

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

Detección automática de las regiones cerebrales asociadas a los desórdenes parkinsonianos en imágenes de resonancia magnética usando técnicas de Aprendizaje Profundo

PROYECTO INTEGRADOR

Previo la obtención del Título de:

Ingeniero/a en Computación

Presentado por:

Ana María Briones Chonillo

Carlos Jiménez Farfán

GUAYAQUIL - ECUADOR

Año: 2022 - 2023

DEDICATORIA

Al concluir esta etapa importante de mi vida académica, quiero dedicar esta tesis como homenaje a mi abuelita Bella y mi inolvidable mamita Olguita. Quienes, aunque ya no están físicamente conmigo, estarán siempre en mi corazón y en mis pensamientos. Su amor incondicional y su dedicación me inspiraron a seguir adelante y alcanzar mis metas. A mis queridos padres, les agradezco por ser mi roca y por estar siempre a mi lado, brindándome su amor y apoyo incondicional. Gracias por creer en mí y por ayudarme a superar los obstáculos en el camino, los amo.

Ana Maria Briones Ch.

DEDICATORIA

El presente proyecto está dedicado principalmente a mi mamá Olga Farfán Vera, mi papá Carlos Jiménez Veloz y hermanos Belén, Angie y Fabiancito, quienes han sido mi motor desde el comienzo de mi vida académica y siempre me han aconsejado e impulsado a salir adelante, cumplir mis objetivos y nunca conformarme. Los amo.

Carlos Jiménez Farfán

AGRADECIMIENTOS

Dios, tu amor no tiene fin, me permites sonreír antes mis logros que son el resultado de tu ayuda, tú me has dado la sabiduría para alcanzar mis objetivos. A Joseph, Davis, Katy, mis sobrinos, Jessy, Sandy, Gerson y David, les quiero expresar mi agradecimiento por ser parte de mi vida y una constante fuente de sabios consejos, alegría y amor. En especial, mi sobrino Elías quien, con su luz y su alegría, ha iluminado mi vida y me ha dado la motivación para seguir adelante. A mis amigos, durante todo este tiempo, han sido una fuente constante de apoyo, no solo en el ámbito académico, sino también emocional. Su presencia ha hecho de este viaje algo mucho más fácil y especial. A la Ing. Rosa Díaz y al Ing. Rafael Bonilla por sus consejos, sus enseñanzas y por haber confiado en mis capacidades.

Ana María Briones Ch.

AGRADECIMIENTOS

Mi más sincero agradecimiento a todos los docentes con los que compartí aulas en mi vida universitaria, definitivamente cada uno de ellos supo agregar un granito más de conocimiento y sabiduría en mí. A mi tutor quien supo guiarnos en cada parte de este proceso académico. Quiero agradecer por sobre todo a mi familia, sin ellos llegar hasta aquí no habría sido posible.

Carlos Jiménez Farfán

DECLARACIÓN EXPRESA

“Los derechos de titularidad y explotación, nos corresponde conforme al reglamento de propiedad intelectual de la institución; Ana Briones Chonillo y Carlos Jiménez Farfán, y damos nuestro consentimiento para que la ESPOI realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”

A handwritten signature in black ink, appearing to read 'Ana Briones', with a large, stylized flourish above the name.

Ana Briones Chonillo

A handwritten signature in black ink, appearing to read 'Carlos Jiménez', with a large, stylized flourish below the name.

Carlos Jiménez Farfán

EVALUADORES

Miguel Andrés Realpe Robalino Ph.D.
PROFESOR DE LA MATERIA

Colón Enrique Peláez Jarrín Ph.D.
PROFESOR TUTOR

RESUMEN

La Enfermedad de Parkinson (EP) es un trastorno neurológico degenerativo, que no tiene cura y no distingue edad. La EP se caracteriza por un daño al sistema nervioso central que provoca la caída de los niveles de dopamina causando temblores, rigidez muscular y pérdida del equilibrio. Para contrarrestar estos síntomas motores se utilizan medicamentos que imitan la acción de la dopamina, sin embargo, se ha demostrado que el consumo prolongado de estos medicamentos puede resultar en efectos secundarios negativos en los pacientes, tales como desórdenes parkinsonianos que incluyen apatía y los trastornos de control de impulsos (ICD). No obstante, aún existe dificultad en determinar las regiones cerebrales asociados a estos desórdenes.

Se desarrolló una metodología que permite detectar automáticamente las regiones del cerebro que permitan llegar a un diagnóstico respecto a la presencia de desórdenes parkinsonianos utilizando imágenes de resonancia magnética (IRM) en conjunto con técnicas de aprendizaje profundo en 3D.

Se realizaron variados experimentos con diferentes arquitecturas del estado del arte de visión por computador (CV) como Swin-Transformer, Compact Convolutional Transformer (CCT) y MeshNet. De los cuales el CCT superó, alcanzando una precisión mayor al 97% para analizar imágenes de resonancia magnética y clasificar los desórdenes parkinsonianos.

Por último, utilizando el algoritmo de interpretabilidad Guided Backpropagation se obtuvo una visualización de las regiones cerebrales activadas en la detección de los desórdenes Parkinsonianos, lo que permite a los médicos e investigadores tener una comprensión más profunda de la anatomía y el funcionamiento cerebral de los pacientes con estos desórdenes.

Palabras Clave: CCT, Transformers, Parkinson, IRM, Guided Backpropagation

ABSTRACT

Parkinson's Disease (PD) is a degenerative neurological disorder, which has no cure and does not distinguish age. PD is characterized by damage to the central nervous system that causes dopamine levels to drop, causing tremors, muscle stiffness, and loss of balance. To counteract these motor symptoms, drugs that mimic the action of dopamine are used, however, it has been shown that prolonged use of these drugs can result in negative side effects in patients, such as parkinsonian disorders that include apathy and mental disorders. impulse control (ICD). However, there is still difficulty in determining the brain regions associated with these disorders.

The methodology developed in this work seeks to automatically detect the brain regions that allow a diagnosis regarding the presence of parkinsonian disorders using magnetic resonance imaging (MRI) and 3D deep learning models.

Several experiments were carried out with different state-of-the-art computer vision (CV) architectures such as Swin-Transformer, Compact Convolutional Transformer (CCT) and MeshNet. Of which the CCT exceeded, reaching an accuracy greater than 97% analyzing MRIs and classifying parkinsonian disorders.

Finally, using the Guided Backpropagation interpretability algorithm, a visualization of the brain regions activated in the detection of Parkinsonian disorders was obtained, which allows clinicians and researchers to have a deeper understanding of the anatomy and brain function of patients. with these disorders.

Keywords: *CCT, Transformers, Parkinson, MRI, Guided Backpropagation*

ÍNDICE GENERAL

EVALUADORES.....	7
RESUMEN	I
<i>ABSTRACT</i>	II
ÍNDICE GENERAL.....	III
ABREVIATURAS.....	VI
ÍNDICE DE FIGURAS	VII
ÍNDICE DE TABLAS	VIII
ÍNDICE DE ECUACIONES	IX
CAPÍTULO 1	1
1. Introducción	1
1.1 Descripción del problema.....	1
1.2 Justificación del problema.....	2
1.3 Objetivos	3
1.3.1 Objetivo General.....	3
1.3.2 Objetivos Específicos	3
1.4 Marco Teórico	3
1.4.1 Enfermedad de Parkinson	3
1.4.2 Trastornos de control de impulsos	4
1.4.3 Ludopatía.....	5
1.4.4 Hipersexualidad (HS)	6
1.4.5 Trastorno por Atracón (TPA).....	7
1.4.6 Apatía	8
1.4.7 Aprendizaje Profundo.....	10
1.4.8 Convolutional Neural Network.....	11
1.4.9 Transformer	12

1.4.10	Vision Transformers (ViT)	13
1.4.11	Swin Transformer	14
1.4.12	MeshNet.....	16
1.4.13	Compact Convolutional Transformer	19
1.4.14	Aprendizaje profundo aplicado a IRM.....	20
2.	METODOLOGÍA.....	21
2.1	Análisis.....	21
2.1.1	Requerimientos	21
2.1.1.1	Requerimientos Funcionales	22
2.1.2	Alcance y limitaciones de la solución	23
2.1.3	Riesgos y beneficios de la solución.....	23
2.1.4	Usuarios de la solución	23
2.2	Prototipado	24
2.2.1	Dataset	24
2.2.2	Módulo de Pre-procesamiento.....	25
2.2.3	Data Augmentation	31
2.2.4	Modelos de Clasificación	32
2.2.5	Algoritmo Guided Backpropagation	33
2.3	Evaluación	34
2.4	Diseño de la Solución	35
2.4.1	Diagrama de Bloques.....	35
2.5	Software	36
2.5.1	Arquitectura	36
2.5.2	Diagramas (Clases, casos de uso, interacción, flujo de procesos)	37
2.5.3	Diseño de la Base de Datos.....	39
2.5.4	Implementación del Prototipo	40
CAPÍTULO 3	41
3.	Resultados y análisis.....	41

3.1	Plan de implementación.....	41
3.2	Métricas de evaluación	45
3.2.1	Accuracy	45
3.2.2	Precision	45
3.2.3	Recall	45
3.2.4	F1 Score	45
3.3	Resultados del mejor modelo.....	46
3.4	Análisis de costos	47
4.	CONCLUSIONES Y RECOMENDACIONES	1
4.1	Conclusiones.....	1
4.2	Recomendaciones	1
5.	Referencias.....	2

ABREVIATURAS

ESPOL Escuela Superior Politécnica del Litoral

CCT Compact Convolutional Transformer

CNN Convolutional Neural Network

EP Enfermedad de Parkinson

IRM Imágenes de resonancia magnética

ICD Impulsive control disorder

CV Computer Vision

ÍNDICE DE FIGURAS

Figura 1.1 IRM de paciente con EP + ICD.....	5
Figura 1.2 IRM de pacientes con EP y Ludopatía.....	6
Figura 1.3. IRM de pacientes con EP+ HS	7
Figura 1.4. IRM de paciente con TPA	8
Figura 1.5. Máscaras ROI para análisis de resonancia magnética. [22].....	9
Figura 1.6. Estructura de una Neurona artificial.....	10
Figura 1.7 Bloques de construcción de una CNN [25]	11
Figura 1.8 Arquitectura del Transformer [26]	12
Figura 1.9 Arquitectura del Vision Transformer [29]	13
Figura 1.10 Arquitectura Swin Transformer [30]	14
Figura 1.11 Diseño de la ventana desplazada. [30].....	14
Figura 1.12 Dos bloques de Swin Transformers consecutivos. [30].....	15
Figura 1.13 Arquitectura MeshNet [33]	17
Figura 1.14 Partes de la cara de MeshNet. [33]	18
Figura 2.1 Diseño de la solución de alto nivel.....	24
Figura 2.2 IMR no procesada de paciente con EP.	26
Figura 2.3 IRM normalizada y co-registrada.	27
Figura 2.4 Extracción de cerebro en IRM	28
Figura 2.5 Corrección de campo de bias en IRM.....	29
Figura 2.6 Normalización de intensidad en IRM	30
Figura 2.7 IRM con detección de bordes	31
Figura 2.8 IRM con técnicas de data augmentation.....	32
Figura 2.9 Resultados del algoritmo Guided Backpropagation.....	33
Figura 2.10 Diagrama de Bloques	35
Figura 2.11 Diagrama de Casos de Uso.....	37
Figura 2.12 Diagrama de Actividades	38
Figura 2.13 Diseño de Base de Datos.	39
Figura 2.14 Prototipo Web con resultados del modelo	40
Figura 3.1 Diagrama semanal de planificación del proyecto	41
Figura 3.2 Matriz de Confusión	42
Figura 3.3 Arquitectura personalizada del CCT-3D utilizado en proyecto integrador	43
Figura 3.4 Interpretabilidad del CCT en el prototipo web.....	44

ÍNDICE DE TABLAS

Tabla 2.1 Híper-parámetros iniciales para los modelos de clasificación.	32
Tabla 2.2 Descripción de librerías usadas	36
Tabla 3.1 Resultados de Data Augmentation	42
Tabla 3.2 Comparación de modelos según la métrica Accuracy	46
Tabla 3.3 Resultado de métricas del modelo CCT-3D.....	46
Tabla 3.4 Costo total del desarrollo del prototipo.....	47

ÍNDICE DE ECUACIONES

Ecuación (3.1) Accuracy	45
Ecuación (3.2) Precision	45
Ecuación (3.3) Recall	45
Ecuación (3.4) F1 Score	45

CAPÍTULO 1

1. INTRODUCCIÓN

El diagnóstico temprano de trastornos de control de impulsos (ICD) o apatía en pacientes con Enfermedad de Parkinson (EP) en etapa temprana puede representar un reto para los médicos. Las imágenes médicas, como la resonancia magnética o la tomografía computarizada, se utilizan actualmente para detectar cambios en la anatomía cerebral de personas con EP y desórdenes parkinsonianos. Hoy en día, los algoritmos de aprendizaje profundo, como las redes neuronales convolucionales (CNN) o los modelos tipo Transformer, se están convirtiendo en una herramienta útil en el campo médico para detectar patrones relevantes para enfermedades en imágenes. En esta sección, se detalla sobre la EP y la dificultad para determinar las regiones cerebrales asociadas a los desórdenes Parkinsonianos. Además, se definen los modelos de aprendizaje profundo que se emplean en la solución del problema.

1.1 Descripción del problema

La Enfermedad de Parkinson es un trastorno neurodegenerativo común que tiene un inicio latente y un progreso lento [1] Las manifestaciones clínicas de la EP en los pacientes son muy heterogéneas presentando síntomas motores como: temblor o rigidez; y no motores como: trastornos del control de los impulsos (ICD), depresión, apatía, entre otros [2]

Los trastornos del control de los impulsos son comportamientos anormales, repetitivos y destructivos donde la persona carece de la capacidad de mantener el autocontrol [3] Aunque la base neurofisiopatológica de los ICD en pacientes con Enfermedad de Parkinson no se comprende bien, se ha planteado la hipótesis de que se produce como resultado de la administración crónica de fármacos dopaminérgicos. El término “trastornos del control de los impulsos” cubre los cuatro ICD principales que ocurren en la enfermedad de Parkinson, los cuales son: Ludopatía, trastorno por atracón, comportamiento sexual compulsivo y trastorno de compra compulsiva. Aunque los síntomas no motores no han sido objeto fundamental de estudio, tienen un gran impacto en la vida de los pacientes mientras la enfermedad va avanzando [4]

En una encuesta multicéntrica de más de 1000 pacientes con EP, el 97% de ellos informaron síntomas no motores [5]. En algunos pacientes, ciertas características no motoras de la EP pueden presentarse antes que los síntomas motores [5]. También se observa que los trastornos del estado de ánimo como la apatía, depresión, ansiedad se clasificaron entre los síntomas no motores más problemáticos en etapa temprana y tardía [6].

Las personas con EP pueden presentar apatía al sentir que han perdido sus capacidades motoras y no pueden encontrar motivación ni entusiasmo, incluso realizando actividades en las que alguna vez se involucraron y disfrutaron [7]. La apatía puede ser un síntoma muy frustrante, no solo para la persona con EP sino también para los amigos y familiares [2]

Los estudios de neuroimagen de la relación entre la EP y los síntomas no motores como ICD y apatía han hecho muchas contribuciones en la exploración de los cambios anatómicos cerebrales en los pacientes. Sin embargo, estudios previos de imágenes de resonancia magnética (IRM) en pacientes de EP con ICD aún muestran resultados discrepantes sobre las áreas de activación cerebral de estos trastornos [8]

Actualmente, existe dificultad en determinar las regiones cerebrales asociadas a ICD y Apatía en IRM de pacientes con EP, poder realizar un diagnóstico a tiempo y seguimiento continuo de trastornos de control de impulsos es esencial para mejorar la calidad de vida y tratamiento temprano en pacientes que padecen Parkinson [9]

1.2 Justificación del problema

Aunque existen varios estudios que analizan los síntomas no motores de la EP mediante IRM, no existe una herramienta de progresión para conocer la tendencia del desarrollo de ICD y apatía dentro de la EP a lo largo del tiempo [9]. Hoy en día, la evaluación de la progresión de la EP en conjunto con ICD y apatía se basa en el autoinforme del paciente o en la calificación cualitativa de un médico [10].

Sin embargo, este enfoque puede resultar en un retraso de hasta diez años en el diagnóstico, ya que sólo se diagnostica una vez que el paciente ya presenta síntomas evidentes. Para mejorar el proceso actual, este proyecto busca detectar el tipo de patrón asociado a esos comportamientos adictivos que se expresan en

los cambios anatómicos cerebrales implementando técnicas avanzadas de aprendizaje profundo. Este enfoque permitirá detectar los trastornos parkinsonianos de manera más temprana y proporcionar una mayor oportunidad para el tratamiento y la gestión de la enfermedad.

1.3 Objetivos

1.3.1 Objetivo General

Construir una herramienta que permita el análisis de IRM para identificar y localizar los cambios anatómicos en regiones específicas del cerebro, asociados a desórdenes Parkinsonianos utilizando técnicas de aprendizaje profundo.

1.3.2 Objetivos Específicos

- Implementar algoritmos de pre-procesamiento y Data Augmentation aplicados a IRM de pacientes con desórdenes Parkinsonianos.
- Diseñar un modelo de clasificación multiclase basado en redes neuronales profundas con arquitecturas tipo Transformer.
- Desarrollar un prototipo web que permita la clasificación de IRM, a partir del mejor modelo de clasificación seleccionado.

1.4 Marco Teórico

1.4.1 Enfermedad de Parkinson

La enfermedad de Parkinson y los trastornos neurodegenerativos parkinsonianos generalmente se desarrollan lentamente durante varios años [11]. Los pacientes afectados y sus familias enfrentan problemas físicos, psicosociales y espirituales cada vez mayores durante este período de tiempo. Los cuidados paliativos tienen un papel importante en estas condiciones a lo largo de la progresión de la enfermedad, y particularmente en las etapas posteriores a medida que se acerca la muerte [11].

