UNIVERSITÀ POLITECNICA DELLE MARCHE

FACULTY OF ENGINEERING

Master's Degree in Biomedical Engineering

C3DKeras for Parkinson-related Dementia Identification

C3DKeras per l'identificazione della demenza relativa al Parkinson

Supervisor

Dr. Agnese Sbrollini

Co-Supervisors

Prof. Laura Burattini

Dr. Selene Tomassini

Master Thesis by

Diana-Maria Rotariu

Academic Year 2022/2023

ABSTRACT

The diagnosis of Parkinson's disease dementia (PDD) still follows the one-year rule, which may be too late for the optimal treatment. Since up to 83% of Parkinson's disease (PD) patients eventually develop dementia later in the disease progression, early identification would allow for timely administration of the appropriate treatment, potentially leading to an increased life expectancy. Additionally, as there is an exponential growth of the aging population in the world the number of people suffering from PD and consequently of PDD is increasing. However, until now no therapeutic method has been discovered for completely treat the subjects affecting by this type of neurodegenerative disease. Therefore, the early detection could be the only way to increase the life expectancy.

This thesis provided a systematic literature review on PDD identification through Machine Learning (ML) algorithms and a novel methodology for the identification of PDD subjects from MRI scans.

A particular 3D CNN, called C3DKeras, pre-trained on large video-clips on sports, was used as core part of the proposed pipeline, whereas the entire methodology was conducted under the structure of ablation studies. Specifically, four ablation experiments were performed, in which a new part was added to the previous one in the pipeline as follows: Experiment 0, no modification from the original C3DKeras; Experiment 1, addition of weighting class; Experiment 2, addition of unfrozen layers; and Experiment 3, addition of rigorous hyperparameter tuning. All the experiments were conducted by taking into account two datasets downloaded from Parkinson's Progression Markers Initiative (PPMI), namely PDD and prodromal PD.

Experiment 3, with an AUC of 86%, SE of 100%, ACC of 73%, and F1-score of 73% for the PDD class; achieved the highest performance among the four experiments and among the literature research results.

Therefore, as the entire methodology was conducted under the structure of ablation studies, it was evident how each fine-tuning step significantly improve the overall performance. Hence, this study provides an innovative approach for the identification and classification of PDD subjects as well as those with PD using the particular CNN called C3DKeras.

CONTENTS

I. *Introduction*

The second most common neurological disorder that prevails among the aging population in the world is considered to be Parkinson's Disease (PD). PD is based on the formation of the so-called Lewy bodies in the substantia nigra of the brainstem. As the disease progresses over time, it could result in Parkinson's disease dementia (PDD). PDD affects the cortex with devastating effects on complex sensory association cortex regions and prefrontal and temporal areas involved in executive brain functions, learning and memory.

The diagnosis of PDD is currently based on the 1-year rule, being in some cases too late for the correct treatment. As up to 83% of patients with PD eventually develop dementia later in the disease course, faster the identification, sooner could be provided the correct treatment, giving the possibility to increase the life expectance.

For the detection of PDD the most widely used diagnostic is the analysis of Magnetic Resonance Imaging (MRI) scans of the brain. The MRI is a non-invasive technique that provide anatomical details about the structure of the brain. However, as the MRI is a 3D structure, it is really difficult for the human eye to analyse the details that distinguished PDD subject's brain from PD.

The advancement of intelligent technologies has made computer-aided detection systems very effective at performing diagnosis of different diseases. In particular, Convolutional Neuronal Network (CNN) excels in tasks like image recognition, object detection, and image classification. Especially pre-trained models, like C3DKeras, are often used for transfer learning. Pre-trained models on large datasets are used to facilitate the learning tasks with limited data.

In scientific literature, only two studies that classify PDD subjects with Machine Learning (ML) applications were found. This is due to the complexity of differentiating PDD from PD subjects. In both studies Positron Emission Tomography (PET) was used as a method for the acquisition of the data. Both PET and MRI are non-invasive techniques, however, PET involves the exposure to ionizing radiation that could be harmful for the patient.

This study aimed to establish an innovative and effective approach for accurately detecting and distinguishing individuals affected by PDD based on structural information derived from

3D MRI data. This innovative methodology sought to contribute to the medical field by providing a more sophisticated approach for the identification and classification of PDD subjects using advanced neural network architectures as C3DKeras.

1. Anatomy and Physiology of the Nervous System

The Nervous System is a highly complex piece of biological machinery that coordinates actions and sensory information by transmitting signals to and from different parts of the body.

1.1. The Cellular Components of the Nervous System

In the early nineteenth century, the cell was recognized as the fundamental unit of all living organisms. However, it was not until the twentieth century, that neuroscientists agreed that nervous tissue, like all other organs, is made up of these fundamental units[1].

The histological studies of Cajal, Golgi, and several successors led to the further accord that the cells of the nervous system can be divided into two broad categories: nerve cells (or neurons), and supporting cells called neuroglia (or simply glia).

1.1.1. Neurons

The neurons are electrically excitable cells, capable of responding to physical and chemical stimuli by modifying the concentration of ions on the faces of their membrane, producing signals in the form of changes in electrical potential, propagating or conducting signals from the point of initialization to other parts of the cell and transmit signals to other cells through synapses[2].

Despite the neurons diversity based on functions performed in different parts of the nervous system, shape, and other characteristics, all of them are composed by a cell body or soma, one or more dendrites, an axon, and the presynaptic terminations (Figure 1). The cell body contains the nucleus and most of the major organelles, the dendrites are instead cellular extensions with many branches and is where the majority of input to the neuron occurs via the dendritic spine while the axon (cable-like projection) carries signals away from the soma and carries information back to it (the part of the axon where it emerges from the soma is called the axon hillock). Each axon is commonly subdivided at its distal end into several preterminal branches of lower calibre and each of these forms a variable number of synaptic boutons that is the site where synapses with other neurons occurs.

Neuronal cell bodies are often grouped together in clusters. The axons of neurons are usually grouped together in bundles. In addition, widespread regions of nervous tissue are grouped together as either grey matter or white matter [3].

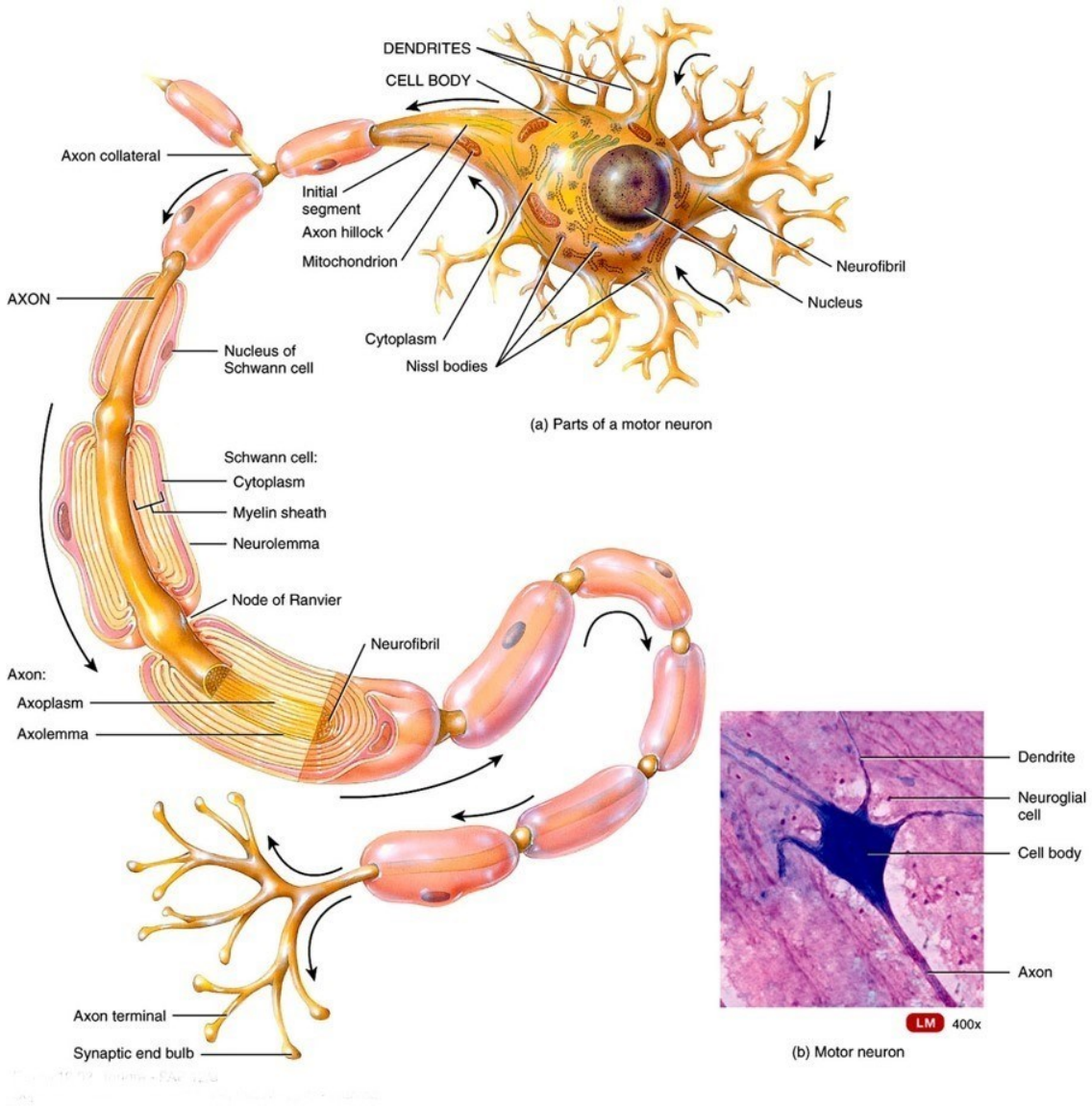


Figure 1.Structure of a multipolar neuron. A multipolar neuron has a cell body, several short dendrites, and a single long axon. Arrows indicate the direction of information flow: dendrites → cell body → axon → axon terminals [3]

There are many different types of neurons in the Nervous System, and they could be classified by several criteria. In accordance with the number of processes possessed by a neuron, the classification of these cells can be as follows:

- Unipolar neurons. They have a single, short axon that emerges from the cell body and then branches into two processes that extend in opposite directions. The process that extends peripherally is known as the peripheral process and is associated with sensory reception.

- Bipolar neurons have two processes, which extend from each end of the cell body, opposite to each other; one is the axon and one the dendrite. They are found in the retina of the eye, the inner ear, and the olfactory area [3]. Particular bipolar neurons are pseudounipolar neurons or T-neurons, whose cellular body is provided with a single extension that at a relatively large distance from the soma is divided into a T [2].
- Multipolar neurons usually have several dendrites and one axon. Most neurons in the brain and spinal cord are of this type, as well as all motor neurons.

In addition to the structural classification, some neurons are named for the histologist who first described them or for an aspect of their shape or appearance; examples include Purkinje cells in the cerebellum and pyramidal cells, found in the cerebral cortex of the brain, which have pyramid-shaped cell bodies [3].

Another type of classification of neurons may be made based on their functions:

- Sensory neurons or afferent neurons either contain sensory receptors at their distal ends (dendrites) or are located just after sensory receptors. Once an appropriate stimulus activates a sensory receptor, the sensory neuron forms an action potential in its axon and the action potential is sent into the Central Nervous System (CNS) through cranial or spinal nerves. Most sensory neurons are unipolar in structure.
- Motor neurons or efferent neurons deliver action potentials away from the CNS to effectors (muscles and glands) in the Peripheral Nervous System (PNS) through cranial or spinal nerves. Motor neurons are multipolar in structure.
- Interneurons or association neurons are mainly located within the CNS between sensory and motor neurons. Interneurons integrate incoming sensory information from sensory neurons and then produce a motor response by activating the appropriate motor neurons. Most interneurons are multipolar in structure [3].

The neurons communicate via a combination of electrical and chemical signals. Within the neuron, electrical signals driven by charged particles allow rapid conduction from one end of the cell to the other. In particular, the dendrites are the part of the neuron that receives incoming signals. Based on the strength of this incoming stimuli the neurons must decide whether to pass the signal along or not. If this stimulation is strong enough the signal is transmitted along the entire length of the axon in a phenomenon called action potential.

Transmission of neuronal signal is entirely dependent on the movement of the ions. Various ions including sodium (Na+), calcium (Ca2+) and potassium (K+) are unequally distributed between the inside and the outside of the cell. When a neuron is not sending a signal, it is considered to be at rest. In a typical neuron in its resting state the concentration of Na+ is higher outside the cell than inside, while the concentration of K+ is the opposite. This separation creates a chemical gradient across the membrane. Furthermore, at rest, there are more positively charged ions outside the cell relative to the inside; this create a difference on charge across the membrane which is called an electrical gradient. The chemical gradient together with the electrical gradient form the electrochemical gradient.

An action potential is a sequence of rapidly occurring events that decrease and reverse the membrane potential and then eventually restore it to the resting state. Membrane potential is defined as the difference between the total charge inside and outside of the cell. At rest, when no signals are being transmitted, neuronal membrane has a resting potential of approximately -70mV.

The action potential has two main phases: a depolarizing phase and a repolarizing phase. Once the cell membrane reaches a certain level termed the threshold (about −55 mV in many neurons) the Na+ channels changes to an open position and Na+ reaches into the cell caused by the electrochemical gradient. During the depolarizing phase, the negative membrane potential becomes less negative, reaches zero, and then becomes positive. As the membrane potential becomes positive, the Na+ channels become inactive, and this stops the flow of Na+ into the cell. The change of the membrane potential causes the opening of the voltage-gated K+ channels and the K+ flow out of the cell. This process is called repolarization. During the repolarizing phase, the membrane potential is restored to the resting state of −70 mV.

The period of time during which a second action potential can be initiated, but only by a larger-than normal stimulus is called the relative refractory period. It coincides with the period when the voltage-gated K+ channels are still open after inactivated Na+ channels have returned to their resting state.

There are two types of propagation of the signal along the axon: continuous conduction and saltatory conduction. In continuous conduction, ions flow through their voltage-gated channels in each adjacent segment of the membrane. Continuous conduction occurs in

unmyelinated ("the myelin sheath is a greatly extended and modified plasma membrane wrapped around the nerve axon in a spiral fashion" [4]) axons and in muscle fibers. Action potentials propagate more rapidly along myelinated axons than along unmyelinated axons. The saltatory conduction is a special mode of action potential propagation that occurs along myelinated axons, it occurs because of the uneven distribution of voltage-gated channels. When an action potential propagates along a myelinated axon, an electric current flow pass through the extracellular fluid surrounding the myelin sheath. The action potential at the first node generates ionic currents in and extracellular fluid that depolarize the membrane to threshold, opening voltage-gated Na+ channels at the second node. The resulting ionic flow through the opened channels constitutes an action potential at the second node. Then, the action potential at the second node generates an ionic current that opens voltage-gated Na+ channels at the third node, and so on. Each node repolarizes after it depolarizes [3].

As describe before the region where communication occurs between two neurons or between a neuron and an effector cell is called synapse. The cell that sends the nerve impulse is named the presynaptic cell and the cell that receives the nerve impulse is called the postsynaptic cell. Synapses are important since some diseases and neurological disorders result from disruptions of synaptic communication.

Structurally, two types of synapses are found in neurons:

1) Electrical synapses: it has a direct physical contact. Electrical transmission is mediated by clusters of intercellular channels called gap junctions that connect the interior of two adjacent cells, and thereby directly enable the bidirectional (axon-axon) passage of electrical currents carried by ions. Since action potentials conduct directly through gap junctions, electrical synapses are faster than chemical synapses; as in the first one the action potential passes directly from the presynaptic cell to the postsynaptic cell. They are usually found in the heart and the gastrointestinal tract [3].
2) Chemical synapses: are the most common type of synapses. To transmit information from one neuron to another or a target organ, a chemical substance termed transmitter is used. Depending on the released transmitter, chemical synapses can have either an excitatory or inhibitory effect downstream cells. Transmitter are synthesized in the soma of the neuron and transported along the axon to the presynaptic terminal. When an action potential reaches the presynaptic membrane,

depolarization occurs. As a result, the voltage-gated calcium channels at the presynaptic terminal open, leading to the influx of calcium cations into the cell. An increase in the concentration of Ca2+ inside the presynaptic neuron serves as a signal that triggers exocytosis of the synaptic vesicles. As vesicle membranes merge with the plasma membrane, neurotransmitter molecules within the vesicles are released into the synaptic cleft (or bouton). Each synaptic vesicle contains several thousand molecules of neurotransmitter. The neurotransmitter molecules diffuse across the synaptic cleft and bind to neurotransmitter receptors in the postsynaptic neuron's plasma membrane. Binding of neurotransmitter molecules to their receptors on ligand-gated channels opens the channels and allows particular ions to flow across the membrane. As ions flow through the opened channels, the voltage across the membrane changes. This change in membrane voltage is a postsynaptic potential. A neurotransmitter causes either an excitatory or an inhibitory graded potential. A depolarizing postsynaptic potential is called an excitatory postsynaptic potential (EPSP). A hyperpolarizing postsynaptic potential is termed an inhibitory postsynaptic potential (IPSP) [3].

One of the transmitters involved in the chemical synapses is the dopamine (Figure 2.). In Parkinson's disease (PD), the dopaminergic neurons of the substantia nigra degenerate, leading to a characteristic motor dysfunction.

In addition, the one that regulates the synaptic vesicle trafficking and subsequently neurotransmitter release is the alpha-synuclein (aSyn). In PD, Parkinson's Disease with Dementia (PDD) and in Dementia with Lewy Body (DLB), insoluble forms of aSyn accumulate as inclusions in Lewy bodies.

"A typical neuron in the CNS receives input from 1000 to 10,000 synapses" [3]. Integration of these inputs involves summation of the postsynaptic potentials that form in the postsynaptic neuron. The greater the summation of EPSPs, the greater the chance that threshold will be reached. There are two types of summation: spatial summation and temporal summation. Spatial summation is summation of postsynaptic potentials in response to stimuli that occur at different locations in the membrane of a postsynaptic cell at the same time. Temporal summation is summation of postsynaptic potentials in response to stimuli that occur at the same location in the membrane of the postsynaptic cell but at different times.

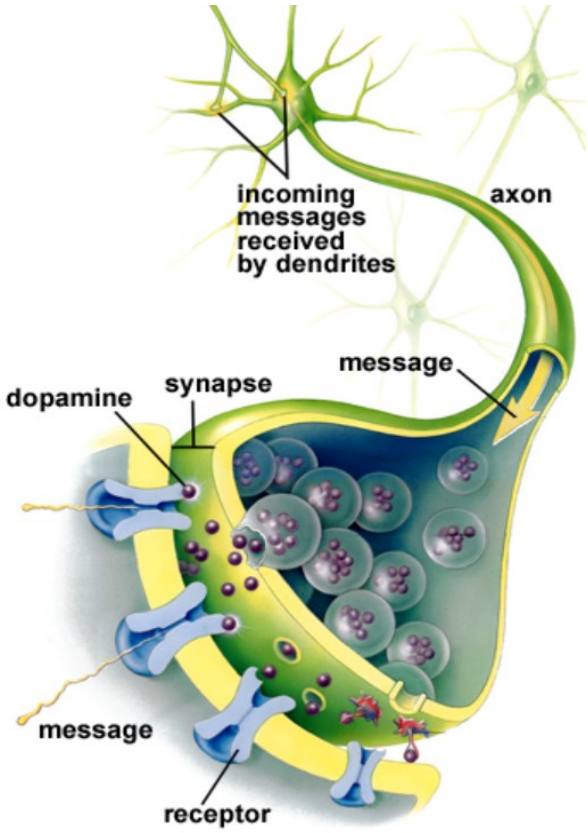


Figure 2. Chemical synapses that use dopamine as a transmitter.

1.1.2. Glia

Glial cells are considered to be supporting cells, and many functions are directed at helping neurons complete their function for communication. Generally, neuroglia are smaller than neurons, and they are 5 to 25 times more numerous. In contrast to neurons, glia do not generate or propagate action potentials. There are six types of glial cells; four of them are found in the CNS (Figure 3.) and two are found in the PNS.

Neuroglia of the CNS can be classified on the basis of size, cytoplasmic processes, and intracellular organization into:

1) ASTROCYTES: star-shaped cells that have many processes and are the largest and most numerous of the neuroglia. The processes of astrocytes make contact with blood capillaries, neurons, and the pia mater (a thin membrane around the brain and spinal cord). Astrocytes contain microfilaments that give them considerable strength, which enables them to support neurons. Processes of astrocytes wrapped around blood capillaries isolate neurons of the CNS from various potentially harmful substances in

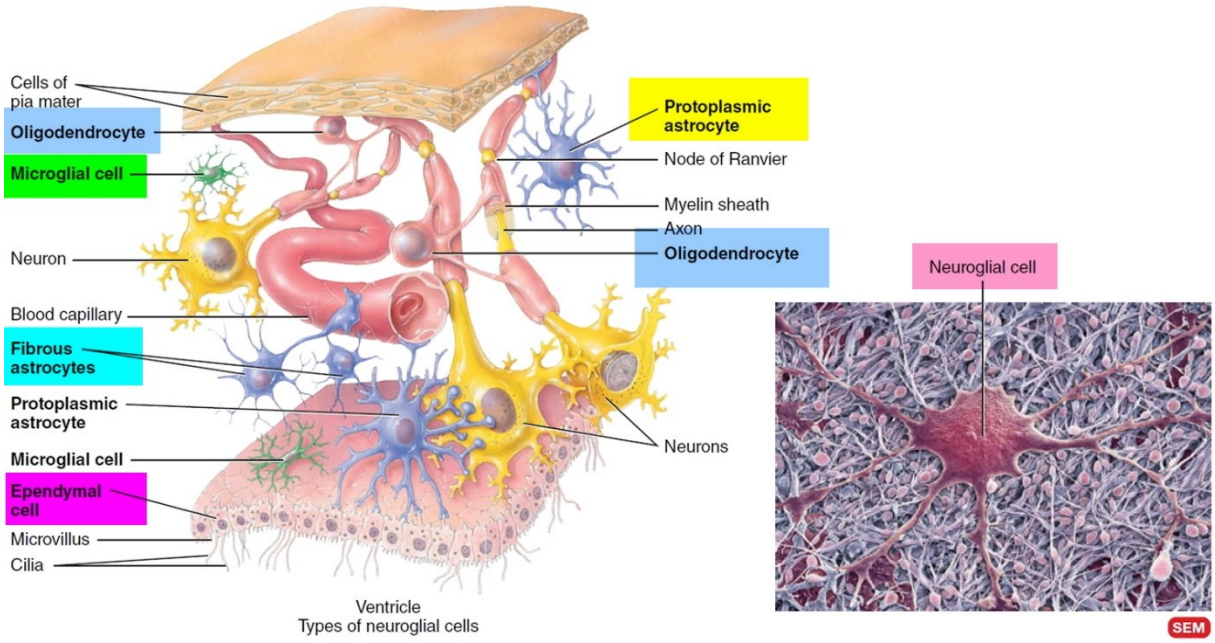


Figure 3. Neuroglia of the central nervous system

blood. Astrocytes help to maintain the appropriate chemical environment for the generation of nerve impulses. For example, they regulate the concentration of important ions such as K+; take up excess neurotransmitters; and serve as a channel for the passage of nutrients and other substances between blood capillaries and neurons. Astrocytes may also play a role in learning and memory by influencing the formation of neural synapses.

2) OLIGODENDROCYTES: they resemble astrocytes but are smaller and contain fewer processes. Processes of oligodendrocytes are responsible for forming and maintaining the myelin sheath around CNS axons. As explained before the myelin sheath is a multilayered lipid and protein covering around some axons that insulates them and increases the speed of nerve impulse conduction; such axons are said to be myelinated.
3) MICROGLIAL CELLS OR MICROGLIA: are small cells with slender processes that give off numerous spinelike projections. Microglial cells function as phagocytes. Like tissue macrophages, they remove cellular fragments formed during normal development of the nervous system and phagocytize microbes and damaged nervous tissue.
4) EPENDYMAL CELLS: are cuboidal to columnar cells arranged in a single layer that possess microvilli and cilia. These cells line the ventricles of the brain and central canal of the spinal cord (spaces filled with cerebrospinal fluid, which protects and nourishes the brain and spinal cord). Functionally, ependymal cells produce, possibly monitor,

and assist in the circulation of cerebrospinal fluid. They also form the blood–cerebrospinal fluid barrier [3].

Neuroglia of the PNS completely surround axons and cell bodies. The two types of glial cells in the PNS are:

1) SCHWANN CELLS: like oligodendrocytes, they form the myelin sheath around axons. A single oligodendrocyte myelinates several axons, but each Schwann cell myelinates a single axon.
2) SATELLITE CELLS: flat cells that surround the cell bodies of neurons of PNS ganglia. Besides providing structural support, satellite cells regulate the exchanges of materials between neuronal cell bodies and interstitial fluid.

White matter is composed primarily of myelinated axons. The white colour of myelin gives white matter its name. The grey matter of the nervous system contains neuronal cell bodies, dendrites, unmyelinated axons, axon terminals, and neuroglia.

1.2. Overall Organization of the Nervous System

A first general functional distinction divides the Nervous System in:

- Sensory Systems: the system that acquire and process information from the environment (the visual system or the auditory system);
- Motor Systems: that responds to the information form the Sensory System by generating movement and other actions.

In addition to this functional distinction, neuroscientists and neurologists have conventionally divided the vertebrate nervous system anatomically into central and peripheral components (Figure 4.).

1.2.1. Central Nervous System

The CNS consist of the brain and the spinal cord. Specially the adult brain consists of four major parts: brainstem, cerebellum, diencephalon, and cerebrum (Figure 5.). The brainstem is continuous with the spinal cord and consists of the medulla oblongata, pons, and midbrain. Posterior to the brainstem is the cerebellum while superior to the brainstem is the diencephalon, which consists of the thalamus, hypothalamus, and epithalamus. Supported on the diencephalon and brainstem is the cerebrum, the largest part of the brain.

