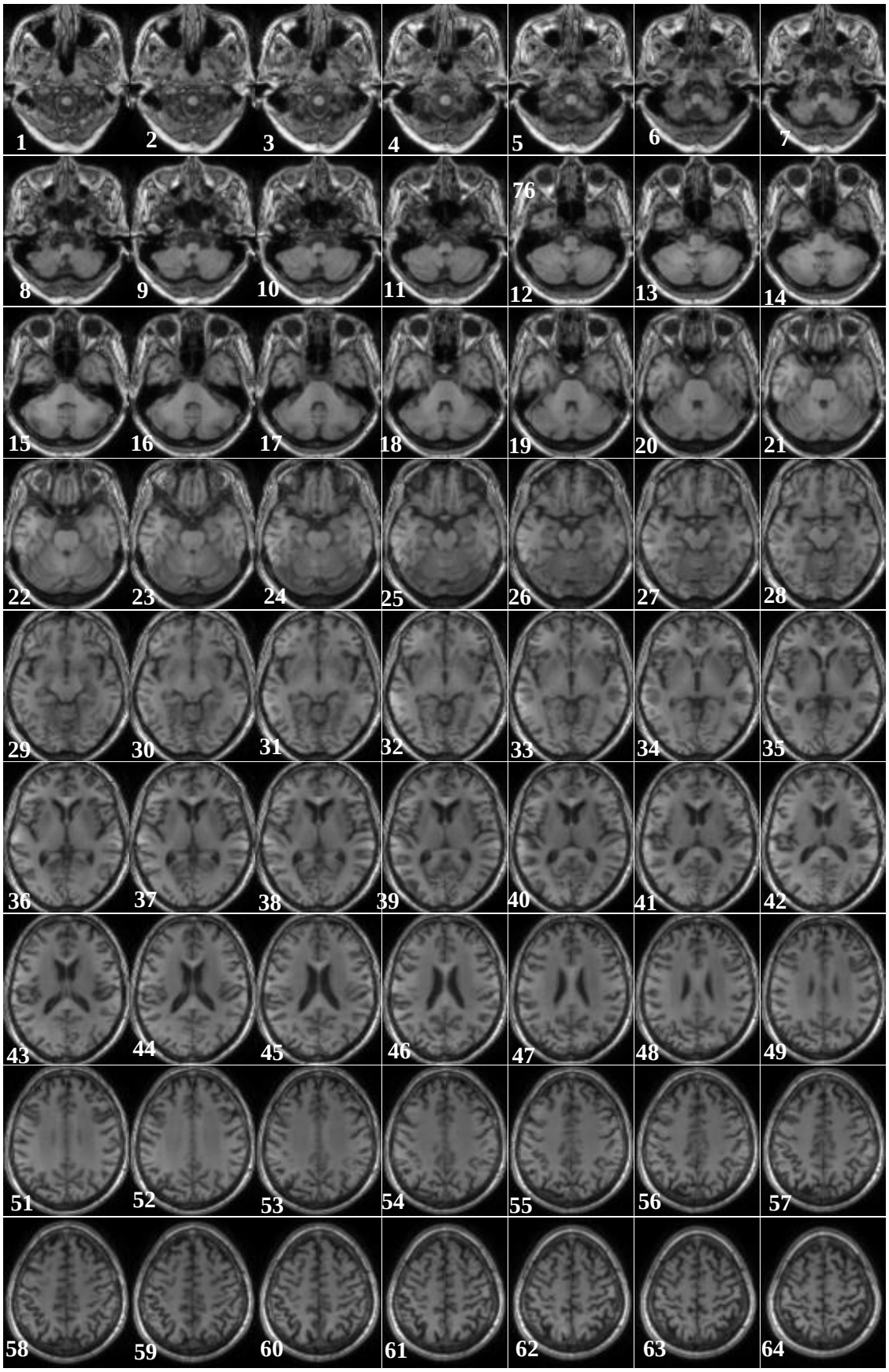
Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg±std (%)
Slices 21_22_23	Total	56.7	40.0	50.0	56.7	40.0	48.7±7.5
	PD	60.0	53.3	66.7	26.7	26.7	46.7±16.9
	SWEDD	53.3	26.7	33.3	86.7	53.3	50.7±20.9
Slices 24-25-26	Total	50.0	56.7	46.7	53.3	56.7	52.7±3.9
	PD	40.0	66.7	33.3	53.3	46.7	48.0±11.5
	SWEDD	60.0	46.7	60.0	53.3	66.7	57.3±6.8
Slices 27-28-29	Total	53.3	60.0	63.3	66.7	60.0	60.7±4.4
	PD	53.3	66.7	46.7	46.7	73.3	57.3±10.8
	SWEDD	53.3	53.3	80.0	86.7	46.7	64.0±16.1
Slices 30_31_32	Total	63.3	63.3	53.3	66.7	63.3	62.0±4.5
	PD	66.7	60.0	53.3	60.0	73.3	62.7±6.8
	SWEDD	60.0	66.7	53.3	73.3	53.3	61.3±7.8
Slices 33_34_35	Total	93.3	73.3	80.0	73.3	80.0	80.0±7.3
	PD	93.3	93.3	73.3	73.3	93.3	85.3±9.8
	SWEDD	93.3	53.3	86.7	73.3	66.7	74.7±14.2
Slices 36_37_38	Total	96.7	93.3	96.7	86.7	93.3	93.3±3.7
	PD	100.0	86.7	93.3	80.0	93.3	90.7±6.8
	SWEDD	93.3	100.0	100.0	93.3	93.3	96.0±3.3
Slices 39_40_41	Total	93.3	83.3	90.0	83.3	80.0	86.0±4.9
	PD	93.3	80.0	93.3	80.0	80.0	85.3±6.5
	SWEDD	93.3	86.7	86.7	86.7	80.0	86.7±4.2
Slices 42_43_44	Total	70.0	63.3	83.3	80.0	83.3	76.0±8.0
	PD	66.7	53.3	100.0	73.3	93.3	77.3±17.2
	SWEDD	73.3	73.3	66.7	86.7	73.3	74.7±6.5
Slices 45_46_47	Total	56.7	66.7	70.0	83.3	76.7	70.7±9.0
	PD	66.7	60.0	66.7	66.7	66.7	65.3±2.7
	SWEDD	46.7	73.3	73.3	100.0	86.7	76.0±17.7