1.4.2 Trastornos de control de impulsos

Los trastornos del control de los impulsos son una clase heterogénea de trastornos que se caracterizan por la realización repetida y excesiva de conductas típicas o inicialmente placenteras [4]. Los comportamientos de ICD pueden comenzar como acciones más motivadas durante el compromiso inicial y volverse menos impulsados por motivaciones placenteras con el tiempo [3]. Los ICD más específicos, agrupados en una categoría del Manual de diagnóstico y estadística (DSM), incluye ludopatía, posiblemente el ICD mejor estudiado, tanto en la población general como entre los pacientes con EP [12]. La categoría ICD en el DSM-IV también incluye una subcategoría no especificada, la cual se puede usar para diagnosticar varias afecciones que se han observado en la EP, incluida la hipersexualidad (posiblemente la ICD identificada más tempranamente en pacientes con EP), trastorno por atracón y las compras compulsivas [13]. Según un estudio donde se evaluaron las alteraciones estructurales y funcionales del cerebro en pacientes con EP-ICD, pacientes con EP sin ICD y pacientes de Control Normal, se pudo determinar que los casos de EP-ICD mostraron adelgazamiento cortical frontal superior y precentral izquierdo como se muestra en la Figura 1.1, además de daño del tracto de sustancia blanca motora y extramotora en comparación con pacientes EP sin ICD [14].

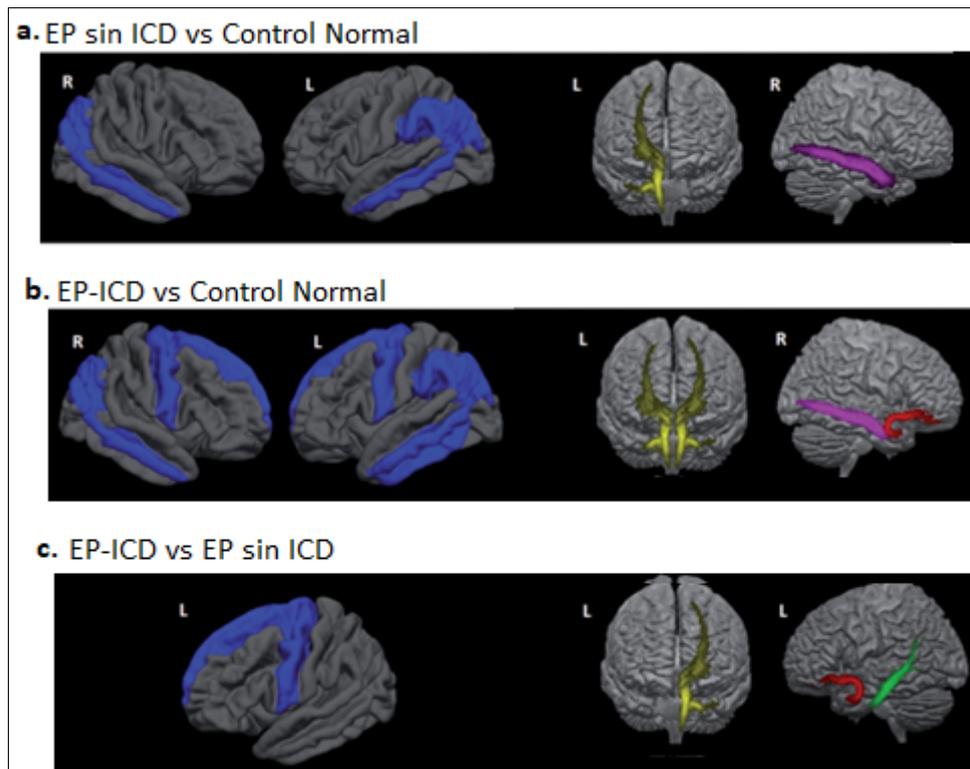


Figura 1.1 IRM de paciente con EP + ICD

[Imagen modificada de *Molecular Psychiatry* (2018) 23, 459-466; doi:10.1038/mp.2017.18]

1.4.3 Ludopatía

El trastorno del juego o ludopatía es un comportamiento de juego desadaptativo persistente y recurrente [12]. El impacto social del juego patológico está aumentando en la actualidad con problemas relacionados que involucran empleo, relaciones personales, solvencia financiera y actividades criminales [15]. En los pacientes con EP se ha observado que la conducta de juego aparece con más frecuencia en los períodos “on” de fluctuaciones motoras y que comienza después del inicio de la enfermedad en la mayoría de los pacientes empeorando con la terapia con levodopa, lo que sugiere que podría representar una manifestación conductual del tratamiento farmacológico [16]. En un estudio donde se realiza la comparación de pacientes sólo con EP y pacientes con EP y Ludopatía, se observa en la Figura 1.2, varias áreas de mayor respuesta que dependen del nivel de oxígeno en sangre (BOLD). Estas áreas se relacionan con señales que se activan en la

corteza cingulada anterior bilateral, las circunvoluciones frontal media y superior, el parietal inferior derecho, lobulillo y cuerpo estriado ventral [15].

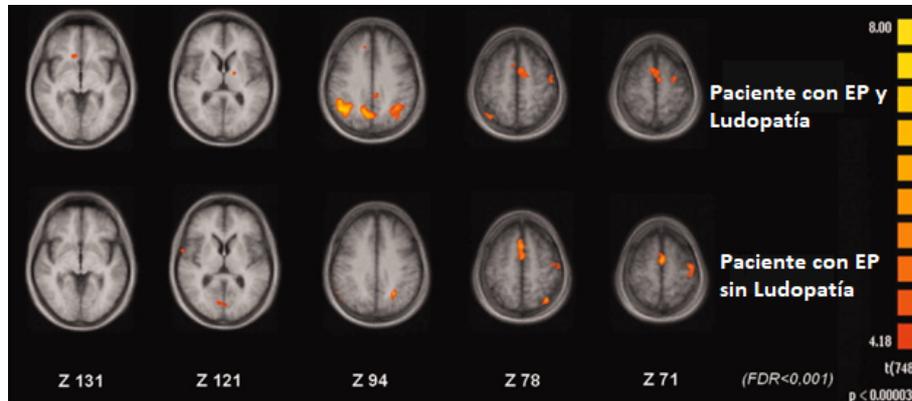


Figura 1.2 IRM de pacientes con EP y Ludopatía

[Imagen modificada de *Movements Disorders* (2010) 25, 2449-2453; doi: 10.1002/mds.23369]

1.4.4 Hipersexualidad (HS)

La hipersexualidad (HS) se considera como el aumento en el interés sexual, la excitación y el comportamiento, lo que tiene consecuencias adversas para el paciente y su pareja o cuidadores [17]. Fue uno de los primeros ejemplos de un ICD cuyo comportamiento se encuentra asociado con el tratamiento de la enfermedad de Parkinson, con una prevalencia estimada de aproximadamente 3.5% [18]. Los datos de imágenes de resonancia magnética funcionales de un estudio que muestra la relación de la HS en pacientes con EP mostraron que el aumento del deseo sexual de los pacientes se correlacionó con activaciones cerebrales mejoradas en el cuerpo estriado ventral y las cortezas cinguladas y orbitofrontales [19]. En la Figura 1.3, se observan IRM de pacientes con EP-HS que ilustran los aumentos de señales en: (A) cuerpo estriado ventral, (B) hipotálamo (C) corteza prefrontal anterior (aPFC), corteza cingulada anterior (ACC), lóbulo parietal superior (SPL) y corteza cingulada posterior (PCC), (D) circunvolución temporal media izquierda y derecha y circunvolución occipital media y corteza prefrontal anterior, y (E y F) corteza orbitofrontal (OFC), lóbulo parietal superior, corteza cingulada posterior y circunvolución temporal media izquierda y derecha y circunvolución occipital media.

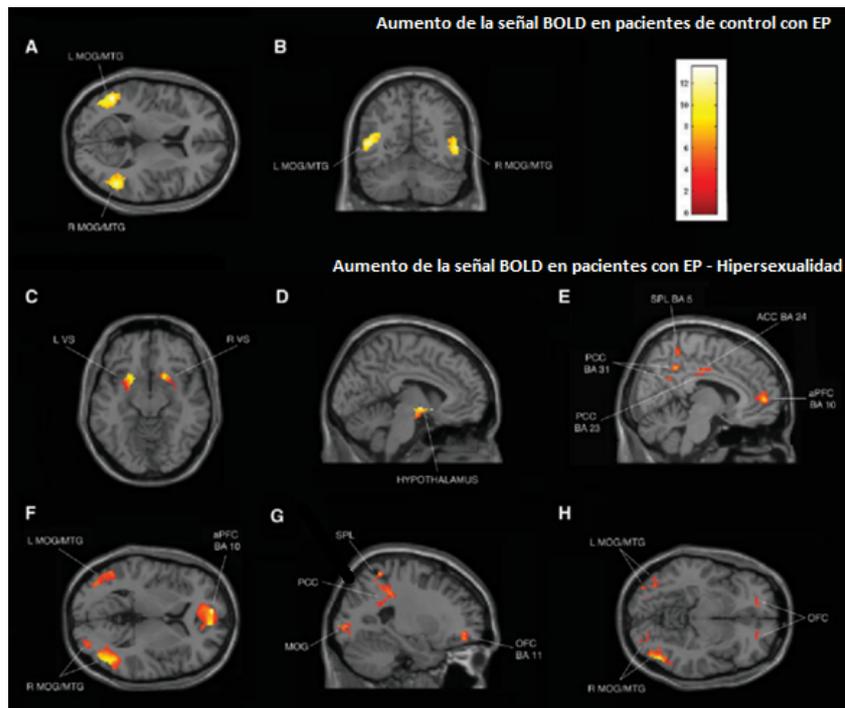


Figura 1.3. IRM de pacientes con EP+ HS

[Imagen modificada de Brain (2013) 136; doi: 10.1093/brain/aws326]

1.4.5 Trastorno por Atracón (TPA)

El trastorno por atracón es un trastorno alimentario grave, potencialmente mortal y tratable que se caracteriza por episodios recurrentes de ingesta de grandes cantidades de alimentos, una sensación de pérdida de control durante el atracón; experimentar vergüenza, angustia o culpa después; y no utilizar regularmente medidas compensatorias poco saludables para contrarrestar los atracones [20]. De 96 pacientes con enfermedad de Parkinson, en el Centro de Trastornos del Movimiento de la Universidad de Florida, el 1% cumplió con los criterios de diagnóstico para el trastorno por atracón y el 8.3% exhibieron TPA subumbral [20]. Según un estudio de identificación de biomarcadores realizado a pacientes que presentan TPA vs pacientes de controles sanos, se implementó el algoritmo SVM para clasificar y mostrar a través de un mapa de distribución de patrones basados en vóxeles, como se observa en la Figura 1.4, las

diferencias estructurales del cerebro según imágenes de resonancia magnética con vista sagital. Esta IRM tiene distribución de patrones, normalizada en un rango entre 0 y 1 según la escala de colores y se superpone a un cerebro estereotáctico estándar para la localización espacial [21].

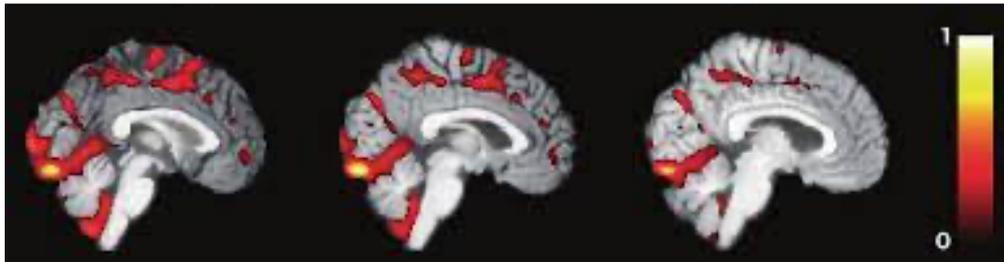


Figura 1.4. IRM de paciente con TPA

[Imagen modificada de Behavioural Neurology (2015), 1-10]

1.4.6 Apatía

La apatía se define como una disminución en la motivación, manifestada por una reducción en la actividad verbal, motora y emocional [7]. Aunque a menudo se presenta junto con la depresión, la apatía también puede ser un síntoma de EP en pacientes que no experimentan depresión [7]. Una característica discriminadora potencialmente útil es el estado de ánimo, que es neutral en apatía y negativo en la depresión [22]. Según un estudio de imágenes de resonancia magnética se muestra como la apatía activa varias señales cerebrales cuando el paciente padece esta enfermedad.

La Figura 1.5 muestra las máscaras de Regiones de Interés (ROI) para identificar la Apatía de la siguiente manera: (A) El polo frontal se muestra en amarillo, la circunvolución frontal superior en rojo, la circunvolución frontal media en verde, la circunvolución frontal inferior pars opercularis en azul y la circunvolución frontal inferior pars triangularis en naranja. (B) Corteza orbital frontal en naranja, división anterior del giro cingulado en luz azul, corteza del lóbulo yuxtaposicional/área motora suplementaria en verde, corteza frontal medial en rojo, giro precentral en amarillo. (C) Amígdala en amarillo

anaranjado, caudado en blanco y rosa, putamen en verde, núcleo accumbens en azul claro, pallidum en rojo. (D) Radiación talámica anterior en púrpura, cápsula interna en azul, fascículo uncinado en verde, fascículo longitudinal superior en amarillo, cuerpo del cuerpo calloso en rojo, cíngulo en luz azul (lado izquierdo).

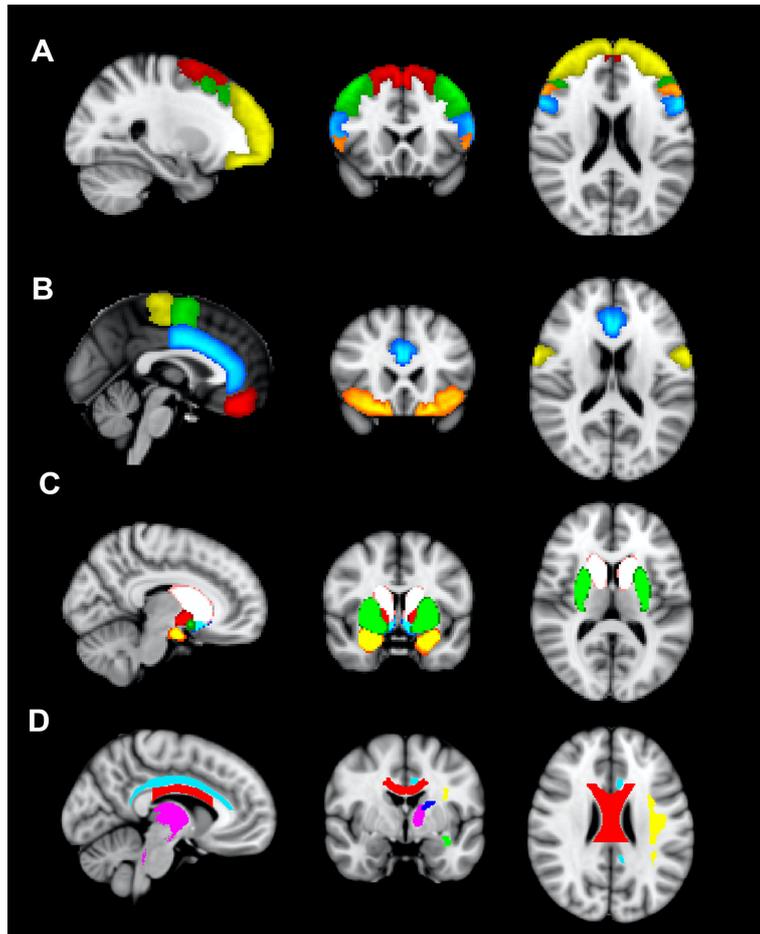


Figura 1.5. Máscaras ROI para análisis de resonancia magnética. [22]

1.4.7 Aprendizaje Profundo

Existen tareas que, debido a su complejidad y la de sus datos, son muy difíciles de resolver mediante técnicas tradicionales de aprendizaje de máquina (Machine Learning), tras esto surge el aprendizaje profundo (Deep Learning) y las redes neuronales profundas [23]. La unidad de estas redes es la neurona artificial, donde cada unidad recibe información de múltiples unidades de la capa previa o la entrada del modelo y calcula un valor propio de activación. El aprendizaje profundo no es un área investigada recientemente, desde 1960 aproximadamente se conocen los primeros acercamientos a la estructura de la neurona artificial pero no ha sido hasta las últimas décadas que debido a la cantidad masiva de datos disponibles y el poder computacional han hecho posible construir y aplicar estas arquitecturas tan complejas [23]. Aplicaciones prácticas de aprendizaje profundo incluyen visión por computador (CV), Procesamiento del lenguaje Natural (NLP), aprendizaje por refuerzo profundo (DRL), entre otros.