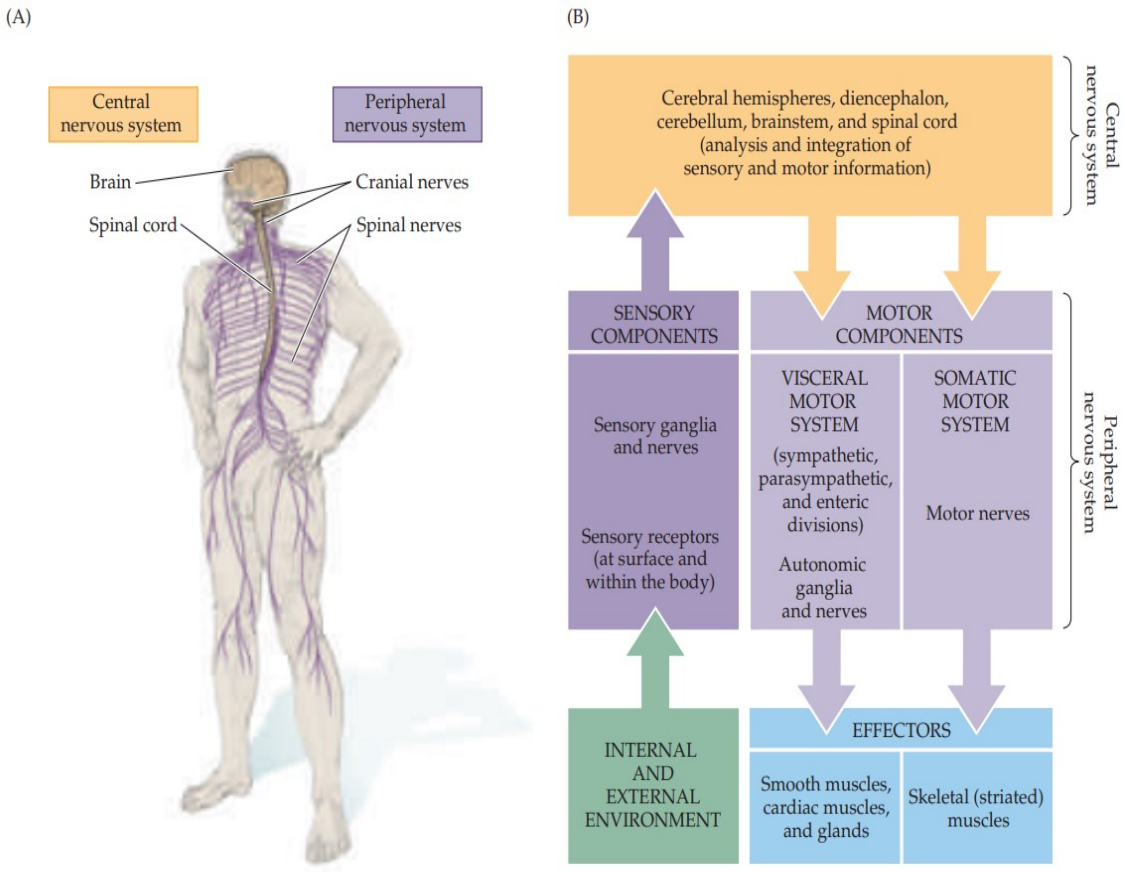


Figure 4. The major components of the nervous system and their functional relationships. (A) The CNS (brain and spinal cord) and PNS (spinal and cranial nerves). (B) Diagram of the major components of the central and peripheral nervous systems and their function relationships [1].

The medulla oblongata, or more simply the medulla, is continuous with the superior part of the spinal cord.

The medulla's white matter contains all sensory (ascending) tracts and motor (descending) tracts that extend between the spinal cord and other parts of the brain. Some of the white matter forms protuberances on the anterior aspect of the medulla. These extensions, called the pyramids are formed by the large corticospinal tracts that pass from the cerebrum to the spinal cord. The corticospinal tracts control voluntary movements of the limbs and trunk. Just superior to the junction of the medulla with the spinal cord, 90% of the axons in the left

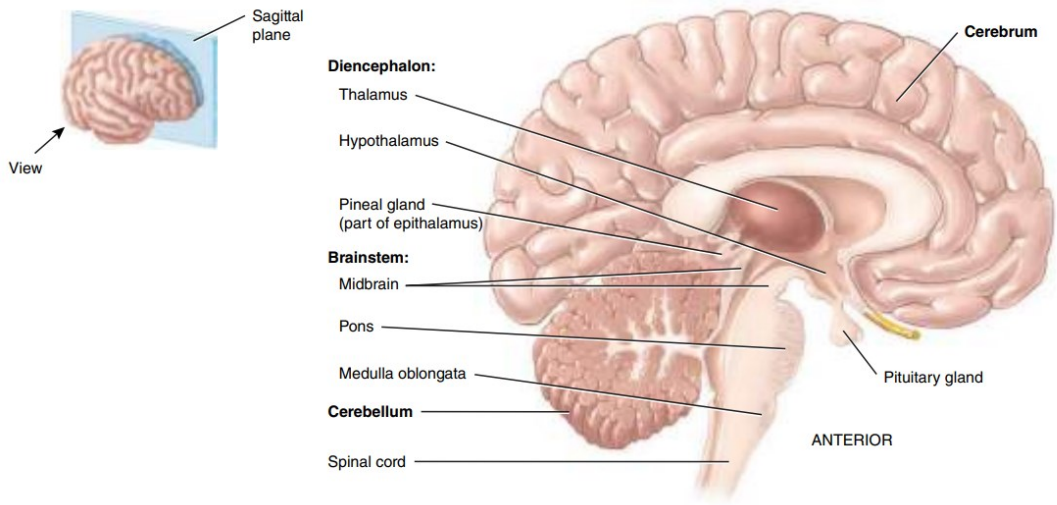


Figure 5. The four principal parts of the brain are the brainstem, cerebellum, diencephalon, and cerebrum [3].

pyramid cross to the right side, and 90% of the axons in the right pyramid cross to the left side. This crossing is called the decussation of pyramids and explains why each side of the brain controls voluntary movements on the opposite side of the body [3].

The pons lies directly superior to the medulla and anterior to the cerebellum. It provides a connection between the cortex of a cerebral hemisphere and that of the opposite hemisphere of the cerebellum. This complex circuitry plays an essential role in coordinating and maximizing the efficiency of voluntary motor output throughout the body. It also contains ascending and descending tracts along with the nuclei of cranial nerves.

The midbrain or mesencephalon extends from the pons to the diencephalon. Like the medulla and the pons, the midbrain contains both nuclei and tracts. Some of the nuclei presented in the midbrain formed the left and right substantia nigra (Figure 6). In particular, neurons that release dopamine, extending from the substantia nigra to the basal nuclei (also called basal ganglia), help control subconscious muscle activities. Loss of these neurons is associated with PD. Furthermore, in the midbrain there are also present the left and right red nuclei, which look reddish due to their rich blood supply in their neuronal cell bodies. Axons from the cerebellum and cerebral cortex form synapses in the red nuclei, which help control muscular movements.

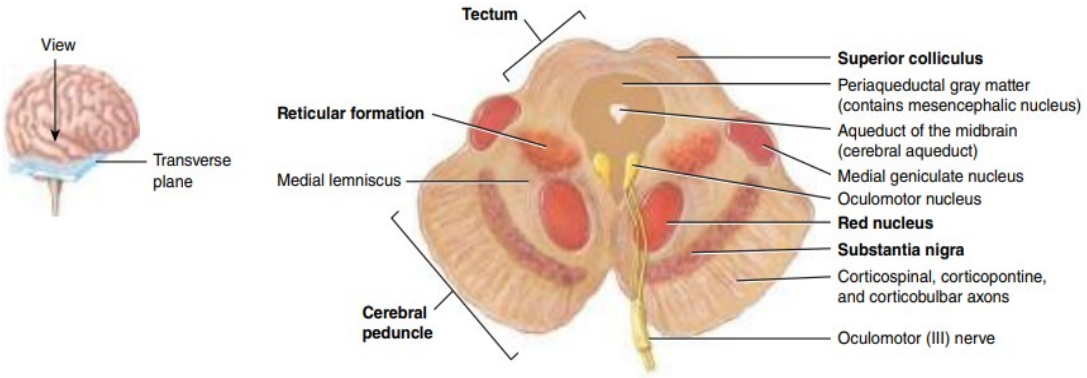


Figure 6. Anterior, transversal section of the midbrain. On the left column there are evidenced the principal parts of the midbrain [3].

The cerebellum is the second in size after the cerebrum and as the last one has a highly folded surface that greatly increases the surface area of its outer grey matter cortex, allowing for a greater number of neurons. The cerebellum contains nearly half of the neurons in the brain. It is posterior to the medulla and pons and inferior to the posterior portion of the cerebrum and it is divided in lobes: the anterior lobe and posterior lobe govern subconscious aspects of skeletal muscle movements; the flocculonodular lobe on the inferior surface contributes to equilibrium and balance.

Three paired cerebellar peduncles attach the cerebellum to the brainstem. The superior cerebellar peduncles contain axons that extend from the cerebellum to the red nuclei of the midbrain and to several nuclei of the thalamus.

The primary function of the cerebellum is to evaluate how well movements initiated by motor areas in the cerebrum are actually being carried out. When movements initiated by the cerebral motor areas are not being carried out correctly, the cerebellum detects the differences, it sends feedback signals to motor areas of the cerebral cortex. The feedback signals help correct the errors, smooth the movements, and coordinate complex sequences of skeletal muscle contractions. Aside from this coordination of movements, the cerebellum is the main brain region that regulates posture and balance. The presence of reciprocal connections between the cerebellum and association areas of the cerebral cortex suggests that the cerebellum may also have nonmotor functions such as cognition and language processing.

Damage to the cerebellum can result in a loss of ability to coordinate muscular movements, a condition called ataxia. Ataxia can also occur as a result of degenerative diseases like PD [3].

The first component of the diencephalon is the thalamus which is the major relay station for most sensory impulses that reach the primary sensory areas of the cerebral cortex from the spinal cord and brainstem. In addition, the thalamus contributes to motor functions by transmitting information from the cerebellum and basal nuclei to the primary motor area of the cerebral cortex. The thalamus also relays nerve impulses between different areas of the cerebrum and plays a role in the maintenance of consciousness.

The second part of the diencephalon is hypothalamus which is located inferior to the thalamus and controls many body activities and is one of the major regulators of homeostasis. The hypothalamus controls and integrates activities of the autonomic nervous system, which regulates contraction of smooth muscle and cardiac muscle and the secretions of many glands.

The epithalamus is a small region superior and posterior to the thalamus, consists of the pineal gland and habenular nuclei.

The cerebrum consists of an outer cerebral cortex, an internal region of cerebral white matter, and grey matter nuclei deep within the white matter. In particular, the cerebral cortex is a region of grey matter that forms the outer border of the cerebrum. Each cerebral hemisphere can be further subdivided into several lobes. The lobes are named after the bones that cover them: frontal, parietal, temporal, and occipital lobes. Specific types of sensory, motor, and integrative signals are processed in certain regions of the cerebral cortex as showed in Figure 7.

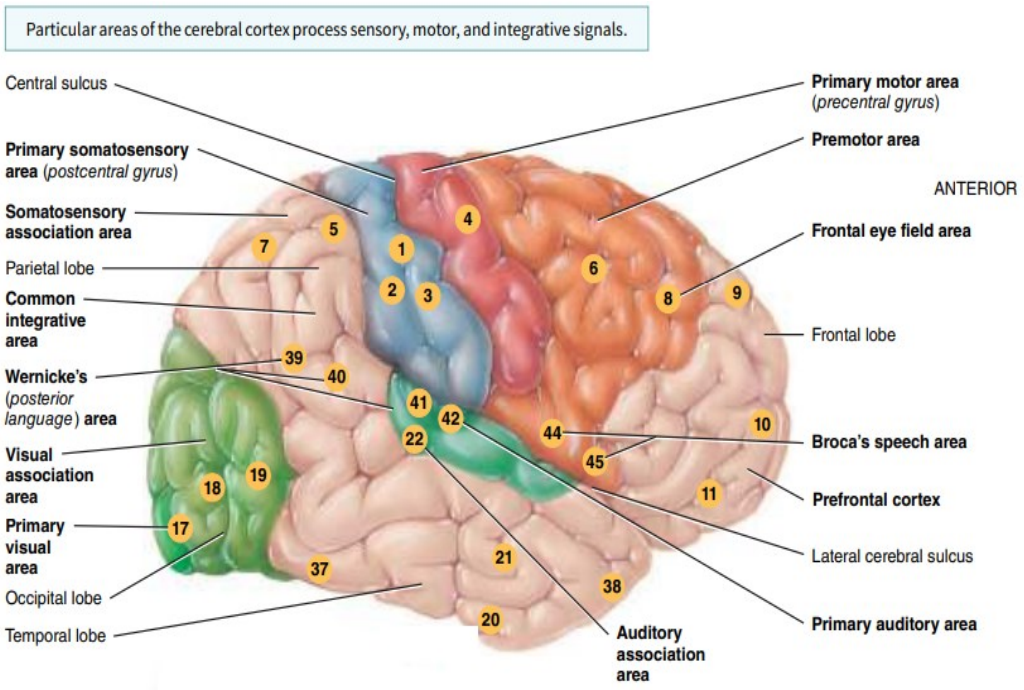


Figure 7. Functional areas of the cerebrum. The numbers, still used today, are from K. Brodmann's map of the cerebral cortex, first published in 1909 [3].

Deep within each cerebral hemisphere are three nuclei (masses of grey matter), that are collectively termed the basal nuclei. Two of the basal nuclei lie beside each other, lateral to the thalamus. They are the globus and the putamen. Together, the globus pallidus and putamen are referred to as the lentiform nucleus. The third of the basal nuclei is the caudate nucleus. Together, the lentiform and caudate nuclei are known as the corpus striatum. Nearby structures that are functionally linked to the basal nuclei are the substantia nigra of the midbrain and the subthalamic nuclei of the diencephalon. Axons from the substantia nigra terminate in the caudate nucleus and putamen.

A major function of the basal nuclei is to help regulate initiation and termination of movements. Activity of neurons in the putamen precedes or anticipates body movements. In addition to influencing motor functions, the basal nuclei have other roles. They help initiate and terminate some cognitive processes, attention, memory, and planning, and may act with the limbic system to regulate emotional behaviours. Disorders such as PD is supposed to involve dysfunction of circuits between the basal nuclei and the limbic system [3].

The basal nuclei suppress unwanted movements and prepare motor neuron circuits for the initiation of movements.

One important functional area is the primary motor area (area 4) that contain a “map” of the entire body. Each region within the area controls voluntary contractions of specific muscles or groups of muscles. Different muscles are represented unequally in the primary motor area. More cortical area is devoted to those muscles involved in skilled, complex, or delicate movement. For instance, the cortical region devoted to muscles that move the fingers is much larger than the region for muscles that move the toes. This distorted muscle map of the body is called the motor homunculus.

Another important functional area that is concerned with the motor association is the premotor area (area 6) that is immediately anterior to the primary motor area. Neurons in this area communicate with the primary motor cortex, the basal nuclei, and the thalamus. The premotor area deals with learned motor activities of a complex and sequential nature, and it also serves as a memory bank for such movements [3].

1.2.2. Peripheral Nervous System

The motor portion of the peripheral nervous system consists of two components. The motor axons that connect the brain and spinal cord to skeletal muscles make up the somatic motor division of the peripheral nervous system, where the cells and axons that innervate smooth muscles, cardiac muscle, and glands make up the visceral or autonomic motor division (Figure 4.).

The PNS can be divided in:

1) Somatic Nervous System (SNS) conveys output from the CNS to skeletal muscles only. Because its motor responses can be consciously controlled, the action of this part of the PNS is voluntary;
2) Autonomic Nervous System (ANS) conveys output from the CNS to smooth muscle, cardiac muscle, and glands. Because its motor responses are not normally under conscious control, the action of the ANS is involuntary. The ANS is comprised of two main branches, the sympathetic nervous system and the parasympathetic nervous system. With a few exceptions, effectors receive innervation from both of these branches, and usually the two branches have opposing actions.

Neural circuits in the brain and spinal cord orchestrate all voluntary movements. Ultimately, all excitatory and inhibitory signals that control movement converge on the motor neurons

that extend out of the brainstem and spinal cord to innervate skeletal muscles in the body. These neurons are known as lower motor neurons (LMNs) because they have their cell bodies in the lower parts of the CNS (brainstem and spinal cord). From the brainstem, axons of LMNs extend through cranial nerves to innervate skeletal muscles of the face and head. From the spinal cord, axons of LMNs extend through spinal nerves to innervate skeletal muscles of the limbs and trunk. Only LMNs provide output from the CNS to skeletal muscle fibres. For this reason, they are also called the final common pathway. Neurons in four distinct but highly interactive neural circuits participate in control of movement by providing input to lower motor neurons (Figure 8.).

The axons of upper motor neurons extend from the brain to lower motor neurons via two types of pathways—direct and indirect. Direct motor pathways provide input to lower motor neurons via axons that extend directly from the cerebral cortex. Indirect motor pathways provide input to lower motor neurons from motor centres in the brainstem. Direct and indirect pathways both govern generation of action potentials in the lower motor neurons, the neurons that stimulate contraction of skeletal muscles.

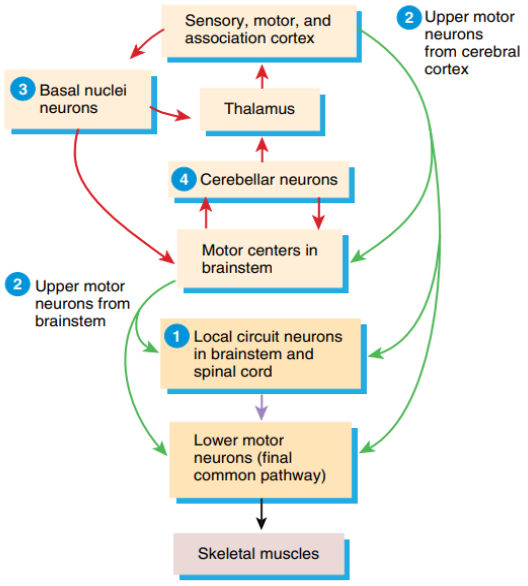


Figure 8. Neural circuits that regulate lower motor neurons. Lower motor neurons receive input directly from 1 local circuit neurons (purple arrow) and 2 upper motor neurons in the cerebral cortex and brainstem (green arrows). Neural circuits involving 3 basal nuclei neurons and 4 cerebellar neurons regulate activity of upper motor neurons (red arrows).

2. *Lewy bodies in neurodegenerative diseases*

Neurological disorders are the leading cause of physical and cognitive disability across the globe, affecting approximately 15% of the worldwide population[5]. In particular, according to the World Health Organization (WHO) disability and death due to PD are increasing faster than for any other neurological disorder. Globally, the prevalence of PD has doubled in the past 25 years with global estimates in 2019 showing over 8.5 million individuals living with PD. Current estimates suggest that, in 2019, PD resulted in 5.8 million disability-adjusted life years, an increase of 81% since 2000, and caused 329,000 deaths, an increase of over 100% since 2000.

Even though PD has a long and rich history, the first complete clinical description of the disease was provided by James Parkinson in 1817 [6]. The most consistent pathological lesion of PD was identified as severe nerve cell loss in the pigmented pars compacta (part of the brainstem) of the substantia nigra with the presence there and in other brainstem regions of what Konstantin Tretiakoff called “corps de Lewy” or Lewy bodies (Figure 9.) [5].

As explained in the first chapter the one that regulates the synaptic vesicle trafficking and subsequently neurotransmitter release in order to guarantee the communication in the neurons is the alpha-synuclein (aSyn). In PD, insoluble forms of aSyn accumulate as inclusions

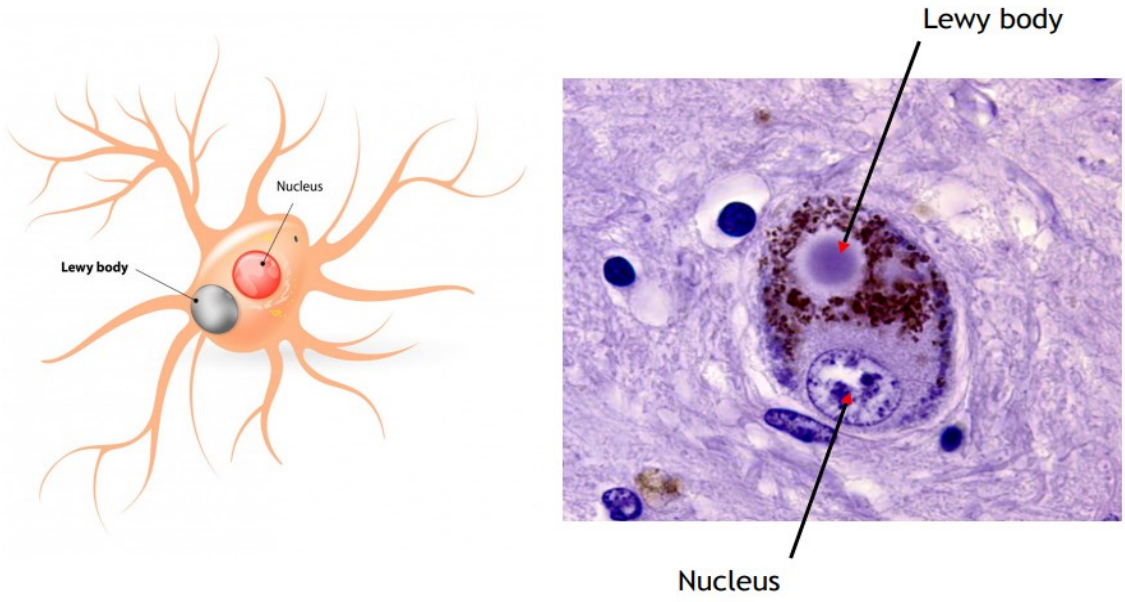


Figure 9. Lewy body in a substantia nigra neuron.

in Lewy bodies. α-Synuclein is a neuronal protein, predominantly located in the presynaptic terminal [7]. Although the process of Lewy body formation is not fully understood a hypothetical scheme can be observed in Figure 10.

As a means of better understanding the progression of the Lewy Bodies pathology, in 2003 the Braak staging system was proposed (Figure 11.). According to this labelling, the lesions begin in the lower brainstem and involve the intermediate reticular zone, anterior olfactory nucleus, and nucleus basalis membrane (NBM), with midbrain regions being preserved (stage 1). It continues progressing into the caudal raphe nuclei (nuclei found in brainstem) and in the gigantocellular reticular nucleus of the medulla oblongata (stage 2). This stage is considered asymptomatic or pre-symptomatic and may explain early non-motor (autonomic and olfactory) symptoms in PD and DLB, preceding much of the motor and sensory dysfunctions. During stage 3 the disease entered in the substantia nigra and Lewy body lesions begin to form in the pars compacta. Finally, stage 4 is characterized by the further damage to the neocortex

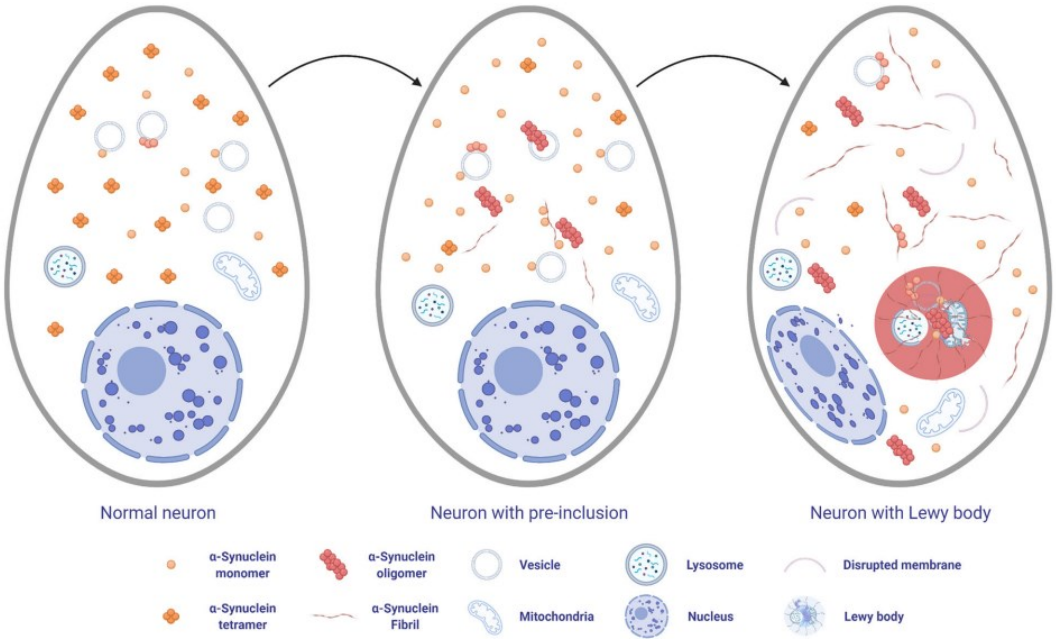


Figure 10. Hypothetical scheme of Lewy body formation in neurons. α-Synuclein exists in equilibrium between monomers and tetramers. Under pathologic conditions, the tetramer-monomer ratio decreases. α-synuclein tend to oligomerize, and toxic oligomers can disrupt membranes. Disrupted membranes, organelles, α-synuclein oligomers and fibrils are involved in the Lewy body formation [8].

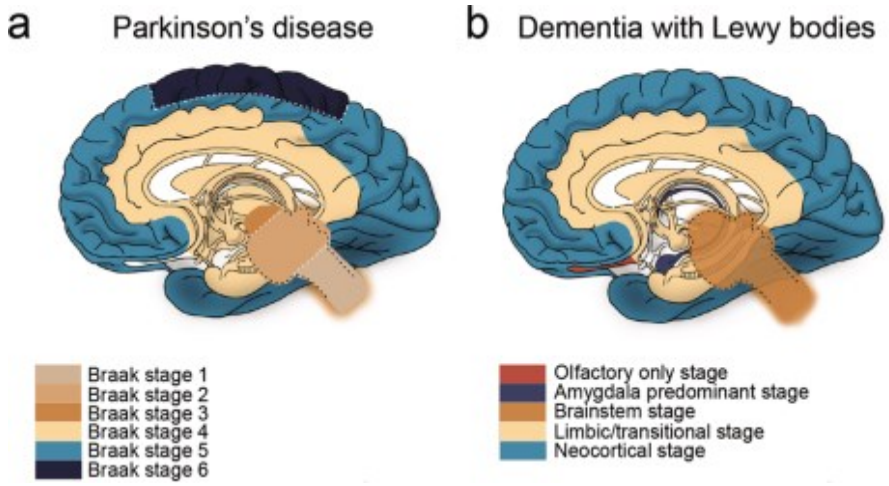


Figure 11. Braak stages for Parkinson's Disease (a) and the different stages for Dementia with Lewy bodies (b).

and can be correlated with the symptomatic stages of PD. Once the abnormal protein inclusions reach stages 5 and 6, it is spread through the cortex with devastating effects on complex sensory association cortex regions and prefrontal and temporal areas (e.g. hippocampus) involved in executive brain functions, and learning and memory[9].