Table B. 6 Accuracy results obtained in the classification of PD vs SWEDD with SPECT slices

Slices	Accuracy	Batch1 (%)	Batch2 (%)	Batch3 (%)	Batch4 (%)	Batch5 (%)	Avg±std (%)
Slices 21_22_23	Total	50.0	56.7	56.7	46.7	46.7	51.3±4.5
	PD	40.0	60.0	60.0	40.0	60.0	52.0±9.8
	SWEDD	60.0	53.3	53.3	53.3	33.3	50.7±9.0
Slices 24_25_26	Total	80.0	80.0	56.7	53.3	76.7	69.3±11.8
	PD	80.0	73.3	46.7	46.7	80.0	65.3±15.4
	SWEDD	80.0	86.7	66.7	60.0	73.3	73.3±9.4
Slices 27_28_29	Total	66.7	76.7	73.3	76.7	73.3	73.3±3.7
	PD	86.7	93.3	66.7	86.7	73.3	81.3±9.8
	SWEDD	46.7	60.0	80.0	66.7	73.3	65.3±11.5
Slices 30_31_32	Total	63.3	46.7	63.3	53.3	63.3	58.0±6.9
	PD	60.0	33.3	60.0	60.0	66.7	56.0±11.6
	SWEDD	66.7	60.0	66.7	46.7	60.0	60.0±7.3
Slices 33_34_35	Total	60.0	56.7	56.7	66.7	70.0	62.0±5.4
	PD	53.3	60.0	60.0	60.0	60.0	58.7±2.7
	SWEDD	66.7	53.3	53.3	73.3	80.0	65.3±10.7
Slices 36_37_38	Total	63.3	53.3	60.0	53.3	70.0	60.0±6.3
	PD	40.0	40.0	40.0	53.3	80.0	50.7±15.5
	SWEDD	86.7	66.7	80.0	53.3	60.0	69.3±12.4
Slices 39_40_41	Total	40.0	60.0	53.3	46.7	73.3	54.7±11.5
	PD	33.3	73.3	60.0	40.0	60.0	53.3±14.6
	SWEDD	46.7	46.7	46.7	53.3	86.7	56.0±15.5
Slices 42_43_44	Total	56.7	53.3	60.0	56.7	56.7	56.7±2.1
	PD	33.3	46.7	53.3	53.3	53.3	48.0±7.8
	SWEDD	80.0	60.0	66.7	60.0	60.0	65.3±7.8
Slices 45_46_47	Total	60.0	50.0	66.7	53.3	56.7	57.3±5.7
	PD	46.7	46.7	60.0	53.3	53.3	52.0±5.0
	SWEDD	73.3	53.3	73.3	53.3	60.0	62.7±9.0