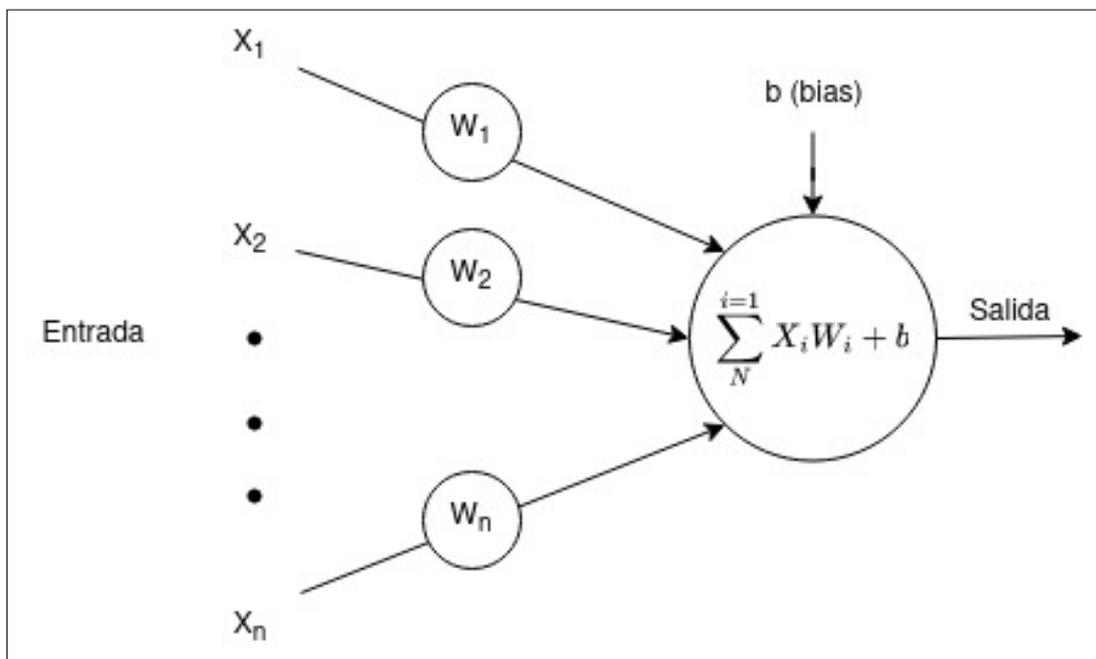


Figura 1.6. Estructura de una Neurona artificial

[Autoría propia]

1.4.8 Convolutional Neural Network

Red neuronal Convolutacional (CNN) es una red neuronal que extrae o identifica una característica en una imagen en particular. Esto forma una de las operaciones más fundamentales en el aprendizaje automático y se usa ampliamente como modelo de base en la mayoría de las redes neuronales como GoogleNet, U-Net, VGG19 y otras, para diversas tareas, como detección de objetos, segmentación, clasificación de imágenes [24]

Como se muestra en la Figura 1.7, los componentes básicos de una CNN son:

- **Convolution:** para detectar características en una imagen
- **ReLU:** para suavizar la imagen y distinguir los límites
- **Pooling:** para ayudar a reparar imágenes distorsionadas
- **Flattening:** o vectorización, convierte la imagen en una representación adecuada
- **Conexión completa:** para procesar los datos en una red neuronal

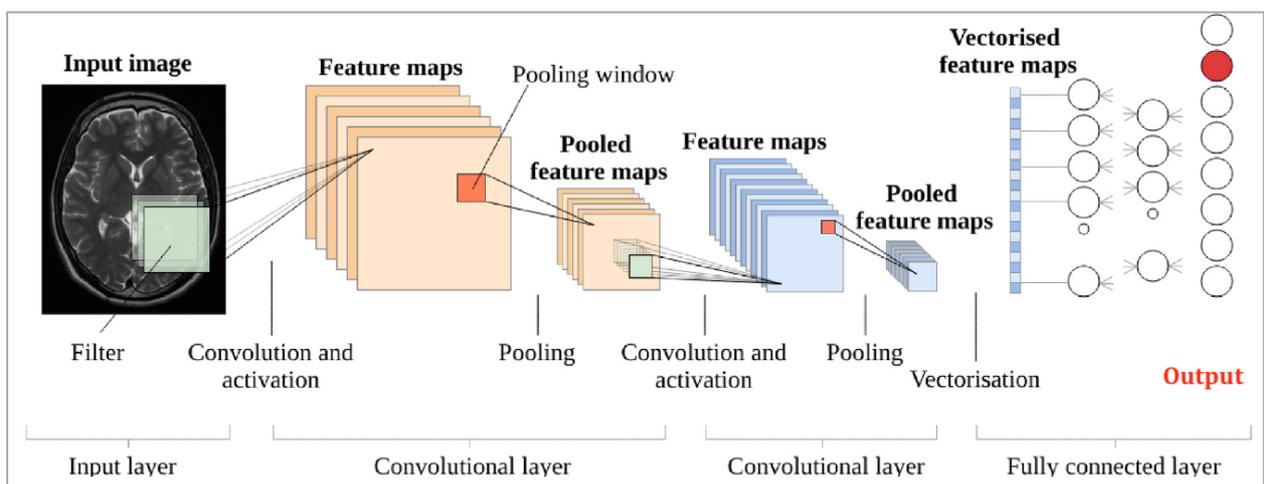


Figura 1.7 Bloques de construcción de una CNN [25]

1.4.9 Transformer

El modelo Transformer es una arquitectura de procesamiento de lenguaje. Este modelo se basa en la atención para procesar secuencias de entrada y se utiliza para una variedad de tareas NLP, como traducción automática, resumen de texto y clasificación de texto [26]. Lo que lo diferencia de otros modelos de procesamiento de lenguaje es su capacidad para procesar secuencias de entrada de manera paralela y secuencial, lo que resulta en una mayor velocidad de entrenamiento y una mejor capacidad de generalización [27]. Además, el uso de atención permite al modelo Transformer considerar dependencias a larga distancia en la entrada, lo que es esencial para una variedad de tareas NLP. [26]

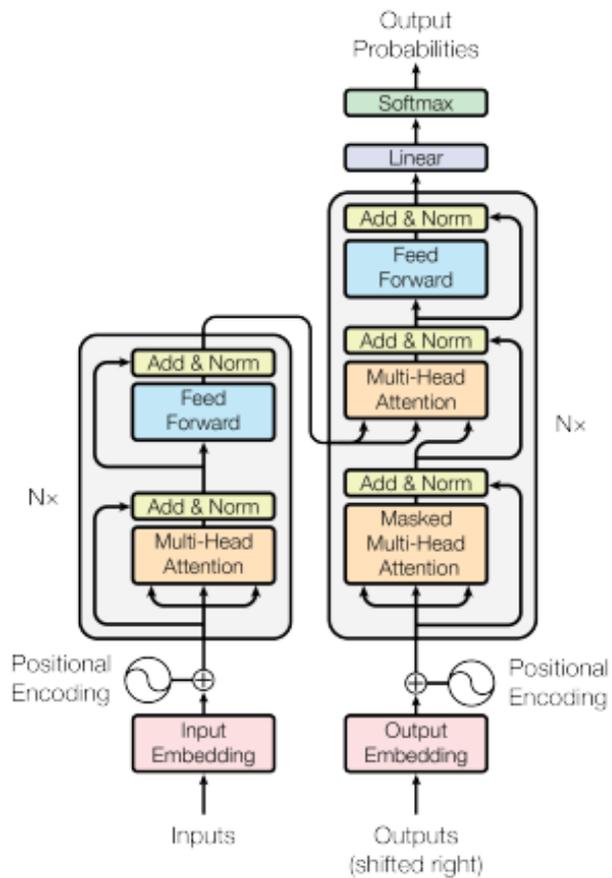


Figura 1.8 Arquitectura del Transformer [26]

1.4.10 Vision Transformers (ViT)

Los ViT son variaciones de la arquitectura Transformers, modelos potentes y precisos aplicados principalmente a NLP [28] que han surgido como alternativa a las CNNs, las cuales han sido por varios años el estado del arte para las tareas de reconocimiento de imágenes. Originalmente, el enfoque radica en interpretar a las imágenes como secuencias o patches de píxeles asemejando a secuencias de texto como el uso original de Transformers para NLP [29]. Por otra parte, existe el enfoque convolucional, el cual procesa las imágenes de entrada a través de n capas convolucionales, produciendo mapas de características (feature maps), los cuales, a su vez se proyectan como la entrada del Transformer [29].

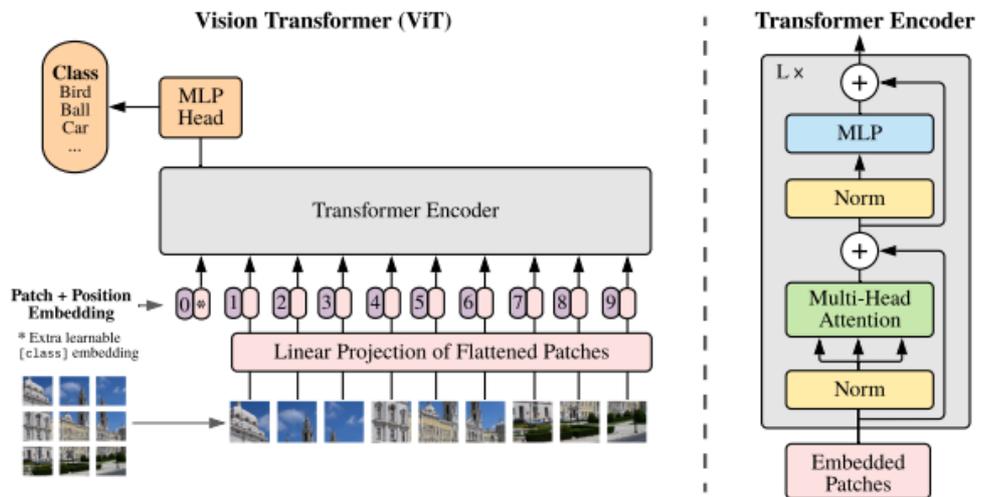


Figura 1.9 Arquitectura del Vision Transformer [29]

1.4.11 Swin Transformer

Liu et al. Propuso Swin (**S**hifted **W**INdow) Transformer [30], un Transformer de propósito general adaptado a tareas de visión por computadora (propósito general significa que es adecuado para diferentes tareas como: segmentación semántica y clasificación de imágenes) y logró mejores rendimientos que el ViT, y arquitecturas basadas en CNN

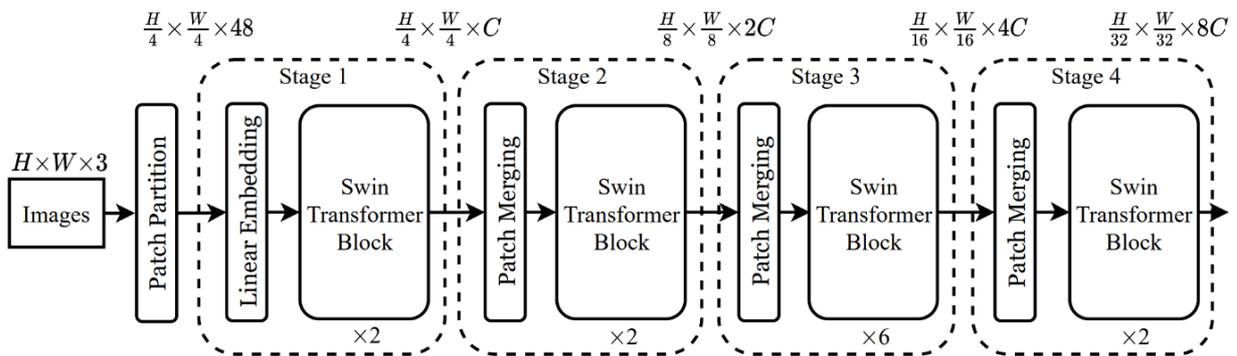


Figura 1.10 Arquitectura Swin Transformer [30]

Al principio, para una imagen de entrada (o un mapa de características), la red crea una ventana local. De forma predeterminada, la dimensión de la ventana garantizaría que sea $\frac{1}{4}$ de la de la imagen (como se muestra en los cuadros rojos en la Figura 1.11).

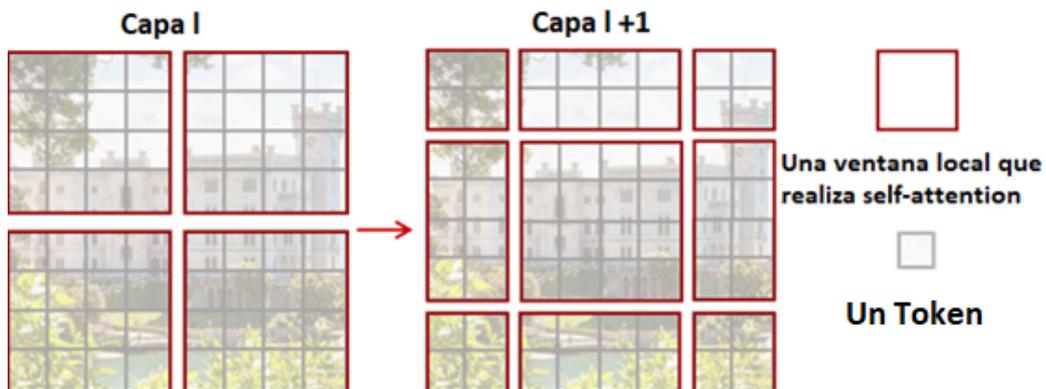


Figura 1.11 Diseño de la ventana desplazada. [30]

Luego, se crean más particiones en una ventana desplazada (shifted window) para crear tokens pequeños. Por lo general, hay $4 \times 4 \times 3 = 48$ tokens, 3 es el número de canales en una imagen RGB. Luego, una capa de embedding lineal convierte cada token en un vector de embeddings, para que el Transformador pueda entenderlo.

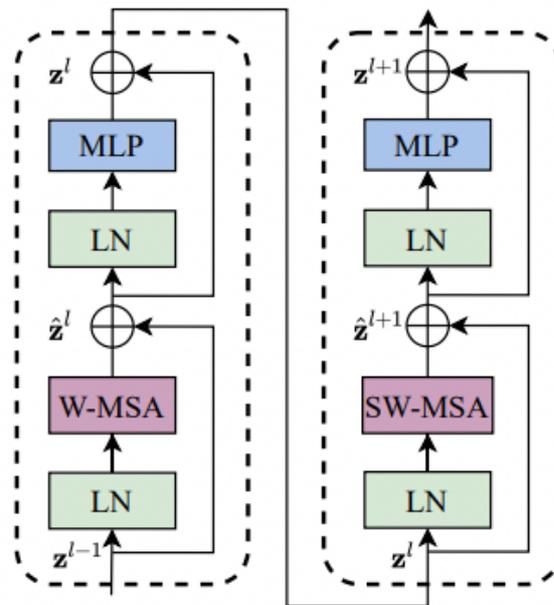


Figura 1.12 Dos bloques de Swin Transformers consecutivos. [30]

A continuación, los embeddings se introducen en dos bloques de Swin Transformer consecutivos, como se muestra en la Figura 1.12. Los módulos abreviados en la figura son [30]:

- **LN:** Normalización de capas.
- **W-MSA:** Autoatención basada en ventanas. Calcula la atención solo dentro de cada ventana lo cual reduce en gran medida el costo computacional, además facilita la comunicación entre cada W-MSA aislado.
- **SW-MSA:** Autoatención de ventana desplazada. Desplaza las ventanas hacia la esquina inferior derecha por un factor de $M/2$, donde M es el tamaño de la ventana. De esta manera introduce conexiones cruzadas importantes entre ventanas y mejora el rendimiento de la red.
- **MLP:** Perceptrón Multicapa.

Las arquitecturas estándar de Transformer buscan relaciones entre un token y todos los demás tokens en la imagen completa, lo que lleva a cálculos extremadamente costosos, especialmente para imágenes grandes [31].

La diferencia con Swin Transformer es que esta arquitectura adopta un diseño de ventana desplazada que solo busca relaciones entre un token y otro en el área de la ventana, utilizando dos configuraciones de ventanas diferentes para establecer conexiones entre ventanas y ventanas desplazadas para encapsular toda la imagen. Por lo tanto, la eficiencia de cálculo mejora drásticamente [31].

1.4.12 MeshNet

MeshNet es una arquitectura que aprende la representación de formas 3D directamente a partir de una colección de vértices, bordes y caras, llamada data mesh, que utiliza las caras como unidad principal [32]. MeshNet ha sido utilizado en aplicaciones de clasificación y recuperación de formas 3D cuyos resultados experimentales y las comparaciones con los métodos más avanzados demuestran que MeshNet puede lograr un rendimiento satisfactorio [33]

Como se observa en la Figura 1.13, la entrada es una lista de caras con valores iniciales, que entran en dos bloques llamados descriptores espaciales y estructurales que generan características espaciales y estructurales iniciales. Luego, las características se agregan con la información vecina a los bloques de convolución etiquetados como “*Mesh Conv*” que alimentan a una función de pooling para generar una característica global. Al final se encuentra una “*MLP*” con los números entre paréntesis que indican la dimensión de las capas ocultas y la capa de salida [33]

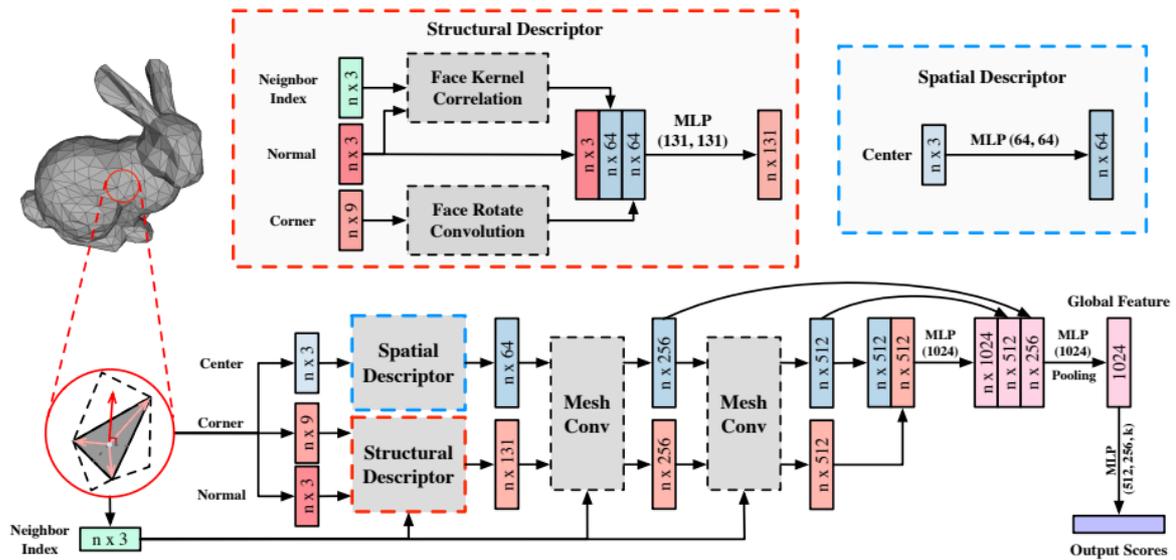


Figura 1.13 Arquitectura MeshNet [33]

La información que contiene una unidad de cara está dividida en dos partes. La primera parte es la información de la cara que consta de: el centro, que contiene la coordenada del punto central; la esquina, es el vector formado por el punto central y tres vértices y la normal, que es el vector normal de la cara (Figura 1.14). La segunda parte contiene la información del vecino, representado a través de un índice vecino [33].