Lewy body disease falls into three major clinicopathologic subtypes: PD, PDD and DLB. Although there are differences between studies, up to 83% of patients with PD eventually develop dementia later in the disease course.

2.1. Parkinson's Disease

Parkinson's disease was first described by Dr. James Parkinson in 1817 as a "shaking palsy." The disease consists of a chronic, progressive neurodegenerative process characterized by both motor and nonmotor symptoms. The term *parkinsonism* is a symptom complex used to describe the motor features of PD, which include resting tremor, bradykinesia, and muscular rigidity. PD is the most common cause of parkinsonism, although a number of secondary causes also exist, including diseases that mimic PD and drug-induced causes [10].

In Parkinson disease, pigmented neurons of the substantia nigra (Figure 12.) and other brainstem dopaminergic cell groups degenerate. Loss of substantia nigra neurons results in reduction of dopamine in the dorsal aspect of the putamen (part of the basal nuclei) and causes many of the motor manifestations of Parkinson disease. Parkinson may have a genetic correlation as about 10% of patients have a family history of this disease [11].

Predominant brainstem Lewy pathology is seen in PD, whereas more diffuse Lewy pathology involving the brainstem, limbic and neocortical regions is typical of DLB and PDD.

The muscular stiffness that occurs in PD is due to degeneration of neurons that release dopamine. Involuntary skeletal muscle contractions often interfere with voluntary movement. As example, the muscles of the upper limb may alternately contract and relax, causing the hand to shake. This shaking, called tremor, is the most common symptom of PD. Also, muscle tone may increase greatly, causing rigidity of the involved body part [3].

Parkinson's disease is the result of neurons in the substantia nigra pars compacta dying. These neurons release dopamine into the corpus striatum. Without that modulatory influence, the basal nuclei are stuck in the indirect pathway, without the direct pathway being activated. The direct pathway is responsible for increasing cortical movement commands. The increased activity of the indirect pathway results in the hypokinetic disorder of PD.

In summary, PD affects many areas of the nervous system and different types of neurons. However, much focus has been put on neurons in a region of the midbrain called the substantia nigra pars compacta. This region forms part of a major pathway in the brain that's critical for facilitating movements. As explained before in PD, dopaminergic neurons in the substantia nigra gradually die, leading to the malfunction of this pathway and the characteristic motor problems. As dopamine neurons are lost microglia, is thought to take up the resulting cellular debris, triggering an immune response. Once activated they release inflammatory cytokines

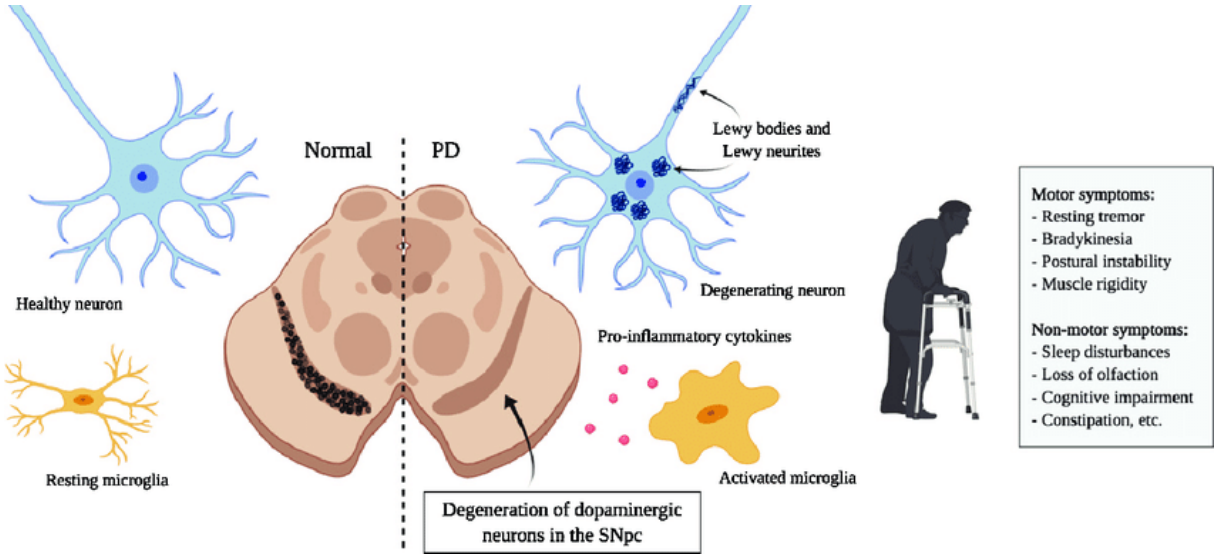


Figure 12. Principal hallmarks of Parkinson's disease[12]

which activate neighbouring microglia and astrocytes. Chemicals released by activated microglia and astrocytes have been shown to injure neurons. With time, more areas of the nervous system develop the pathology. In particular is thought that the Lewy bodies pass from one neuron to another.

2.2. Dementia with Lewy Body

Dementia with Lewy bodies is a heterogeneous neurodegenerative disease in which alpha-synuclein (the precursors of the Lewy Bodies) is the main pathological trait. However, concomitant Alzheimer's disease (AD) and cerebrovascular disease are common in DLB, contributing to disease heterogeneity [13].

In the clinical setting, essential for a diagnosis of DLB is dementia, defined as a progressive cognitive decline of sufficient magnitude to interfere with normal social or occupational functions, or with usual daily activities. Prominent or persistent memory impairment may not necessarily occur in the early stages but is usually evident with progression. Deficits on tests of attention, executive function, and visuo-perceptual (visuospatial) ability may be especially prominent and occur early. Hallucinations in DLB may occur spontaneously, independent of visuospatial and perception impairment, and possibly related to Lewy Bodies in the temporal lobe[14]. However, this disease is often misdiagnosed and consequently, patients are prone to non-beneficial or even harmful cures [15].

Research suggests that median survival after DLB diagnosis is 3-4 years, but there are rapidly and slowly progressive forms. Most individuals with DLB die of complications of the disease. Comparing the rates of detection of DLB in autopsy studies to those diagnosed while in clinical care indicates that as many as one in three diagnoses of DLB may be missed [16]. The extent of Lewy body neuronal damage is a key determinant of dementia in the Lewy body disorders.

In contrast to AD, in which the hippocampus is among the first brain structures affected, and episodic memory loss is typically the earliest symptom, memory impairment occurs later in DLB (Figure 13.).

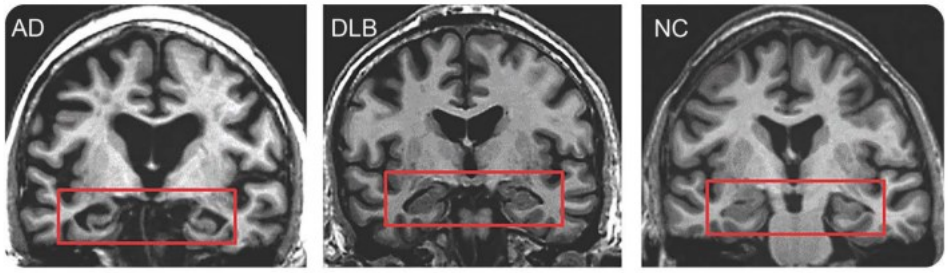


Figure 13. The medial temporal lobe volume is relatively preserved in DLB, which is similar to normal controls (NC), whereas atrophy is obvious in AD [17] .

2.3. Parkinson's Disease Dementia

The condition in which PD patients without dementia have cognitive deficits is called mild cognitive impairment (PD-MCI) and usually the patients affected by it are developing dementia. Early cognitive deficits PD may represent the earliest manifestation of dementia, providing an optimum intervention period to slow or even prevent the manifestation of dementia.

PDD subjects are those identified as patients having motor symptoms for a 12-month period or more prior to the development of features of DLB.

Patients with dementia demonstrate impaired attention, quantified by cognitive reaction time and vigilance, and accompanied by fluctuations in attention.

In DLB, cognitive decline either antedates or occurs simultaneously with parkinsonism, whereas in PDD it follows the constellation of parkinsonism.

In particular PDD may be considered as the last two stages of Braak, and it affects the cortex with devastating effects on complex sensory association cortex regions and prefrontal and temporal areas involved in executive brain functions and learning and memory. The subjects affected by PDD have an expected life from 5 to 7 years, although this can vary from patient to patient.

2.4. Parkinson's Disease and Lewy Body Disease Spectrum

Currently, the neurological community is facing a classification problem in the group of intraneuronal synucleinopathies, which covers the proper clinical and differential diagnosis among PD, PDD, and DLB.

Lewy body disease (LBD) is a spectrum of disorders characterized pathologically by alpha-synuclein inclusions in the brainstem, subcortical nuclei, limbic and neocortical areas, and clinically by attentional disturbance, Parkinsonism, dementia, and visual hallucinations. Two clinical diagnoses within the LBD spectrum are DLB and PDD. Since the two syndromes present considerable clinical overlap, it has been argued that DLB and PDD may represent the same disease entity [18].

The pathological delineation of PD compared to PDD/DLB lies in the stage of progression of pathology through the brain; PD patients have inclusions restricted to the brainstem and limbic regions, whilst in PDD and DLB patients Lewy body pathology extends to the neocortex. While a significant difference between the DLB and PDD is the location of the Lewy bodies in the brain. In Parkinson's they are found mainly in the substantia nigra which is in the midbrain, whereas in DLB they are more widely distributed throughout the cerebral cortex. Clinically DLB closely resembles PDD. If the physical symptoms precede the cognitive symptoms by one year, a diagnosis of PDD will be made; if the onset of cognitive symptoms precedes or starts at the same time the physical symptoms start it is considered to be DLB.

Parkinson's disease dementia would be the diagnosis when PD is well established before the dementia occurs (*one-year rule*).

In some cases, PDD exhibited bilateral loss of grey matter in the occipital lobe compared to PD. The different propagation of Lewy bodies in PD and PDD can be observed in Figure 14.

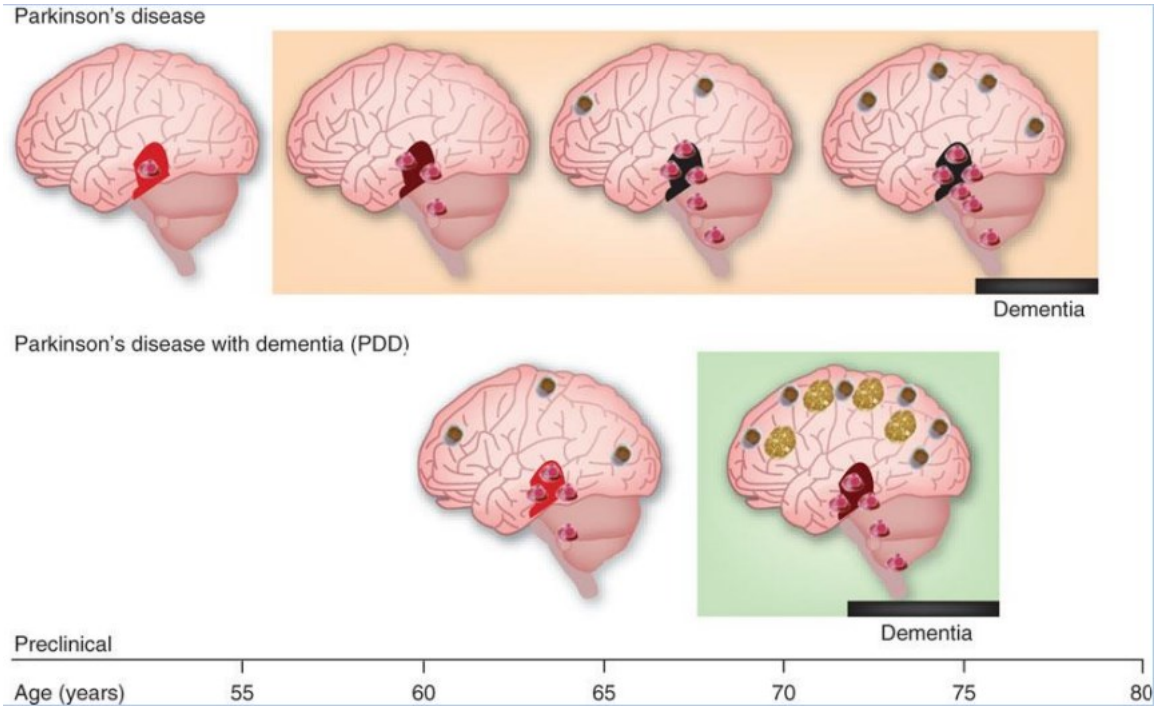


Figure 14. Propagation of Lewy Bodies in PD and PDD. The age represented on the axe is just an illustration of the majority of cases, there are many younger subjects affected by these diseases.

3. Magnetic Resonance Imaging

Humans are multicellular organisms with high complexity. The structure and function of the human body have been a matter of great interest for millennia and over the last few centuries, considerable progress has been made in the understanding of these aspects. It is partly due to several new techniques that have made possible the increase of the knowledge concerning the human body. One technique that has improved considerable the understanding of the human being is the Magnetic Resonance Imaging (MRI), a non-invasive imaging method that has recently been widely used in medical imaging.

Comparing MRI to other useful imaging modalities such as planar x-ray, x-ray computed tomography (CT), positron emission tomography (PET) and single photon emission computed tomography (SPECT), MRI does not use ionizing radiation to form an image. Magnetic Resonance Imaging uses the magnetic properties of water (over 60% of the human body) to form an image. The properties associated with water are influenced by the environment and change sensitively with degeneration and pathology; and in many cases, MRI can be used to detect these alterations.

Structural brain imaging of PD subjects via magnetic MRI may provide a robust in vivo method to identify macroscopic cerebral atrophy. As the contribution to cognitive decline in PD includes alpha-synuclein (as explained above: precursor of the Lewy Bodies) accumulation, which can be associated with structural brain changes.

3.1. Basic Physic

Medical Magnetic Resonance imaging uses the signal from the nuclei of hydrogen atoms (H) (from the water molecules) for image generation. A hydrogen atom consists of a nucleus containing a single proton and a single electron. Apart from its positive charge, the proton possesses spin (Figure 15.), an intrinsic property of nearly all elementary particles making the proton to rotate around its axis.

As a rotating mass (m), the proton has a magnetic moment (B) (that makes it behaves like a small magnet) and an angular momentum. The net resulting magnetization (M) is the sum of all the elementary moments.

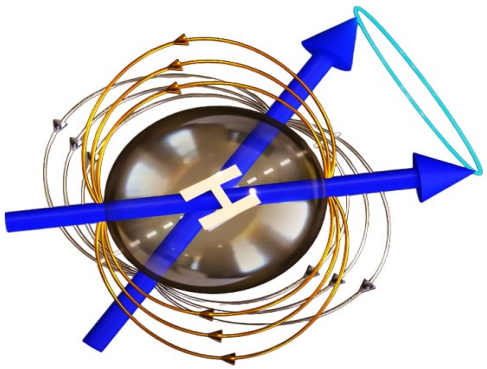


Figure 15. The spin of a Hydrogen atom.

From a macroscopic point of view, no resulting field can be observed directly since each spin has its own, independent, random orientation. However, when placed in a powerful external magnetic field (B0), the spin directions will line up parallel or anti-parallel to the primary magnetic field with a small majority aligning with the direction of the primary magnetic field. More precisely, each spin rotates within a cone around B0 (spin precession). The frequency of rotation, called the Larmor frequency, is related to the magnetic field B0 through the gyromagnetic ratio γ unique to each element, by the Larmor equation (1), where ω_o is the angular precession frequency.

$$\omega_o = \gamma B_o \qquad (1)$$

With the protons aligned with the main magnetic field they can be influenced by using externally applied radio frequency (RF) or RF pulses: when this take place, the protons are putted down into an alternate plane and also precessed together in phase. The angle depends on the strength and duration of the RF pulse. Putting the protons down into another plane makes a change in their longitudinal magnetization. Normally the majority of protons are going with the flow and following the direction of the external magnetic field but applying an extra energy called excitation, protons have the ability to go against the current and instead orient themselves in the opposite direction against that of the magnetic field: this is called anti-parallel. The protons will also process together in phase and this synchronization is called the transverse magnetization of the protons.

The net magnetization M can be decomposed into two components (Figure 16.): a longitudinal component (Mz), parallel to B0 and a transverse component (Mxy), orthogonal to B0. At

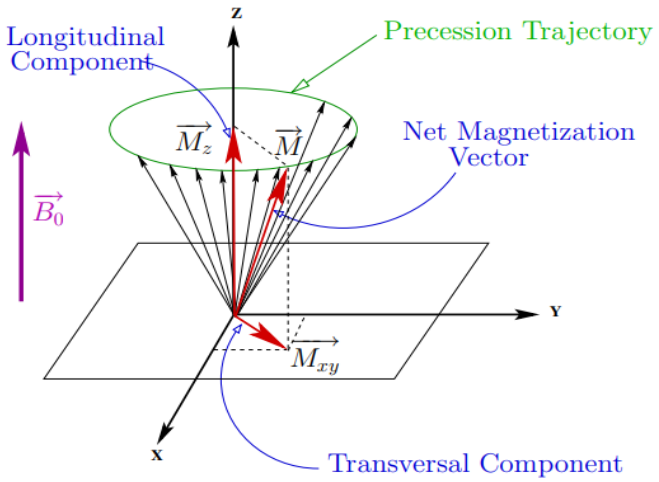


Figure 16. The net magnetization vector M, decomposed into a longitudinal component Mz and a transverse component Mxy.

equilibrium, after a sufficient exposure time to B0, the transverse component Mxy vanishes. All the individual spins are indeed precessing, but they are all out of phase with each other.

It is possible to perturb the difference in the number of atoms between the two energy states by applying an oscillating electromagnetic (RF) pulse toward the area of the body to be examined. This method involves applying a much weaker field than B0 at the Larmor frequency of the target nuclei and to apply it through a rotating reference frame orthogonal to B0. This causes the particles in that area to spin in a different direction and move from a lower (parallel direction) to a higher (antiparallel direction) energy state as a result of absorbing energy. After the excitation the energy source is removed, and some protons will return to the low energy state (parallel direction) releasing RF energy at the same frequency. This emitted energy is detected by highly sensitive antennas to capture the MR signal.

Application of RF excitation pulse (B1) synchronized to the precessional frequency of the protons causes absorption of energy and displacement of the sample magnetic moment from equilibrium conditions. The resonance frequency corresponds to the energy separation between the protons in the parallel and antiparallel directions. Protons oriented parallel and antiparallel to the external magnetic field are separated by an energy gap ΔE. These protons will go from the low to the high energy level only when the RF pulse is equal to the precessional frequency, so Mz changes from the maximal positive value at equilibrium, through zero, to the

maximal negative value. The degree of Mz rotation by the B1 field as it is applied along the x-axis or along the y-axis perpendicular to Mz is also called the flip angle. A torque is applied on Mz, rotating it from the longitudinal direction into the transverse plane. The rate of rotation occurs at an angular frequency equal to:

$$\omega_1 = \gamma B1 \qquad (2)$$

as for the Larmor equation. When a RF pulse (B1 field) is applied over a time t, the magnetization vector displacement angle ϑ is determined as:

$$\vartheta = \omega_1 t = \gamma B1 t \qquad (3)$$

and the product of the pulse time and B1 amplitude determines the displacement of Mz. So flip angles describe the rotation of Mz away from the z-axis, small flip angles (less than 45 degrees) produce small transverse magnetization and large flip angles (75 to 90 degrees) produce large transverse magnetization.

The flip angle induced by an RF pulse depends on the strength and duration of the pulse and as explained above, the vector flipped is the net magnetization vector. At the beginning of a standard spin echo sequence usually there is the application of a 90-degree pulse: this means that after the RF pulse has been applied the net magnetization vector is perpendicular to its original orientation, this orientation is achieved by eliminating longitudinal magnetization and generating a transverse magnetization vector by synchronizing proton precession. During recovery longitudinal magnetization increases and transverse magnetization decreases the protons diphase: this looks like a spiralling of the net magnetic vector along the z axis (Figure 17.). This spiralling of the net magnetization vector induces an electrical signal by a process called free induction decay. The recovery of the longitudinal magnetization of a proton occurs exponentially the point at which 63 percent of the longitudinal magnetization has been recovered is called the T1 time; the time at which 63 percent of the transverse magnetization has been lost is called the T2 time. The T1 and T2 time is unique to each tissue type and with this the MRI sequences can be adjusted to highlight the differences (weighting).

Careful manipulation of magnetic field gradients and radiofrequency pulses make it possible to construct extraordinarily detailed images of the brain at any location and orientation with good resolution. The strong magnetic field and radiofrequency pulses used in MRI scanning

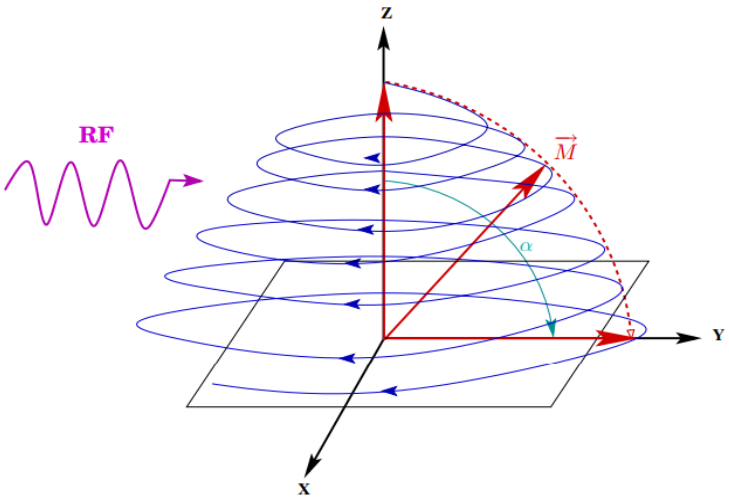


Figure 17. Excitation phase: the energy given by the RF pulse flips the net magnetization vector M of an angle α (here α = 90◦).

are harmless, making this technique completely non-invasive. Interesting aspect is the fact that by changing the scanning parameters, images based on a wide variety of different contrast mechanisms can be generated. For example, conventional MR images take advantage of the fact that hydrogen in different types of tissue (e.g., gray matter, white matter, cerebrospinal fluid) have slightly different realignment rates, meaning that soft tissue contrast can be manipulated simply by adjusting when the realigning hydrogen signal is measured. Different parameter settings can also be used to generate images in which gray and white matter are invisible but in which the brain vasculature stands out in sharp detail [1].

3.1.1. T1: Spin-lattice interaction

The spin lattice relaxation is based on the energy exchange between protons and surrounding molecules. This energy dissipation is characterized by the restoration of the longitudinal component to its equilibrium value. This recovery process is modelled by an exponential function characterized by a time constant T1 (Figure 18.).

For solid and slowly moving structures, the hydration layer permits only low-frequency molecular tumbling frequencies so there is almost no spectral overlap with the Larmor frequency. For unstructured tissues and fluids in bulk water, there is only a small spectral overlap with the tumbling frequencies. In each of these situations, release of energy is constrained and T1 relaxation time is long. For moderately sized proteins and fatty tissues,

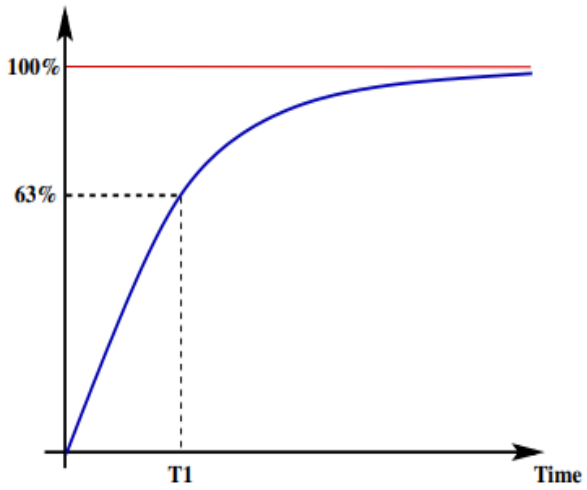


Figure 18. Spin lattice relaxation describes the longitudinal component recovery as a function of time and is characterized by the T1 constant.

molecular tumbling frequencies are most conducive to spin-lattice relaxation because of a larger overlap with the Larmor frequency and result in a relatively short T1 relaxation time. The T1 time is strongly dependent on the physical characteristics of the tissues and their associated hydration layers with values in the range of 0.1 to 1 s for soft tissues, and 1 to 4 s in aqueous tissues. When there is an increase of the main magnetic field strength, there is also an increase in the Larmor precessional frequency which causes a decrease in the overlap with the molecular tumbling frequencies and a longer T1 recovery time[19].

T1-weighted image is one of the basic pulse sequences in MRI and demonstrates differences in the T1 relaxation times of tissues. A T1-weighted image relies upon the longitudinal relaxation of a tissue's net magnetization vector. Basically, not all tissues get back to equilibrium equally quickly, and a tissue's T1 reflects the amount of time its protons spins realign with the B0. Fat quickly realigns its longitudinal magnetization with the B0, and it therefore appears bright on a T1-weighted image. Conversely, water has much slower longitudinal magnetization realignment after an RF pulse, and therefore has less transverse magnetization after a RF pulse. Thus, water has low signal and appears dark. If T1-weighted images didn't have short time of repetition (TR), then all the protons would recover their alignment with the main magnetic field and the image would be uniformly intense. Selecting

a TR shorter than the tissue's recovery time allows one to differentiate them, such as with tissue contrast.

3.1.2. T2: Spin-spin interaction

Spin-spin relaxation refers to the loss of net magnetization in the transverse plane related to protons dephasing. Spins do not only give up their energy to surrounding molecules but also to other neighbouring nonexcited spins. This process is also modelled by an exponential function characterized by a time constant T2.

