Sistema de clasificación de los niveles de cognición normal, afectación leve y enfermedad de Alzheimer a partir de Imágenes de Resonancia Magnética

JAIME ANDRÉS VILLAREJO GARCÍA
Cod. 201631283

UNIVERSIDAD DEL VALLE
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA
PROGRAMA ACADÉMICO DE INGENIERÍA ELECTRÓNICA - 3744
SANTIAGO DE CALI, COLOMBIA
2025

Sistema de clasificación de los niveles de cognición normal, afectación leve y enfermedad de Alzheimer a partir de Imágenes de Resonancia Magnética

JAIME ANDRÉS VILLAREJO GARCÍA
Cod. 201631283

Trabajo presentado para optar por el título de Ingeniero Electrónico

Director:
HUMBERTO LOAIZA CORREA, Ph. D.

Grupo de Investigación Percepción y Sistemas Inteligentes

UNIVERSIDAD DEL VALLE
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA
PROGRAMA ACADÉMICO DE INGENIERÍA ELECTRÓNICA - 3744
SANTIAGO DE CALI, COLOMBIA
2025

CONTENIDO

Pág.

NOTA ACEPTACIÓN

_____
_____
_____
_____
_____

_____
Firma

_____
Firma

_____
Firma

AGRADECIMIENTOS

Quisiera poder agradecer a todas las personas que me apoyaron durante todo el desarrollo de mi carrera profesional, tanto desde la cercanía como desde la lejanía. Un especial agradecimiento a mis padres y familiares que siempre estuvieron al tanto y animándome a concluir este trabajo. A mis padres Celmira García y Jaime Villarejo por el apoyo incondicional que me brindaron durante mi desarrollo profesional incluso en mis momentos de mayor duda y conflictivos. A mis amigos más cercanos Luis Córdoba y Santiago Duque por el apoyo y el acompañamiento emocional que me brindaron durante toda la carrera universitaria. A Laura Colmenares por ser un apoyo emocional en la recta final de ese proyecto de vida.

Un especial agradecimiento a mi director de trabajo de grado Humberto Loaiza por la paciencia y apoyo durante el desarrollo de este trabajo. Finalmente, agradezco a cada una de las personas que me impulsan día a día a mejorar personal y profesionalmente.

LISTA DE TABLAS

RESUMEN

Este trabajo de grado implementó dos máquinas de aprendizaje automático para la clasificación de tres niveles de cognición cerebral CN (cognición normal), LMCI (afectación leve cognitiva), y AD (enfermedad de Alzheimer) en imágenes de resonancia magnética del plano sagital en secuencia T1. La base de datos utilizada fue obtenida desde ADNI (Iniciativa de Neuroimágenes de la Enfermedad de Alzheimer) [1].

La base de datos utilizada contiene imágenes en formato médico DICOM a las cuales se le realizó un preprocesamiento previo al entrenamiento de los modelos de clasificación, que consistió en aplicar técnicas de filtrado espacial, reducción de ruido y realce de las zonas de interés de las imágenes, y cambio al formato jgp. Estas imágenes procesadas constituyeron la entrada a la etapa de extracción de características que son utilizadas en la clasificación.

Se emplearon dos modelos de clasificación: Random Forest y una red neuronal convolucional VGG-16. Se utilizaron 14 sujetos de cada clase para los datos de entrenamiento y 3 sujetos por clase para los datos de test, obteniendo un total de 3922 y 840 imágenes respectivamente. Para el modelo Random Forest se realizó la extracción de características morfológicas utilizando las técnicas *CLAHE* para el realce de contraste, un filtrado gaussiano para suavizar las imágenes y el método *OTSU* para la segmentación del cerebro; también se utilizaron los patrones locales binarios (LBP) y el histograma de gradientes orientados (HOG) como métodos para la extracción de características de forma complementaria. Este último permitió alcanzar el mayor desempeño en la clasificación, con una precisión de 74,57% para los datos de test. En cuanto a la red neuronal, se basó en la técnica transfer learning haciendo uso de una red convolucional pre-entrenada VGG-16. Para esta red no fue necesario la extracción manual de características debido a que esta red proporciona un total de 14’715.714 parámetros, de las cuales se establecieron 7’080.450 parámetros entrenables. Con esta configuración se obtuvo una precisión de 67,26% utilizando la técnica *fine tuning* y un 46,67% complementando la red con la técnica de *Data Augmentation*. Además, se realizó la clasificación utilizando un modelo en cascada (CN vs AD + LMCI y AD vs LMCI) con el fin de obtener un mejor desempeño. Se obtuvo una precisión de 77% para la arquitectura Random Forest, y de 71% utilizando la red VGG-16

El desempeño del modelo Random Forest desarrollado fue mayor al obtenido en [2] para la clasificación de tejido blando (72,5%); aunque en [2] también se clasifiquen otras clases como aire (92,8%) y hueso (85,4%), estas no son de gran

relevancia en el caso de estudio de este trabajo. Para la red neuronal VGG-16 se obtuvo un desempeño significativamente menor comparado con la literatura [3] y [4]; estos últimos alcanzando precisiones mayores al 90%. En ambos modelos, la clasificación de la clase que representa enfermedad de Alzheimer fue la más precisa con 97,89% y 93,48% respectivamente, mientras que la distinción entre alteración cognitiva leve y Cognición normal presentaron una complejidad mayor que podría resolverse con una mayor cantidad de datos.

Por último, se realizó una interfaz gráfica sencilla que permite ingresar resultado de la clasificación. Todo el desarrollo fue realizado en Python.

Palabras clave: Alzheimer, Imágenes de Resonancia Magnética, Random Forest, VGG-16, Data Augmentation, Procesamiento de imágenes.

ABSTRACT

This undergraduate thesis implemented two machine learning models for the classification of three levels of brain cognition CN (normal cognition), LMCI (mild cognitive impairment), and AD (Alzheimer's disease) using T1-weighted sagittal plane magnetic resonance images. The dataset used was obtained from ADNI (Alzheimer's Disease Neuroimaging Initiative) [1].