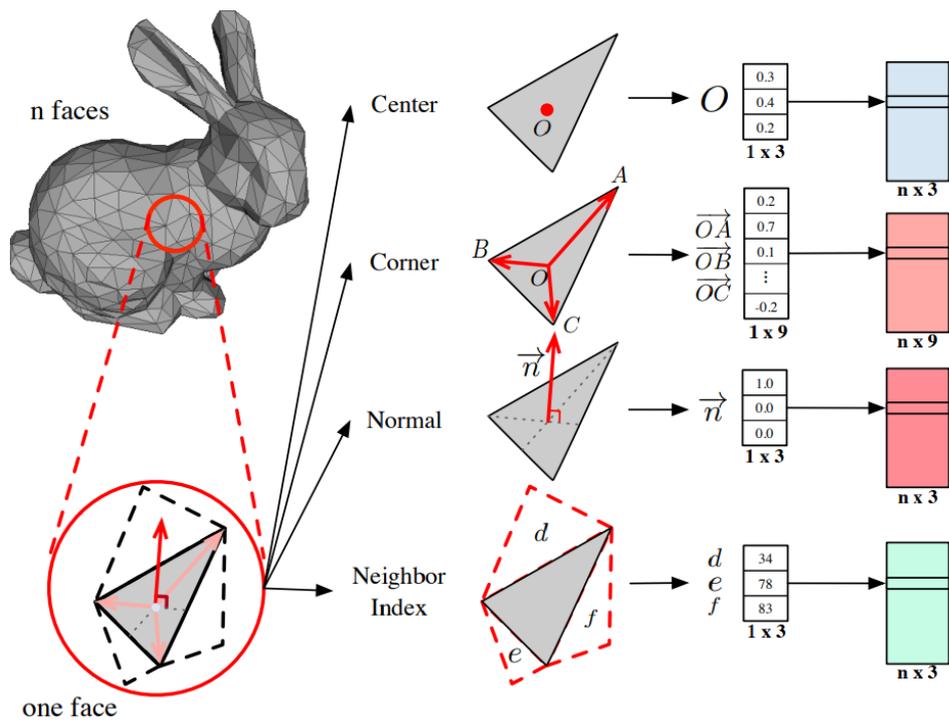


Figura 1.14 Partes de la cara de MeshNet. [33]

La arquitectura de la red consta de dos bloques denominados descriptores espaciales y estructurales que permiten aprender las características iniciales, un bloque de convolución *Mesh* que agrega características vecinas. Luego se emplea una función de *pooling* para generar una característica global. Finalmente, una MLP se aplica para fusionar aún más las características vecinas y generar una nueva característica estructural [33].

El descriptor espacial permite conocer la posición espacial de las caras aplicando una MLP compartida al centro de cada cara. Mientras que el descriptor estructural, contiene la información de la forma y estructuras locales [33].

1.4.13 Compact Convolutional Transformer

Compact Convolutional Transformer (CCT) es una arquitectura basada en Vision Transformer (ViT) [34]. Fue desarrollado con el objetivo de ser un modelo más eficiente y que utilice muchos menos datos de entrenamiento sin sacrificar precisión [35]. El CCT logra este objetivo por medio de dos cambios al ViT original, utiliza un *generador de tokens convolucional* en lugar de un *generador de tokens secuencial* [35] de esta forma aprovecha las ventajas de las CNNs al insertar un “*inductive bias*” y permitirle al CCT detectar mejores características locales en las imágenes [36]. Por otro lado, introduce una capa a la que denominan “*Sequence pooling (SeqPool)*”, la cual elimina a los [cls] tokens del ViT que la red debe aprender para clasificar, agrupando toda la secuencia de tokens que son output del “*Transformer encoder*” en una sola capa mejorando así el rendimiento y eficiencia del modelo [35]

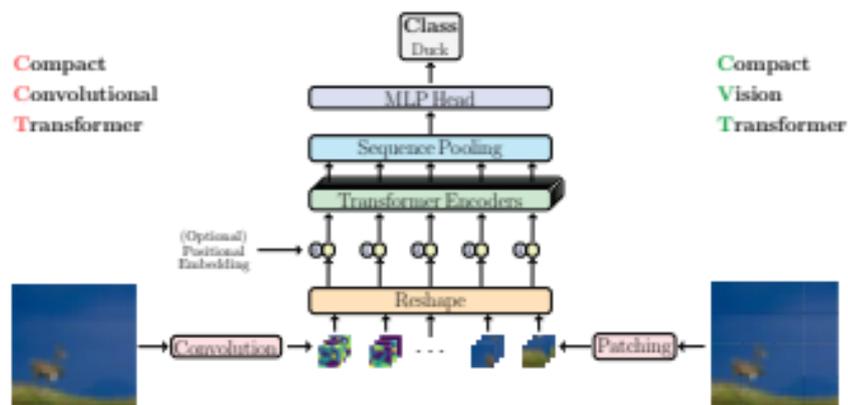


Figura 1.15 Arquitectura original del CCT [35]

1.4.14 Aprendizaje profundo aplicado a IRM

Los rápidos avances tecnológicos y computacionales han permitido el aumento del uso de aprendizaje profundo en el área médica, particularmente en el análisis de imágenes médicas con tareas como segmentación, clasificación, detección y localización [37].

Con el auge de AlexNet en 2012 y GoogLeNet en 2014 los cuales fueron posibles debido a los avances en computación paralela [37], se vio un auge significativo a la investigación en aprendizaje profundo aplicado a imágenes médicas, entre ellas aplicaciones en la detección de cáncer en imágenes de rayos X, detección e identificación de lesiones cerebrales en IRM y tomografías (CT), entre otros [38].

Una arquitectura particularmente exitosa en las tareas de segmentación utilizando datos médicos como IRM es la U-Net [39] y a partir de esta otras han aparecido adaptando arquitecturas más cercanas al estado del arte como Swin-UNETR, como se aprecia en la Figura 1.16 [40].

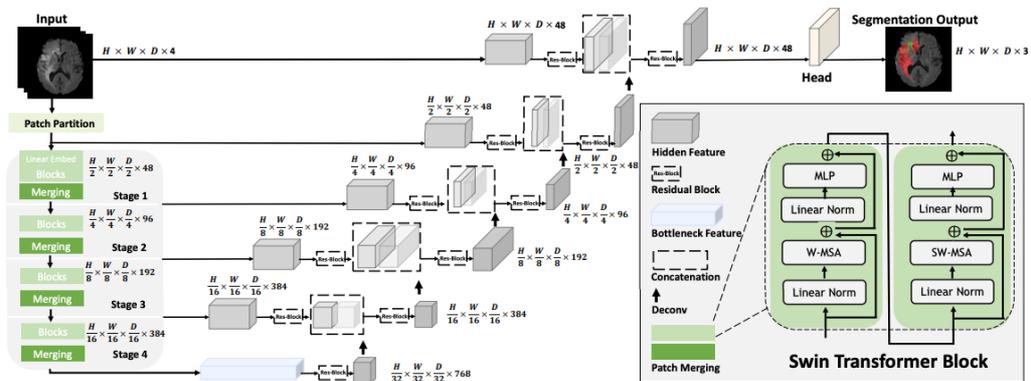


Figura 1.16 Arquitectura Swin-UNETR [40]

CAPÍTULO 2

2. METODOLOGÍA

En esta sección se analizan los requerimientos necesarios para construir un prototipo web como solución al problema planteado en el primer capítulo. Mediante el uso de un módulo de pre-procesamiento y técnicas de data augmentation se preparan los datos para el modelo. Para la clasificación de los desórdenes Parkinsonianos, se emplearon diferentes modelos de aprendizaje profundo y se compararon sus resultados. Finalmente, presentamos el diseño de la solución final y el software utilizado para implementarla.

El prototipo consiste en una aplicación web que permite visualizar el patrón anatómico del cerebro que ha sido afectado en pacientes con EP que presenten algún tipo de ICD o apatía. El prototipo consta de un módulo de preprocesamiento de IRM para su estandarización el cual alimenta un modelo de clasificación de aprendizaje profundo. El modelo Compact Convolutional Transformer (CCT) fue elegido de tres arquitecturas del estado del arte de visión por computadora: Swin-T, CCT y MeshNet.

2.1 Análisis

Para conocer el alcance del proyecto, así como los resultados que esperaba obtener y sus limitantes, se realizó reuniones semanales con el cliente.

Después de definir los requerimientos mínimos del proyecto se procedió a revisar la documentación de las versiones previas, junto con el estado del arte de imágenes de resonancia magnética en 3D, la enfermedad de Parkinson, Trastornos de Control de Impulsos y arquitecturas tipo Transformers.

2.1.1 Requerimientos

Durante las primeras reuniones se especificaron las características y requerimientos que se esperan de la aplicación, tomando en cuenta las limitaciones y necesidades de los usuarios. A continuación, se listan los requerimientos funcionales y no funcionales del sistema:

2.1.1.1 Requerimientos Funcionales

REQ.1. El sistema permite la importación de archivos de formato NifTI.

REQ.2. El sistema realiza el debido pipeline de preprocesamiento de la imagen la cual servirá como entrada del modelo de clasificación (Transformer).

REQ.3. El sistema muestra el resultado de la clasificación de modelo Transformer.

REQ.4. El sistema muestra de forma interactiva las imágenes de IRM en 3D.

REQ.5. El sistema permite la observación del área afectada de las imágenes cerebrales.

REQ.6. El sistema permite el ingreso a dos roles: administrador y médico.

REQ.7. El sistema permite que el administrador observe las imágenes importadas por los médicos.

REQ.8. El sistema permite que el administrador cree otros usuarios médicos.

REQ.9. El sistema permite que médicos realicen una retroalimentación de los resultados mostrados por el sistema.

2.1.1.2 Requerimientos No-Funcionales

REQ.10. Se requiere emplear el lenguaje de programación Python en conjunto de librerías para el manejo de datos masivos.

REQ. 11. El sistema debe operar en el servidor web proxy HTTP nginx 1.22.

REQ. 12. El sistema requiere el uso del framework web Django.

REQ. 13. El sistema requiere el un sistema de bases de datos MySQL para albergar la data de administradores y médicos, así como las retroalimentaciones de las imágenes.

REQ. 14. Se requiere emplear un algoritmo de interpretabilidad como Guided Backpropagation para la visualización del área donde el modelo identifica la ubicación del patron de interés.

2.1.2 Alcance y limitaciones de la solución

Con respecto a la base de datos, la mayor dificultad fue encontrar IRM para cada una de las clases de ICD y la clase de apatía. Por ello, se requirió efectuar data augmentation de los datos.

Para mejorar la capacidad del modelo, se tuvo que limitar el tamaño del batch en el entrenamiento de los modelos de aprendizaje profundo. Esto permite cargar menos instancias de las imágenes en memoria por época, por lo tanto, se pudo aprovechar de mejor forma los recursos computacionales.

2.1.3 Riesgos y beneficios de la solución

Debido a que la solución se basa en un prototipo, no se requiere seguridad exhaustiva de los datos.

Se busca mejorar la calidad de vida de personas que tengan desórdenes Parkinsonianos a través de una herramienta que ayude a identificar patrones anatómicos de ICD y apatía para que los pacientes puedan obtener un tratamiento en etapa temprana.

2.1.4 Usuarios de la solución

2.1.4.1 Médico

El usuario de tipo médico es el encargado de ingresar las IRM para la predicción correspondiente.

2.1.4.2 Administrador

El usuario administrador se encarga de la creación y eliminación de usuarios. Además, puede visualizar las imágenes ingresadas por los médicos con su respectiva evaluación.

2.2 Prototipado

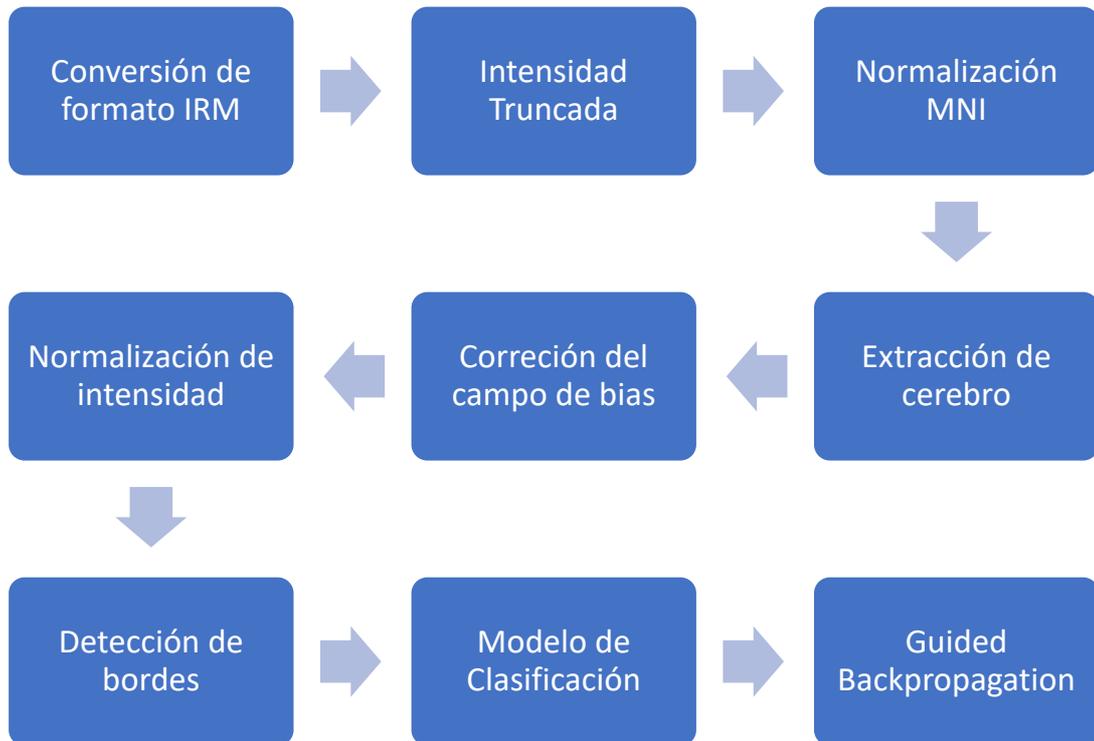


Figura 2.1 Diseño de la solución de alto nivel.

[Autoría propia]

La Figura 2.1 describe un proceso que comprende los siguientes pasos: (1) Conversión de formato IRM, (2) Intensidad Truncada, (3) Normalización MNI, (4) Extracción de cerebro, (5) Corrección del campo de bias, (6) Normalización de intensidad, (7) Detección de bordes, (8) Modelo de Clasificación y (9) Guided Backpropagation. Este proceso es utilizado para el desarrollo de la solución que es una combinación de técnicas de procesamiento de imágenes de resonancia magnética y aprendizaje profundo.

2.2.1 Dataset

Como base de datos de entrenamiento se utilizó IRM de diferentes pacientes etiquetadas como una de las 4 clases (Sólo Párkinson, Apatía, Apatía-ICD, ICD), estas imágenes fueron descargadas directamente desde el sitio oficial de Parkinson's Progression Markers Initiative (PPMI).

2.2.2 Módulo de Pre-procesamiento

Las imágenes requirieron un procesamiento previo al entrenamiento de los modelos. Se lo realizó con el fin de garantizar la estandarización de imágenes. Usualmente, en las IRM se aprecian detalles como los ojos y tronco encefálico, por lo que se descartó dichas estructuras, manteniendo solamente el área del cerebro. También, hay que tomar en consideración que cada imagen puede variar en tamaño, por lo que se debe realizar su correcta transformación de dimensiones sin afectar las características físicas del cerebro, ya que se utilizarán para la detección de trastornos.

Para el pre-procesamiento de los datos el prototipo realiza las siguientes operaciones:

2.2.2.1 Conversión de formato de IRM

El formato comúnmente usado en imágenes de resonancia magnética es DICOM, se caracteriza porque contiene un conjunto de archivos en formato DICOM, siendo cada archivo un corte diferente. El conjunto de cortes forma la imagen en 3D. No obstante, posee mucha más información y entorpece su lectura al ser un conjunto de datos. Por ello, se seleccionó el formato NifTI por su forma compacta de almacenar la IRM, así como por su tamaño y cantidad de cortes como se muestra en la Figura 2.2.