T2-weighted scans differentiate anatomical structures mainly on the basis of T2 values, for example the scanning parameters are set (long TR/long time of echo (TE)) to minimize T1 relaxation effects.

This dephasing is actually further increased by local magnetic field inhomogeneities, since the Larmor frequency will also be nonuniform throughout the region. A time constant slightly different to T2, T2*, is therefore used. The transverse component induces a current in a coil, known as Free Induction Decay (FID). The T2* constant can be evaluated through the convex envelop of the FID curve.

3.1.3. Data acquisition

The image obtained with MRI is built thanks to a repetition of a series of events that allow to sample the volume of interest. For this process there is the need of the: TR that is the period between B1 excitation pulses. During the TR interval, T2 decay and T1 recovery occur in the tissues. Then there is the TE, that is the time between the excitation pulse with the B1 RF pulse and the appearance of the peak amplitude of an induced echo, which is determined by applying a 180-degree RF inversion pulse or gradient polarity reversal at a time equal to TE/2; the time of inversion (TI), that is the time between an initial inversion/excitation (180 degrees) RF pulse that produces maximum tissue saturation, and a 90-degree pulse. During the TI, Mz recovery occurs. The record pulse converts the recovered Mz into Mxy, which is then measured with the formation of an echo at time TE.

Saturation is a state of tissue magnetization. At equilibrium, the protons in a material are unsaturated with full Mz amplitude. The first excitation (B1) pulse in the sequence produces the largest transverse magnetization and recovery of the longitudinal magnetization occurs at the T1 time constant over the TR interval. Because the TR is less than at least five times the T1

of the sample, Mz recovery is incomplete and consequently less Mxy amplitude is generated in the second excitation pulse. The steady state equilibrium is reached after the third pulse. In the steady state equilibrium, the amount of Mz recovery and Mxy signal amplitude are constant, and the tissues achieve a state of partial saturation. Tissues with long T1 experience a greater partial saturation than do tissues with short T1.

It is possible to acquire data with different pulse sequences. The principal sequences are three: gradient echo (GE), inversion recovery (IR), spin echo (SE). The combination of one sequence with a spatial localization method provides “contrast-weighted” images.

The relaxation of both longitudinal and transverse magnetization in different tissues provides the basic contrast for image formation in MRI. The FID contains this information, but it is never measured directly. Instead “echoes” are produced using a combination of RF pulses and magnetic field gradients. Spin echoes (produced by RF pulses) and gradient echoes (produced by magnetic field gradients) both rephase spins that have lost phase coherence. When spins are refocused and come into phase with each other, an echo is produced. This echo signal is recorded and used for image reconstruction.

In the MRI the data acquired during the scanning are data points that fill the k-space and will form the final image with specific mathematical tool. The acquisition of the data is done filling up the k-space. The k-space is a matrix of numbers representing spatial frequencies signals (a measure of how often sinusoidal components, determined by the Fourier transform, repeated themselves per unit of distance) acquired during the MRI scanning, during the evolution and decay of the echo. The k-space represents what is acquired with the MRI, without any mathematical process: it is only with the use of the inverse Fourier transformation that is possible to reconstruct the image. Overall, it is possible to see the cells, which make up the matrix, putted on a rectangular grid with principal axes kx (along the rows) and ky (along the columns). Each axis is symmetric in the centre of k-space, ranging from -f to +f along the rows and the columns. The individual points (kx, ky) contain information about spatial frequency and phase information of every pixel in the final image, so they do not correspond one-to-one with the individual pixels. When k-space is full (at the end of the scan) the data are mathematically processed with the inverse Fourier transformation to produce a final image.

Field of View (FOV) and k-space are strictly connected, and both define the resolution of the image. FOV is the distance in mm or cm in which the image is displayed, and it is divided into hundreds of elements called pixels. Both FOV and pixels width determine the number of digitized samples in the k-space that must be obtained to reconstruct an image with the desired resolution. FOV is inverse proportional to the spaces between the samples in k-space.

The greyscale image is obtained performing the inverse two-dimensional Fourier transformation, by sequentially applying the one-dimensional transforms for each row, and then for each column. The time to acquire an image is determined by the data needed to fill the fraction of k-space that allows the image to be reconstructed.

Most of the MRI scans involve the construction of a three-dimensional image from a set of two-dimensional slices with specific thickness, a three-dimensional image acquisition is performed, in which the anatomical structures are split into a number of volume elements, the voxels. It is possible to regulate the thickness of each slices: an increment of the slice thickness will cause the collection of more different tissues in the 2D slice, this will cause blurring which will decrease the spatial resolution in the image. In the end, there are no more pixels that compose the k-space (pixels have a 2D dimensions), but 3D elements, called voxels. Three-dimensional image acquisition requires the use of a “slab-selective” RF pulse to excite a large volume of protons simultaneously; then two-phase gradients are applied in the slice encode and phase encode directions.

An inverse three-dimensional Fourier transformation is applied for each column, row, and depth axis in the image matrix. Volumes that are obtained can be either isotropic, voxels have the same size in all three directions, or anisotropic where at least one dimension is different in size.

3.1.4. MRI components

A tube-like structure, the bore, holds the components of the MRI machine:

- a magnet, which produces a uniform and STATIC magnetic field $B0$;
- gradient coils that produce a gradient of the static magnetic field;
- RF coils that transmit and receive signal from the body.

Spatial localization is essential for creating MR images and determining the location of discrete sample volumes for MR spectroscopy. This is achieved by superimposing linear magnetic field variations on the main magnetic (B0) field to generate corresponding position-dependent variations in precessional frequency of the protons. Inside the magnet bore, three sets of gradients reside along the logical coordinate axes (x, y, and z) and produce a magnetic field variation determined by the magnitude of the applied current in each coil set (Figure 19.).

3.2. The role of MRI in dementia

Neuroimaging techniques aimed at studying structural changes of the brain may provide useful information for the diagnosis and the clinical management of patients with dementia[20].

The role of neuroimaging in dementia nowadays extends beyond its traditional role of excluding neurosurgical lesions. Radiological findings may support the diagnosis of specific neurodegenerative disorders and sometimes radiological findings are necessary to confirm the diagnosis. As a matter of fact, the MRI technique was widely put in application to investigate patterns of gray matter (GM) atrophy. [14]

Nowadays, the majority of the researcher utilising MRI focus their attention on Alzheimer's disease or confronting this last one with other types of diseases. With nearly as high number

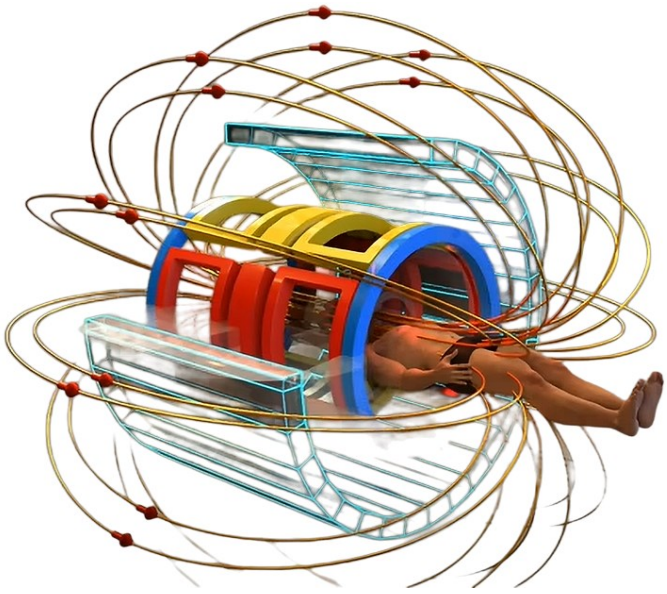


Figure 19. *MRI scanner with 3 gradient coils: X (red), Y (yellow) and Z (turquois) direction, to select the slice, column and row of the patient's anatomy.*

as the studies on the AD there is also an increase of the application of this technique of interest on the patients with PD and PDD (Figure 20.).

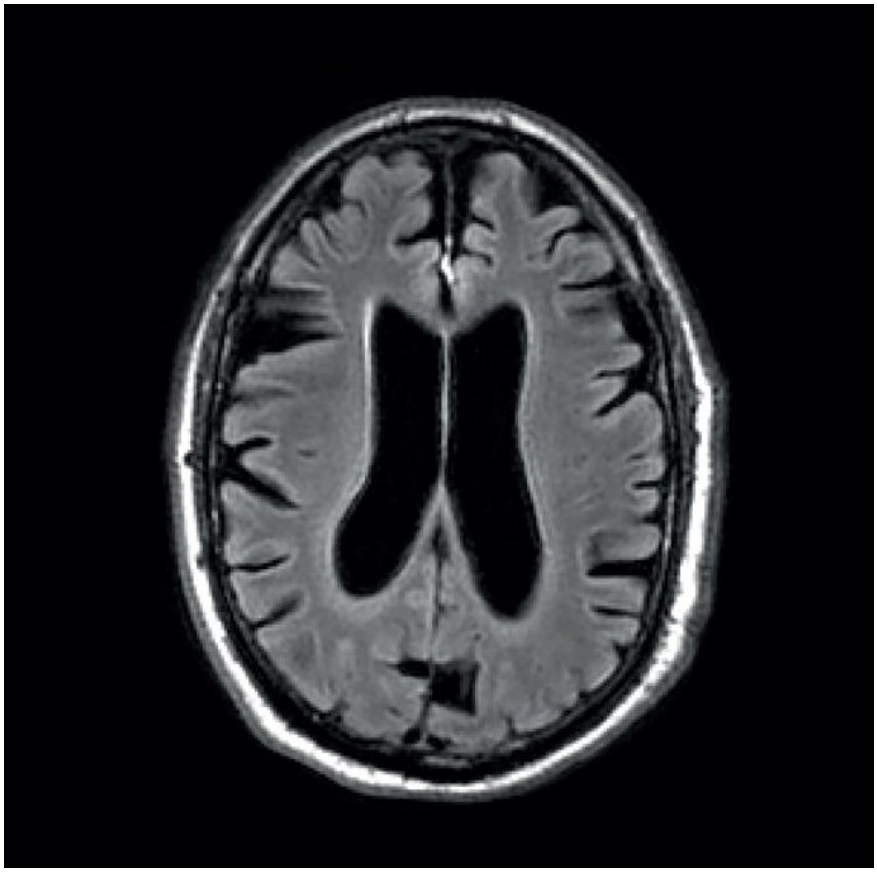


Figure 20. Transaxial MRI fluid-attenuated inversion recovery image showing global atrophy and enlarged ventricles in a Parkinson's disease with dementia subject[21].

4. *Machine Learning*

Machine learning (ML) is a subset of artificial intelligence (AI) that focuses on the development of algorithms and statistical models that enable computers to learn and make predictions. Machine learning plays a significant and increasingly important role in the field of medicine: it can be used by medical professionals to develop better diagnostic tools to analyse medical images. This type of machine learning algorithm could potentially help doctors make quicker, more accurate diagnoses leading to improved patient outcomes.

Machine Learning models are based on training a machine by providing it large quantities of data. The machine will follow a set of rules, known as an algorithm, to analyse and draw conclusion from the data. The more data the machine analyses, the more accurate it becomes in carrying out tasks such as make classification, making future predictions and helping in decision making.

Over the past years, the ML subfield, the Deep Learning (DL) has gained remarkable popularity, as it has superior results in analysing unstructured data such as medical images, text data, and audio data [22]. The relationship between AI, ML and DL can be observed in Figure 21.

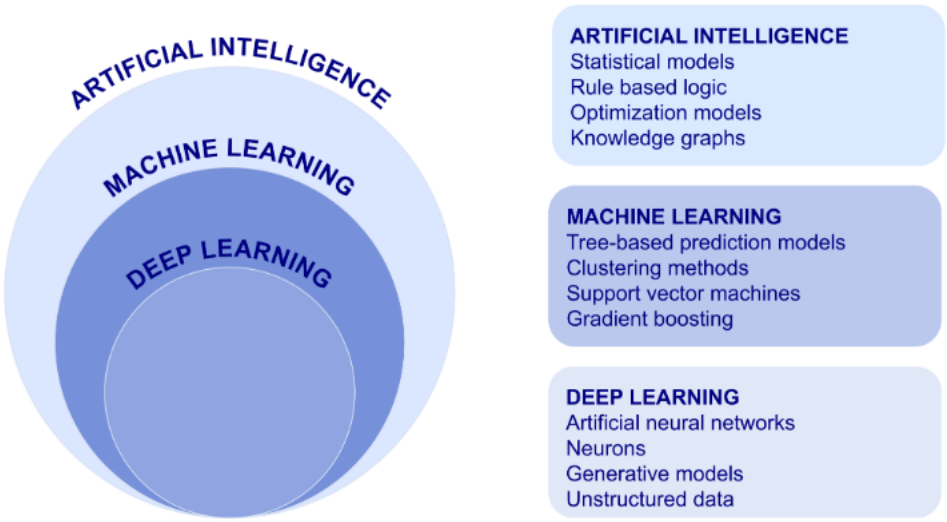


Figure 21. Hierarchy of Artificial Intelligence. On the left side of the figure there are three lists representing typical algorithms for each category [22].

4.1. Classifiers in Machine Learning

A classifier in machine learning is an algorithm that automatically orders or categorizes data into one or more of a set of classes.

A classifier is a function that takes the values of various *features* (independent variables or predictors, in regression) in an *example* (the set of independent variable values) and predicts the *class* that the specific example belongs to (the dependent variable). In a neuroimaging setting, the features are the voxels, and the class could be the type of stimulus (for an fMRI session); at example, the subject was looking at when the voxel values were recorded [23]. A classifier has several parameters that have to be learned from training data (a set of examples used to fit the parameters of the model). The learned classifier is essentially a model of the relationship between the features and the class label in the training set.

Depending on the process that have to be followed and the data to be analysed, there are different types of classification algorithms as:

- Decision Tree
- Naïve Bayes Classifier
- K-Nearest Neighbors (kNN)
- Support Vector Machines (SVM)
- Artificial Neural Networks (ANN)
- Logistic Regression (LR)

The decision tree classifier creates the classification model by building a decision tree. Each node in the tree specifies a test on an attribute, each branch descending from that node corresponds to one of the possible values for that attribute. Each leaf represents a class labels. Instances in the training set are classified by navigating them from the root of the tree down to a leaf, according to the outcome of the tests along the path [24]. One particular type of decision trees is Gradient-Boosted Decision Trees (GBTs).

A Naive Bayes classifier is a probabilistic machine learning model that's used for classification task. The core of the classifier is based on the Bayes theorem. It assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

The simplest classification procedure is called Nearest Neighbour (kNN), and it doesn't even involve explicitly learning a classification function. Classification of a test example is done by finding the training set example that is most similar to it by some measure and assigning the label of this nearest neighbour to the test example. Nearest Neighbour classification can work very well if there is a small number of features, it tends to fare worse in situations where there are many features and only a few are informative. In the typical MRI study, generally there are many more features than examples. The effect of this is that it will generally be possible to find a function that can classify the examples in the training set well, without this necessarily meaning that it will do well in the test set: this phenomenon is called overfitting.

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space (where N is the number of features) that distinctly classifies the data points. To separate the classes of data points, there are many possible hyperplanes that could be chosen. Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Furthermore, the dimension of the hyperplane depends upon the number of features. The objective is to find a plane that has the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement in order for the future data points to be classified with more confidence. The SVM is one of the most utilise classifier in the clinical field.

Logistic Regression is used when the dependent variable (target) is categorical and is based on the application of a sigmoid function in order to return the probability of a label.

4.2.Artificial Neuronal Network

Artificial Neuronal Network (ANN) are virtual network structures that are designed, in their topology and behaviour, in ways that resemble an extremely simplified model of neuron cells and their connections in the biological brain (Figure 22.). A biological neuron contains dendrites that receive signals from other neurons and a cell body that sums those signals. As explained in the first chapter once sufficient input is received, the cell fires and transmits a signal through its axon to other cells. The artificial neuron models the parts of the biological neurons as input (X_i), weight (W_i), bias (B), summation function (∑), activation function (*f*) and corresponding output signal (y).

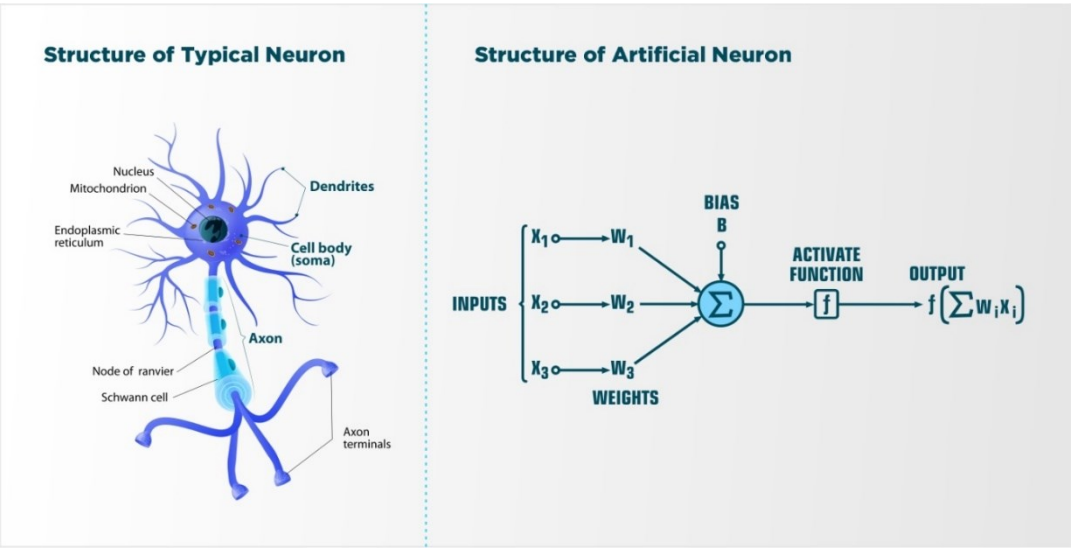


Figure 22. Comparation between a biological neuron and an artificial neuron.

The ANN consists of nodes: artificial neurons, organized into layers, with weighted connections between neurons in the layers immediately preceding and following them. The neurons in input layers receive the raw data, and the output layers produce the resulting outcome. In between there are typically a series of hidden layers that process the data. They are activated to different degrees based on the data features that the previous layer observed and propagated further. As a result, the output layer can provide some estimated output. The output layer in general uses a different activation function from the hidden layers.

Typically, a network with more than tree hidden layers is considered to be a deep neural network.

Artificial Neuronal Network also utilize more advanced techniques to improve their performance, such as the backpropagation algorithm which retroactively adjusts the weights between neurons by comparing the output of labelled data with the initial input in an attempt to minimize errors and improve the performance of the neural network.

The backpropagation algorithm primarily purpose is updating network weights with the objective of reducing the cost function (network error). One of the most used backpropagation algorithms is the gradient descent method, by which network parameters are incrementally adjusted to reduce a cost function. The approach involves multivariate calculus, since the cost function to be minimized is a combination of weights, biases, and tasks.

Similar to the classical machine learning classifier, the ANN learn form a training dataset, validate the first results by applying the algorithm on a validation dataset in order to prevent overfitting, and at the end it is applied on the testing dataset.

Being able to perform well on previously unobserved inputs is called generalization. A model is said to underfit the data when it can neither model the training data nor generalize to new data, which means there are not enough neurons or epochs (a single pass through all of the training data). Alternately, a model might be able to fit a training set but not do a good job at fitting test sets. A model that fits training but not test data is said to overfit the data, meaning that there are too many neurons or epochs.

4.3. Deep Learning

Deep learning is a subset of machine learning that uses artificial neural networks to mimic the learning process of the human brain.

Deep learning distinguishes itself from classical machine learning by the type of data that it works with and the methods in which it learns. It eliminates some of data pre-processing that is typically involved with machine learning. These algorithms can ingest and process unstructured data, like text and images, and it automates feature extraction, removing some of the dependency on humans.

Machine learning and deep learning models are capable of different types of learning as well such as: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning uses labelled datasets to categorize or make predictions requiring some human intervention to label input data correctly. In contrast, unsupervised learning does not require labelled datasets, and instead, it detects patterns in the data, clustering them by any distinguishing characteristics. Reinforcement learning is a process in which a model learns to become more accurate for performing an action in an environment based on feedback in order to maximize the reward.

Deep neural networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimize the prediction or categorization. This progression of computations through the network is called forward propagation. The input and output layers of a deep neural network are called *visible* layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or

classification is made. Deep learning algorithms are complex and there are different types of neural networks to address specific problems or datasets.

Two types of ANNs used in deep learning, are convolutional neural networks (CNN) and recurrent neural networks (RNN).

- Convolutional neural network: These neural networks are composed of many layers, each layer breaking down the input data into simple information, such as points. Then, through the different intermediate convolutional levels, information is aggregated to identify structured information such as edges or borders. Gradually the information is composed and recognized as structured objects. These neural networks are used to analyse images and extract information such as the presence or absence of specific objects;
- Recurrent neural network: It can store certain pieces of information and consider the time dimension during the learning phase. They are employed to keep track of the intrinsic knowledge contained within a sequence or time series.

4.4. Convolutional Neuronal Network

CNN is a feed-forward neural network (forward direction of the flow of information, through the hidden nodes (if any) and to the output nodes, without any cycles or loops) that learns by itself via kernel (grid-structured) optimization.

Convolutional neural networks led to major improvements in image recognition tasks, especially between 2011 and 2015 as the top-5 error-rate decreased from over 25% to around 4% in the ImageNet challenge[25]. This good result is given by the fact that for *each* neuron in the fully-connected layer 10,000 weights would be required for processing an image sized 100 × 100 pixels. However, applying cascaded *convolution* (or cross-correlation) kernels, only 25 neurons are required to process 5x5-sized tiles.

“Convolutional neural networks are designed to work with grid-structured inputs, which have strong spatial dependencies in local regions of the grid” [25]. The most obvious example of grid-structured data is a 2-dimensional image. This type of data also exhibits spatial dependencies because adjacent spatial locations in an image often have similar colour values of the individual pixels. An additional dimension captures the different colours, which creates a 3-dimensional input volume. Therefore, the features in a convolutional neural network have

dependencies among one another based on spatial distances. The vast majority of applications of convolutional neural networks focus on image data, although one can also use these networks for all types of temporal, spatial, and spatiotemporal data.

In its basic form convolution means that many inputs are provided and transformed into one output via two functions. Mathematically there is a function providing a smoothed estimate of the input x at time point t (Eq. 4)[26]. The dimension of the input space to a convolutional layer can be defined as nq × nq × dq for the q^{th} layer. The kernel, also called filter, in the q^{th} layer always has the same depth dq as the input space. The width and height of the input space are usually the same, so that the input space, also called spatial input field, is a square. This is also common for the kernel where the dimension can be defined as Fq × Fq × dq . Common values for F are three or five [25].

$$s(t) = (x \cdot w)(t) \tag{4}$$

The kernel is applied to every possible position of the input space. Therefore, the number of possible positions defines the width and height of the output space which is the next hidden layer. The width and height of the next layer are defined as can be seen in Eq. 5.

$$nq + 1 = nq - Fq + 1 \tag{5}$$

One result of the convolution operation is that the width and height of the next hidden layer are dependent on the number of possible positions of the filter which means that the dimension of the next layer depends mainly on the dimension of the filter. There are two possible parameters that can be set to change the dimension of the next hidden layer. These are stride and padding.

Stride is defined as the jump of the kernel from the current to the next pixel centre. This means that the location of the kernel centre changes by S in both dimensions for an image [25]. The higher the stride is the smaller the resulting feature map will be. In Figure 23. a convolution operation is shown with a 3×3 kernel and a stride of 2. The convolution operation would usually result in a feature map of dimension 3 × 3. The effect of a stride of 2 not 1 is that the dimension of the feature map is reduced to 2 × 2.

Padding is used when it is not beneficial to reduce the dimension of the feature map because this means a loss of information along the borders of the input space. To increase the

dimension of the feature map pixels are added around the input space. Under the assumption that the dimension of the input space and the next hidden layer should be the same, the number of pixels that are added around the input space is defined by (Fq −1)/2. This increases the width and height by Fq − 1 and is exactly the loss of dimension due to the convolution operation. The values of the pixels is 0 which is called zero padding. In Figure 24. the dimension of the feature map increases to 5 × 5 if zero padding and a stride of 1 is used.

Another typical part of convolutional networks are pooling layers. The pooling function summarises a neighbourhood of the dimension n × n into one value. It is supposed to make outputs invariant to translations in the input space, so that smaller changes in the input space do not change the output [26].

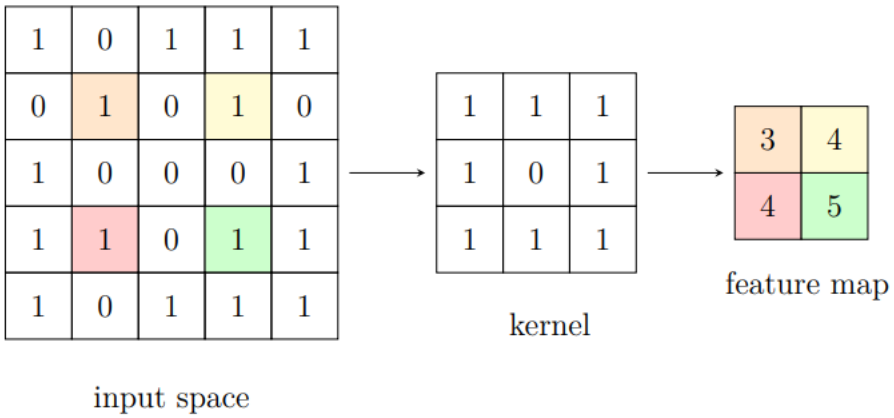


Figure 23. Example of a convolution operation using a 3x3 kernel and a stride of 2 with. The numbers are randomly selected. The kernel centres are coloured to the depending feature map output.

The goal is also to cancel out less crucial information. For that a filter with n × n dimension is defined and a stride similar to the stride used in the convolution operation. Possible pooling methods are maximum, average and global pooling. Max-pooling is used most of the times and takes the maximum value of a n × n region. Pooling is commonly used with a stride larger than 1 so that the dimension of each activation map is reduced. Max-pooling is defined in Eq. 6. where a_i are the pixel values and μ(n, n) defines the window function with the dimension n × n.