The dataset consisted of medical DICOM format images, which were preprocessed before training the classification models. Preprocessing involved applying spatial filtering techniques, noise reduction, enhancement of regions of interest, and conversion to JPG format. These processed images were then used as input for the feature extraction stage, which fed into the classification models.

Two classification models were employed: Random Forest and a VGG-16 convolutional neural network. The training set included data from 14 subjects per class, and the test set included 3 subjects per class, yielding a total of 3,922 and 840 images, respectively. For the Random Forest model, morphological feature extraction was carried out using techniques such as CLAHE for contrast enhancement, Gaussian filtering for image smoothing, and the OTSU method for brain segmentation. Additionally, Local Binary Patterns (LBP) and the Histogram of Oriented Gradients (HOG) were used as complementary feature extraction methods. HOG provided the best classification performance, achieving an accuracy of 74.57% on the test set.

The neural network model was based on transfer learning using a pre-trained VGG-16 convolutional neural network. Manual feature extraction was not necessary, as the network provides a total of 14,715,714 parameters, of which 7,080,450 were trainable. With this configuration, an accuracy of 67.26% was achieved using fine-tuning, and 46.67% when combined with Data Augmentation. Additionally, a cascaded classification model was implemented (CN vs AD + LMCI and AD vs LMCI) to improve performance. An accuracy of 77% was obtained using the Random Forest architecture, and 71% using the VGG-16 network.

The performance of the developed Random Forest model exceeded that reported in [2] for soft tissue classification (72.5%). Although [2] also classifies other categories such as air (92.8%) and bone (85.4%), these are less relevant for the context of this study. In contrast, the VGG-16 network showed significantly lower performance compared to results found in the literature [3] and [4], where accuracies above 90% were reported. For both models, the Alzheimer's disease class was the

most accurately classified, achieving 97.89% and 93.48% respectively. However, distinguishing between mild cognitive impairment and normal cognition proved to be more complex, which could be improved with a larger dataset.

Finally, a simple graphical interface was developed to input classification results. The entire implementation was carried out in Python.

Keywords: Alzheimer's, Magnetic Resonance Imaging, Random Forest, VGG-16, Data Augmentation, Image Processing.

Capítulo 1 Introducción general

Este capítulo se describe brevemente la definición del problema, justificación y objetivos del del desarrollo de un sistema de clasificación automática para tres niveles de cognición a partir de imágenes de resonancia magnética. Adicionalmente, explica los conceptos básicos de sistemas de clasificación automática, señales biomédicas relacionadas con la enfermedad Alzheimer y el tratamiento de datos por medio de algoritmos de procesamiento, como lo son: el filtrado Gaussiano, segmentación OTSU, realce de contraste CLAHE, entre otras.

1.1. Definición del problema

La Enfermedad de Alzheimer (AD, Alzheimer's disease) es una enfermedad neurodegenerativa que afecta las habilidades cognitivas, comportamentales y sociales del paciente. Estas limitaciones, que son progresivas, no permiten que las personas afectadas sean capaces de velar por sí mismas debido al deterioro cognitivo que presenta. Esta enfermedad representa aproximadamente el 60% de todos los casos de demencia en el mundo, afectando a cerca de 46,8 millones de personas [5], [6]

Entre las principales técnicas de diagnóstico de esta enfermedad se encuentra la interpretación de imágenes de resonancia magnética (MRI, Magnetic Resonance Image) del cerebro por parte de un experto en el área de salud, basándose en sus conocimientos, experiencia y su capacidad de entendimiento de la imagen. Las herramientas que se han desarrollado para apoyar cuantitativamente a la detección de fases tempranas de la enfermedad han ido variando con el tiempo. En la actualidad se están empleando máquinas de inteligencia artificial debido a su capacidad de extracción de información y procesamiento de altos volúmenes de datos.

Las MRI son utilizadas tanto para la detección como la clasificación de diferentes estadios de la AD. Las MRI suministran información de la composición y estado de las diferentes regiones del cerebro, como por ejemplo la materia gris, materia blanca y líquido cefalorraquídeo. Debido a que la AD se encuentra estrechamente relacionado con la pérdida de tejido de materia gris del cerebro, las MRI suponen ser una fuente de información importante para la detección y clasificación de diferentes fases de la enfermedad [8]

Generalmente, las imágenes MRI son analizadas visualmente por parte de los profesionales de la salud para diagnóstico, detección y seguimiento de la evolución de patologías[6]. Este método de análisis manual de las imágenes es altamente subjetivo pues depende de la experticia, agudeza visual y fatiga del profesional de salud. Este método también puede conducir a mayores tiempos para la elaboración de diagnósticos y por consiguiente a retrasos en los tratamientos para las personas afectadas.

Considerando que el análisis de imágenes MRI para la detección y clasificación del estado de avance de la AD es un tema abierto en investigación, se desarrolló este trabajo como respuesta a la pregunta problematizadora: ¿Cómo implementar un sistema de reconocimiento de patrones con máquinas de aprendizaje automático para clasificar tres niveles de cognición (normal, afectación leve y enfermedad de Alzheimer) a partir de información extraída de imágenes de resonancia magnética del cerebro de personas?

1.2. Justificación

El Alzheimer es una enfermedad (AD, Alzheimer Disease) neurodegenerativa progresiva que afecta a millones de personas en todo el mundo, siendo el mayor causante de demencia del mundo [6]. La detección temprana y el conocimiento del estado de avance del Alzheimer es crucial para realizar una intervención temprana, brindar un tratamiento eficiente y hacer seguimiento al paciente.

Los sistemas de clasificación y detección de la enfermedad de Alzheimer utilizan principalmente imágenes de resonancia magnética debido a que es una señal biomédica que muestra el tejido blando del cerebro de manera clara, característica afectada por el Alzheimer. Debido a que estas imágenes no poseen patrones simples de identificar en los diferentes sujetos, las redes neuronales convolucionales sugieren ser una poderosa herramienta para la clasificación de diferentes niveles de deterioro cognitivo.