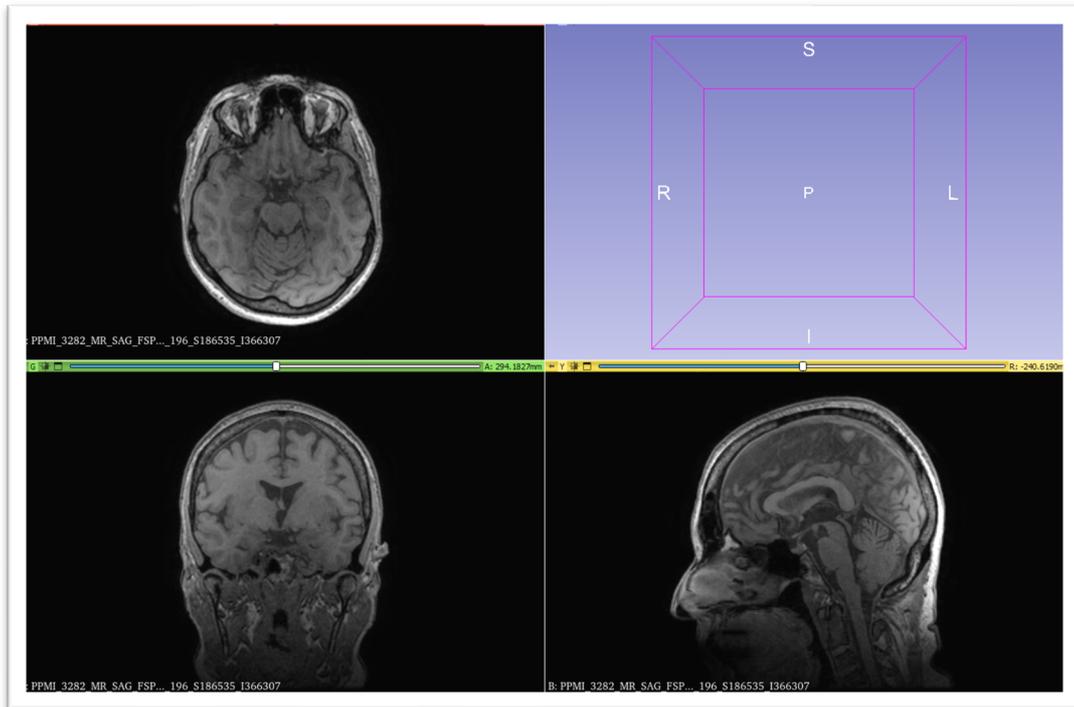


Figura 2.2 IMR no procesada de paciente con EP.

[Autoría propia]

2.2.2.2 Intensidad truncada

Esta es una técnica que limita el rango numérico de los vóxeles determinados por cuantiles, con el propósito de regular la iluminación y contraste de la imagen. [41]

2.2.2.3 Normalización y Co-registro de la imagen médica (MNI)

La normalización y co-registro ayuda a reajustar el cerebro del paciente para que coincida con la forma y el tamaño de una imagen de plantilla. Este es un paso crucial requerido para los análisis a nivel grupal. Se introdujeron plantillas especializadas que permiten aplicar algoritmos de normalización para todas las imágenes de los distintos grupos de ICD y apatía, como se observa en la Figura 2.3 [41].

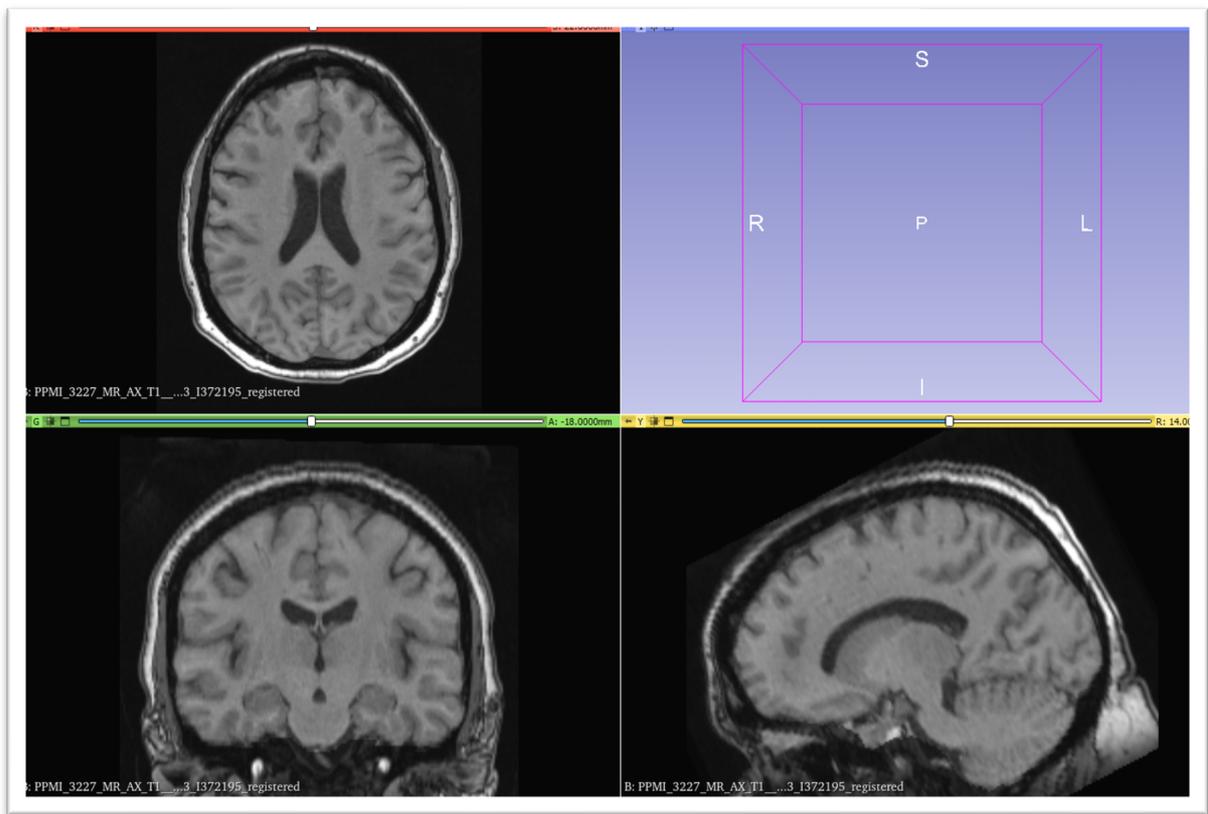


Figura 2.3 IRM normalizada y co-registrada.

[Autoría propia]

2.2.2.4 Extracción de cerebro

Las imágenes de resonancia magnética suelen tener estructuras adicionales al cerebro, como es la presencia de cráneo, ojos y cuello. Para tener una imagen que tenga netamente el cerebro se requiere el uso de algoritmos de extracción, en este caso, se empleó la función *brain extraction* de la librería Ants para la extracción de cerebro. [41]



Figura 2.4 Extracción de cerebro en IRM

[Autoría propia]

2.2.2.5 Corrección del campo de bias

La corrección de campo de bias es un método que permite corregir el gradiente de intensidad que adquieren las IMR, lo cual hace que algunas partes de la imagen sean más brillantes que otras, ya que el gradiente puede influir erróneamente en algoritmos de segmentación [42].

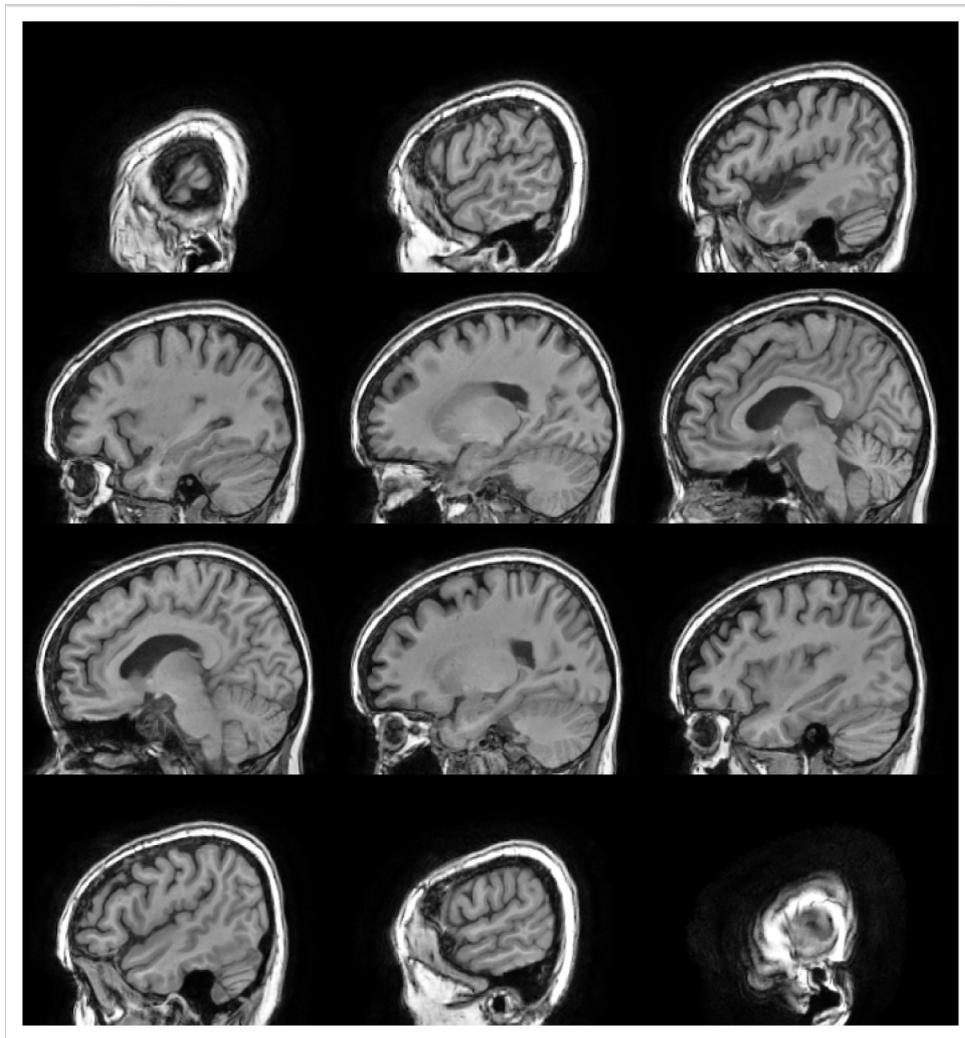


Figura 2.5 Corrección de campo de bias en IRM

[Autoría propia]

2.2.2.6 Normalización de intensidad

La normalización de la intensidad en IMR tiene como objetivo corregir las variaciones dependientes del escáner, lo cual es esencial para un análisis de resonancia magnética preciso. Es una técnica ampliamente aplicada para reducir la varianza de los datos para la cual existen varios métodos disponibles que van desde la transformación de uniformidad hasta la ecuación de histogramas [43].

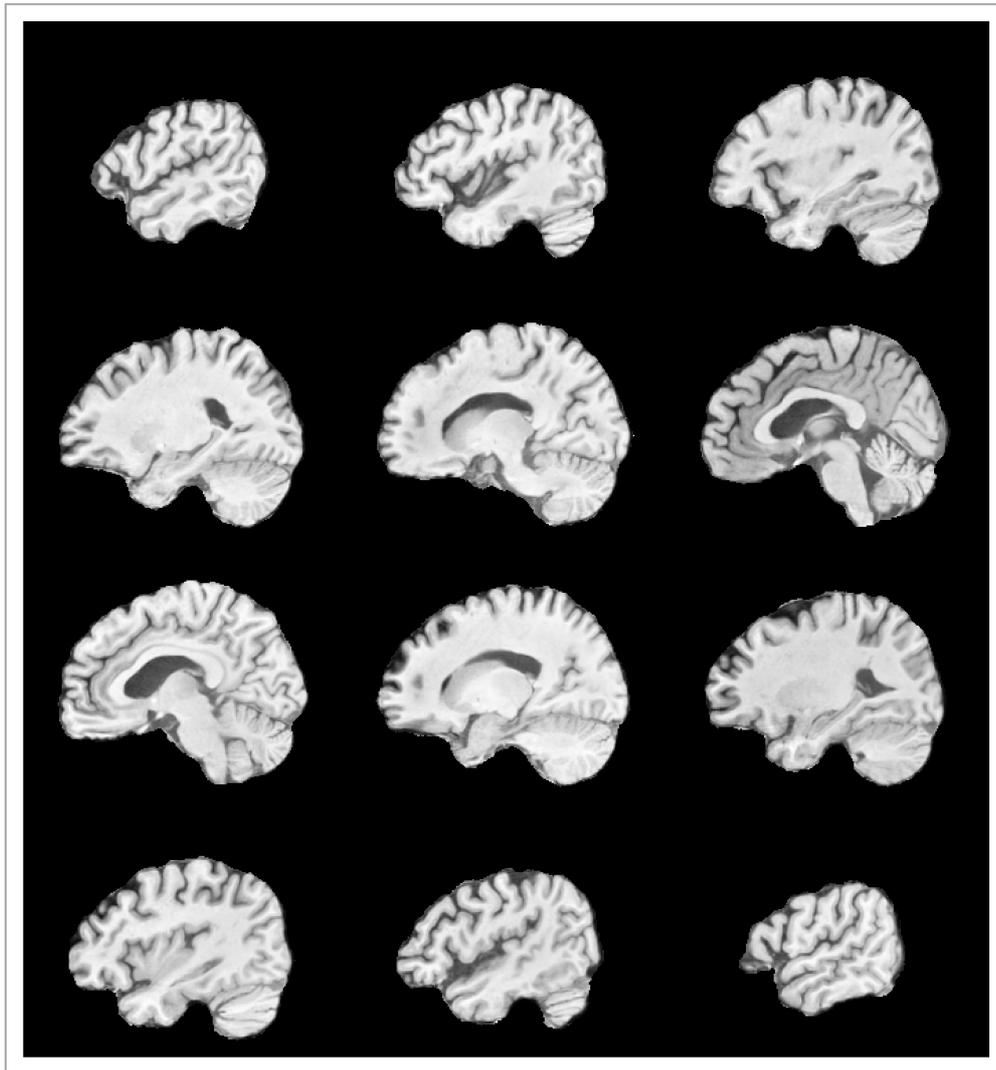


Figura 2.6 Normalización de intensidad en IRM

[Autoría propia]

2.2.2.7 Detección de bordes

Se utilizó la detección de bordes que es una técnica de procesamiento de imágenes utilizada para identificar puntos en una imagen digital con discontinuidades; es decir, cambios bruscos en el brillo de la imagen. Estos puntos en los que el brillo de la imagen varía considerablemente se denominan bordes (o límites) de la imagen [44].

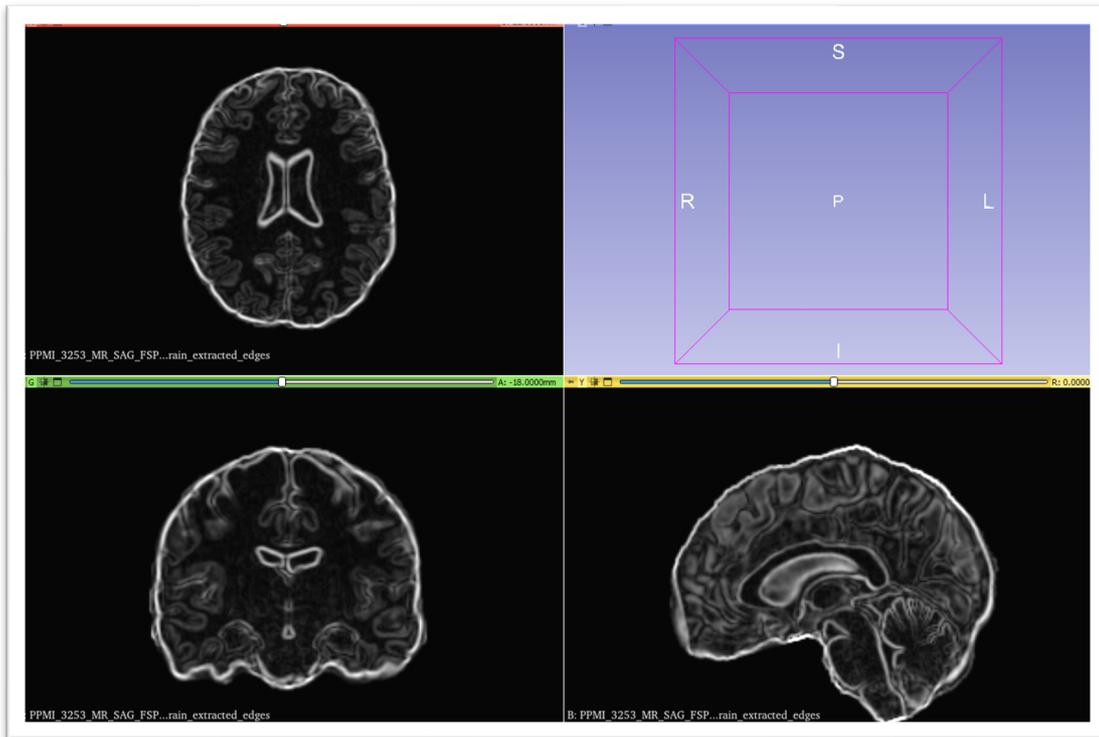


Figura 2.7 IRM con detección de bordes

[Autoría propia]

2.2.3 Data Augmentation

Se generan diferentes versiones de la data para aumentar su tamaño. Los modelos de visión por computadora y procesamiento de lenguaje natural (NLP) utilizan la estrategia de aumento de datos para manejar la escasez de datos y la diversidad de datos insuficiente. En la Figura 2.8 se muestra la implementación de las técnicas: denoising y rotación, en 6 grados distintos (-15,-10,-5,5,10,15) [45].