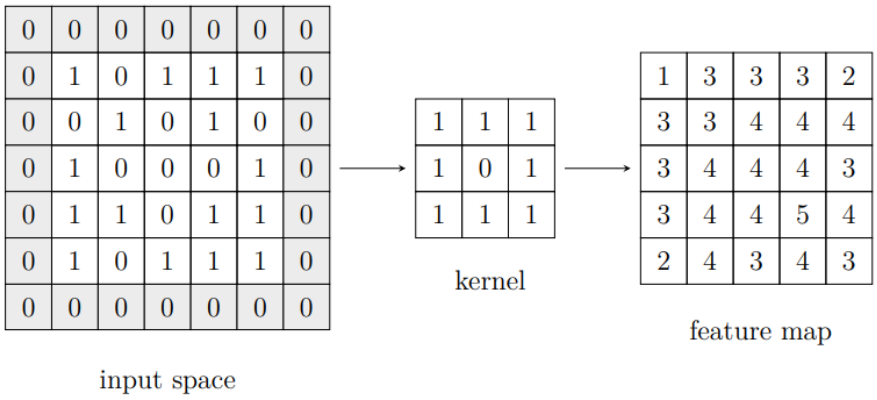


Figure 24. Example of a convolution operation using a 3x3 kernel with padding zero padding and a stride of 1.

$$a_j = \max_{N \times N}(a_i^{n \times n} \mu(n,n)) \quad (6)$$

CNNs are particularly useful because they provide highly accurate results, especially when a lot of data is involved as it is in the case of MRI. An example of CNN and the process it follows can be observed in Figure 25.

A particular CNN is the 3D Convolutional Neural Network (3D CNN) that is a type of neural network architecture designed to handle three-dimensional data, such as volumetric images or video data. Unlike 2D CNN (the classical type of CNN with the characteristics explained above) that operates on 2D images (height x width), 3D CNN handles three-dimensional data (height x width x depth). This can include volumetric medical images (es. MRI), video sequences, or any data with spatiotemporal properties. Figure 26 depicts the 3D CNN's basic design.

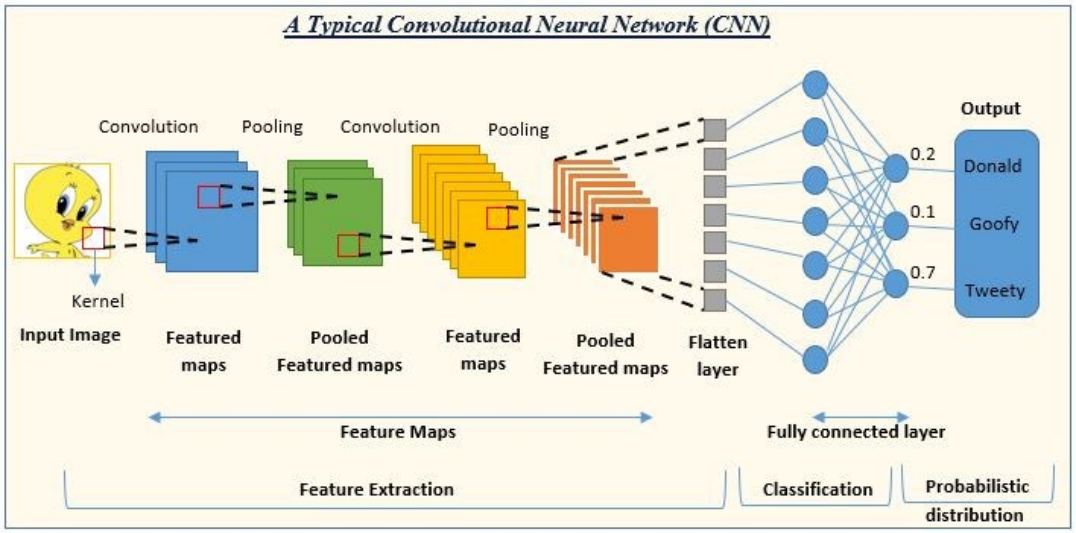


Figure 25.Example of an entire process of an CNN.

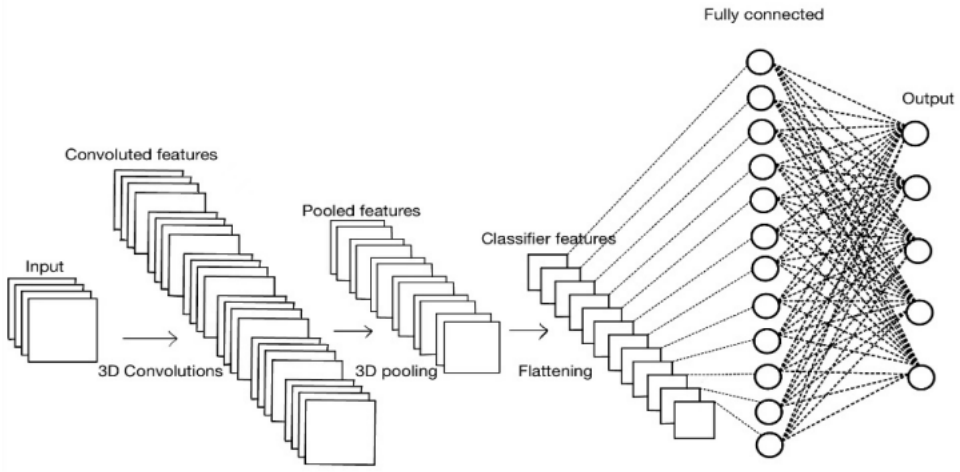


Figure 26. Simple 3D CNN Architecture [27].

5. *Literature review*

Two systematic literature searches were performed in order to have an overview of the newest as well as the oldest aspects of the three pathologies explained above. These searches were also conducted as a mean to understand how machine learning was used to classify the pathologies and which classifier emerged as the most effective. The main focus on both searches was the Lewy bodies, being this the aspect in common of all of three diseases, as can be observed from Figure 27.

The first search focussed on the difference between DLB and PDD/PD. Due to the lack of databases access for DLB, it made it necessary to perform the second search that identify the Lewy bodies in PD and PDD/DLB.

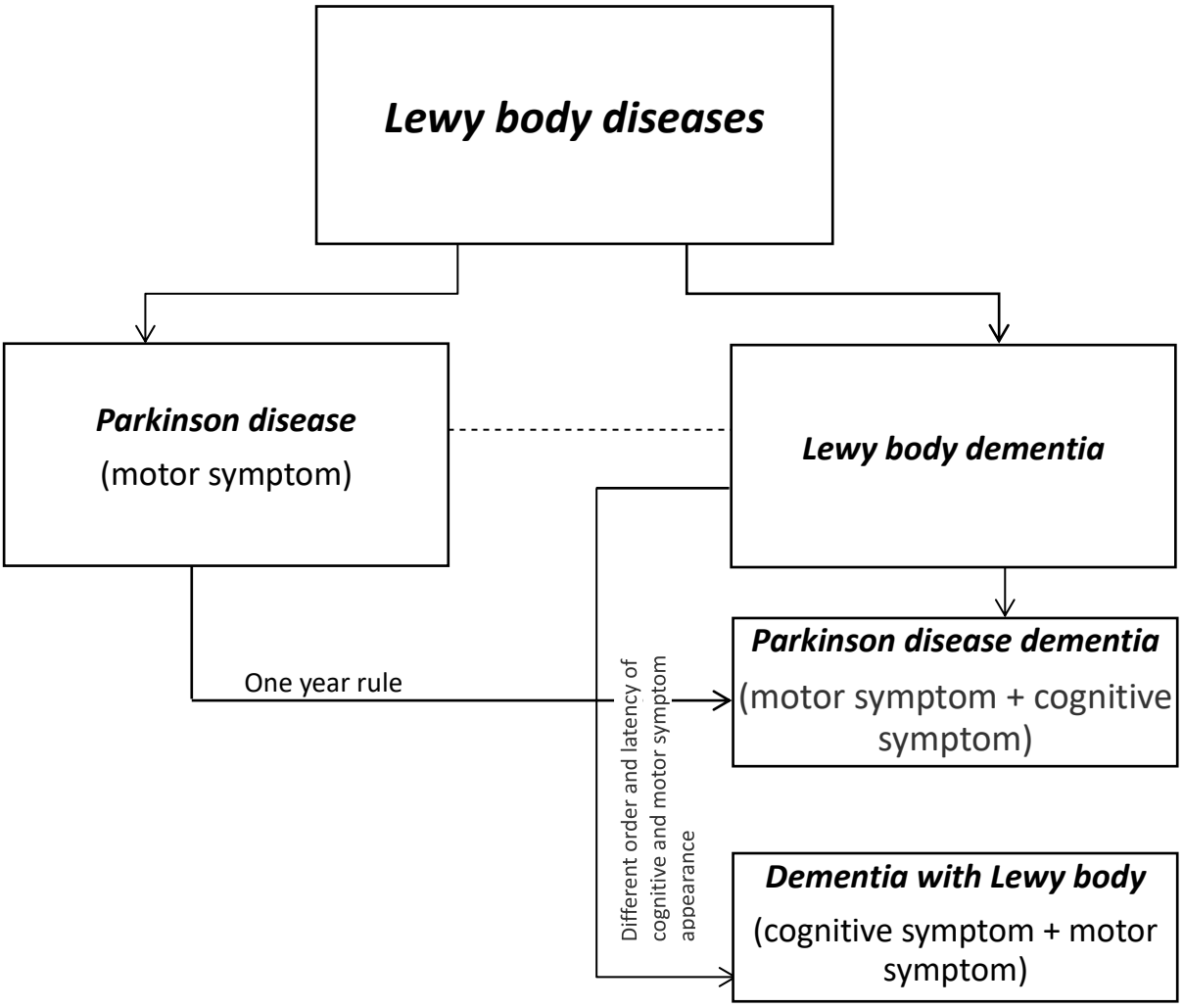


Figure 27. Main differences and similarity between PD, PDD and DLB.

5.1. Literature Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework[28] was used to guide this literature search process. Four electronic databased were consulted, namely: SCOPUS, PubMed, Sage Journals and ScienceDirect.

5.1.1. DLB and PDD/PD

The search was structured into 4 main concepts described by the roots "Dementia", related to the general terminology of all the disease that have in commune the decline in cognitive abilities, " Lewy bodies", "Lewy body", "LB" , "DLB", linked to the first disease of interest, "Parkinson", "PDD", related for the second one and "Machine learning" ,"ML" for the tool utilized for the classification of the two diseases. The Boolean operators "OR" and "AND" were used to combine the terms within and between the 4 concepts, while "Title/Abstract" was used as a search filed limit. Initially, no time constrains, or open access restrictions were applied. The last research was performed on the 1st of September 2023.

5.1.2. PDD and PD

The search was structured into 4 main concepts described by the roots "Lewy bodies" and "Lewy body", related to the abnormal aggregations of protein that develop inside nerve cells affected by PD, PDD and DLB, "Parkinson" and "PD", linked to the first disease of interest, "Parkinson's disease dementia" and "PDD", related for the second one and "Classification" in order to difference the two disease. The Boolean operators "OR" and "AND" were used to combine the terms within and between the 4 concepts, while "Title/Abstract" was used as a search filed limit. Initially, no time constrains, or open access restrictions were applied. The last research was performed on the 3rd of October 2023.

5.2. Selection process, data analysis and study quality

In order to remove duplicates and closed access articles, the database search results were exported to Zotero reference management software.

5.2.1. DLB and PDD/PD

The following inclusion criteria for title, abstract and full-text analysis were considered:

- Studies that examined a population composed exclusively of alive humans;

- Studies that considered the “ML” abbreviation as “Machine learning” and not “millilitre”;
- Studies that explored the differences between DLB and PDD(or PD) with or without considering differences between other diseases and not both DLB/PDD (considered as a unique disease) compared to other diseases;
- Studies that not considered only the pharmacological aspect between the two diseases;
- Studies that examined both pathologies in the classification results.

A table in Microsoft Excel was created to collate the data. The records were organized based on study design, participants characteristics (number of subjects, age, sex), data collection technique, algorithm, strength point, weak point. The age data was reported as mean ± standard deviation (SD). The sex was indicated as a count. Quality of appraisal was performed according to Joanna Briggs Institute Critical Appraisal tools checklist [29].

5.2.2. PDD and PD

The following inclusion criteria for title, abstract and full-text analysis were considered:

- Studies that examined a population composed exclusively of alive humans;
- Studies that make a classification using Machine Learning Classifiers;
- Studies that explored the differences between PDD and PD with or without considering differences between other diseases and not both PD/PDD compared to other diseases;
- Studies that not considered only the pharmacological aspect between the two diseases;

5.3. Literature Search Results

5.3.1. DLB and PDD/PD

As shown in Figure 28., 84 studies were collected through database search. After the removal of 10 duplicates and 32 closed access articles, just 42 studies were eligible for title screening. Subsequently, 6 articles were excluded for title and 25 were removed by abstract, leaving 10 studies for the full-text examination. Lastly, only 4 studies were included in literature review for the quality of appraisal.

Table 1. reports a description of the 4 selected studies. Overall, 691 subjects were examined, out of which 377 (55%) were males and 314 (45%) females. Two studies had approximately the same number of males and females ([30],[31]), the other two considered nearly the double of males than females.

The age variability was between 58.4±11.6 ([32]) and 75±8 ([30]). [31] and [32] used FDG-PET as data collection technique, while [30] used SPECT. Only [32] did not perform the data collection using nuclear medicine imagining modalities. There were different machine learning algorithms performed in all 4 studies.

As can be observed in Figure 29 and Figure 30. there were a variety of different diseases among the subjects. Only 12% of the total number of subjects suffered of DLB, from which [33] had the higher number. 18% of the total number of subjects suffered of PD, while only 11% had PDD, among this [30] and [32] didn't contain any subject suffering of PDD. Also subjects suffering of Alzheimer or Alzheimer disease had a valuable percentage (13%). Only [30] didn't contain health control subjects and the percentage of the total number of subjects was 15%. The higher percentage was given by other disease and/or unknown disease (31%), the majority

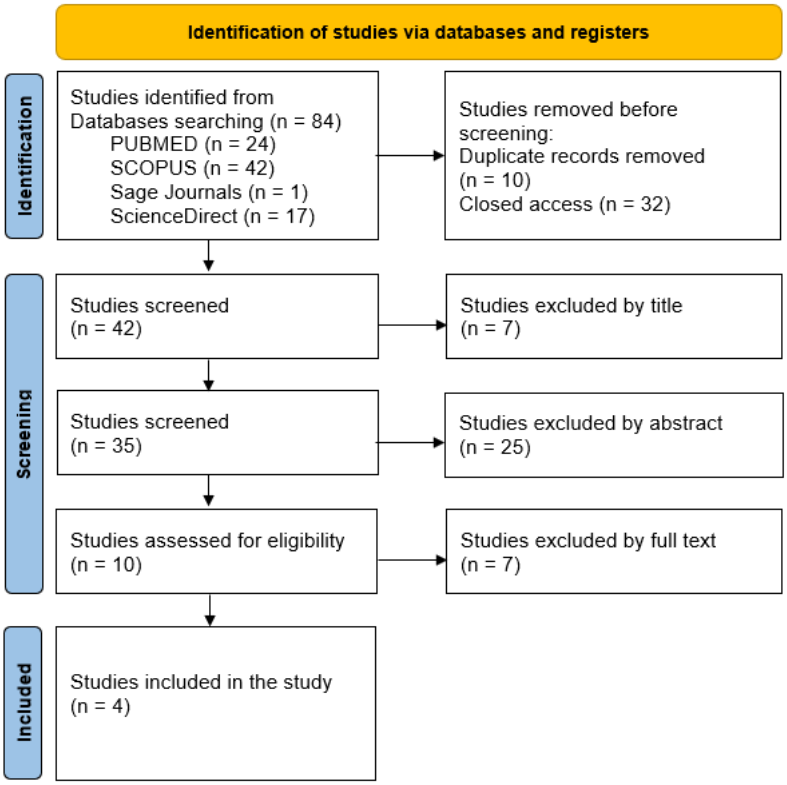


Figure 28. Flowchart of the literature search for DLB and PDD/PD.

Table 1. Summary of the eligible studies. The records are organized based on study design, participants characteristics (number of subjects, age, sex), method for data collection, Machine Learning algorithm, strength point, weak point. Continuous feature (age) is reported as mean ± Standard Deviation (SD). Categorical variable (sex) is indicated as counts.

Ref.	Study design	Number of subjects	Age (years)	Sex M/F	Method for Data collection	Machine Learning algorithm	Strength Point	Weak Point
[30]	Cohort	239	75±8 67±15	107/132	SPECT	GBTs LR KNNs	Specificity of abnormalities	The total amount of information used for ML training is a limitation
[32]	Cohort	189	58.4±11.6 72.1±7.6	114/76	FDG-PET	SSM/PCA + GMLVQ	Good interpretability	PD and DLB were often confused
[33]	Cohort	98	-	74/24	EEG	KNNs LR SVM	Classification between DLB and PDD (not DLB and PD)	Focused more on the AD-DLB than PD-DLB
[31]	Case-control	165	59.72±6.89 73.78±5.83	82/83	FDG-PET	SVM	Prediction of PDD	No specific distinction between DLB and PDD but just an end test to understand if the classifier is also sensitive to DLB.

being contained in [30].

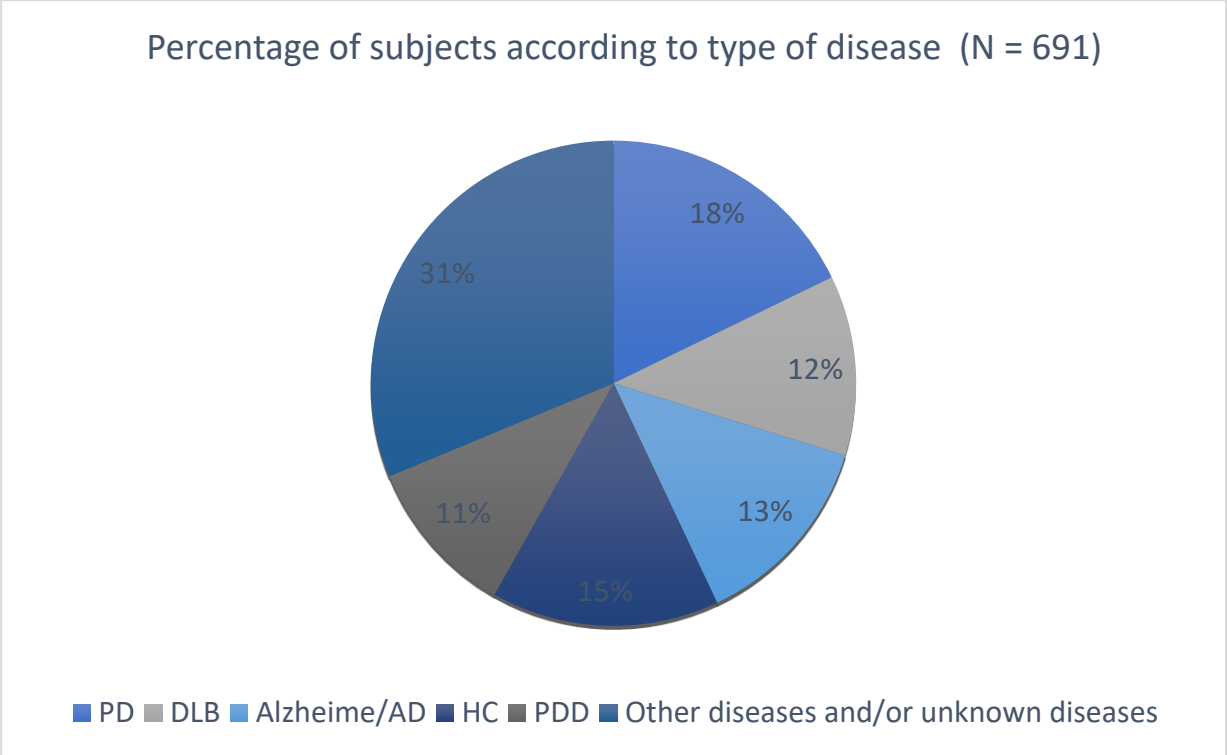
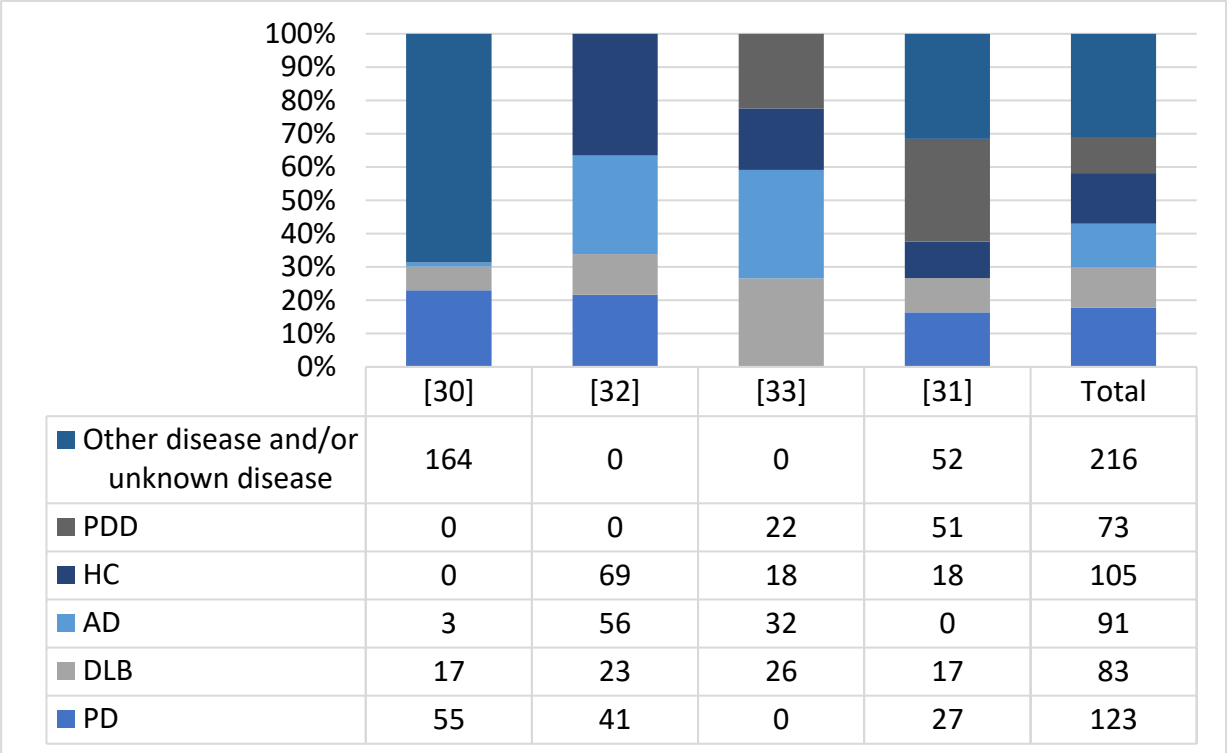


Figure 29. Percentage of subjects according to type of disease



	[30]	[32]	[33]	[31]	Total
Other disease and/or unknown disease	164	0	0	52	216
PDD	0	0	22	51	73
HC	0	69	18	18	105
AD	3	56	32	0	91
DLB	17	23	26	17	83
PD	55	41	0	27	123

Figure 30. Column diagram of diseases with counts and cumulative frequencies for each study and for the cumulative population.

5.3.2. *PDD and PD*

As shown in Figure 31., 64 studies were collected through database search. After the removal of 14 duplicates and 28 closed access articles, just 22 studies were eligible for title screening. Subsequently, 4 articles were excluded for title and 12 were removed by abstract, leaving 6 studies for the full-text examination. Lastly, only 2 studies were included in literature review for the quality of appraisal. Table 2. reports a description of the 2 selected studies. Overall, 832 subjects were examined, out of which 465 (55,89%) were males and 367 (44,11%) females. One study had approximately the same number of males and females ([31]), the other one considered more males than females.

The age variability was between 46 ([34]) and 73.78±5.83 ([31]). Both studies used FDG-PET as method for data collection. There were different machine learning algorithms performed in all 4 studies. Both of the studies used SVM as Machine Learning algorithm for classification, while [34] used more than one classifier.

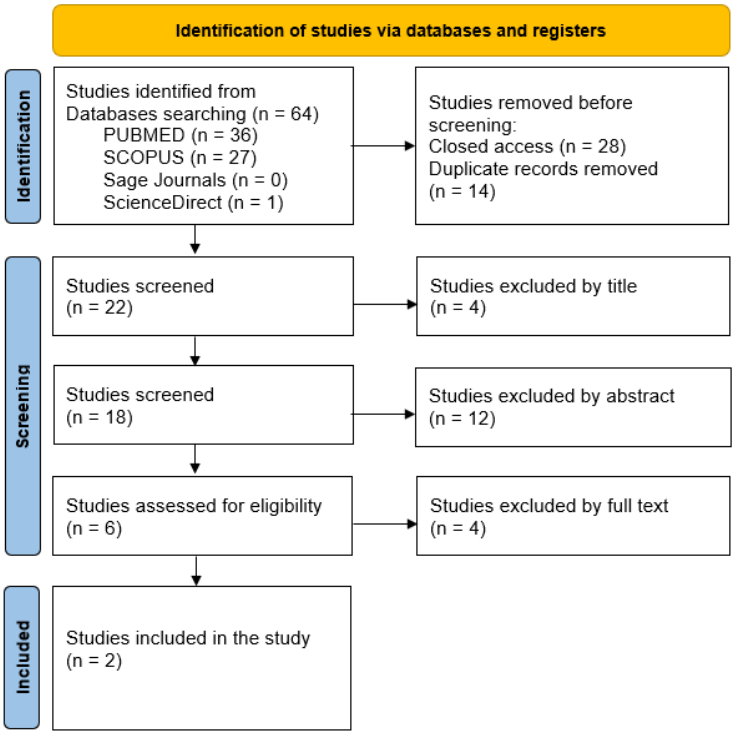


Figure 31. Flowchart of the literature search.

Table 2. Summary of the eligible studies. The records are organized based on study design, participants characteristics (number of subjects, age, sex), method for data collection, Machine Learning algorithm, strength point, weak point. Continuous feature (age) is reported as a range is not normally distributed and as mean ± Standard Deviation (SD) in the other cases. Categorical variable (sex) is indicated as counts.