Avg: Average; std: Standard Deviation

Appendix C: Example of MRI and SPECT slices



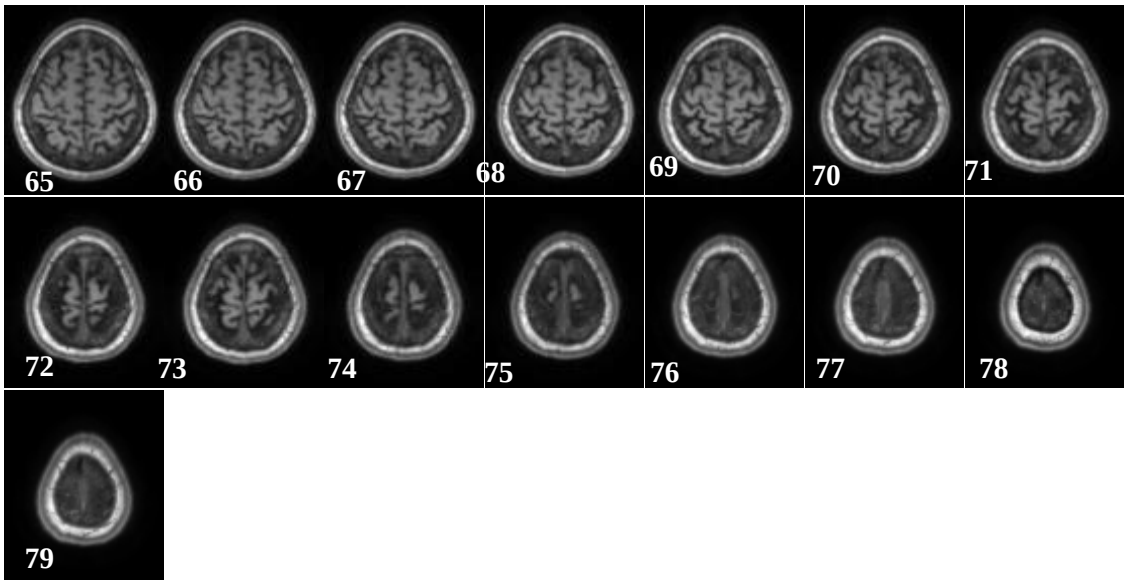
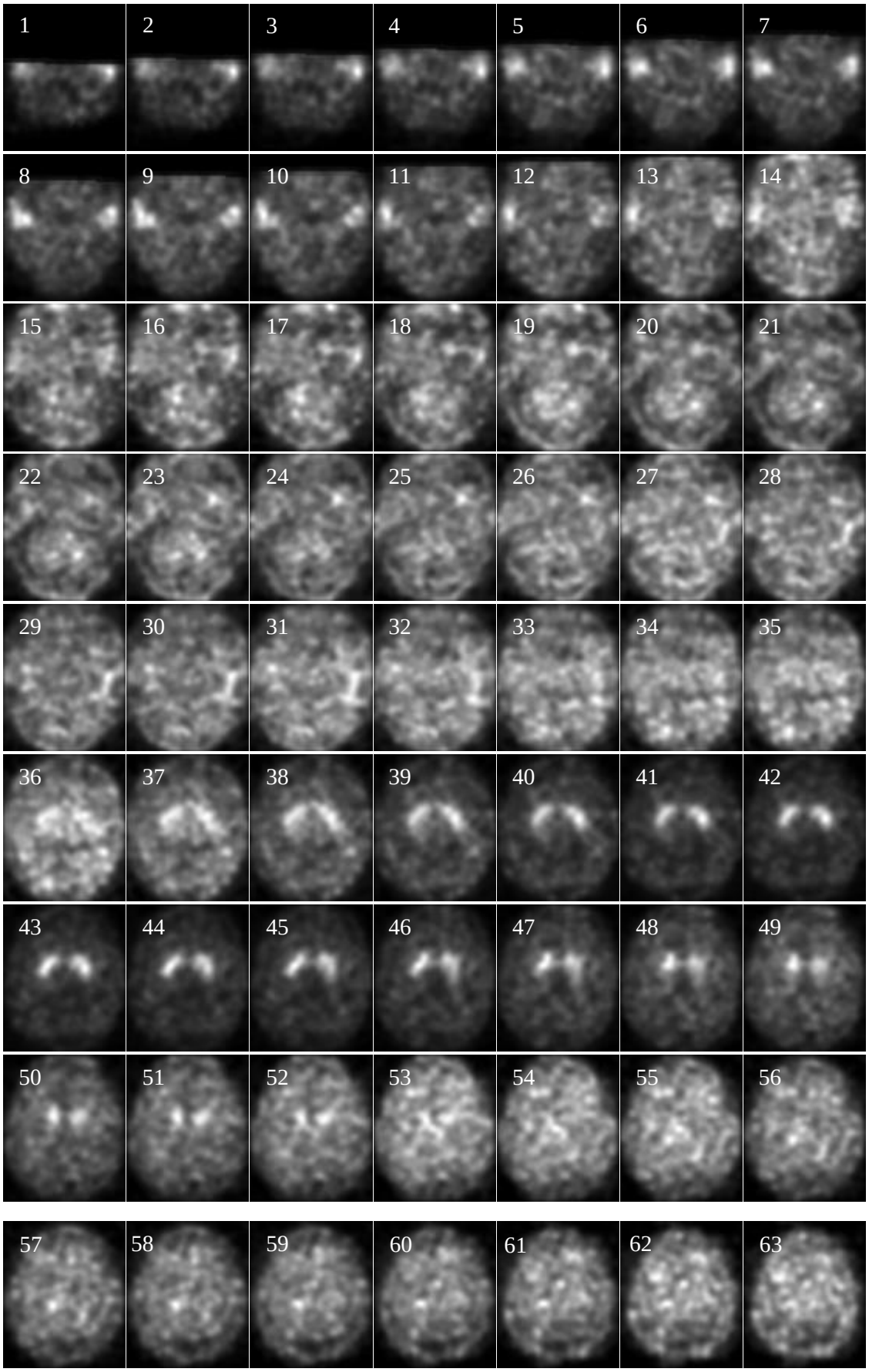