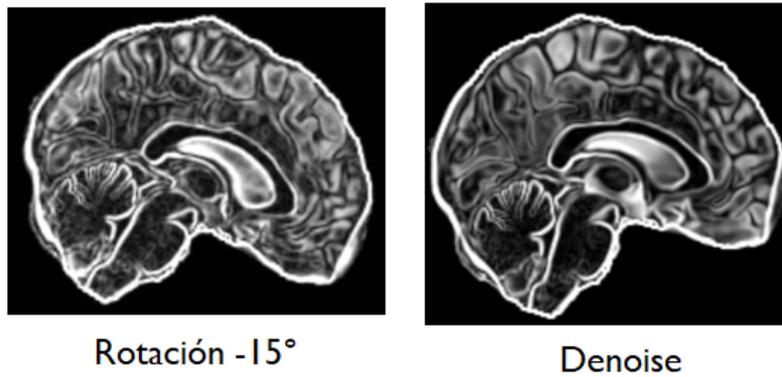


Figura 2.8 IRM con técnicas de data augmentation

[Autoría propia]

2.2.4 Modelos de Clasificación

Para los modelos de clasificación se partió de los hiper-parámetros propuestos originalmente y se procedió a realizar grid search fine tuning para encontrar la mejor configuración para nuestra tarea de clasificación de IRM. Debido a que se analizan IMR estos modelos fueron modificados a 3D. Los valores de partida se observan en la Tabla 2.1.

Tabla 2.1 Híper-parámetros iniciales para los modelos de clasificación.

Modelo	Dimensión de entrada	Épocas	Batch size	Optimizer	Scheduler	Learning rate	Weigth decay
Swin Transformer	192x192x189	300	1024	AdamW	cosine	0.001	0.05
CCT	224x224	300	1024	AdamW	cosine	-	-
MeshNet	192x192x189	100	64	SGD	-	0.01	0.0005

2.2.5 Algoritmo Guided Backpropagation

El algoritmo Guided Backpropagation se basa en el algoritmo de Backpropagation, que es el proceso utilizado para entrenar las redes neuronales. La principal diferencia es que Guided Backpropagation utiliza una técnica de "re-tracing" para mostrar solo los caminos en la red que contribuyen positivamente a la activación de una neurona específica [46]. Este algoritmo de visualización se utilizó para analizar cómo el modelo está tomando decisiones para la clasificación multiclase, ya que, como se ve en la Figura 2.9, muestra qué partes de la IRM están siendo utilizadas para activar una determinada neurona o capa de la red.

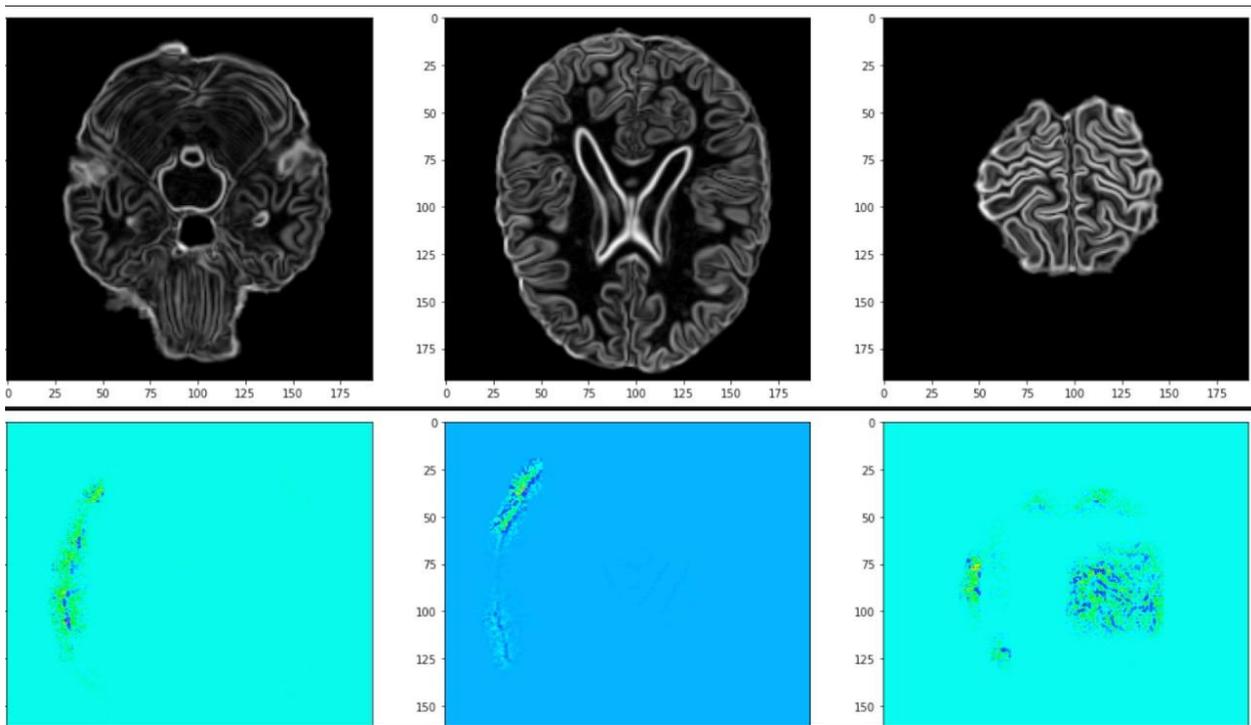


Figura 2.9 Resultados del algoritmo Guided Backpropagation

[Autoría Propia]

2.3 Evaluación

En el proceso de aumentar el número de datos en el dataset para la detección y clasificación de desórdenes parkinsonianos, como se observa en la Figura 2.8, se aplicaron diferentes técnicas de data augmentation como rotaciones en los ejes X e Y , reducción de ruido y contraste en las imágenes. Debido a que se requiere evitar la distorsión del cerebro en la imagen, no se recomienda la utilización del eje Z para las rotaciones en data 3D [47]. Estas técnicas permitieron aumentar significativamente el tamaño del dataset, pero también generaron un problema de desequilibrio de datos en la clase "Solo Parkinson", lo que puede causar overfitting en la validación del modelo. Se propuso el uso de técnicas de regularización como oversampling o undersampling para solucionar este problema.

También, la técnica de contraste resultó perjudicial para la precisión del modelo, y una vez que estas IRMs se eliminaron de la lista de entrenamiento, la precisión del modelo aumentó en un 21% aproximadamente.

El módulo backend del prototipo web recibe las imágenes en formato matricial, realiza el preprocesamiento, la transformación y la clasificación, y devuelve los mapas de atención también en formato matricial. Con esta configuración, el tiempo de respuesta en un servidor con solo CPU para el procesamiento es de aproximadamente 2.5 minutos, lo cual no cumple con los requerimientos de tiempo de la solución. Por lo tanto, se buscan formas de mejorar la eficiencia en el manejo de los datos y aprovechar al máximo el CPU.

2.4 Diseño de la Solución

2.4.1 Diagrama de Bloques

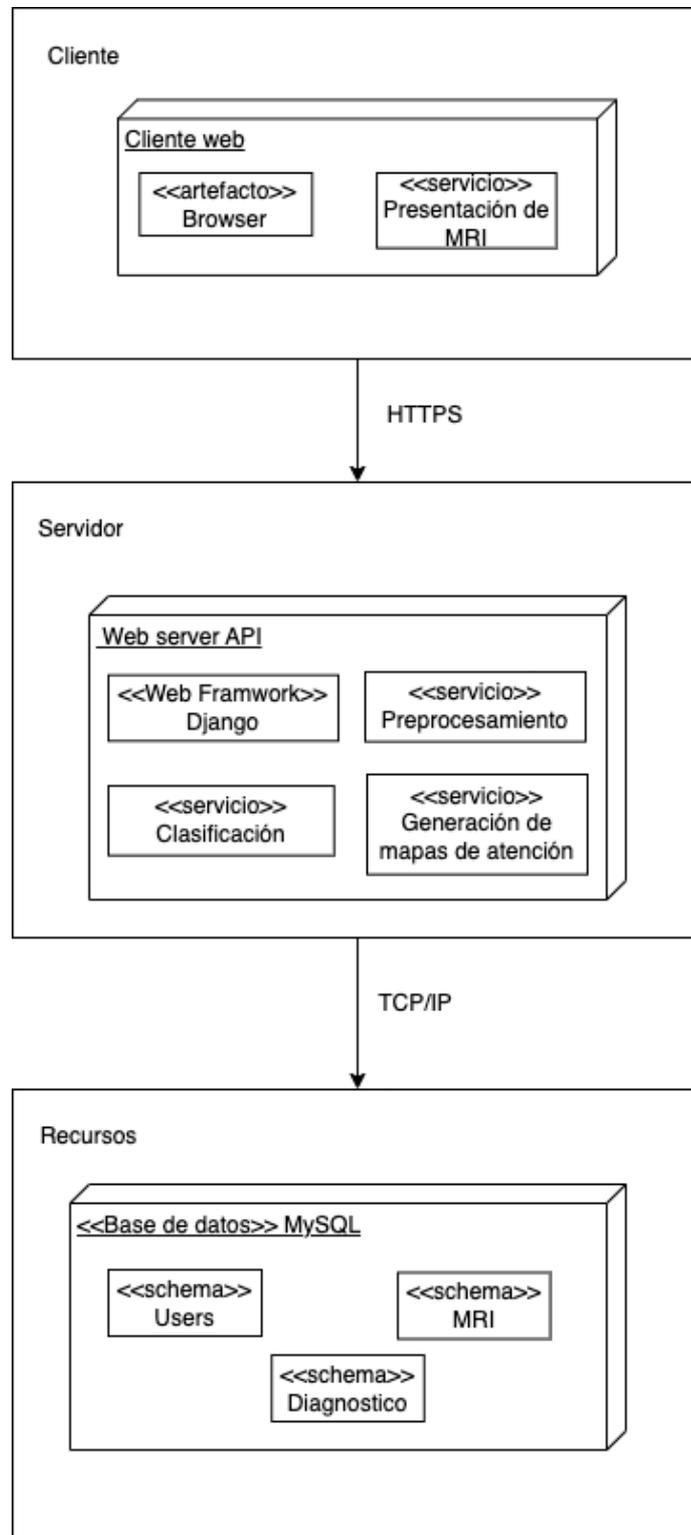


Figura 2.10 Diagrama de Bloques

[Autoría propia]

Para la solución se planteó una aplicación en 3 bloques como se aprecia en la Figura 2.10. El primer bloque consta del cliente web; el segundo es el servidor en el backend donde se utilizó Django para la ejecución de los scripts. Por último, el bloque de recursos consta de la base de datos MySQL.

2.5 Software

2.5.1 Arquitectura

2.5.1.1 Bloque de Aplicación

Para el preprocesamiento y predicción de la presencia de desórdenes Parkinsonianos se requirió diversas herramientas para el manejo de matrices y características de las IRM. Las librerías empleadas se detallan en la Tabla 2.2, usadas en el ambiente Python 3.8.10

Tabla 2.2 Descripción de librerías usadas

[Autoría propia]

Librerías	Versión	Descripción
AntsPy	0.3.4	Librería enfocada en modificación de imágenes médicas como IMR. En esta librería se encuentran los algoritmos de eliminación de ruido, intensidad y ajuste de orientación.
ANTSPyNet	0.1.7	Librería que emplea arquitecturas de aprendizaje profundo adaptadas al lenguaje Python y herramientas para el procesamiento básico de imágenes médicas. Basado en keras y tensorflow
numpy	1.19.5	Librería de Python especializada en el cálculo y análisis de datos masivos, como es la manipulación de matrices y arreglos.
PyTorch	1.12.0 +CU116	Es una librería de tensores optimizada para el aprendizaje profundo mediante GPU y CPU.
tensorflow	2.10.0	Librería enfocada en la inteligencia artificial. Se la emplea con el objetivo de crear redes neuronales y modelos de machine learning.
dicom2nifti	2.3.0	Librería de conversión de formatos de imágenes médicas. Convierte una carpeta con archivos formato dicom (.dcm) a formato NifTI (.nii).
SimpleITK	2.2.0	Librería gráfica principalmente orientada para segmentación, registro y manipulación de imágenes médicas como el uso de detección de bordes.

2.5.2 Diagramas (Clases, casos de uso, interacción, flujo de procesos)

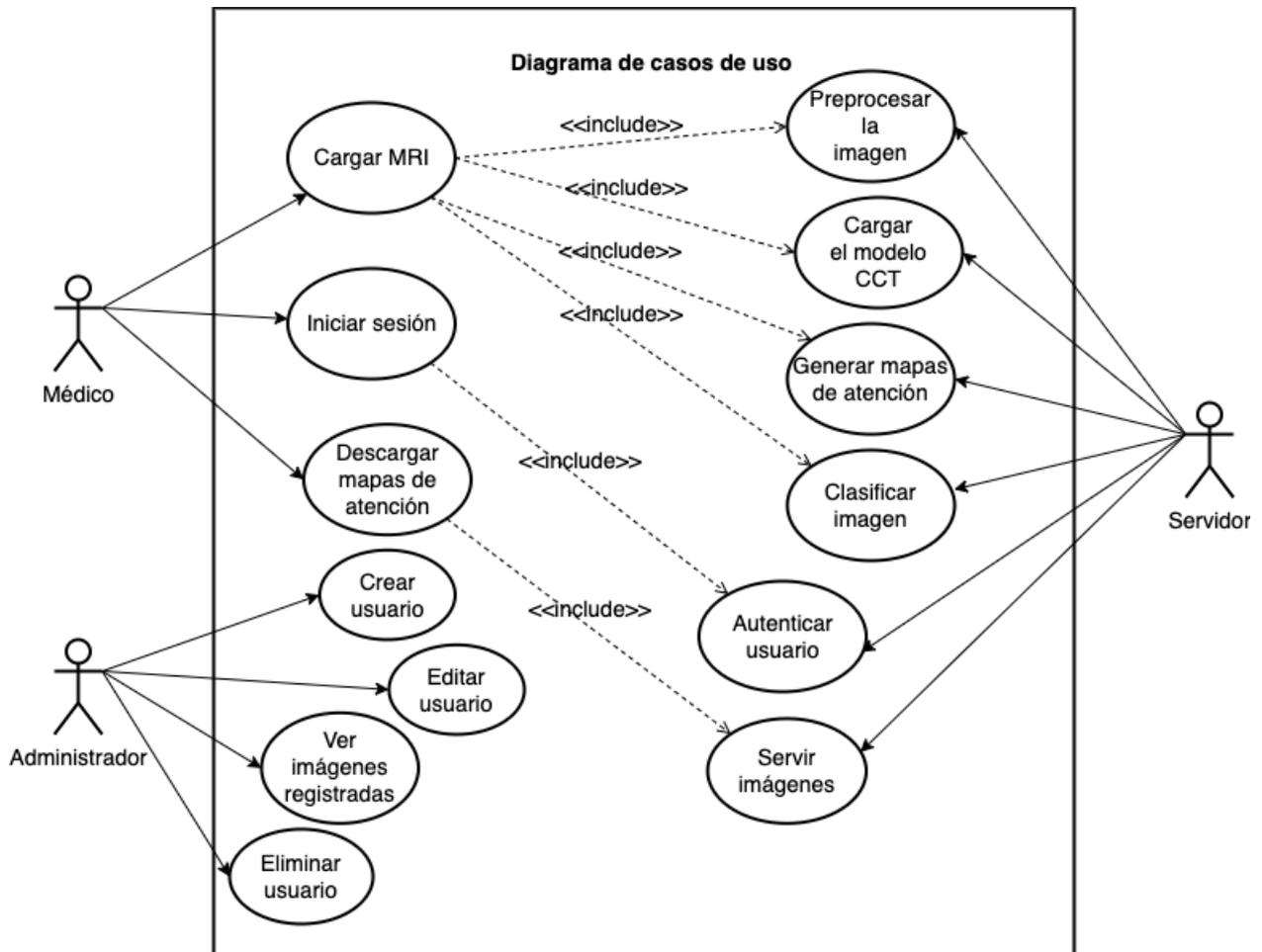


Figura 2.11 Diagrama de Casos de Uso

[Autoría propia]

El diagrama de casos de uso, como se aprecia en la Figura 2.11, muestra los diferentes roles que interactúan con el prototipo propuesto. Los dos principales son el administrador, el cual puede crear, editar y eliminar usuarios. El médico, que tiene permitido iniciar sesión, cargar las IRM y descargar los mapas de atención. El servidor es el encargado de recibir y preprocesar las IRM, esto incluye normalización, extracción de cerebro y detección de bordes; posteriormente alimenta al CCT-3D para la clasificación y generación de mapas de atención.