Ref.	Study design	Number of subjects	Age (years)	Sex M/F	Method for data collection	Machine Learning algorithm	Strength Point	Weak Point
[34]	Cross-sectional	667	46-83	383/284	FDG-PET	GLM SSM/PCA SVM	SVM predicted PDD from PD	Focus on AD
[31]	Case-control	165	59.72 ±6.89 73.78 ±5.83	82/83	FDG-PET	SVM	Prediction of PDD	Small sample size used for SVM training, may have resulted in overfitting

As can be observed in Figure 32. and Figure 33. there were a variety of different diseases among the subjects. Only 8% of the total number of subjects suffered of PD, from which [34] had the higher number. 15% of the total number of subjects suffered of PD/DLB. Also subjects suffering of AD had a valuable percentage (23%), while [31] didn't contain any patient suffering of AD. 18 % of the total were health controls. The higher percentage was given by other disease and/or unknown disease (36%), the majority being contained in [34].

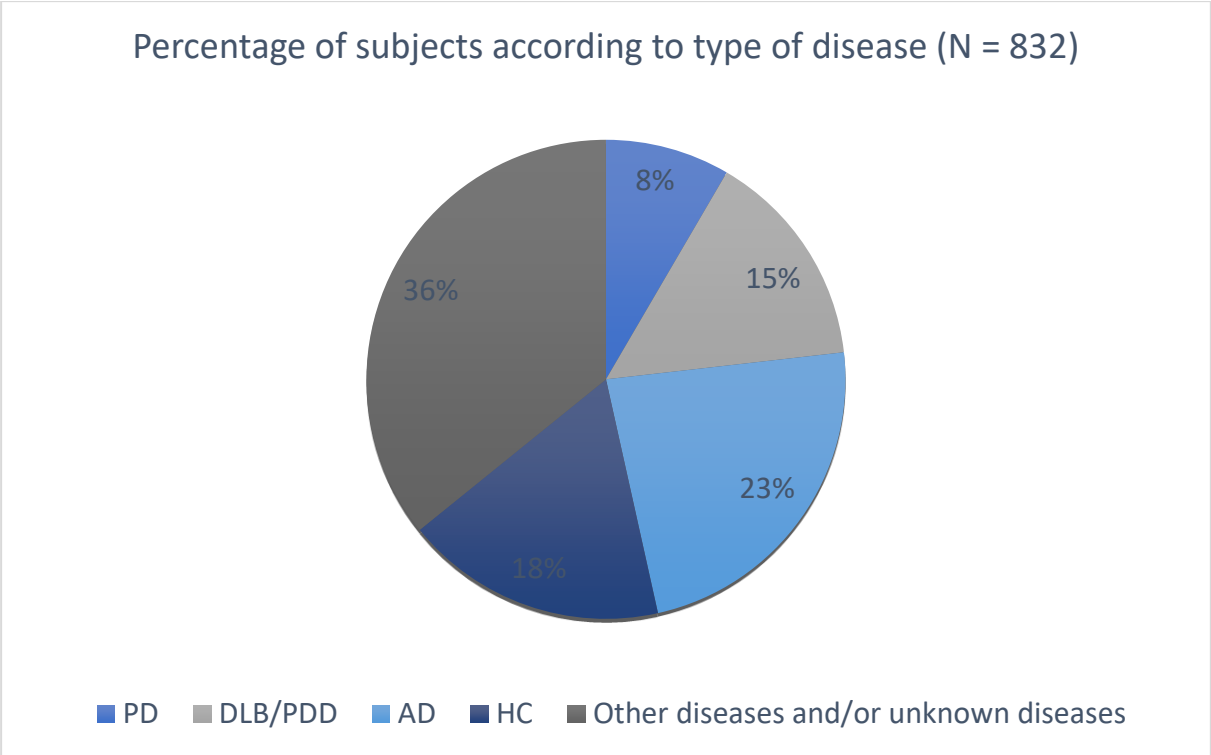
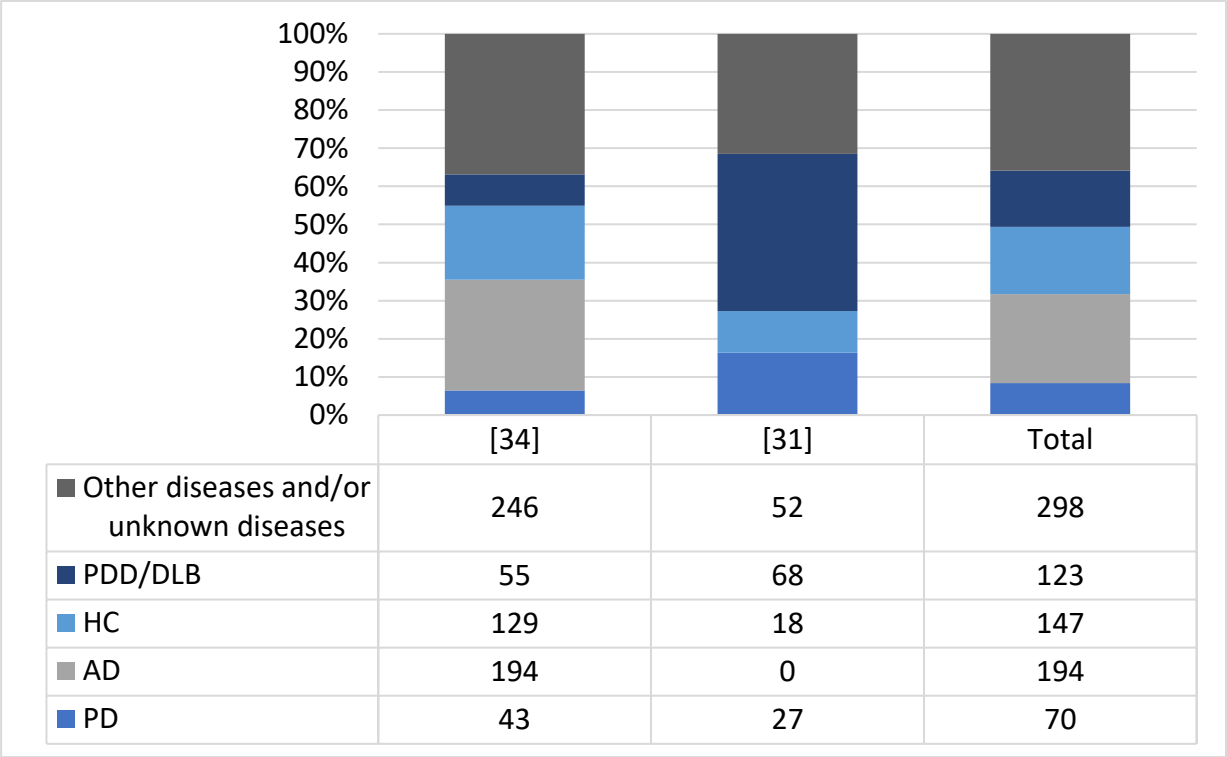


Figure 32. Percentage of subjects according to type of disease



	[34]	[31]	Total
Other diseases and/or unknown diseases	246	52	298
PDD/DLB	55	68	123
HC	129	18	147
AD	194	0	194
PD	43	27	70

Figure 33. Column diagram of diseases with counts and cumulative frequencies for each study and for the cumulative population.

5.4. Discussion

From both searches emerged that researchers are more focused on Alzheimer than other types of dementia as can be observed in Figure, Figure, Figure and Figure.

Even though the terms utilised for the searches were not the same, [31] was considered in the results of DLB and PDD/PD and, also in the results of PD and PDD.

As the two searches were concentrated on distinct pathologies there was the necessity to consider different weak point and inclusion criteria.

It emerged that despite the pathologies analysed the most exploited classifier was SVM, being more effective in high dimensional spaces.

5.4.1. DLB and PDD/PD

This literature review chapter, systematically investigate the scientific data relative to DLB and PDD (or PD) classified by different Machine learning algorithms. Overall, 4 studies were included. Initially, no time constrains, or open access restrictions were applied. This was done in order to understand how many articles were published regarding this subject of interest. Even though, no time restriction was imposed, all of the 4 records considered at the end were published in 2022, highlighting the fact that the classification of different neurological disease using Machine Learning tools has just recently gain interest. The typically constrain regarding English language (being English the language mostly used in science) was not applied as just one article was written in another language (study that was eliminated in Zotero as it was a closed access article).

The alive human criterion was considered as the intention of this study is to classify and offer faster the treatments to subjects affected by dementia in a clinical environment even though most of the data collection regarding DLB are performed post-mortem.

In order to explore the differences between PDD (or PD) and DLB, and since some studies considered DLB and PDD as a unique disease (LBD), only studies that examined the two diseases separately were considered.

Additionally, many studies focused their attention on finding new treatments for dementia. However, these articles were not taken into account as this study aimed to find (and maybe create) an improved classification before focusing on the pharmacological aspect.

Data form 691 subjects, with almost uniformly distribution among males (55%) and females (45%), were analysed (Table 1, Figure 29., Figure 30.). Only [31] and [32] had subjects suffering of PDD. In particular [30] had an accuracy of 0.61 ± 0.16. Three of these articles considered had a cohort study design demonstrate that an accurate diagnostic needs a long period of follow up of the patients just one used retrospective chart review study (case-control).

The data collected from nuclear medicine imagining modalities are more easily interpretable even for the machine learning algorithms, in fact only [32] did not perform the data collection using this type of data collection technique.

Another interesting aspect that emerged from this literature review was the fact that two of the study ([30],[31]) were done in Asia, from which one was done in collaboration with Canada and the other two([31],[32]), were performed in Europe. Figure 34. shows the map that evidence the places where the studies were conducted.

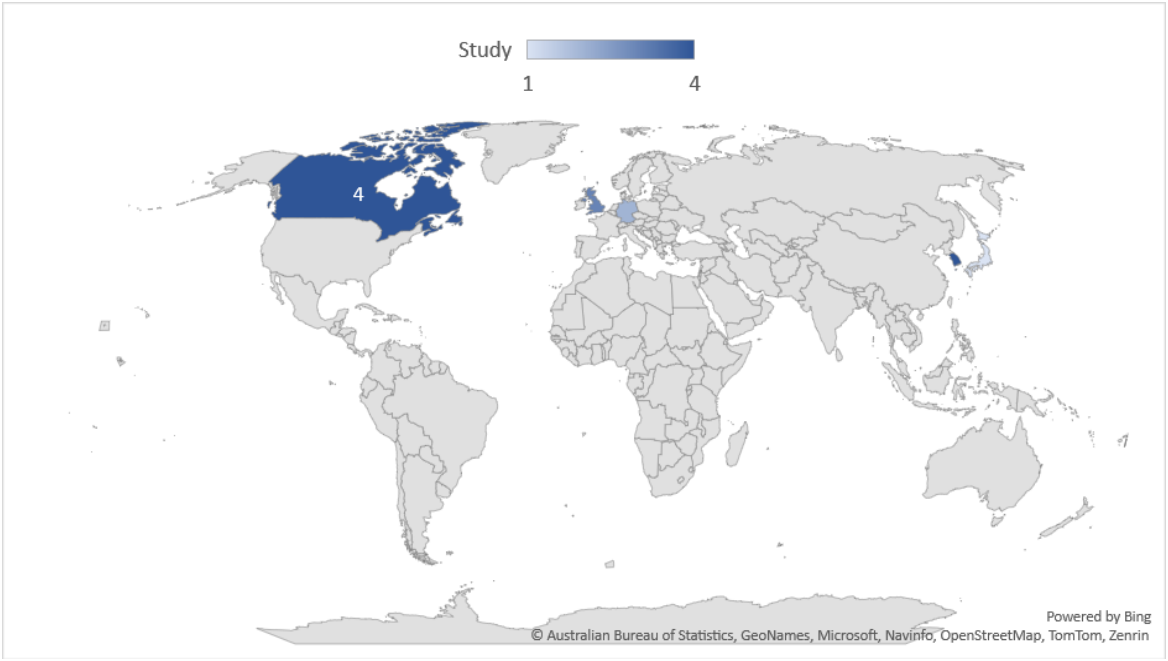


Figure 34. Map of the different places where the studies were performed. The lighter blue represent the study [30], followed by [31],[32],[33].

5.4.2. PDD and PD

This literature review chapter, systematically investigate the scientific data relative to PD and PDD/DLB classified by different Machine learning algorithms. Overall, 2 studies were included. Exactly as the first search, no time constrains, or open access restrictions were applied. This was done in order to understand how many articles were published regarding this subject of interest. Even though, no time restriction was imposed, one of the articles was published in 2022 ([31]), and the other was published in 2018 ([34]) highlighting the fact that the classification of different neurological disease using Machine Learning tools has just recently gain interest. The typically constrain regarding English language (being English the language mostly used in science) was not applied as just one article was written in another language (study that was eliminated by title).

The alive human criterion was considered as the intention of this study is to classify and offer faster the treatments to subjects affected by dementia in a clinical environment.

In order to explore the differences between PD and PDD, only studies that examined the two diseases separately were considered.

Additionally, many studies focused their attention on finding new treatments for dementia. However, these articles were not taken into account as this study aimed to find (and maybe create) an improved classification before focusing on the pharmacological aspect.

Data form 832 subjects, with almost uniformly distribution among males (55,98%) and females (45,11%), were analysed (Table 2., Figure 32., Figure 33.).

The data collected from nuclear medicine imagining modalities are more easily interpretable even for the machine learning algorithms, in fact both studies analysed had collected the data with FDG-PET.

The following part shows the most relevant aspects of the four studies analysed.

The first article found from the literature search is the one written by K. Nakajima et all. The authors concentrated their attention on building “a multivariable model incorporating machine learning (ML) that could accurately differentiate abnormal profiles on ^{123}I-ioflupane images and diagnose Parkinson syndrome or disease and dementia with Lewy bodies (PS/PD/DLB)” [30].

The type of study design applied in this article was a retrospective one, considering ^{123}I-ioflupane single-photon emission computed tomography images acquired from 239 patients at two hospitals. Patients with suspected neurodegenerative diseases and dementia including PD, PS and DLB were classified as having Parkinson syndrome or disease and DLB.

The machine learning algorithms used in this article are GBTs, LR, and KNNs, for the classification of ^{123}I-ioflupane images. The articles also highlights the importance of combining age with image features for improved diagnosis.

The article focuses on the identification of image features related to abnormality and the explainable results obtained from four contributing variables. The diagnostic accuracy for a diagnosis of PS/PD/DLB was 0.86 ± 0.04 for SE.

The second study that was considered in this chapter of the literature review is the one done by Rick van Veen et all. The aim of this study was to apply a Generalized Matrix Learning Vector Quantization (GMLVQ) to FDG-PET scans of healthy controls, and patients with AD, PD and DLB in order to discriminate between these neurodegenerative conditions. Rick van Veen et all. determined the diagnostic performance by performing ten times repeated ten fold cross-validation and then analysed the validity of the classification system by inspecting the GMLVQ space[32].

The results of this study are shown in Figure 35.

The third study to be examined is the one done by Jennings et all. Different form the previous two, this study extracted spectral properties from EEG signals. Their aim was to create an accurate diagnostic approaches for clinical use capable of differentiation of Lewy body dementia from other common dementia [33].

80 dementia patients were recruited for the study; this included 32 AD patients (22 male, 10 female), 26 DLB patients (21 male, 5 female) and 22 PDD patients (20 male, 2 female). Along with these dementia patients, 18 age-matched healthy controls (11 male, 7 female) were also recruited for between-group comparisons.

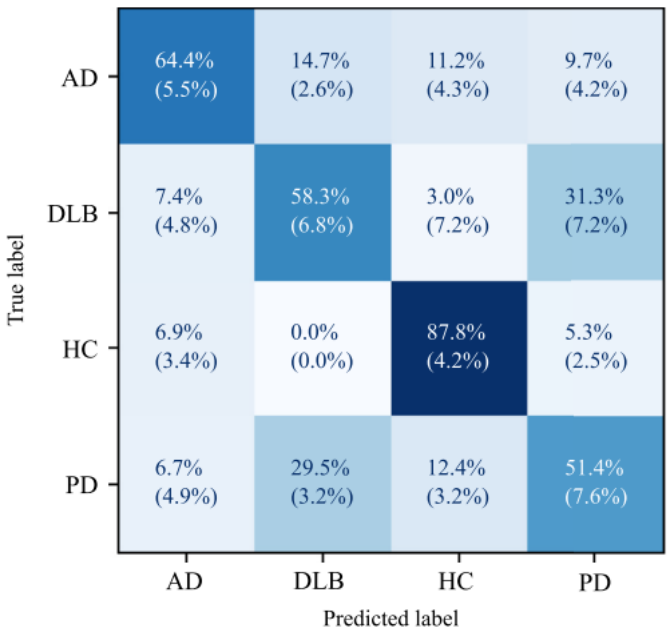


Figure 35. Average testing confusion matrix of the differential diagnostics problem between (AD, DLB, HC and PD). Averages and standard deviations are extracted from the cross-validation procedure[31].

"For this study, participants underwent in-depth neurological and neuropsychiatric testing. The Cambridge Cognition Examination (CAMCOG) and Mini-Mental state exam (MMSE) were used to assess cognitive function in patients, with both tests being commonly used for assessing the extent of a participant's dementia symptoms. Additionally, the Neuropsychiatric inventory test was performed to assess the severity and frequency of hallucinations for participants (NPI hal)."[32]

"In total, 150 s of resting state EEG was acquired using a 128 sintered Ag/AgCl electrode Waveguard cap (ANT Neuro, The Netherlands) placed in a 10-5 positioning system for each participant. Channels were recorded at a sample rate of 1024 Hz with an electrode impedance of no more than 5kΩ. Before the analysis of qEEG features, all EEG pre-processing and cleaning was carried out blinded to group membership using EEGLAB MATLAB functions (R2012; MathWorks, Natick Massachusetts) for both EC and EO recordings."[32]

From the original 98 subjects, 65 were used for analysis, including 15 HC, 12 AD, 21 DLB and 17 PDD subjects. The other 33 patients (2 HC, 13 AD, 4 DLB, 4 PDD) were removed from the dataset due to participants not having at least 20 s of combined resting state eyes closed or

eyes open EEG after cleaning. EEG segments were analysed over 5 cortical regions: frontal (F), central (C), temporal (T), parietal (P) and occipital (O).

Several supervised machine learning methods were evaluated using the k-fold cross-validation method in MATLAB and WEKA for the classification. Feature selection was also performed in these two programming language using neighbourhood component analysis Selected features were used to train machine learning classifiers utilising 10-fold cross-validation. Table n+7 summarized the classification results between AD and DLB patients using the k-nearest neighbour algorithm, logistic regression, and a quadratic support vector machine.

Jennings et all. also investigated the capability of a multiclass classification approach to differentiate HCs from non-healthy participants and then differentiate each dementia type from other participants determined to be non-healthy. They were also analysing the differences between diseases; in particular the accuracy for DLB-PDD was of 0.61±0.16.

The fourth study that has been selected in the literature review process is the one written by Booth S. et all. Their aim was to develop a metabolic pattern that predicts future dementia development in patients with PD-MCI, given the fact that approximately 90% of patients with PD-MCI progress to PDD. The data used was from FDG-PET of 165 patients with PD-MCI, DLB or other diseases, who were followed for a period of up to 8 years. The FDG-PET method was used as “brain imaging techniques have the potential to show changes in brain function that occur in patients in the early stages of PDD and can thus be used as a predictive biomarker”[31]. All FDG-PET image preprocessing was carried out using the standard procedure implemented in Statistical Parametric Mapping 12 (SPM) software [31].

Booth S. et all. used the SVM to build a classifier that predicts whether patients will maintain stable MCI or progress to PDD based on FDG-PET scans at baseline. The classifier achieved good sensitivity (86.96%) and specificity (85.00%) with k-fold cross-validation.

The last article considered in this chapter is the one by Katako et all. where general linear model, scaled subprofle modeling and SVM were examined. Among the tested classification methods, SVM with Iterative Single Data Algorithm produced the best performance, with a sensitivity of 0.84 and a specificity of 0.95. Theye have applied the same classification algorithm to four different datasets from ADNI, Health Science Centre (Winnipeg, Canada), Dong-A University Hospital (Busan, S. Korea) and Asan Medical Centre (Seoul, S. Korea). The

data analysed confirmed that the SVM showed the best performance in prediction of future development of AD from the prodromal stage (mild cognitive impairment), and that it was also sensitive to other types of dementia such as Parkinson's Disease Dementia and Dementia with Lewy Bodies [34].

The main weak point of this article was the fact that it focused on the AD not on the differences between PD and PDD/DLB.

6. Materials and methodology

6.1. Data Collection and Preprocessing

6.1.1. Data Collection

In this study, two datasets from the Parkinson's Progression Markers Initiative (PPMI) [35] database were collected. PPMI database for neuroimages is considered to be a landmark, international, and multicenter study to research the biomarkers that are responsible for Parkinson's Disease progression[36]. PPMI aims to provide to the broad research community a comprehensive, standardized, longitudinal data set and biosample library to speed breakthroughs and enable validation toward clinical application of new findings.

Data collected by PPMI are derived from a wide range of imaging modalities and formats, such as MRI and Digital Imaging and Communications in Medicine (DICOM) format. According to the standards committee, DICOM is the international standard for transmitting, storing, retrieving, printing, processing, and displaying medical imaging information [37]. A DICOM file contains a header and image data sets combined into one file. The header consists of tags such as patient demographics, including the patient's name, date of birth, age, gender, and it can contain study parameters such as image dimensions, acquisition parameters, pixel intensity, and matrix size.

The MRI scans selected for the study were based on particular imaging protocols described in Table 3.

The first dataset collected contains 29 subjects affected by prodromal PD, while the second one includes 30 subjects with PDD. The prodromal PD refers to the stage at which individuals do not fulfill diagnostic criteria for PD (bradykinesia and at least 1 other motor sign) but do exhibit signs and symptoms that indicate a higher than average risk of developing motor symptoms and a diagnosis of PD in the future. Some participants in the prodromal cohort received a diagnosis of PD at the baseline visit.

Out of 59 subjects, 32 (54,24%) were males and 27 (45,75%) were females. In Figure 36. it can be observed the percentage of sex of the total number of subjects, while in Figure 37. is showed the percentage of sex based on the two classes.

Table 3. Imaging protocol for the scans of the two datasets.

Imaging Protocol	Values
Modality	MRI
Research Group	PD and PDD
Visit	Baseline-Month12-Month24
Acquisition Plane	AXIAL
Acquisition Type	3D
Field Strength	1.5-3.0 Tesla
Flip Angle	1-8-9-15-20-30 Degree
Scanner Manufacturer	Philips-Siemens-GEMedicalSystems
Pixel Spacing	0.5-1.0 mm (X & Y)
Slice Thickness	0.9-2.0 mm
Weighting	T1

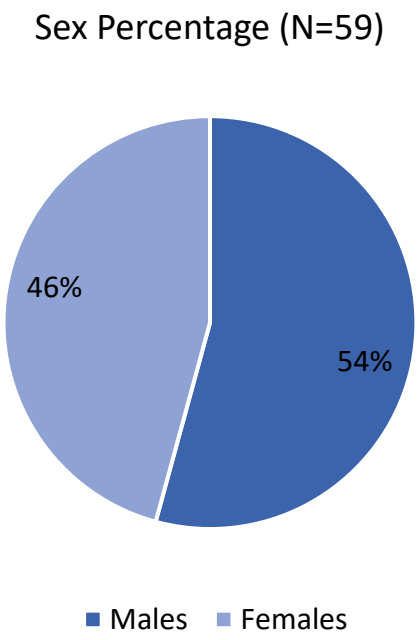


Figure 36. Percentage of sex based on males and females.

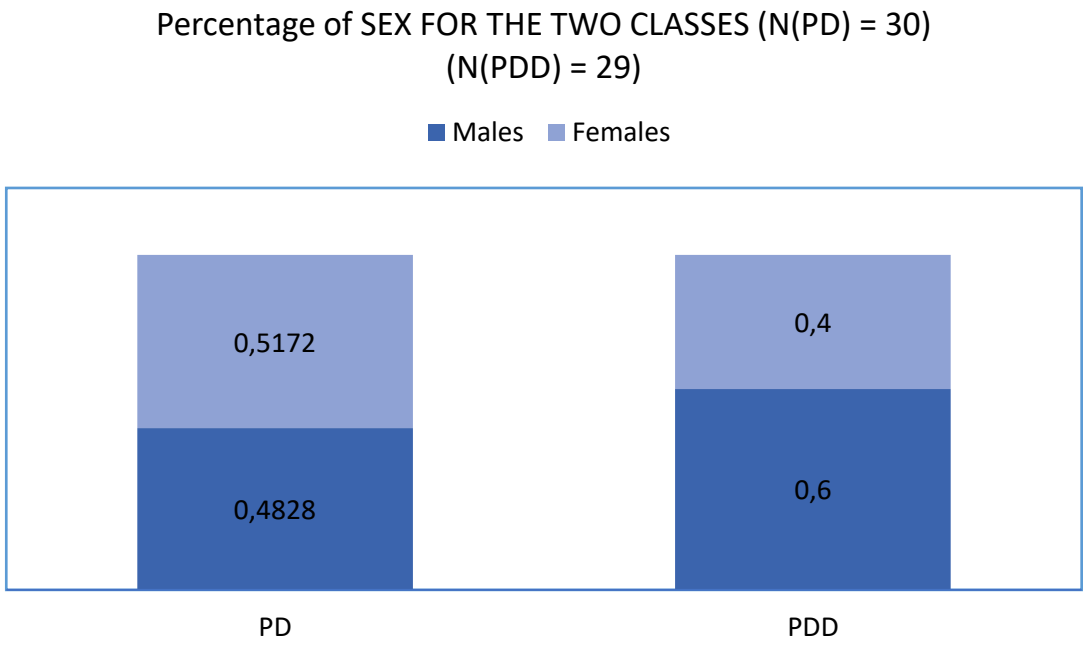


Figure 37. Sex percentage for the two classes.

The age variability was between 45,06 years and 77,02 years for Prodromal PD subjects, and between 33,5 years and 80,6 years for PDD subjects.

In this study, the oldest image considered was acquired on 06/01/2011 on a PDD subject, while the most recent image was acquired on 08/09/2023 on a subject with prodromal PD.

6.1.2. Data Preprocessing

After data selection, all the MRI images were pre-processed in Colab short for Colaboratory, a product from Google Research. More technically, Colab is a hosted Jupyter notebook service that requires no setup to use, while providing free access to computing resources including GPUs. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to ML, data analysis and education. Its main advantage is the on-cloud nature.