Las estadísticas indican que el 15% de la población, a nivel mundial, mayores de 65 años ya se encuentran afectados por algún deterioro cognitivo leve. En las últimas tres décadas, las investigaciones y las imágenes de resonancia magnética han permitido, en gran medida, un aporte significativo en el estudio y el entendimiento del cerebro, el mapa funcional del cerebro y redes cerebrales en estado de reposo [9]

El análisis de las imágenes de MRI presentan unas grandes limitaciones con respecto a la objetividad, debido a que los diagnósticos son realizados visualmente por profesionales de la salud. Aunque éstos estén capacitados para la extracción de información de forma manual, sigue siendo un método que no es constante en el tiempo debido al desgaste visual, fatiga y experiencia del propio profesional de la salud. Por esta razón, y la falta de equipos para el diagnóstico asistido por computadora en países en vía de desarrollo; se hace crucial el desarrollo de estas herramientas con el fin de aportar una solución a esta problemática mundial.

1.3. Objetivos generales y específicos

En este trabajo de grado se establecieron los siguientes objetivos:

1.3.1 Objetivo General

Clasificar automáticamente los niveles de cognición normal, afectación leve y enfermedad de Alzheimer mediante la utilización de técnicas de visión por computador y reconocimiento de patrones aplicadas a imágenes de resonancia magnética del cerebro

1.3.2 Objetivos específicos:

- Seleccionar una base de datos con imágenes MRI del cerebro que contengan casos de personas con cognición normal, afectación leve y enfermedad de Alzheimer.
- Implementar una técnica de procesamiento para adecuación y mejoramiento de las imágenes de MRI que potencialicen la detección y clasificación automática de los niveles de cognición normal, afectación leve y enfermedad del Alzheimer.
- Aplicar una técnica de reconocimiento de patrones con máquinas de aprendizaje para clasificación automática de niveles de cognición a partir de MRI.
- Desarrollar una interfaz para la visualización de las imágenes MRI y los resultados del clasificador desarrollado.
- Elaborar un protocolo de pruebas para establecer alcances y limitaciones del sistema de clasificación de los niveles de cognición normal, afectación leve y enfermedad de Alzheimer a partir de imágenes MRI.

1.4. Descripción del trabajo

A lo largo de este documento se realizó la selección y construcción de los modelos de clasificación que se evaluaron para la identificación de los tres niveles de cognición de interés, utilizando las técnicas *CLAHE* para el realce de contraste, un filtrado gaussiano para suavizar las imágenes y el método *OTSU* para la segmentación del cerebro; también se utilizaron los patrones locales binarios (LBP) y el histograma de gradientes orientados (HOG) como métodos para la extracción de características de forma complementaria. Luego se presenta un breve acercamiento teórico a los conceptos utilizados durante el desarrollo de los sistemas de clasificación. Se incluye información sobre la base de datos que se utiliza para la clasificación de los niveles de cognición. Se plasmó el paso a paso utilizado para la preparación de datos e imágenes que se utilizarían en el sistema de clasificación Random Forest, así como la red neuronal VGG16 seleccionada; y finalmente, se encuentran los resultados obtenidos de ambos clasificadores, con una precisión de 74% para el modelo Random Forest y 67,73% con la red VGG-16. También se plasmaron sus comparaciones, conclusiones, limitaciones y una interfaz básica para la visualización de resultados.

1.5 Organización de los capítulos

Este documento cuenta con un total de 7 capítulos. Cada capítulo aborda conceptos necesarios para el desarrollo de la solución propuesta de este proyecto. En el capítulo 1 se desarrolló la problemática y la justificación para el desarrollo de sistemas de clasificación automático para la detección de tres niveles de cognición. También se mencionan los objetivos y se presenta una descripción general de la metodología empleada para la implementación de la solución propuesta.

En el segundo capítulo se plasman los antecedentes nacionales e internacionales que se usaron como referencia para el desarrollo de la solución. Además, se plasman todos los conceptos teóricos utilizados a lo largo del documento. El tercer capítulo describe la metodología utilizada en los algoritmos de las máquinas de aprendizaje seleccionados; se plasmaron las características claves para establecer la arquitectura de las máquinas de aprendizaje que se usaron. El capítulo 4 muestra las pruebas y resultados de los algoritmos establecidos anteriormente. Se presenta la descripción de los datos que se utilizaron para el entrenamiento y test. También

se muestran las diferentes variaciones de hiperparámetros y métodos de extracción de características utilizados para obtener el mejor resultado.

El capítulo 5 muestra la interfaz gráfica simple desarrollada para el uso y evaluación de las máquinas de aprendizaje ya entrenadas. Se desarrolla la interfaz y sus funcionalidades. Pasando al capítulo 6, se plasmaron las conclusiones, alcances, limitaciones y las oportunidades de mejora para trabajos futuros Con base en los algoritmos presentados en este trabajo. Finalmente, en el último capítulo se encuentran las referencias bibliográficas utilizadas a lo largo del documento.

Capítulo 2 Antecedentes y Marco Teórico

2.1 Antecedentes

Debido a los avances tecnológicos que se han presentado en las últimas décadas, con respecto a la obtención de información y a la calidad de esta ha permitido el desarrollo de herramientas automatizadas, como las máquinas de aprendizaje automático, que suponen una ayuda cuantitativa para el diagnóstico de patologías y/o alteraciones en la salud. Con la finalidad de seguir contribuyendo al desarrollo de estas habilidades, y basándose en las referencias obtenidas en la propuesta de este proyecto, se realizó una investigación más detallada y específica, de documentación comprendida entre 2018 y 2024 que se centran en el uso de ML para la clasificación de los tres niveles de cognición.

Específicamente en el caso de las imágenes médicas se han encontrado varios estudios como [10], [11] y [12] entre otros, se realizan evaluaciones de resultados con base en máquinas de aprendizaje superficial como lo son las máquinas de soporte vectorial (SVM, por sus siglas en inglés), o el algoritmo de Árboles de decisión aleatorio (Random Forest o RF, por sus siglas en inglés) evaluado en [2].