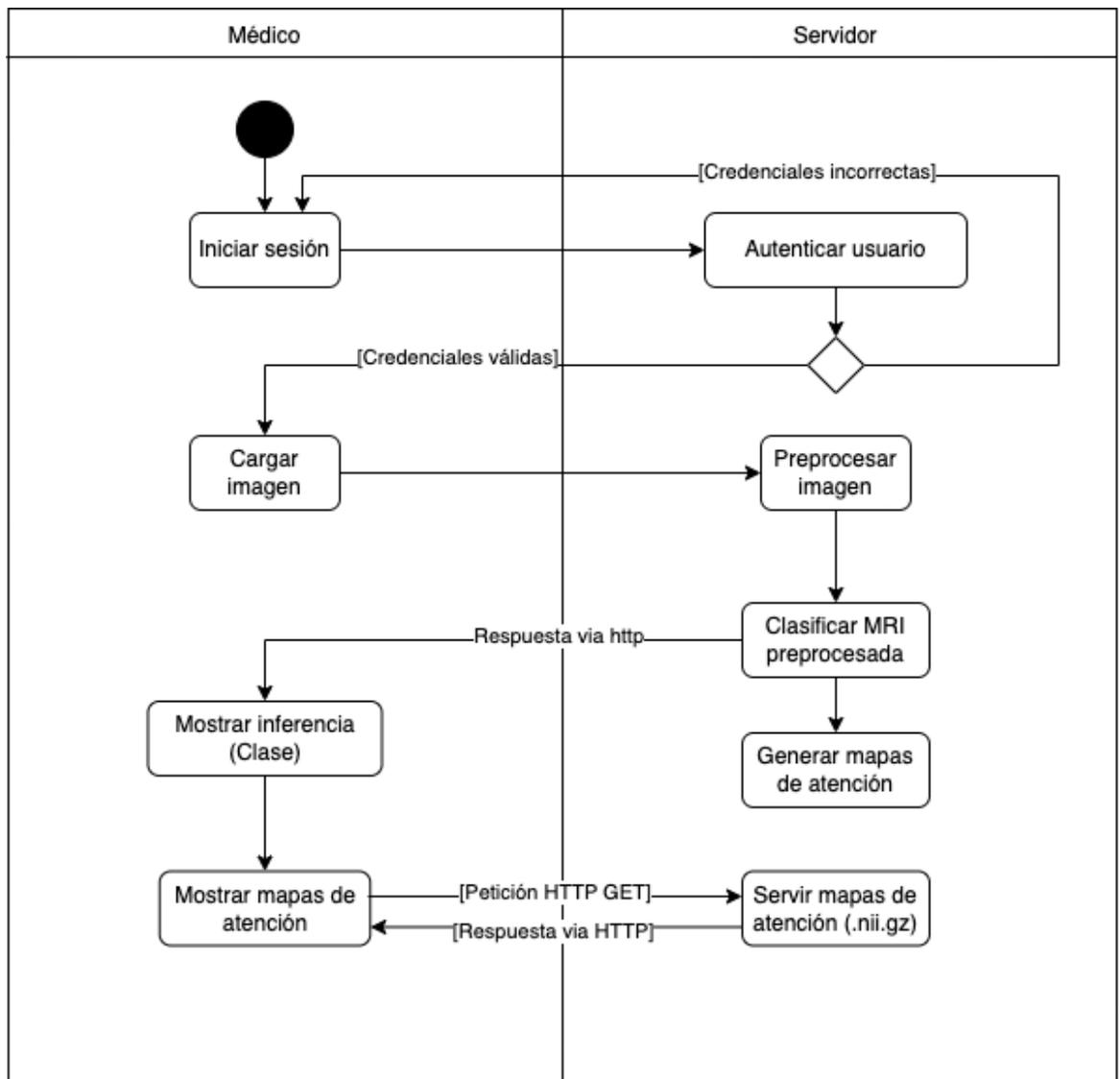


Figura 2.12 Diagrama de Actividades

[Autoría propia]

La Figura 2.12 muestra la secuencia de actividades para los dos roles principales que interactúan con el prototipo. Después de iniciar sesión, el médico carga una IRM la cual es enviada al servidor, una vez aquí es preprocesada, el cerebro es extraído y se detectan bordes. Luego, son alimentadas al modelo CCT-3D el cual realiza la clasificación la cual es retornada al usuario médico al mismo tiempo que guarda los mapas de atención generados en formato NifTI para poder ser servidos mediante API GET.

2.5.3 Diseño de la Base de Datos

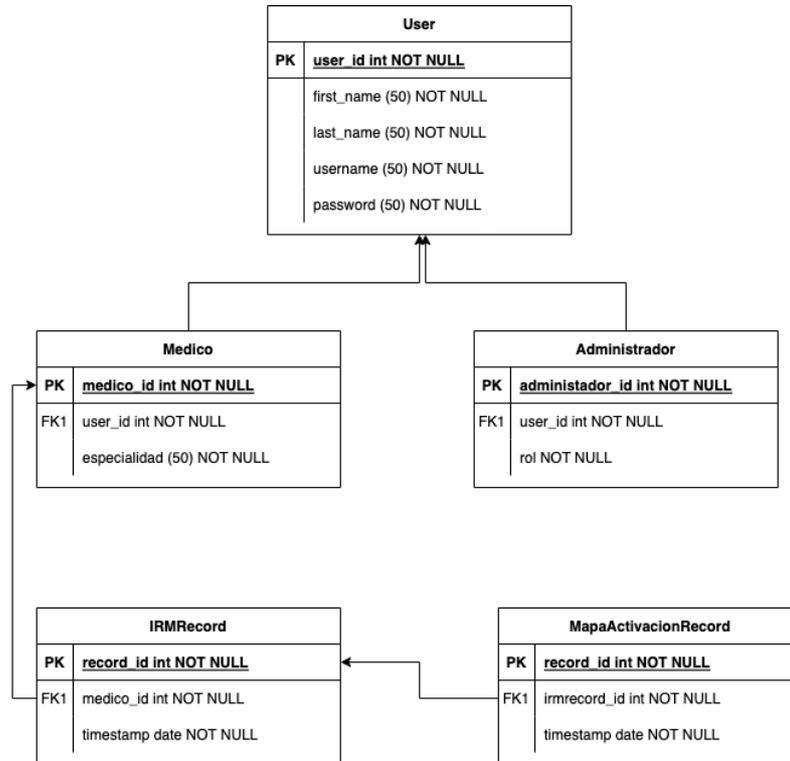


Figura 2.13 Diseño de Base de Datos.

[Autoría propia]

Se implementó una base de datos relacional como indica la Figura 2.13, utilizando MySQL como DBMS. Para el diseño se tomó en cuenta las relaciones entre los diferentes roles del prototipo así como la capacidad de almacenar tanto las IRM como los mapas de activación.

2.5.4 Implementación del Prototipo

El prototipo en cuestión es una plataforma web de arquitectura cliente-servidor diseñada para mostrar la clasificación y visualización de las diferentes regiones del cerebro asociadas a los trastornos Parkinsonianos. La plataforma cuenta con dos roles distintos, el rol de administrador y el rol médico. El rol de administrador se encarga de la creación y eliminación de usuarios, mientras que el rol médico se encarga de subir imágenes en formato NifTI y observar el resultado de la clasificación a través de la interfaz web. Además, esta plataforma permite a los médicos tener una mejor comprensión de la distribución de los desórdenes Parkinsonianos en el cerebro, lo que podría ayudar en el diagnóstico y tratamiento de esta enfermedad.

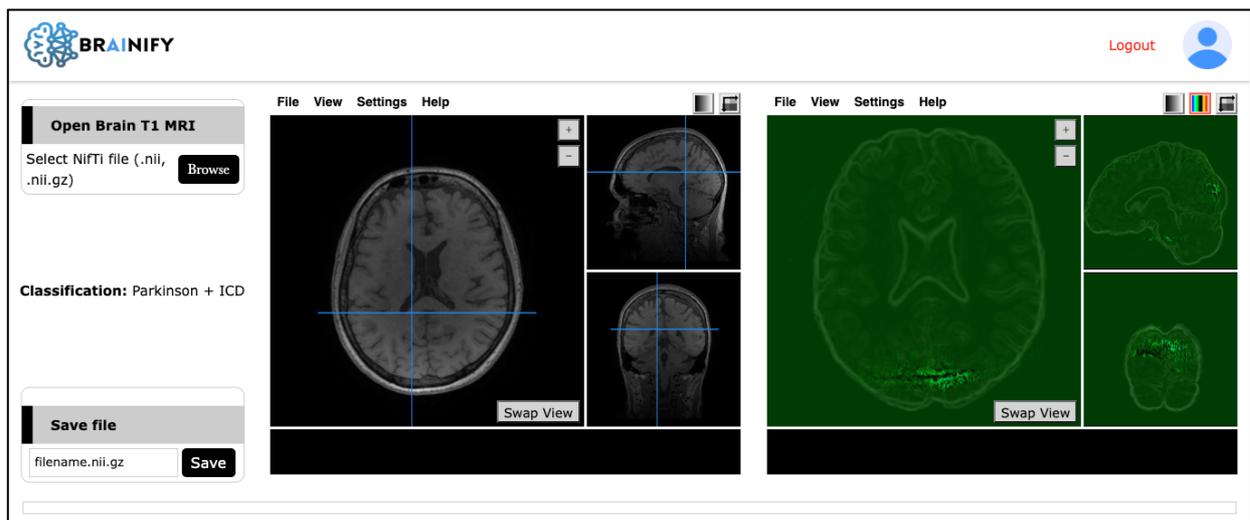


Figura 2.14 Prototipo Web con resultados del modelo

[Autoría propia]

CAPÍTULO 3

3. RESULTADOS Y ANÁLISIS

3.1 Plan de implementación

En el plan de implementación se establecieron todas las tareas necesarias para llevar a cabo la implementación del prototipo final.

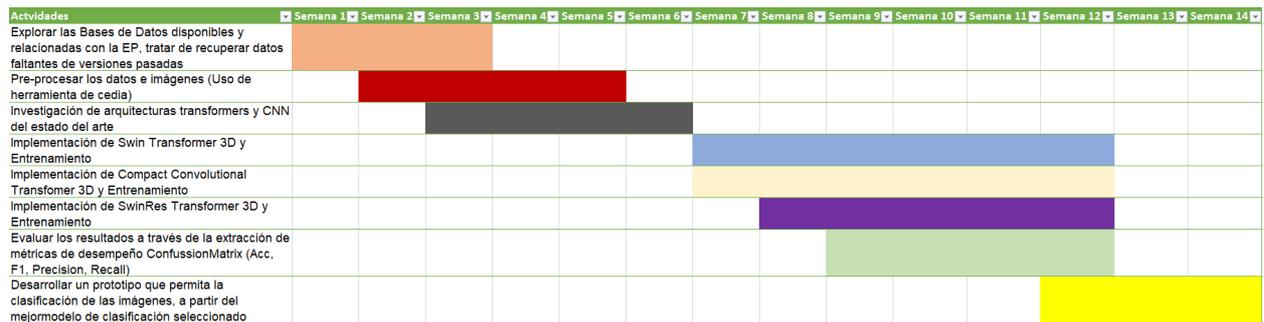


Figura 3.1 Diagrama semanal de planificación del proyecto

[Autoría propia]

Para facilitar la gestión de las actividades, éstas han sido divididas en tres etapas distintas:

- **Etap 1:** Después de aplicar el módulo de preprocesamiento, se implementó el uso de la técnica de data augmentation, lo cual tuvo un impacto significativo en el tamaño de los datasets. En la Tabla 3.1 se observa: el dataset de "Sólo Parkinson", el cual incrementó de 260 a 7456 ejemplos, el dataset de "Apatía", incrementó de 49 a 1372 ejemplos, el dataset de "ICD", incrementó de 64 a 1397 ejemplos y, el dataset de "Apatía más ICD" incrementó de 23 a 644 ejemplos. Luego, se redujo el tamaño de la clase "Solo Parkinson" a 1456 como estrategia para mejorar la representación de los datos y evitar el sobre entrenamiento del modelo, con esto la precisión del modelo aumentó en aproximadamente un 65%. Se pudo notar que al utilizar ángulos de rotación relativamente pequeños (menores a 15 grados) en la técnica de rotación en el eje X e Y, las IRM no sufrieron una distorsión significativa, lo que permitió obtener más datos para reducir el sobreajuste sin afectar la precisión del modelo.

Según la evaluación del modelo implementado, se utilizó una división de 80 a 20 en cuanto al tamaño de los datos de entrenamiento y prueba. Los resultados obtenidos indicaron que el modelo tuvo una efectividad prácticamente del 100% en todas las clases, excepto en la clase de apatía e ICD, como se observa en la Figura 3.2. Este resultado demuestra cómo el uso de data augmentation puede ser efectivo para incrementar el tamaño y mejorar la calidad de los datos utilizados en modelos de aprendizaje automático aplicados en el área de medicina.

Tabla 3.1 Resultados de Data Augmentation

CLASE	DATASET INICIAL (IRM)	POST DATA AUGMENTATION (IRM)
<i>Sólo Parkinson</i>	260	1456
<i>Apatía</i>	49	1372
<i>ICD</i>	64	1397
<i>Apatía + ICD</i>	23	644

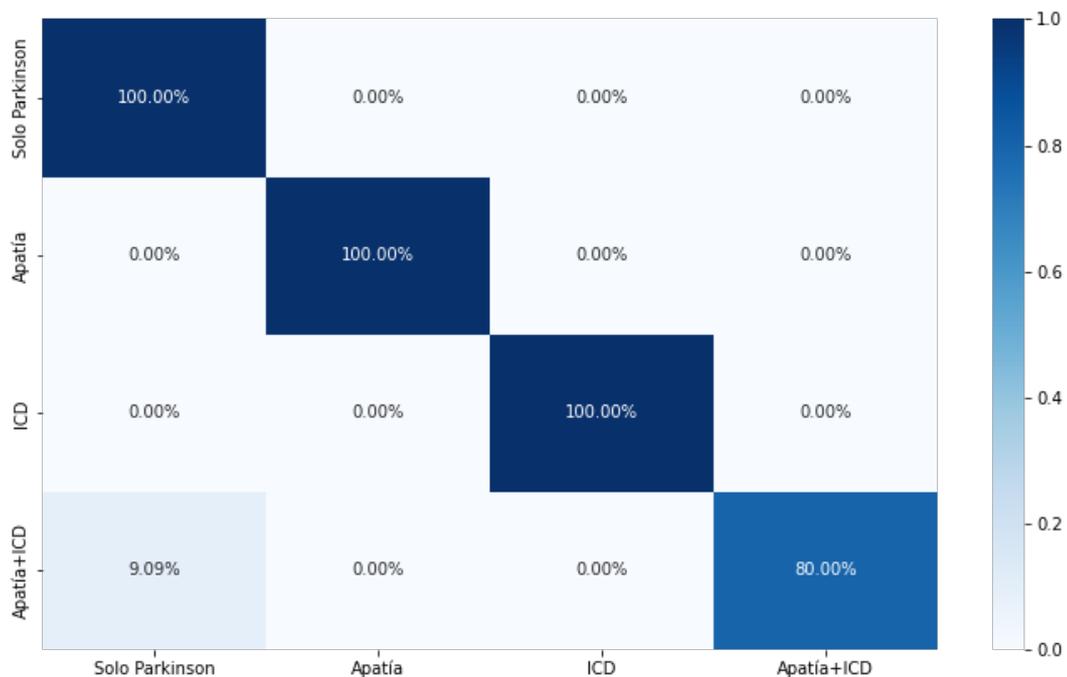


Figura 3.2 Matriz de Confusión

[Autoría Propia]

Etapa 2: Para el análisis de las imágenes, se implementó un modelo Compact Convolutional Transformer con hiperparámetros específicos. El tamaño de las imágenes fue de 192x192x189, el modelo se entrenó durante 200 épocas, con un tamaño de lote de 8 imágenes por iteración, utilizando el optimizador AdamW y un programador de tasa de aprendizaje lineal. La tasa de aprendizaje establecida fue de 1e-5; además, el modelo incluyó 2 capas de atención con una cabeza de atención cada una y 2 capas de procesamiento de imágenes en el tokenizador, como se puede apreciar en la Figura 3.3. Los resultados del modelo mostraron una precisión del 97.14% en las métricas de accuracy, precision, F1 score y recall, tal como se muestra en la sección 3.2 – Métricas de evaluación. Este proceso se basó en la revisión previa de la literatura sobre el uso de Transformers en el procesamiento de imágenes en 3D.

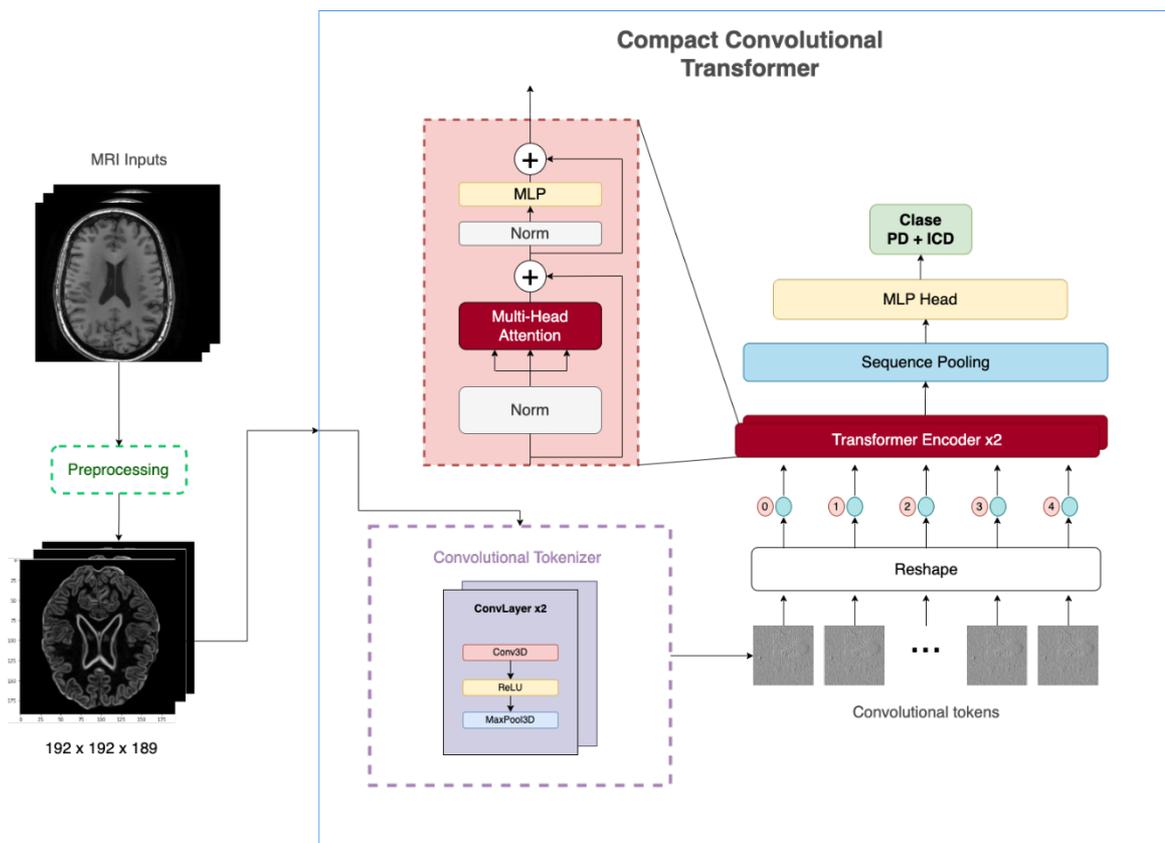


Figura 3.3 Arquitectura personalizada del CCT-3D utilizado en proyecto integrador

[Autoría propia]

Los resultados obtenidos del modelo indican su efectividad en la identificación y clasificación de patrones en imágenes cerebrales. La Figura 3.4 ilustra las regiones cerebrales más activas durante la clasificación multiclase, representadas como puntos brillantes en diferentes áreas. La evaluación de estas áreas de activación es fundamental porque brinda a los investigadores la oportunidad de recopilar y analizar datos más detallados y precisos en relación con los desórdenes Parkinsonianos.