Figure C. 1 Example from one control subject of the 79 slices obtained with 3D T1-Weighted MRI images division in Axial plane.


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63

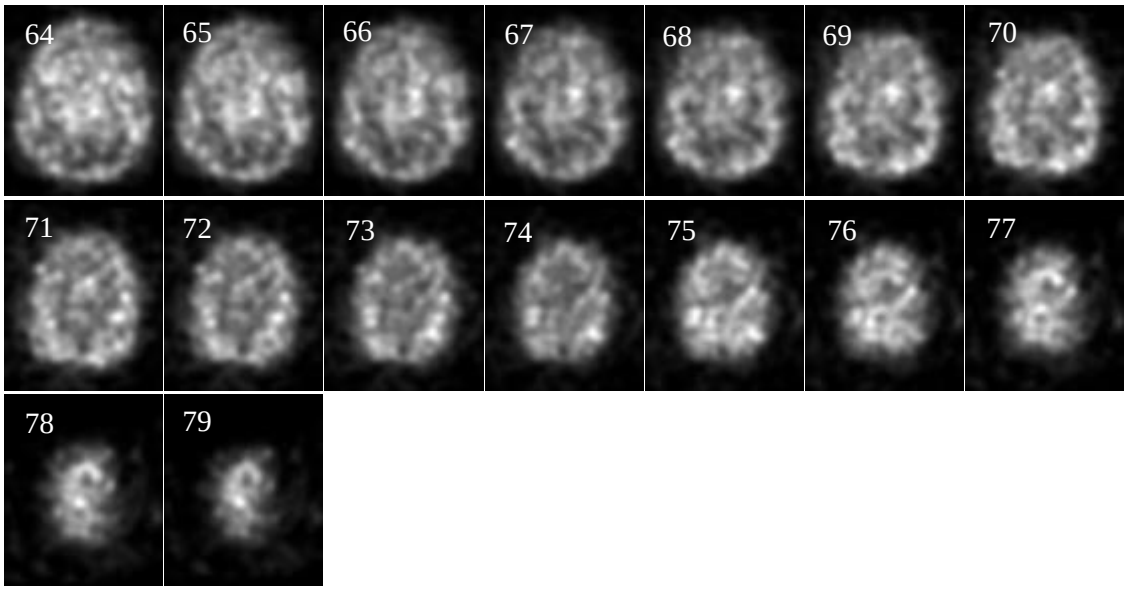


Figure C. 2 Example from one control subject of the 79 slices obtained with DatTscan SPECT images division in Axial plane.