2.1.1 Antecedentes Nacionales

En la universidad tecnológica de Pereira [12] se realizó una evaluación del mejor modelo para la clasificación de tres etapas del deterioro cognitivo utilizando una red convolucional llamada “DenseNet121”. La investigación se basó en la variación del número de cortes del plano coronal de cada imagen PET. El resultado del estudio con respecto a la precisión del modelo de clasificación arrojó los siguientes valores para las etapas NC (Cognición normal), MCI (Deterioro cognitivo leve) y AD respectivamente: 75,72%, 50,62% y 78,66%.

2.1.2 Antecedentes internacionales

En [6] se propone un estudio comparativo entre cuatro niveles de cognición, utilizando las redes neuronales VGG-16 y VGG-19 para la clasificación de personas que poseen NC, EMCI (Deterioro cognitivo leve prematuro), LMCI (Deterioro cognitivo leve tardío) y personas con AD. Para la clasificación de los estadios mencionados, se aplicó técnicas de segmentación a las MRI para localizar regiones con materia gris. La clasificación alcanzó una precisión del 97,89% con la red VGG-16 y 98,47% con VG-19. Las MRI utilizadas fueron obtenidas de la base de datos pública de la Iniciativa de Neuroimagen de la Enfermedad de Alzheimer (ADNI).

En [14]se realizó un estudio de comparación entre diversos modelos de redes neuronales con diferentes arquitecturas para la clasificación de tres niveles de cognición: NC, MCI y AD. La comparación se realizó entre los modelos 2D-DCNN, AlexNet y VGG-16, obteniendo una precisión de 99,89%, 98,97% y 99,31% respectivamente.

También en [4] se realiza una comparación entre dos modelos de red convolucional VGG-16 y AlexNet con la finalidad de realizar la clasificación binaria entre sujetos diagnósticos con la enfermedad de Alzheimer y sin afectaciones de cognición. Este estudio comparativo demostró una clara ventaja de la red VGG-16 con un 93,48% de precisión frente a 89,44% de precisión de la red AlexNet.

2.1.3 Discusión

En los antecedentes referenciados se aprecia que, aunque las redes neuronales convolucionales se usan cada vez más en los entornos de clasificación, aún es común encontrar modelos superficiales como “Random Forest” debido a su bajo costo computacional. Los diferentes estudios presentan diferentes procesos tanto de selección de parámetros como de preprocesamiento de las imágenes.

Las técnicas de procesamiento utilizadas en los estudios varían dependiendo de las características y cantidad de imágenes diagnósticas y los estados de la AD que se desean clasificar. Cada estudio realizó las etapas de procesamiento utilizando diferentes algoritmos a su elección como lo fueron *DICOM* para el cambio de formato de imágenes, *MANGO tool kit* para el ajuste de contraste de imágenes, remoción de ruido entre otras.

En Tabla 2.1 se consigna, de manera reducida, las principales características de los sistemas de clasificación utilizados en los antecedentes investigados,

destacando la precisión de los algoritmos, las clases que identifican y el tipo de arquitectura que se usaron. Con base en la información recolectada de cada uno de los estudios, se decide realizar una comparación entre los métodos “Random Forest” y la red neuronal convolucional VGG-16. La alta precisión de ambos algoritmos de clasificación sugieren ser una buena solución a la pregunta problematizadora.

Ref.	Tipo de máquina de aprendizaje	Arquitectura	Dataset	Clases	Sujetos	Precisión
[2]	Shallow Learning	Random Forest	Imágenes MRI y CT	Aire, Tejido Blando y hueso	14	92,8%, 72,5% y 85,4%
[12]	Deep Learning	DenseNet21	Imágenes FDG-PET ADNI	NC, MCI, AD	352, 394 y 283	75,72%, 50,62%, 78,66% (NC, MCI, AD)
[6]	Deep Learning	VGG-16 y VGG-19	Imágenes MRI ADNI 3 class y OASIS	NC, MCI, AD	2633, 1512 y 2480 (imágenes)	97,12% y 98,47% (VGG16, VGG-19)
[14]	Deep Learning	2D-DCNN (Red Neuronal Convolucional de dos dimensiones Deep ConvNet), AlexNet y VGG-16	Imágenes MRI ADNI	NC, MCI, AD	53, 62 y 45	99,89%, 98,97% y 99,31%
[4]	Deep Learning	VGG16 y AlexNet	Imágenes MRI	NC, y AD	380 sujetos (1474 imágenes / sujeto)	93,48% y 89,44%

Tabla 2.1 Características principales de los métodos investigados en las referencias investigadas.

2.2 Marco Teórico

El objetivo general de este trabajo es desarrollar un sistema que clasifique los niveles de cognición normal, afectación leve y enfermedad de Alzheimer mediante la utilización de técnicas de visión por computador y reconocimiento de patrones aplicadas a MRI. En este apartado se presentan los conceptos básicos que son necesarios para el desarrollo de este trabajo.

2.2.1 Enfermedad Alzheimer

La enfermedad de Alzheimer es un trastorno neurológico que provoca la muerte de las células del cerebro que desencadena en un avance paulatino, normalmente asociado al envejecimiento común. Con el avance de la enfermedad, el deterioro

cognitivo se va agravando, afectando la capacidad de recordar acontecimiento, la toma de decisiones y, en etapas avanzadas, la capacidad motora de la persona. La AD también puede desarrollarse junto con un cambio de la personalidad y dificultad de realizar actividades cotidianas [15].

Esta enfermedad representa el 50% de la demencia a nivel mundial siendo, con diferencia, la más común de estas. En el mundo se calcula un aproximado de 22 millones de personas que sufren de esta enfermedad y se espera que en las siguientes tres décadas este número aumente al doble. Aunque no es común, la Asociación de Alzheimer Internacional sugiere que la enfermedad podría comenzar desde los 50 años, además no existe cura aún.