The first step, after uploading the data from the source folder, was to convert the DICOM files in *Neuroimaging Informatics Technology Initiative* (NIfTI) files. NIfTI is sponsored by the US National Institute of Mental Health and the National Institute of Neurological Disorders and Stroke and defines a file format for neuroimaging data that is meant to meet the needs of the MRI research community. In particular, NIfTI was developed to support inter-operability of tools and software through a common file format. Out of the initial 63 DICOM files (some

subjects had more MRI scans), only 55 were correctly converted to NIfTI files as some DICOM files may not contain the necessary image data or metadata required for a meaningful conversion. From this two scans were discarded for the inhomogeneity with the atlas used for the registration.

As explained above, PPMI is a multicenter study, therefore, the imaging scans acquired in the study contained temporal and spatial differences. In order to maintain a constant modality between all the scans it was required that all the scans needed to be in the same space such as Montreal Neurological Institute (MNI) or individual brain atlases using statistical parametric mapping (IBASPM)[38]. Therefore, to transform the PPMI MRI data that has been collected from multiple centers across the globe to a fixed coordinate system, an image registration procedure was performed. Image registration is a process where a traversal is performed on a fixed image (atlas) to find the alignment parameters and coordinated so that an unknown or an unseen image can be aligned similarly to fixed one. The MRI scans obtained from the PPMI database were considered as the moving image while the MNIPD25-T1MPRAGE-1 mm atlas [39] was considered as the fixed image. The MNI-PD25 was obtained from NeuroImaging and Surgical Technologies Lab (NIST). In the NIST lab there is the development of computer vision image processing algorithms for analysis of medical images that are focused on registration and segmentation. These techniques are applied to different research projects that include: image guided neurosurgery, disease diagnosis, and prognosis and quantification for diseases such as multiple sclerosis, epilepsy, schizophrenia and degenerative diseases such as Alzheimer's dementia and Parkinson [40].

The MNI-PD25 is a collection of 6 multi-contrast brain MRI atlases, accompanied by the associated probabilistic maps for three main brain tissue types, segmented labels for 8 subcortical nuclei, and a co-registered histology-based atlas. Derived from 3T MRI scans of a cohort of 25 Parkinson's disease patients, the atlases were obtained by nonlinearly co-registering each patient's anatomy to a common space. The atlases are in MNI PCBM152 stereotactic space, with 3 image resolutions available: $1 \times 1 \times 1$ mm^3, $0.5 \times 0.5 \times 0.5$ mm^3 and $0.3 \times 0.3 \times 0.3$ mm^3. The one applied in this study was $1 \times 1 \times 1$ mm^3 (MNIPD25-T1MPRAGE-1 mm atlas). The specifications of the MNIPD25-T1MPRAGE-1 mm atlas are described in Table 4.

Table 4. Specification of MNIPD25-T1MPRAGE-1 mm atlas

Image Parameters	Values
Dimensions	$193 \times 229 \times 193$ pixels
Interslice Gap	0.0 mm
Slice Thickness	1.0 mm
Spacing	$1 \times 1 \times 1$ mm^3
Plane	Sagittal

Before the registration, the MNIPD25-T1MPRAGE-1 mm was reoriented from Left-to-right, Posterior-to-Anterior, Inferior-to-Superior (LPI) to Right-to-Left, Anterior-to-Posterior, Inferior-to-Superior (RAI) in order to have the same orientation as PPMI detests.

The registration as well as the orientation of the MRI scans was performed using one of the most effective normalization tools known as advanced normalization tools Python (ANTsPy) [41]. ANTsPy is used particularly in the field of imaging research for extracting important information from complex imaging datasets to perform preprocessing on MRI. The registration of the acquired MRI scans with the MNIPD25-T1MPRAGE-1 mm atlas was performed using translation normalization.

The second step was to normalize the scan. Normalize the data refers to the process of standardizing the intensity values of the image to a common scale or range. This is done to make the data more comparable across different subjects, imaging sessions, or scanners. The method applied in this study was Z-Score Normalization, also known as standardization, a method that scales the intensity values to have a mean of 0 and a standard deviation of 1.

The last step of the preprocessing was the bias field correction that refers to the process of removing intensity variations caused by non-uniformities in the image acquisition process. The bias field, known as the shading or inhomogeneity field, can result from factors such as variations in radiofrequency (RF) coil sensitivity, magnetic field inhomogeneities, or other imperfections in the imaging system. The presence of a bias field can affect the accuracy of quantitative analyses and image interpretation. Bias field correction aims to restore the original intensity distribution of the tissue by compensating for these non-uniformities. The bias field correction process involves iteratively estimating and correcting the bias field. This can be computationally intensive, especially for high-resolution images. To speed up the

process shrink factor of 4 was applied. Figure 38. shows the pipeline followed to perform the dataset pre-processing with the scans with before and after each step.

6.2. Materials and Methods

The main premise of the study focusses on the classification of MRI scans in PDD and PD using C3D Model for Keras (C3DKeras) [42]. The methodology was primarily divided into four stages: MRI scan acquisition from the PPMI database; data preprocessing; C3DKeras and finally the results and performance evaluation of the C3D Model for Keras architecture based on classification matrices.

6.2.1. C3DKeras

C3DKeras, as the name suggests, is an adaptation of a convolutional 3D neural network originally developed for Caffe in the paper by Du et al. [42] in order to make it compatible with Keras. This model is in turn a modification of the BVLC Caffe model, which was trained on the Sports-1M dataset which contains video clips of various sports in order to recognize the type of sport contained in each video clip.

C3DKeras is an effective approach for spatiotemporal feature learning. 3D ConvNets are more suitable for spatiotemporal feature learning compared to 2D ConvNets. In particular, a homogeneous architecture with small 3 × 3 × 3 convolution kernels in all layers is among the best performing architectures for 3D ConvNets. Compared to 2D ConvNet, 3D ConvNet has the ability to model temporal information better owing to 3D convolution and 3D pooling operations. In 3D ConvNets, convolution and pooling operations are performed spatio-temporally while in 2D ConvNets they are done only spatially.

The first request of C3DKeras was to have 112 × 112 images, in order to do so the MRI scans where cropped considered the measured requested.

As explained above the C3DKeras was pre-trained on sports video clips that were split into non-overlapped 16-frame clips which were then used as input to the networks. In order to maintain the characteristics of C3DKeras, the MRI scans considered in this study were also split in blocks of 16 slices. For those scans that could not be divided in groups of 16 slices the zero padding was applied. For zero padding is intended that process in which there are added black

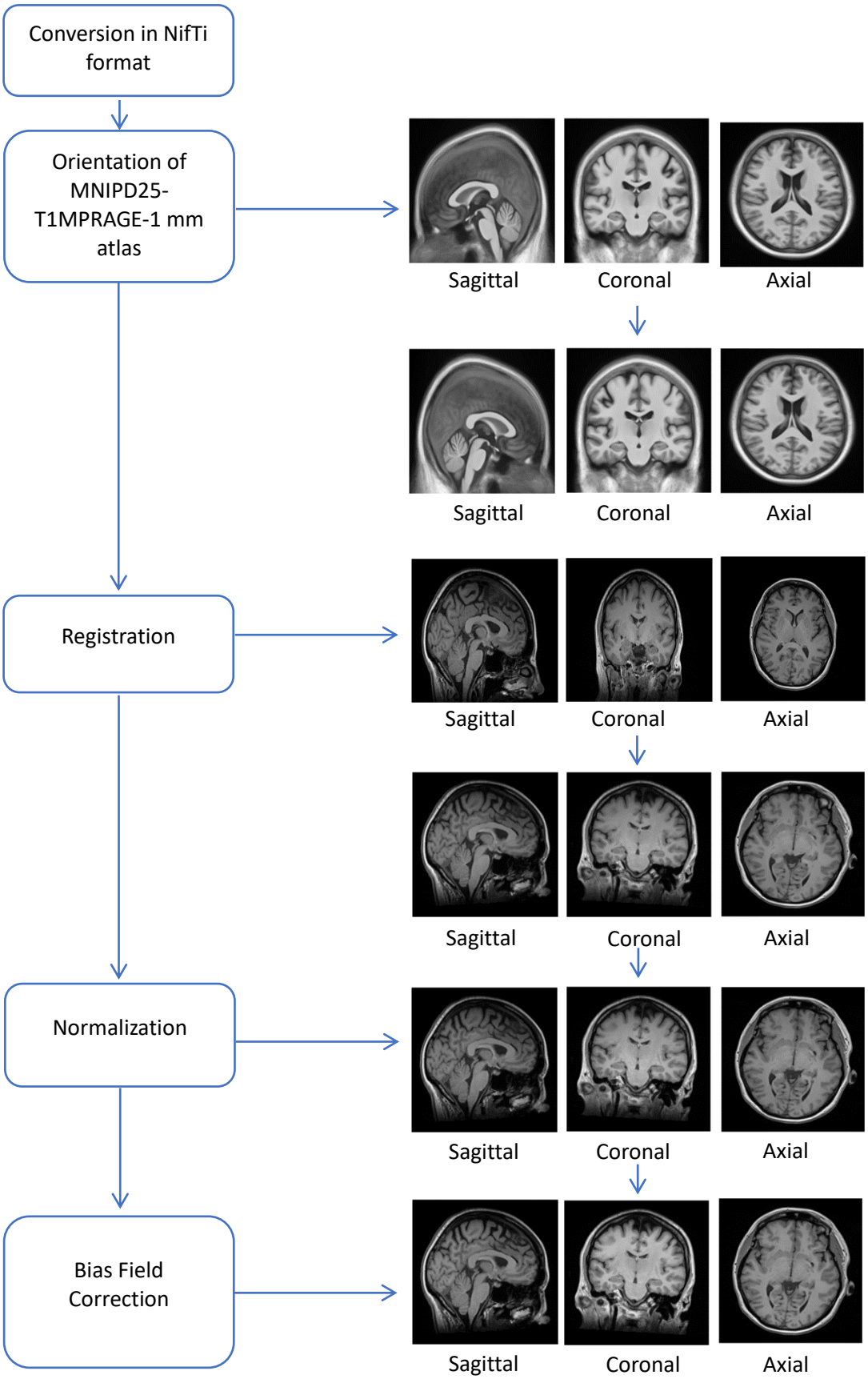


Figure 38. Pipeline of preprocessing the MRI data.

slices in order to fulfil the requirement of the number for each block. The black and white images were after, converted in coloured images.

The networks have 5 convolution layers and 5 pooling layers (each convolution layer is immediately followed by a pooling layer), one flatten layer, 2 fully-connected layers and a softmax loss layer to predict action labels. The summary of the CD3Keras architecture can be observed in Table 5.

Table 5. Summary of C3DKeras model showing layer names, output shapes and parameters.

Layer	Type	Output Shape	Parameters
conv1	Convolutional 3D	(None, 16, 112, 112, 64)	5248
pool1	Max Pooling 3D	(None, 16, 56, 56, 64)	0
conv2	Convolutional 3D	(None, 16, 56, 56, 128)	221312
pool2	Max Pooling 3D	(None, 8, 28, 28, 128)	0
conv3a	Convolutional 3D	(None, 8, 28, 28, 256)	884992
conv3b	Convolutional 3D	(None, 8, 28, 28, 256)	1769728
pool3	Max Pooling 3D	(None, 4, 14, 14, 256)	0
conv4a	Convolutional 3D	(None, 4, 14, 14, 512)	3539456
conv4b	Convolutional 3D	(None, 4, 14, 14, 512)	7078400
pool4	Max Pooling 3D	(None, 2, 7, 7, 512)	0
conv5a	Convolutional 3D	(None, 2, 7, 7, 512)	7078400
conv5b	Convolutional 3D	(None, 2, 7, 7, 512)	7078400
zeropad5	Zero Padding 3D	(None, 2, 8, 8, 512)	0
pool5	Max Pooling 3D	(None, 1, 4, 4, 512)	0
flatten_2	Flatten	(None, 8192)	0
fc6	Dense	(None, 4096)	33558528
dropout 6	Dropout	(None, 4096)	0
fc7	Dense	(None, 4096)	16781312
dropout 7	Dropout	(None, 4096)	0
fc8	Dense	(None, 487)	1995239

6.2.2. *Training strategy*

Once pre-processed, the entire dataset (100%) was split into 80% of data for training and 20% for testing. Subsequently, from the training dataset 20% was reserved for the internal validation useful for monitoring the training in order to evaluate the necessity or not of tuning the hyperparameters.

The optimization algorithm used for training of C3DKeras and modify the model parameters to minimize loss is Adam (short for Adaptive Moment Estimation), described in the paper of Kingma [43] in 2015. Adam is an adaptive learning rate algorithm designed to improve training speeds in DL and reach convergence quickly. It customizes each parameter's learning rate based on its gradient history, and this adjustment helps the neural network learn efficiently as a whole [43].

6.2.3. *Statistics*

Once trained, the classifier can be used to determine whether the features used contain information about the class of the example. If the classifier is correctly implemented, it ought to be able to predict the classes of examples it hasn't seen before (test dataset). The typical assumption for classifier learning algorithms is that the training (and testing) examples are independently drawn from an example distribution. An evaluation of a classifier is based upon its performance on any test set from the same distribution as shown in Figure 39. The performance of a classifier is given by some evaluation functions like Sensitivity (Eq. 7), Accuracy (Eq. 8), F1-score (Eq. 9) where: True Positive (TP) is when the ML predicted a person disease when is actually suffering form that disease, True Negative (TN) is when the classifier predicted as healthy a person who is actually healthy; the False Negative (FN) is when the classifier has found absence of a disease when the disease is present; the False Positive (FP) is when ML predicted the presence of a disease when the disease is absent.

For the purpose of providing a visual representation of the performance of the classifier, the AUC (Area Under the Curve) ROC (Receiver Operating Characteristics) curve can be used. AUC - ROC curve is a performance measurement for the classification problems at various threshold settings. ROC is a probability curve and AUC represents the degree or measure of separability; it tells how much the model is capable of distinguishing between classes. Higher the AUC, the better the model is at distinguishing between patients with the disease and no disease. The

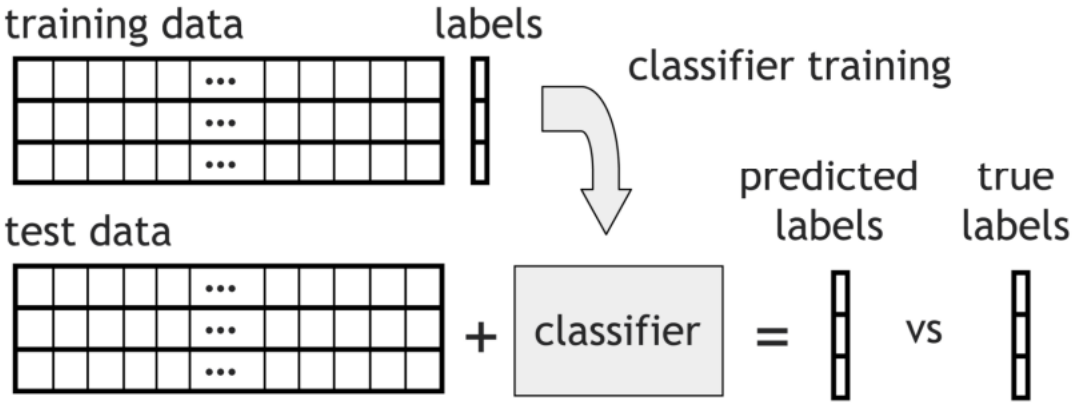


Figure 39. "A classifier is learned from the training set, examples whose labels it can see, and used to predict labels for a test set, examples whose labels it cannot see. The predicted labels are then compared to the true labels and the accuracy of the classifier – the fraction of examples where the prediction was correct – can be computed." [23]

$$\mathrm{SE} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{7}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

$$F1-score = \frac{2 \cdot \frac{TP}{TP+FP} \cdot \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}} \tag{9}$$

ROC curve is plotted with True Positive Rate (TPR) against the False Positive Rate (FPR) where TPR is equal to the sensitivity and FPR equal to one minus the Specificity.

6.3. Ablation Study

The original meaning of "Ablation" is the surgical removal of body tissue. The term "Ablation Study" has its roots in the field of experimental neuropsychology of the 1960s and 1970s, where parts of animal's brains were removed to study the effect that this had on their behaviour.

In the context of ML and DL, "Ablation Study" has been adopted to describe a procedure where certain parts of the network are removed (or changed or inserted), in order to gain a better understanding of the importance of each step.

With the aim of improving the accuracy, the sensitivity, and the area under the curve of the C3DKeras, 4 main experiments were conducted as can be observed in Figure 40.

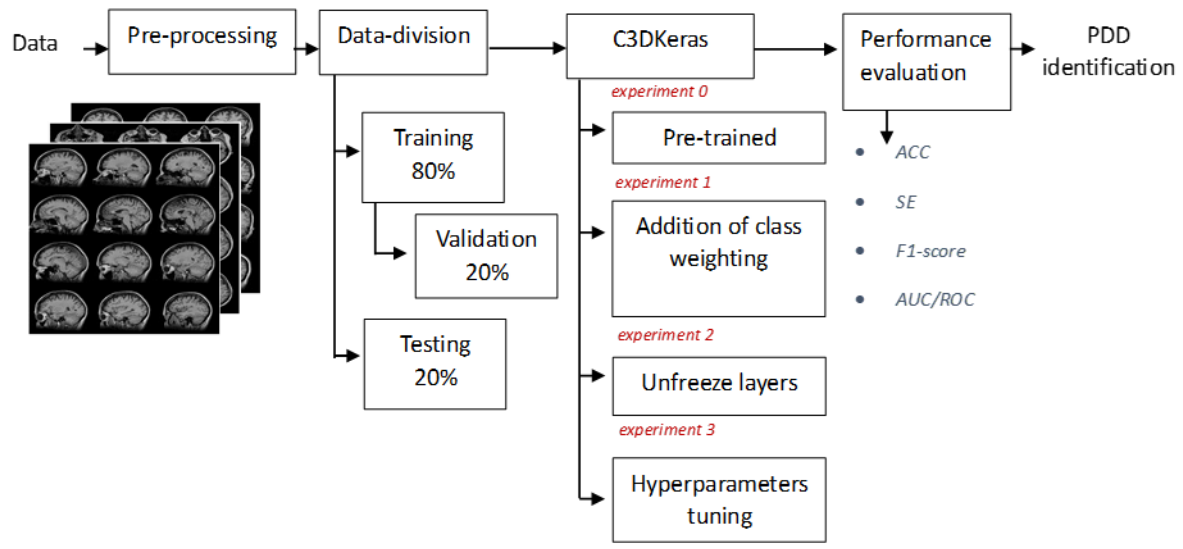


Figure 40. The workflow of improving the performance of the identification of PDD subjects: four experiments were conducted; in the experiment 0 neither class weighting nor hyperparameters were performed; in the experiment 1 class weights were add; in the experiment 2 some layers were unfrozen and in the experiment 4 the hyperparameters tuning was executed.

6.3.1. Experiment 0: Pre-trained C3DKeras

The first experiment also called the "Experiment 0" is intended as the one in which neither class weighting nor hyperparameters were performed. This experiment was executed in order to understand if this type of pretrained -on video clips on sports- CNN, could provide effective results in the case of medical images and if not, where should the attention be directed in order to improve the classification process.

6.3.2. Experiment 1: Addition of class weighting

Since the original dataset was imbalanced, for "Experiment 1", weights associated with each class were calculated and taken into account in both training and testing phase. Class imbalance is an important topic in ML since the learning algorithms tend to exhibit bias towards the majority class and in extreme cases might pay no attention to the minority class altogether.

The weights were calculated considering the total number of scans divided by the product of the number of classes by the number of scans for each class. This formula aims to make the class weight inversely proportional to the respective class frequency. The idea is to assign higher weight to less frequent class, compensating for the class imbalance.

During the training process, these weights were applied in order to assist the algorithm in learning from less frequent class as well as was doing from the more frequent one.

6.3.3. Experiment 2: Unfreeze layers

In the "Experiment 0" as well as in "Experiment 1" the layers were frozen, meaning that the weights and biases were blocked from being updated during the training process. The frozen layers retained the knowledge it has been gained from the pre-trained process without accessing to the data of this study. In order to make the C3DKeras to learn from the actual data, the last two layers were unfrozen, giving to the model the possibility to improve the accuracy as in this way it has information on the real problem.

6.3.4. Experiment 3: Addition of hyperparameter tuning

The study most important contribution is indeed the application of C3DKeras. However, the components that need to be considered carefully for creating the learning algorithm is choosing the right set of hyperparameters for optimizing the accuracy of the classification.

In this study, the primary aspect that lies in the optimization algorithm is to minimize the testing error of the model. In order to do so, different hyperparameters may be applied.

There are three hyperparameters that usually influence the model performance: learning rate, dropout rate, and batch size.

The learning rate hyperparameter controls the rate or speed at which the model learns. It represents the step size taken by the optimization algorithm during the weight update process. A higher learning rate may lead to faster convergence but risks overshooting the optimal solution, while a lower rate may improve stability but prolong training. In the Adam optimizer, there is no need to manually set a global learning rate as it would be in traditional gradient descent. Instead, could be needed to tune hyperparameters specific to Adam, such as:

- Learning Rate**:** Although Adam adapts the learning rates for each parameter, it might be possible to change the initial learning rate. The default is set to 0.001.

- Beta1 (β1) and Beta2 (β1): The hyper-parameters β1 and β2 of Adam are initial decay rates used when estimating the first and second moments of the gradient, which are multiplied by themselves (exponentially) at the end of each training step (batch) [43]. Decreasing β1 and β2 will make the learning slower. The default values are often set to 0.9 and 0.999.

- Epsilon (epsilon): A small constant added to improve numerical stability. The default is $1e^{-7}$.

In this study, using the Adam optimizer, only the learning rate was tuned, considering the following values: [0.001,0.0005,0.0001,0.00005,0.00001].

The second global hyperparameter that can be changed is the dropout. The main idea behind dropout is that at each training step, every node is either "dropped" (ignored) with a probability p or "kept" with a probability 1 – p. In the case of drop both the node and its connecting edges are temporarily removed from the model and the training occurs on the remaining neurons. This is done in order to prevent overfitting; with dropout rate small the accuracy will gradually increase, and the loss will gradually decrease. The initial value of the dropout was 0.5; the following values were used in the dropout tuning: [0.5,0.4,0.3,0.2].

The batch size is the hyperparameter that defines the number of training examples used at one time during training to work through before updating the internal model parameters. The choice of batch size is a crucial consideration that affects both computational efficiency and memory usage. Smaller batch sizes result in more frequent weight updates during training iterations, but this can increase the number of epochs required for convergence. On the other hand, larger batch sizes can accelerate the training process but may compromise the model's ability to generalize effectively. Choosing the batch size needs a balance between the level of detail in weight updates and the overall efficiency of the training process. After the initial employing of 10 batch size the number applied in this study has been decreased as follow: [10, 8, 6, 4, 2, 1].

7. Results

7.1. Experiment 0: Pre-trained C3DKeras

The Experiment 0 provides an accuracy of 0.55 (Figure 40.), however a high misdiagnosis of the prodromal PD subjects also emerged. The overall ACC and SE, F1-Score and AUC for PDD class reported in percentage are found in Table 6.

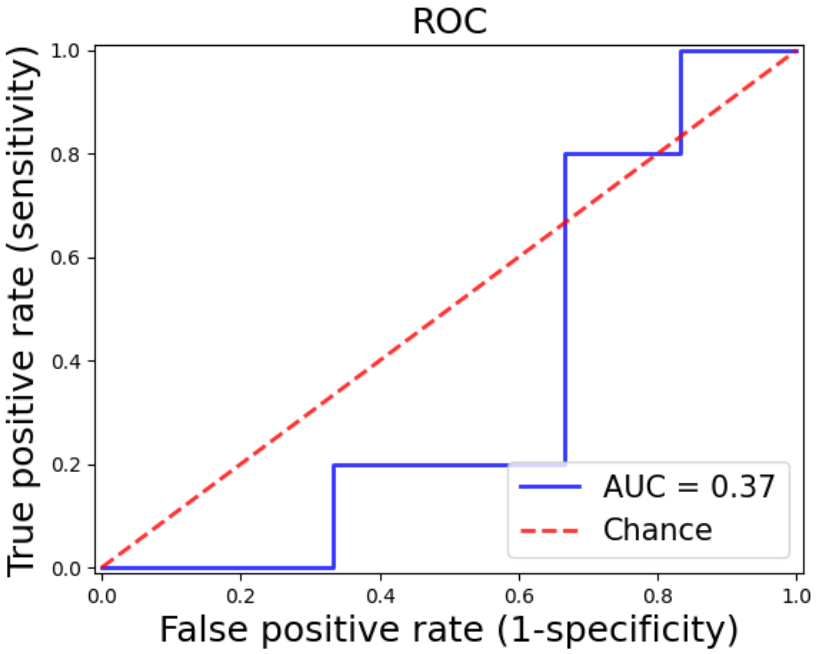


Figure 41. ROC – AUC of Experiment 0.

Table 6. Performance for the identification PDD class in Experiment 0.

ACC (%)	SE (%)	F1-score (%)	AUC (%)
55%	100%	71%	37%

7.2. Experiment 1: Addition of class weighting

The second experiment concerning the addition of weighting for the class imbalance gives an overall ACC of 64% and an AUC of 50% (Figure 42.). The SE, F1-Score and AUC for PDD class, as well as the overall ACC for Experiment 1 are found in Table 7.