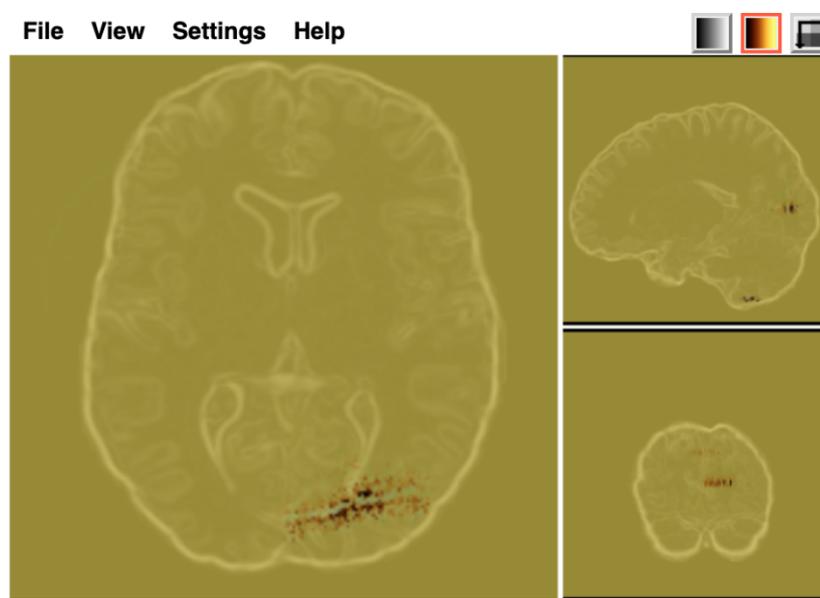


Figura 3.4 Interpretabilidad del CCT en el prototipo web

[Autoría propia]

- **Etapa 3:** En la tercera fase, se implementó una aplicación web con dos roles de usuario diferenciados. Uno de ellos es para el médico especialista, que tiene la capacidad de cargar imágenes de resonancia magnética y luego visualizar la clasificación y regiones cerebrales asociadas. El otro rol es para el administrador, que puede agregar, editar y eliminar usuarios, asignándoles el rol apropiado. Todos los requerimientos tanto funcionales como no funcionales de la aplicación web fueron cumplidos y satisfactoriamente implementados.

3.2 Métricas de evaluación

En la Tabla 3.3 se muestran los resultados obtenidos por el modelo CCT, que obtuvo el mejor rendimiento. Para evaluar su desempeño, se utilizó varias métricas basadas en la matriz de confusión, incluyendo la precisión, recall, F1 score y accuracy. De esta manera, se pudo obtener una evaluación precisa y detallada de la clasificación del modelo.

3.2.1 Accuracy

Proporciona una medida general del desempeño del modelo [48].

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (3.1)$$

3.2.2 Precision

La precisión mide la capacidad del modelo de clasificar correctamente las muestras positivas [48].

$$precision = \frac{TP}{TP + FP} \quad (3.2)$$

3.2.3 Recall

El recall evalúa la habilidad del modelo para encontrar todas las muestras positivas [48].

$$recall = \frac{TP}{TP + FN} \quad (3.3)$$

3.2.4 F1 Score

El F1 Score combina precisión y recall en una sola métrica [48].

$$F1\ score = \frac{2 \times precision \times recall}{precision + recall} \quad (3.4)$$

3.3 Resultados del mejor modelo

Se evaluaron tres modelos diferentes para la identificación y clasificación de desórdenes Parkinsonianos en IRM: el modelo Swin Transformer 3D, el modelo MeshNet 3D y el modelo Compact Convolutional Transformer (CCT 3D), como se muestra en la Tabla 3.2.

Los resultados indicaron que el modelo CCT 3D fue el que obtuvo una mayor precisión en la detección de regiones cerebrales asociadas a desórdenes Parkinsonianos, con un accuracy del 0.8355 al utilizar IRM sin técnica de bordes y un accuracy del 0.9714 al aplicar la técnica de detección de bordes.

Por otro lado, el modelo Swin Transformer 3D tuvo un accuracy del 0.6435 con IRM normales y un accuracy del 0.7556 al aplicar la técnica de bordes. Mientras que el modelo MeshNet 3D tuvo un accuracy del 0.7500 con IRM normales y un accuracy del 0.7945 al aplicar la técnica de bordes en las imágenes. De modo que, la aplicación de la técnica de detección de bordes puede mejorar la precisión de los resultados.

La Tabla 3.3 muestra que el modelo CCT 3D ha logrado un desempeño consistente y uniforme en todas las métricas de evaluación, lo que indica una alta capacidad de predicción y una baja tasa de falsos positivos o falsos negativos. Sin embargo, es importante destacar que los modelos Swin Transformer 3D y MeshNet 3D también presentaron un desempeño aceptable, pero en menor medida en comparación con el modelo CCT 3D.

Tabla 3.2 Comparación de modelos según la métrica Accuracy

Modelo	Mejor accuracy	
	Normal	Edges
Swin Transformer 3D	0.6435	0.7556
CCT 3D	0.8355	0.9714
MeshNet 3D	0.7500	0.7945

Tabla 3.3 Resultado de métricas del modelo CCT-3D

Modelo	Accuracy	Precision	Recall	F1 Score
CCT-3D (Edges)	0.9714	0.9714	0.9714	0.9714

3.4 Análisis de costos

El costo de la herramienta para la detección y clasificación de desórdenes parkinsonianos se calculó en base a los siguientes parámetros: el precio por hora-hombre, según el Decreto Ejecutivo 611 en vigencia [49], es de \$3.75. Se consideraron 80 días laborables con un total de 480 horas trabajadas, 6 horas al día. En conclusión, el precio total de la herramienta es de \$3.600,00.

Tabla 3.4 Costo total del desarrollo del prototipo

Precio Hora-Hombre	Días laborables	Total Horas Hombre	Horas por día	Precio neto
\$3,75	80 días (4 meses)	480	6	\$3.600,00

CAPÍTULO 4

4. CONCLUSIONES Y RECOMENDACIONES

4.1 Conclusiones

- El modelo CCT permite clasificar con una tasa de error menor al 4%.
- Entrenar modelos de tipo Vision Transformer en 3D es computacionalmente muy costoso, por lo que requiere mucho tiempo y hardware especializado.
- El desarrollo de un prototipo web que permite visualizar regiones cerebrales de pacientes con Parkinson más ICD y Apatía, es útil para médicos e investigadores ya que permite tener una comprensión más detallada y visual de las regiones cerebrales lo que facilita identificar patrones o correlaciones entre las regiones cerebrales afectadas proporcionando una mayor comprensión de la enfermedad y apoyo en el desarrollo de nuevos tratamientos
- Se están utilizando técnicas avanzadas en Computer Vision y Machine Learning para mejorar el procesamiento y comprensión de datos en 3D, especialmente en el campo de la clasificación de IRM.

4.2 Recomendaciones

Este trabajo representa una oportunidad importante para futuros desarrollos en el reconocimiento de patrones en imágenes cerebrales relacionados con trastornos Parkinsonianos. Sin embargo, actualmente el modelo solo es capaz de procesar imágenes T1, que es una configuración específica de la resonancia magnética. Para obtener una mayor cantidad de información relevante, se sugiere agregar otras configuraciones de la resonancia magnética. Además, es importante tener en cuenta que para entrenar modelos de aprendizaje profundo en 3D demanda tiempo de entrenamiento prolongado y requiere hardware especializado (GPUs, TPUs) el cual es costoso.

5. REFERENCIAS

- [1] M. N. G. K. Ray Chaudhuri, «Parkinson's Disease: An Overview,» *Journal of Parkinson's Disease* , 2011.
- [2] American Parkinson Disease Association, «APDA,» 01 enero 2021. [En línea]. Available: <https://www.apdaparkinson.org/article/apathy-and-parkinsons-disease/>. [Último acceso: 20 10 2022].
- [3] W.-N. D. a. L.-N. W. Jia-Yi Dong, «Impulse control disorders: clinical features, pathophysiology and management,» *Journal of Mental Health*, 2015.
- [4] J. K. M. N. P. A. D. S. M. S. V. V. J. W. G. R. W. A. E. L. Daniel Weintraub 1, «Impulse Control Disorders in Parkinson's Disease,» *The American Journal of Psychiatry*, vol. 176, nº 1, p. 6, 01 enero 2019.
- [5] D. Z. J. S. V. R. N. Rizos, «The prevalence and impact of non-motor symptoms in Parkinson's disease,» *Journal of Neurology*, 2014.
- [6] M. Kelvin L Chou, «UpToDate,» 03 03 2022. [En línea]. Available: <https://www.uptodate.com/contents/clinical-manifestations-of-parkinson-disease?sectionName=NONMOTOR%20SYMPTOMS&search=Parkinson's%20Disease%20apathy&topicRef=4>. [Último acceso: 20 10 2022].
- [7] M. R. W. M. P. B. G. A. K. N. Okai, «Apathy in Parkinson's disease: an overview,» *Journal of Parkinson's Disease*, 2012.
- [8] A. S. D. W. D. Antonini, «Functional magnetic resonance imaging studies in patients with impulse control disorders and Parkinson's disease: a systematic review,» *Journal of Parkinson's Disease*, 2013.
- [9] E. A. e. al., «Apathy and impulse control disorders in Parkinson's disease: neuroimaging and neuropsychological perspectives,» *Journal of Parkinson's Disease*, 2015.
- [10] C. T. e. al., «Assessing the Progression of Parkinson's Disease with Impulse Control Disorders and Apathy: A Systematic Review,» *Journal of Parkinson's Disease*, 2019.
- [11] G. W. W. P. Eduardo Tolosa, «The diagnosis of Parkinson's disease,» *The Lancet Neurology*, vol. 5, nº 1, pp. 75-86, enero 2006.
- [12] A. E. e. al., «Pathological gambling in Parkinson's disease: a critical review,» *Journal of Clinical Psychiatry*, 2006.

- [13] X.-X. W. Y. F. R. F. J. J. y. Y.-C. W. J.-F. Zhang, «Impulse Control Disorders in Parkinson's Disease: Epidemiology, Pathogenesis and Therapeutic Strategies,» *Frontiers in Psychiatry*, vol. 12, febrero 2021.
- [14] F. A. F. C. Imperiale, «Brain structural and functional signatures of impulsive–compulsive behaviours in Parkinson's disease - Molecular Psychiatry,» *Mol Psychiatry*, vol. 23, p. 459–466, 2018.
- [15] D. Frosini, «Parkinson's Disease and pathological gambling: Results from a functional MRI study Movement Disorders,» *Movement Disorders*, vol. 25, nº 14, p. 2449–2453, 25 octubre 2010.
- [16] A. E. e. al., «Pathological gambling and other impulse control disorders in Parkinson's disease,» *Journal of Neurology, Neurosurgery & Psychiatry*, 2006.
- [17] Parkinson's Foundation, «Navegando el control de impulsos y la hipersexualidad con la enfermedad de parkinson,» [En línea]. Available: <https://www.parkinson.org/library/fact-sheets/impulsos-hipersexualidad> .
- [18] A. Weintraub, «Pathological gambling in Parkinson's disease: a review,» *Movement Disorders*, 2005.
- [19] M. Politis, «Neural response to visual sexual cues in dopamine treatment-linked hypersexuality in Parkinson's disease,» *Brain*, vol. 136, nº 2, pp. 400-411, 31 enero 2013.
- [20] L. B. Zahodne, «Binge Eating in Parkinson's Disease: Prevalence, Correlates and the Contribution of Deep Brain Stimulation,» *Journal of Neuropsychiatry*, vol. 23, nº 1, pp. 56-62, 01 febrero 2011.
- [21] A. Cerasa, «Biomarkers of Eating Disorders Using Support Vector Machine Analysis of Structural Neuroimaging Data: Preliminary Results,» *Behavioural Neurology*, vol. 2015, pp. 1-10, 2015.
- [22] O. Lucas-Jiménez, «Apathy and brain alterations in Parkinson's disease: a multimodal imaging study,» *Annals of Clinical and Translational Neurology*, vol. 5, nº 7, p. 803–814, 25 mayo 2018.
- [23] W. Zhang, G. Yang, Y. Lin, C. Ji y M. M. Gupta, «On Definition of Deep Learning,» *2018 World Automation Congress (WAC)*, pp. 1-5, 2018.
- [24] D. M. U. M. J. G. L. & S. J. (. F. H. P. C. N. N. f. I. C. I. J. C. o. A. I. Ciresan, «Flexible, High Performance Convolutional Neural Networks for Image Classification.,» *International Joint Conference on Artificial Intelligence*, 2011.
- [25] ELSTER LLC, «QUESTIONS AND ANSWERS IN MRI,» 2021. [En línea]. Available: <https://mriquestions.com/neural-network-types.html>.
- [26] A. S. N. P. N. U. J. J. L. G. A. K. Ł. P. I. Vaswani, « Attention Is All You Need,» *In Advances in Neural Information Processing Systems*, pp. 5998-6008, 2017.

- [27] J. C. M. L. K. & T. K. Devlin, «BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,» *In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, pp. 4171-4186, 2018.
- [28] F. Shamshad, «Transformers in Medical Imaging: A Survey.,» 2022.
- [29] A. Dosovitskiy, «An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.,» 2020.
- [30] Y. L. Y. C. H. H. Y. W. Z. Z. S. L. B. G. Ze Liu, «Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,» *arXiv*, 2021.
- [31] H. H. Y. L. Z. Y. Z. X. Y. W. J. N. Y. C. Z. L. D. F. W. B. G. Ze Liu, «Swin Transformer V2: Scaling Up Capacity and Resolution,» *arXiv*, 2022.
- [32] Z. Wu, «MeshNet: 3D Shape Representation Learning using Graph Convolutional Networks,» *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
- [33] Y. F. H. Y. X. Z. Y. G. . Yutong Feng, «Mesh Neural Network for 3D Shape Representation,» *arXiv*, 2018.
- [34] J. Cai, «Efficient Vision Transformers with Compact Convolutional Layers,» *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2021.
- [35] S. W. N. S. A. A. J. L. H. S. Ali Hassani, Escaping the Big Data Paradigm with Compact Transformers, *arXiv preprint*, 2021.
- [36] I. G. a. Y. B. a. A. Courville, Exploring Deep Learning Robustness: Inductive Biases in Vision Transformers and MLP-Mixers, MIT Press, 2016.
- [37] L. W. S. G. H. G. P. P. a. B. G. S. P. Singh, «3D Deep Learning on Medical Images: A Review,» 2020.
- [38] C. G. Z. L. I. R. Z. Y. W. J. Y. G. Q. W. J. Z. D. S. Kelei He, «Transformers in Medical Image Analysis: A Review,» 2022.
- [39] P. F. T. B. Olaf Ronneberger, U-Net: Convolutional Networks for Biomedical Image Segmentation, vol. 9351, Springer, 2015, pp. 234--241.
- [40] V. N. Y. T. D. Y. H. R. D. X. Ali Hatamizadeh, «Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images,» *arXiv*, 4 Enero 2022.
- [41] AntsPy, «ANTsPyNet,» 2019. [En línea]. Available: <https://antsx.github.io/ANTsPyNet/docs/build/html/utilities.html#applications>.
- [42] D. o. R. a. B. Imaging, «Yale School of Medicine,» [En línea]. Available: <https://medicine.yale.edu/bioimaging/suite/manual/guide/correction/>.

- [43] A. D. D. T. S. L. G. J. R. R. D. G. Nina Jacobsen, «Analysis of intensity normalization for optimal segmentation performance of a fully convolutional neural network,» vol. 19, nº 128-138, Mayo 2019.
- [44] Great Learning, «Great Learning,» 13 Diciembre 2022. [En línea]. Available: <https://www.mygreatlearning.com/blog/introduction-to-edge-detection/>.
- [45] J. Brownlee, «Machine Learning Mastery,» 12 Abril 2019. [En línea]. Available: <https://machinelearningmastery.com/how-to-configure-image-data-augmentation-when-training-deep-learning-neural-networks/>.
- [46] A. D. T. B. M. R. Jost Tobias Springenberg, «Striving for Simplicity: The All Convolutional Net,» *arXiv*, 2014.
- [47] D. A. A. E. Marco Domenico Cirillo, «What is the best data augmentation for 3D brain tumor segmentation?,» *IEEE International Conference on Image Processing*, 2021.
- [48] S. M. Mohammad Hossin, «A Review on Evaluation Metrics for Data Classification Evaluations,» *International Journal of Data Mining & Knowledge Management Process*, vol. 5, nº 2, 2015.
- [49] B. Zapata, «La hora extraordinaria sube,» *El Universo*, 5 Diciembre 2022.