La clasificación de la enfermedad se presenta inicialmente por la edad en la que comienza, siendo una patología temprana si comienza antes de los 65 años, y tardía si comienza después de esa edad. Así mismo, dentro de esas dos categorías se establecen dos subcategorías que son: Familiar: si hay historial familiar con este padecimiento, y esporádica, si no hay antecedentes familiares [15].

2.2.2 Imágenes de resonancia magnética

Las imágenes de resonancia magnética o MRI se considera un gran avance tecnológico en el ámbito de la salud para el prediagnóstico de diferentes patologías debido al detalle de los tejidos en el cuerpo obtenidos mediante esta técnica. Las MRI utilizan un campo magnético externo para alinear los átomos de más presencia en el cuerpo humano (hidrógeno) y obtener imágenes detalladas de los tejidos blandos.

Existen los tejidos tipo 1 y tipo 2 que permiten obtener información diferente del cuerpo humano. Ambos se pueden detectar utilizando MRI con cierta variación en el procedimiento de la toma de imágenes. Los tejidos de tipo 1 son aquellos que permiten ver patologías que se basan en la visualización de grasa o materia tangible, mientras que los tejidos tipo 2 muestran de forma más detalladas los materiales líquidos del cuerpo, utilizados para la detección de patologías que afectan o que se basan en el contenido de agua de cierta parte del cuerpo [16]. En la Figura 2.1 se aprecia un ejemplo de una MRI para cada secuencia obtenida en [17].

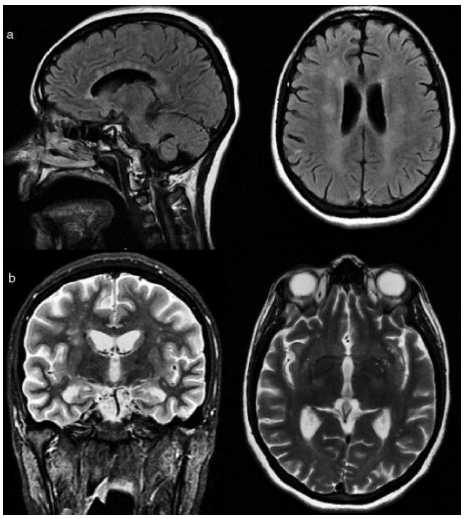


Figura 2.1 Imagen de resonancia magnética del cerebro. A) Cortes sagital y axial en secuencia T1. B) Corte coronal y axial en secuencia T2.

Las imágenes de interés para el desarrollo de este trabajo son las MRI que muestran los tejidos de tipo T1, debido a que este tipo de imágenes refleja los sólidos de color blanco luminoso permitiendo identificar deficiencias estructurales en el cerebro, característica particular de la enfermedad de Alzheimer y sus diferentes etapas previas [18] .

El banco de imágenes que se utilizó para el entrenamiento de los modelos de aprendizaje automático fue obtenido desde la plataforma ADNI debido a su gran número de imágenes MRI en formato médico. Las imágenes que se utilizaron se descargan originalmente en formato DICM, un formato médico que contiene información variada y parámetros para la visualización de imágenes de contienen un diferente número de cortes y/o sub-imágenes de la materia gris, que juntas reconstruyen todo el cerebro.

2.2.3 Máquinas de aprendizaje automático

Las máquinas de aprendizaje automático presentan dos grandes variantes: máquinas de aprendizaje superficial (Shallow Learning Machine, SLM) y máquinas de aprendizaje profundo (Deep Learning Machine, DLM). Las principales diferencias entre las SLM y las DLM radica en que, en las primeras, la extracción y selección de características la realiza el experto con base en su conocimiento, y las

segundas utilizan más capas en su arquitectura para realizar la extracción y selección de características de manera automática [7].

En particular los algoritmos de aprendizaje profundo han supuesto un gran avance en las tecnologías de clasificación debido a su gran precisión. Estos algoritmos, aunque presentan varias ventajas también presentan dificultades y es la necesidad de recursos tecnológicos para el entrenamiento y el tener una gran cantidad de información etiquetada y seleccionado para el aprendizaje [19].

Para el desarrollo de las máquinas de aprendizaje superficial y profundo, se realizó una investigación a profundidad en busca de los mejores parámetros y características que supondrían una mejor respuesta de los clasificadores. En este capítulo se explica los diferentes métodos de extracción de características utilizados y la razón de su aplicación.

2.2.3.1 Extracción de características

Las imágenes MRI presentan dificultades particulares para la segmentación y caracterización de zonas de interés debido a su bajo nivel de nitidez, el ruido, la variación de parámetros y calidad del dispositivo de adquisición, y en particular, la dificultad de diferenciar tejidos de manera automática debido a la posible similitud de valores de intensidad de los pixeles en la imagen. En [20] se realiza una revisión bibliográfica y aplicación de diversos métodos para la descripción de MRI.

La selección de los descriptores se debe hacer dependiendo del tipo de imagen que se desea estudiar, y también la información de interés que aporta al caso de estudio. En el caso de este trabajo, donde se busca analizar las diferencias de la corteza cerebral que pueda haber entre tres diferentes fases de cognición, se plantean una serie de características morfológicas, clasificadas en: de forma, texturales y de intensidad.

Descriptores de forma

Las características morfológicas son las permiten describir o representar una imagen plenamente identificable de otras. Dichas características, según [20], deben cumplir con ser invariantes a traslación, rotación y escala; resistentes al ruido e invariante a la ocultación, es decir que una parte de la imagen así se encuentre oculta o segmentada, no altera las características del resto de la imagen. En la Tabla 2.2 se describen algunos descriptores de forma de interés.