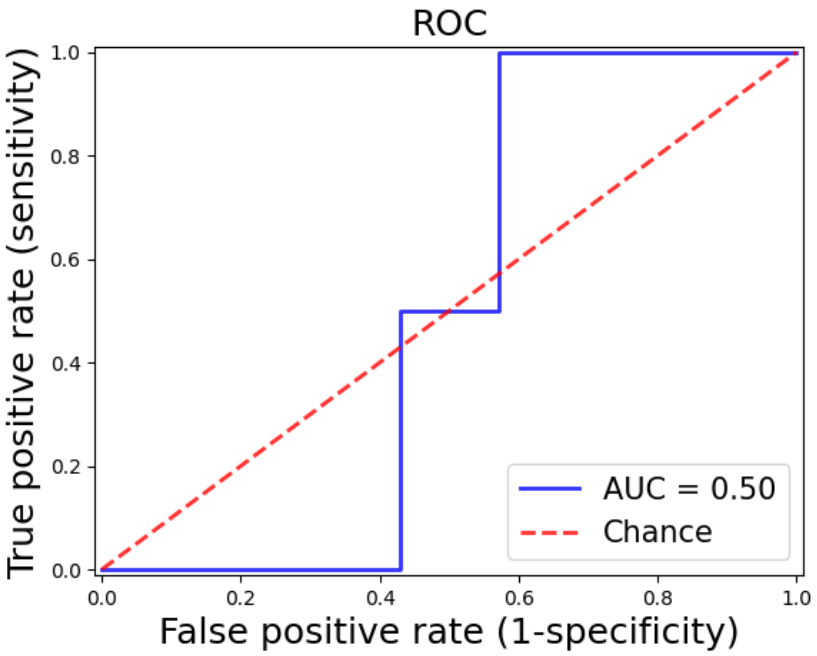


Figure 42. ROC – AUC of Experiment 1.

Table 7. Performance for the identification of the PDD class in Experiment 1.

ACC (%)	SE (%)	F1-score (%)	AUC (%)
64%	100%	78%	50%

7.3. Experiment 2: Unfreeze layers

In the Experiment 2, concerning the addition of the capacity of C3DKeras to learn from the actual data, gives an overall ACC of 45% and an AUC related to the identification of PDD of 63% (Figure 43.). The performances related to this experiment are found in Table 8.

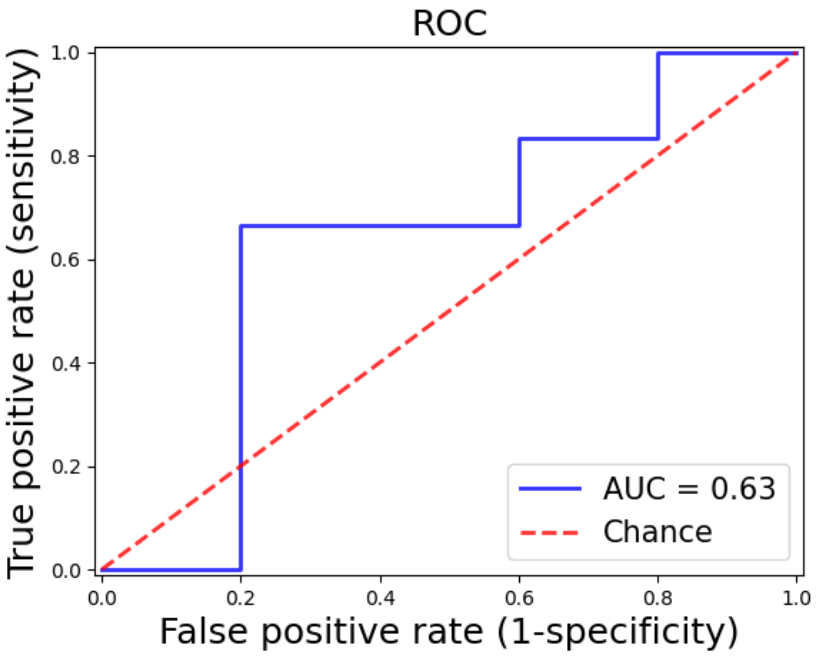


Figure 43. ROC – AUC of Experiment 2.

Table 8. Performance for the identification of the PDD class in Experiment 1.

ACC (%)	SE (%)	F1-score (%)	AUC (%)
45%	100%	62%	63%

7.4. Experiment 3: Addition of hyperparameter tuning

After the addition of the weighting class and the unfrozen layers some experiments for the hyper-parameter tuning were performed.

The best-performing combination for PDD identification turned out to be the following:

- Best dropout rate value: 0.2;
- Best learning rate value: 0.00001;
- Best batch size value: 4.

The overall ACC and the SE, F1-score and AUC for the above hyper-parameters can be observed in Table 9, while the ROC-AUC is shown in Figure 44.

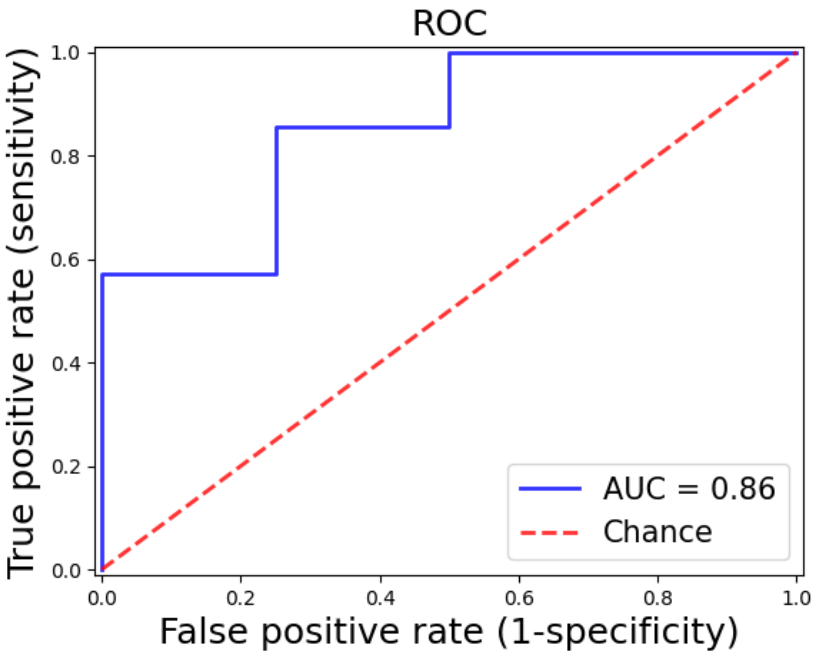


Figure 44. ROC – AUC of Experiment 3.

Table 9. Performance for the identification of the PDD class in Experiment 3.

ACC (%)	SE (%)	F1-score (%)	AUC (%)
73%	100%	73%	86%

7.5. Overall results

In Table 10. are shown the performance of all the experiments. As it can be observed the best performance is given by the Experiment 3, having an overall performance accuracy of 73%, a sensibility of PDD class of 100% and an area under the curve of 86%.

In Table 11., it can be observed the differences of performances between the experiment 3 of this study with the ones found in literature.

Since Experiment 3 was the one with the higher performances, the Confusion matrix of this experiment is also reported in Figure 45.

Table 10. Performance for the identification of the PDD class in all 4 experiments.

Experiment	ACC (%)	SE (%)	F1-score (%)	AUC (%)
0	55%	100%	71%	37%
1	64%	100%	78%	50%
2	45%	100%	62%	63%
3	***73%***	***100%***	***73%***	***86%***

Table 11. Performance for the identification of the PDD class from literature and from Experiment 3.

Study	ACC (%)	SE (%)	F1-score (%)	AUC (%)
[34]	-	50%	-	50%
[31]	73.7%	67%	-	73%
Experiment 3	***73%***	***100%***	***73%***	***86%***

8. Discussion

8.1. Materials, methodology and results

The study concerns the application of C3DKeras a 3D convolutional neural network architecture for the detection of PDD from MRI scans.

After the acquisition, the initial step was to convert the DICOM files to NIfTI files. Out of the total number of files, only 55 were suitably converted, as some of the DICOM files did not comprise pertinent image data for useful conversion. Additionally, two files had to be excluded to ensure homogeneity between scans as the registration process could not be applied to them. Consequently, only 53 MRI scans were utilized in the last phases of the preprocessing pipeline. Out of the 53 scans, 28 were of individuals with PDD and 25 were of those with prodromal PD, resulting in an imbalance between the two groups.

As previously mentioned, the PPMI is a multicenter study, hence, the acquired MRI scans had spatial and temporal differences. In order to bring all the MRI scans to the same space, an image registration process was performed over all the MRI scans. The registration of images was performed using MNIPD25-T1MPRAGE-1 mm atlas. Since the MRI scans had a different orientation from the MNIPD25-T1MPRAGE-1 mm atlas, on this last one was performed a change in orientation from LPI to RAI. Thus, the translation for the moving images (the MRI scans) to the static one (MNIPD25-T1MPRAGE-1 mm atlas) could correctly be performed as both were on the same plane, hence the process of registration could be performed properly.

The second step of the preprocessing was to normalize the data in order to have a common scale across different subjects and imaging session.

Lastly, as a way to facilitate the process of identification PDD subjects, the bias field correction was applied in the preprocessing pipeline.

After preprocessing the MRI scans were insert in the C3DKeras. As explained above, the C3DKeras is an adoption of a 3D CNN. A 3D CNN is designed to handle three-dimensional data, such as volumetric images as the MRI scans or video data, as the ones on which was firstly applied C3DKeras. The architecture of the C3DKeras consists of eight convolutional layers, five pooling layers, and three fully connected layers. C3DKeras was trained using labeled video datasets for tasks such as action recognition. The training involved optimizing the parameters

to discriminative spatiotemporal features for accurate classification [42]. The pre-trained C3DKeras was applied on the MRI scans for the identification of PDD subjects. This is the first study that applied a 3D CNN on identification of PDD subjects from MRI scans.

This study was conducted as an Ablation study. Ablation studies are particularly important in machine learning, precisely, in evaluating how the removal of certain components affects performance. Four experiments were performed, with the aim of improving the overall accuracy, the sensitivity and the area under the curve.

In Experiment 0, no class weighting or hyper-parameter adjustments were made. This experiment was performed as a way to evaluate the different steps of the ablation study. The results in Table 6 show that both the overall ACC and AUC were low, which is why the other three experiments were conducted. C3DKeras is a pre-trained CNN and in order to have good performances for the actual task, alterations are necessary.

In Experiment 1 the class weighting was added, as C3DKeras displayed bias towards the PDD class. The outcome of this trial showed an improvement of nearly 20% in AUC, indicating the significance of class weighting for imbalanced datasets.

In Experiment 2, unfrozen layers were added. Specifically, in this study the last two convolutional layers were unfrozen, giving the possibility of the network to study directly from the actual data. Unfreezing layers is helpful when the distribution of the new dataset differs from that of the pre-training dataset, allowing the model to adjust its parameters more effectively to the actual task. This step unfortunately, lower the ACC of the previous experiment. However, it increases the AUC of 10%. A higher AUC score indicates that the test has good discriminatory power, assisting doctors in making informed decisions about patient care.

In the last experiment, more than 50 combination of parameter tuning were performed. The best combination was given by a dropout rate value of 0.2, a learning rate value of 0.00001 and a batch size value of 4. As a consequence of applying these parameters, the AUC has increased by over 20% and the ACC has increased by nearly 10%, which confirms the significance of the tuning step.

The ablation study conducted validates the significance of each step in applying a pre-trained CNN. Table 10 highlights the effect of each step, demonstrating an improvement in the overall performance of PDD identification.

As it can be observed in Table 11. the results from this study are similar to the one found in literature. Specifically, the performances of Experiment 3 can be compared to [31], as this last one was the only article explicitly focused on the identification of PDD subjects. On the other hand, the focus of [34] was on identifying Alzheimer's subjects, with only a secondary test attempting to examine the performance of different pathologies such as PDD and PD. The outcome of the secondary test was however low, pointing out the importance of providing more study that concentrate only on the Parkinson related dementia.

8.2.Advantages

Based on the literature search, it emerged that no study used MRI data or applied a CNN to identify PDD subjects. However, MRI is a non-invasive technique, providing a high level of information about brain structure while ensuring the safety of the subjects. Furthermore, the CNN demonstrated high performance in different fields as well as the medical one. In fact, one of the advantages of the CNN is that it automatically learns features from data without the need for manual feature extraction. This is particularly useful in tasks where extracting relevant features is challenging, as the one of identification of PDD subjects. This study merges the benefits from the acquisition technique as well as the excellent performance of the CNN, in particular of the C3DKeras (a pre-trained 3D CNN), for classification.

This is the first study that use 3D CNN for the identification of PDD subjects from MRI scans, that unlike 2D CNN that process slices independently, it considers the entire three-dimensional structure of the MRI volume. This helps in preserving spatial relationships and capturing intricate details that may be crucial for accurate disease classification.

The diagnose of PDD is still done with the 1-year rule, being in some cases too late for the correct treatment. As up to 83% of patients with PD eventually develop dementia later in the disease course, faster the identification, sooner could be provided the correct treatment, giving the possibility to increase the life expectance.

This is the first study with high AUC, ACC and high SE for both classes. In fact, differently from [31], that has an SE of the PD class around 12%, in this study the SE of the PD class arrive at around 57%, providing great prospect in inserted this kind of medical tool in the clinical field.

A total of four ablation experiments were performed to determine the step that most impacted the classification performance. It emerged that the Experiment 3 not only had the best performance but was also the one with the highest increase in performance among the other three experiments. It turned out that the step of hyperparameter tuning is essential to archive optimal model performance.

8.3. Limits

The major limitation was the dramatic scarcity of open-access data due to the serious challenge of distinguish between prodromal PD and PDD subjects, as the two pathologies are deeply correlated; this was reflected in the results as well.

As the literature research chapter shows, just a few studies focused on the classification of the Parkinson related dementia. As this type of disease is less studied, it emerged that there are also few data sets available, increasing again the complexity of the task. In fact, for a better identification of PDD, it would be required a higher dataset than the one used in this study.

Even though the 3D CNN could give out better performance than the 2D CNN, it also requires a high computational effort. Improving the length of the dataset could also require a higher period for the correct identification of the diseases.

8.4. Future perspectives

In order to improve the overall performance of C3DKeras in identifying PDD subjects, future studies should use a larger dataset. Additionally, as in the medical field is essential to have an interpretation of the results, the future studies should, also focus on the development of a 3D class activation map (3D CAM) that could help to determine whether the model is selecting the appropriate areas or regions for PDD detection

II. Conclusion

In the proposed study, a 3D MRI analysis was performed for the detection of PDD using C3DKeras. The MRI data used in this study was collected from Parkinson's Progression Markers Initiative (PPMI).

PPMI is an international and multicenter study, therefore some pre-processing steps were performed. All the pre-processing steps as well as the application of the C3DKeras were performed on Colab. The pre-processing steps were as follows: orientation of the template, registration, normalization and bias field correction.

The entire methodology has been carried out within the structure of an ablation study. In Deep Learning (DL) the terms "ablation study" has been adopted to describe a procedure where certain parts of the network are removed (or changed or inserted), in order to determine which parameter improves the overall performance. Four ablation experiments were performed: Experiment 0, Experiment 1, Experiment 2 and Experiment 3. In Experiment 0 also called Pre-Trained C3DKeras, no class weighting nor hyperparameter tuning were performed. This gave out a low overall performance. In Experiment 1 weighing associated with each class were added, improving considerately the performance. Experiment 2 focused on unfreezing the last two convolutional layers, while Experiment 3 focused on the hyperparameter tuning. At the end the best performance was obtained by the Experiment 3, having an AUC of 86% and SE of 100% for the PDD class.

Confronting the results with those found in literature research it emerged that, with one of the studies the comparison could not be completely performed as it concentrated on Alzheimer identification, making only a secondary test on the identification of PDD. Therefore, only one study could be fully confronted with the one presented in this research. Differently from the study found in literature this is the only study with high performances for both PD and PDD classes.

This study provides an excellent medical tool for identification the subjects affecting by PDD from the PD subjects. In particular, as it was realized as an ablation study, it could be clearly observed how the essential fine-tuning steps could highly improve the overall performance.

To conclude, the outcome of the proposed study is very motivating. However, it would be required an interpretation tool in order to provide insights into whether the model accurately selected the relevant region of interest within the MRI data. The inclusion of such an interpretation tool enables clinicians and researchers to understand and validate the decisions made by the model, thereby improving the overall reliability and clinical utility of the developed methodology for PDD diagnosis.

Figures

Tables

III. Bibliography

[1] Purves, D. (2004). *Neuroscience*. Sinauer Associates Incorporated.

[2] G. Ambrosi, D. Cantino, P. Castano, S. Correr, L. D'Este, R.F. Donato, G. Familiari, F. Fornai, M. Gulisano, A. Iannello, L. Magaudda, M.F. Marcello, A.M. Martelli, P. Pacini, M. Rende, P. Rossi, C. Sforza, C. Tacchetti, R. Toni, G. Zummo, Anatomia dell'uomo, Edi.Ermes s.r.l, Milano, 2006, Parte quinta Embriologia e organogenesi

[3] Tortora, G. J., & Derrickson, B. H. (2016). *Principles of anatomy and physiology*. Wiley Global Education.

[4] Morell P, Quarles RH. The Myelin Sheath. In: Siegel GJ, Agranoff BW, Albers RW, et al., editors. Basic Neurochemistry: Molecular, Cellular and Medical Aspects. 6th edition. Philadelphia: Lippincott-Raven; 1999. Available from: https://www.ncbi.nlm.nih.gov/books/NBK27954/

[5] Van Schependom, J., & D'haeseleer, M. (2023). Advances in Neurodegenerative Diseases. *Journal of clinical medicine*, *12*(5), 1709. https://doi.org/10.3390/jcm12051709

[6] Goedert, M., Jakes, R., & Spillantini, M. G. (2017). The Synucleinopathies: Twenty Years On. *Journal of Parkinson's disease*, *7*(s1), S51–S69. https://doi.org/10.3233/JPD-179005

[7] Iwai A, Masliah E, Yoshimoto M, Ge N, Flanagan L, de Silva HA, Kittel A, Saitoh T. The precursor protein of non-A beta component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. Neuron. 1995;14:467–75.

[8] Koga, Shunsuke & Sekiya, Hiroaki & Kondru, Naveen & Ross, Owen & Dickson, Dennis. (2021). Neuropathology and molecular diagnosis of Synucleinopathies. Molecular Neurodegeneration. 16. 10.1186/s13024-021-00501-z.

[9] Rocha Cabrero, F., & Morrison, E. H. (2023). Lewy Bodies. In *StatPearls*. StatPearls Publishing

[10] DeMaagd, G., & Philip, A. (2015). Parkinson's Disease and Its Management: Part 1: Disease Entity, Risk Factors, Pathophysiology, Clinical Presentation, and Diagnosis. *P & T : a peer-reviewed journal for formulary management*, *40*(8), 504–532.

[11] https://www.msdmanuals.com/en-pt/professional/neurologic-disorders/movement-and-cerebellar-disorders/parkinson-disease

[12] Castonguay, Anne-Marie & Gravel, Claude & Levesque, Martin. (2020). Treating Parkinson's Disease with Antibodies: Previous Studies and Future Directions. Journal of Parkinson's Disease. 11. 1-22. 10.3233/JPD-202221.

[13] Inguanzo, A., Poulakis, K., Mohanty, R. *et al.* MRI data-driven clustering reveals different subtypes of Dementia with Lewy bodies. *npj Parkinsons Dis.* **9**, 5 (2023). https://doi.org/10.1038/s41531-023-00448-6

[14] Jellinger, K. A., & Korczyn, A. D. (2018). Are dementia with Lewy bodies and Parkinson's disease dementia the same disease?. *BMC medicine*, *16*(1), 34. https://doi.org/10.1186/s12916-018-1016-8

[15] Yousaf, T., Dervenoulas, G., Valkimadi, P. E., & Politis, M. (2019). Neuroimaging in Lewy body dementia. *Journal of neurology*, *266*(1), 1–26. https://doi.org/10.1007/s00415-018-8892-x

[16] Armstrong M. J. (2021). Advances in dementia with Lewy bodies. *Therapeutic advances in neurological disorders*, *14*, 17562864211057666. https://doi.org/10.1177/17562864211057666

[17] McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J. P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G., Bayston, A., Beach, T. G., Blanc, F., Bohnen, N., Bonanni, L., Bras, J., Brundin, P., Burn, D., Chen-Plotkin, A., Duda, J. E., … Kosaka, K. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. *Neurology*, *89*(1), 88–100. https://doi.org/10.1212/WNL.0000000000004058

[18] Sanchez-Castaneda, C., Rene, R., Ramirez-Ruiz, B., Campdelacreu, J., Gascon, J., Falcon, C., Calopa, M., Jauma, S., Juncadella, M., & Junque, C. (2009). Correlations between gray matter reductions and cognitive deficits in dementia with Lewy Bodies and Parkinson's disease with dementia. *Movement disorders : official journal of the Movement Disorder Society*, *24*(12), 1740–1746. https://doi.org/10.1002/mds.22488

[19] Mahesh M. (2013). The Essential Physics of Medical Imaging, Third Edition. *Medical physics*, *40*(7), 10.1118/1.4811156. https://doi.org/10.1118/1.4811156

[20] Pantano, P., Caramia, F., & Pierallini, A. (1999). The role of MRI in dementia. *Italian journal of neurological sciences*, *20*(5 Suppl), S250–S253. https://doi.org/10.1007/s100729970006

[21] Petrou, M., Kotagal, V., & Bohnen, N. I. (2012). An update on brain imaging in parkinsonian dementia. *Imaging in medicine*, *4*(2), 201–213. https://doi.org/10.2217/iim.12.10

[22] Merlijn van Breugel, Rudolf S. N. Fehrmann, Marnix Bügel, et al. Current state and prospects of artificial intelligence in allergy. *Authorea.* April 21, 2023.

[23] Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*, *45*(1 Suppl), S199–S209. https://doi.org/10.1016/j.neuroimage.2008.11.007

[24] Capozzoli, Alfonso & Cerquitelli, Tania & Piscitelli, Marco. (2016). Enhancing energy efficiency in buildings through innovative data analytics technologies. 10.1016/B978-0-12-803663-1.00011-5.

[25] Aggarwal, C. C. (2018). Neural networks and deep learning. In *Springer eBooks*. https://doi.org/10.1007/978-3-319-94463-0

[26] https://www.deeplearningbook.org/

[27] C. Wang, "A Review on 3D Convolutional Neural Network," 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 2023, pp. 1204-1208, doi: 10.1109/ICPECA56706.2023.10075760.

[28] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical research ed.), 372, n71. https://doi.org/10.1136/bmj.n7

[29] Aromataris E, Munn Z. Chapter 1: JBI Systematic Reviews. In: Aromataris E, Munn Z (Editors). JBI Manual for Evidence Synthesis. JBI, 2020.

[30] Nakajima, K., Saito, S., Chen, Z., Komatsu, J., Maruyama, K., Shirasaki, N., Watanabe, S., Inaki, A., Ono, K., & Kinuya, S. (2022a). Diagnosis of Parkinson syndrome and Lewy-body disease using 123I-ioflupane images and a model with image features based on machine learning. Annals of Nuclear Medicine, 36(8), 765–776. https://doi.org/10.1007/s12149-022-01759-z

[31] Booth, S., Park, K. W., Lee, C. S., & Ko, J. H. (2022). Predicting cognitive decline in Parkinson's disease using FDG-PET-based supervised learning. *The Journal of clinical investigation*, *132*(20), e157074. https://doi.org/10.1172/JCI157074

[32] Van Veen, R., Meles, S. K., Renken, R., Reesink, F. E., Oertel, W. H., Janzen, A., De Vries, G., Leenders, K. L., & Biehl, M. (2022). FDG-PET combined with learning vector quantization allows

classification of neurodegenerative diseases and reveals the trajectory of idiopathic REM sleep behavior disorder. Computer Methods and Programs in Biomedicine, 225, 107042. https://doi.org/10.1016/j.cmpb.2022.107042

[33] Jennings, Jack & Peraza, Luis & Baker, Mark & Alter, Kai & Taylor, John-Paul & Bauer, Roman. (2022). Investigating the power of eyes open resting state EEG for assisting in dementia diagnosis. Alzheimer's Research & Therapy. 14. 10.1186/s13195-022-01046-z.

[34] Katako, A., Shelton, P., Goertzen, A. L., Levin, D., Bybel, B., Aljuaid, M., Yoon, H. J., Kang, D. Y., Kim, S. M., Lee, C. S., & Ko, J. H. (2018). Machine learning identified an Alzheimer's disease-related FDG-PET pattern which is also expressed in Lewy body dementia and Parkinson's disease dementia. *Scientific reports*, *8*(1), 13236. https://doi.org/10.1038/s41598-018-31653-6

[35] https://www.ppmi-info.org/

[36] Chakraborty, S., Aich, S., & Kim, H. C. (2020). Detection of Parkinson's Disease from 3T T1 Weighted MRI Scans Using 3D Convolutional Neural Network. *Diagnostics (Basel, Switzerland)*, *10*(6), 402. https://doi.org/10.3390/diagnostics10060402

[37] https://www.dicomstandard.org/

[38] Alemán-Gómez, Y. IBASPM: Toolbox for automatic parcellation of brain structures. In Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping, Florence, Italy, 11–15 June 2006

[39] Xiao, Y., Fonov, V., Bériault, S., Al Subaie, F., Chakravarty, M. M., Sadikot, A. F., Pike, G. B., & Collins, D. L. (2015). Multi-contrast unbiased MRI atlas of a Parkinson's disease population. *International journal of computer assisted radiology and surgery*, *10*(3), 329–341. https://doi.org/10.1007/s11548-014-1068-y

[40] https://nist.mni.mcgill.ca/

[41] Miikkulainen, R.; Liang, J.; Meyerson, E.; Rawal, A.; Fink, D.; Francon, O.; Raju, B.; Shahrzad, H.; Navruzyan, A.; Duffy, N.; et al. Evolving deep neural networks. In Artificial Intelligence in the Age of Neural Networks and Brain Computing; Academic Press: Cambridge, CA, USA, 2019; pp. 293–312

[42] Tran, Du & Bourdev, Lubomir & Fergus, Rob & Torresani, Lorenzo & Paluri, Manohar. (2015). Learning Spatiotemporal Features with 3D Convolutional Networks. 4489-4497. 10.1109/ICCV.2015.510.

[43] Kingma, Diederik & Ba, Jimmy. (2014). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations.

IV. Acknowledgments

I would like to thank my supervisor Dr. Agnese Sbrollini and my co-supervisor Prof. Laura Burattini for giving me the possibility to work on a subject that fascinates me and for believing in the capacity of each of us -that pass through their classroom- in creating and improving this world.

And last but not least, I am grateful for the support of my second co-supervisor Dr. Selene Tomassini, that was always present and answered each of my silly questions during the process of writing this thesis. She brought back in me, my lost love for the brain and its complexity and I will not have enough words to thank her for this. I am truly grateful for her willingness to share her knowledge with me, leading me to the answers I sought. I would be really, and infinitely delighted if each of the future students could have a dedicated and knowledgeable professor as she is.