NOMBRE	DESCRIPCIÓN
Área	Número de pixeles que tiene la región de interés.
Perímetro	Longitud del borde la región de interés.
Circularidad	Describe qué tan similar es la región a un círculo.
Excentricidad	Mide qué tan lejos esta la región de parecerse a una circunferencia.
Longitud de los ejes	Describe el tamaño de los ejes principales de la elipse que mejor se ajusta a la región de interés.
Solidez	Describe qué tan compacta es la región de interés.
Extensión	Describe que tan rellena es la región de interés.
Número de euler	Mide la irregularidad de la región y que tantos vacíos posee.

Tabla 2.2 Contiene una breve descripción de cada descriptor de forma de interés.

Algunas de las características morfológicas se obtienen con el uso de métodos pertenecientes a la función regionprops. La solidez, característica que mide la densidad de la región de interés usando la relación entre el área de la región y el área de su envolvente convexa. Un valor de 1, indica que tiene pocas irregularidades y es compacta; por otro lado, si se acerca a cero significa que la imagen es muy irregular y presenta múltiples concavidades. La ecuación para el cálculo de ese parámetro se aprecia en [2.1].

$$Solidez = \frac{Área\ de\ la\ región\ de\ interés}{Área\ de\ la\ envolvente\ convexa} \quad [2.1]$$

La extensión hace referencia a la proporción del área de la región de interés con respecto al área de su caja delimitadora ***bounding box***. Si bien, el área envolvente convexa que se utilizaba en solidez es el polígono más pequeño que puede rodear completamente la forma de la región de interés, la caja delimitadora es el rectángulo más pequeño que puedo encerrar por completo la región de interés.LA ecuación para el cálculo de la extensión es [2.2]

$$Extención = \frac{Área\ de\ la\ región\ de\ interés}{Área\ del\ bounding\ box} \quad [2.2]$$

Otras de las características morfológicas que se utilizaron en el estudio fue el número de Euler, siendo una característica topológica que, utilizando la conectividad de las regiones segmentadas, describe la presencia de vacíos o "huecos" en la imagen. Este método se utiliza con imágenes binarias que solo se encuentran tonalidades de grises, por lo general pixeles que representan vacío o "huecos" en la imagen.

El número de Euler se define como la diferencia entre los componentes conexos y los vacíos o "huecos" que existen en la imagen. Los componentes conexos son regiones de pixeles que se encuentran conectados entre sí, puede ser de forma vertical, horizontal y/o diagonal. Se expresión matemática se aprecia en [2.3]

$$Número\ Euler = Componentes\ conexos - Vacíos \quad [2.3]$$

La circularidad como característica morfológica, nos indica que tanto se acerca la imagen a un círculo perfecto. El cálculo de esta característica se hace utilizando el perímetro de la imagen utilizando la función ***regio.parimeter***. El valor de la circularidad entre más cercano a 1, más perfecta es la circunferencia de la región evaluada, por otro lado, un valor cercano a cero indica regiones más alargadas o irregulares. LA ecuación matemática que representa esta característica se muestra en [2.4].

$$Circularidad = \frac{4\pi \times \text{Área}}{Perimetro^2} \qquad [2.4]$$

Descriptores texturales

Las características texturales de una imagen hacen referencia a la propiedad superficial externa de los objetos. En el procesamiento de imágenes se puede relacionar la textura con los niveles de intensidad de los píxeles de una imagen, describiendo zonas irregulares que podrían significar la presencia de lesiones. La aplicación de este tipo de algoritmos en MRI supone un posible determinante de los niveles de cognición del paciente debido a que el deterioro cognitivo está estrechamente relacionado con la degradación de la materia gris [20]. En la Tabla 2.3 se describen algunos descriptores texturales de interés.

NOMBRE	DESCRIPCIÓN
Contraste	Mide la variabilidad de los valores de los píxeles (niveles de gris)
Disimilitud	Mide la variabilidad de los niveles de gris de una imagen, donde se le asigna un peso extra entre las diferencias más grandes.
Homogeneidad	Medida de qué tan similares son los niveles de gris de la zona de interés.
Energía	Representa la uniformidad de la GLCM*, donde valores altos significan una variación menor.
Correlación	Mide la relación lineal entre pares de píxeles.

*Tabla 2.3 Contiene una breve descripción de cada descriptor textural de interés. *Abreviatura en inglés para la matriz de coocurrencia de niveles de grises.*

Descriptores de intensidad

Las características de intensidad de una imagen MRI hace referencia a los valores de los píxeles o niveles de gris que se encuentran en la imagen. Estas variaciones de niveles de gris pueden significar cambios de tejidos, lesiones, ruido en secciones particulares e incluso el relieve de una zona o región de interés. El cálculo de métricas de intensidad algunas veces puede ser de gran ayuda en la clasificación de imágenes; sin embargo, no se puede desconocer que son valores muy afectados por el ruido de una imagen. En la se describen algunos descriptores de intensidad de interés.

NOMBRE	DESCRIPCIÓN
Media	Describe la intensidad promedio de la región.
Desviación estándar	Describe la variación de los valores de intensidad de la región.
Skewness	Describe la asimetría de la distribución de intensidades.
Kurtosis	Describe el pico de la distribución de intensidades.

Tabla 2.4 Contiene una breve descripción de cada descriptor de intensidad de interés.

Segmentación usando Local Binary Pattern (LBP)

Existen diferentes algoritmos de segmentación de imágenes dependiendo de la finalidad de la segmentación que sea necesaria. Incluso algunos tipos de problemas de detección de objetos en imágenes son referenciados como problemas de segmentación [21].

En este trabajo se desarrolló la segmentación utilizando el algoritmo *Local Binary Pattern* (LBP, por sus siglas en inglés) como el utilizado en [22] donde se utilizó para la visualización de cambios en la sangre por la absorción de luz laser incidente en el dedo índice; sin embargo, en este trabajo se optó por ese método debido a que ese algoritmo también funciona bastante bien para imágenes MRI debido a las diferencias de contraste de estas.

El método LBP se utiliza para la descripción de texturas en imágenes. Es un método comúnmente utilizado en visión por computadora y son muy útiles para la detección de rostros, la clasificación de texturas y la segmentación de imágenes, que es la finalidad de su uso en este caso de estudio.

Este algoritmo recibe como parámetros de entrada; la imagen original, el vecindario *P* y el radio *R*. El vecindario *P* hace referencia al número de píxeles que se toman alrededor de un píxel central, generalmente en forma de un círculo. El radio *R*, también es una distancia calculada en número de pixeles, pero en este caso hace referencia al radio de la circunferencia que encierra al píxel central.

El histograma resultante de la aplicación del método LBP brinda información sobre la frecuencia de aparición de valores concretos de pixeles en toda la imagen; es decir, cuenta cuantas veces aparece cada valor de la LBP en la imagen de entrada. Este histograma resultante es un vector normalizado que muestra la distribución de patrones LBP, y se utilizó posteriormente para la extracción de características y entrenamiento de la máquina de aprendizaje Random Forest.

Histograma de Gradientes Orientados

El método Histograma de Gradientes Orientados o HOG, por sus siglas en inglés, es un descriptor de características que busca la descripción de texturas e identificación de bordes en las imágenes. Este método es ampliamente utilizado en problemas de detección de objetos e incluso de rostros, debido su simplicidad y su bajo costo computacional.

Aunque HOG sea un método simple, cabe resaltar es un algoritmo robusto para la caracterización de bordes y texturas de imágenes basados en la intensidad de sus píxeles. Este algoritmo divide la imagen en celdas, medidos en píxeles, donde se calculan los gradientes de cada celda; así mismo, estas celdas se agrupan en bloque más grandes cuya finalidad es normalizar los gradientes de las celdas para reducir la variación que pueda haber entre cambios de intensidades.

En [23] se desarrolló un sistema basado en SVM y el algoritmo HOG para la detección de rostros. La ventaja de utilizar el método HOG, es la sencillez del código además de la simplicidad de la información en términos globales, es decir, el principal motivo que se utiliza este algoritmo es porque permite la representación de una imagen completa con un único vector de características, en lugar de tener una tener muchos vectores de características por zonas locales de una imagen.

2.2.3.2 Preprocesamiento

Teniendo en cuenta que las MRI son imágenes médicas obtenidas por un proceso no invasivo, algunas veces se usa sustancias de alto contraste para obtener una mayor claridad de la imagen por parte del experto clínico. De la misma manera, una imagen más clara y precisa permite un diagnóstico con menos probabilidad de error. La calidad de este tipo de imágenes que presentan diferentes tonalidades expresando tejido blando, duro y hasta líquidos, pueden variar por muchos factores como lo son: el equipo médico utilizado, los parámetros de la máquina de resonancia e incluso por la cantidad de contraste aplicado al paciente, de ser el caso.

Es común que este tipo de imágenes les sean aplicadas algoritmos de ajuste de contraste, buscando una imagen más clara y limpia que permita un clasificador más robusto y con menos probabilidad de errores. Este proceso se realiza antes de la extracción de características de las imágenes con la finalidad de que sea más sencillo y mucho más claro para el modelo de clasificación la segmentación de zonas y/o regiones de interés.

Debido a que existe una gran variedad de técnicas de radiología para la obtención de imágenes como la Tomografía por emisión de positrones (PET, por sus siglas en inglés), la Tomografía computarizada (CT, por sus siglas en inglés), la imagen de resonancia magnética funcional (fMRI, por sus siglas en ingles), y la misma MRI; se debe realizar una correcta selección de algoritmos que permita la extracción más eficiente y de las características más importantes para cada tipo de imagen, que permita una generación de información médica objetiva que ayude a un diagnóstico temprano [11].

Aunque en diferentes estudios se utilicen diferentes algoritmos de preprocesamiento, dependiendo del tipo de imágenes que se utilicen (PET, CT, fMRI entre otras) hay una generalidad en el tratamiento previo a las imágenes de resonancia magnética. La remoción del ruido es vital para tener una imagen clara y que el sistema pueda diferenciar más fácilmente las características de cambios en la escala de grises.

Realce de contraste

Existen diferentes algoritmos que buscan solventar los inconvenientes mencionados anteriormente, y mejorar el contraste de la imagen como lo son *Histograma de Ecualización* (HE, por sus siglas en inglés), *Ecualización Adaptativa del Histograma* (AHE, por sus siglas en ingles) y *Ecualización adaptativa del histograma limitada por contraste* (CLAHE, por sus siglas en inglés). En [24] se aplican estos tres algoritmos a MRI con la finalidad de comparar cuál algoritmo es más fiable y da mejores resultados, obteniendo finalmente que CLAHE es la opción que menos amplifica el ruido de las imágenes al aumentar el contraste de forma local.

Filtrado y Suavizado

El método de suavizado gaussiano se basa en la campana de Gauss o distribución normal, donde los parámetros de la ecuación 3.1 x y y corresponden a las coordenadas del píxel que se está avaluando, y σ es la desviación estándar que controla el grado de suavizado, a un valor de σ más alto, mayor será el desenfoque de la imagen.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad [2.5]$$

La distribución gaussiana asigna más peso a los pixeles que se encuentran cerca a la ventana de convolución o kernel establecido, eso significa que los pixeles más alejados representan un menor peso. Un kernel más grande (7x7), filtra más valores y podría generar una pérdida de detalle de la imagen, mientras que uno pequeño (3x3) podría ignorar algunas regiones con ruido, pero mantendría mayor detalle en la imagen. La ecuación [2.5] se muestra la expresión matemática que define al filtro.

Método Otsu

El algoritmo de Otsu se basa en un histograma obtenido de la imagen basándose en la distribución de intensidades de los píxeles. Se utiliza una binarización de los valores de intensidad de la imagen utilizando un umbral t, diviendo la imagen dos clases w, la clase 1 son píxeles con intensidades menores o iguales a t, y la clase dos las que son mayores a t.

El umbral t que minimiza la varianza intra-clase, es el umbral óptimo según Otsu de forma que maximice la diferencia entre las clases obteniendo una segmentación más precisa. Para el cálculo de la varianza intra-clase se utiliza la probabilidad de cada clase $P_1(t)$ y $P_2(t)$. La ecuación [2.6] expresa el cálculo de este parámetro.

$$\sigma_w^2(t) = P_1(t)\sigma_1^2(t) + P_2(t)\sigma_2^2(t) \qquad [2.6]$$

2.2.4 Random Forest

Este método de aprendizaje superficial es un algoritmo que se basa en la división de los datos de entrenamiento en un número *k* de subconjuntos, creando *k* árboles de decisión. Este método se conoce como *Bootstrapping* y consiste en entrenar diferentes modelos individualmente, y finalmente combinar las probabilidades de cada árbol de decisión y clasificar basándose en la moda de los resultados de cada algoritmo [25].

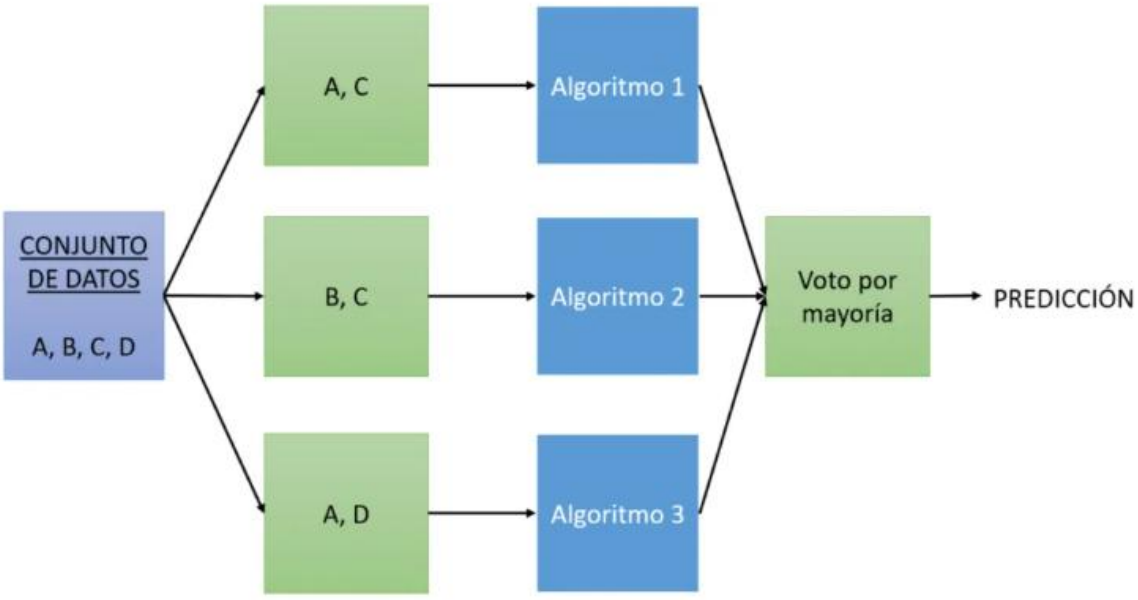


Figura 2.2 Diagrama de la estructura de un algoritmo Random Forest. Fuente: [25]

Una de las principales ventajas del método de clasificación Random Forest es que el número de árboles de decisión no es un parámetro crítico que afecte el sobre entrenamiento; esto es debido a que por mucho que aumente el número de algoritmos de decisión (árboles) el valor del *test error* se estabiliza y la precisión del modelo llega a su límite [26].

2.2.5 Red Neuronal Convolucional

Las redes neuronales convolucionales son ampliamente utilizadas en el ámbito médico para la clasificación de imágenes y extracción de características de patologías comunes. Este tipo de redes aprenden de manera supervisada, basándose en parámetros previamente calculados para el ajuste de la misma red.

Las redes convolucionales, más específicamente el modelo VGG-16, cuenta con diferentes niveles o capas que busca la mayor robustes en la clasificación de imágenes. Las capas convolucionales son parte fundamental del proceso de extracción de características debido a que, mediante la aplicación de filtros que recorren toda la imagen, se generan mapas de características. Filtros pequeños de 3x3 son los encargados de recorrer toda la imagen, y extraer y/o reconocer patrones importantes locales como lo pueden ser los bordes, colores, anomalías, entre otras. La Figura 2.3 ejemplifica el proceso de reconocimiento de patrones locales que se hace con los filtros.

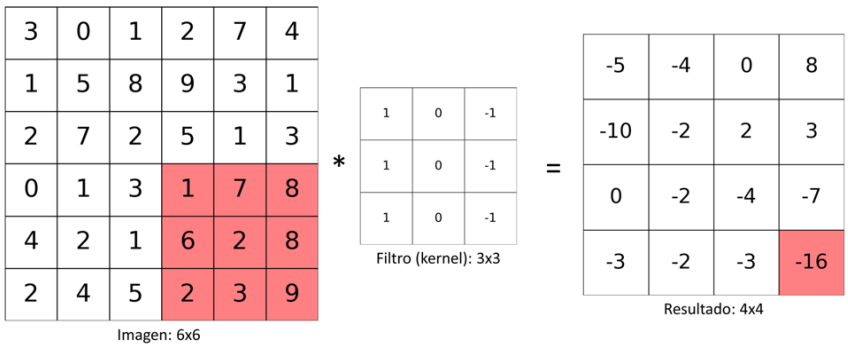


Figura 2.3 Iteración n de la aplicación de un filtro 3x3 a una matriz nxn demostrando la función de reconocimiento de patrones o importancia de una subsección de la imagen original. Imagen tomada de: https://codificandobits.com/blog/convolucion-redes-convolucionales

Posterior al filtrado, la red VGG-16 utiliza también capas ***"Max Pooling"*** 2x2 con la finalidad de reducir el tamaño de la imagen original a la mitad. Esta capa se aplica justo después de una capa de filtrado porque la finalidad es reducir el tamaño de la imagen utilizando únicamente los patrones más característicos de la misma.

La red VGG-16 consta de 13 capas convolucionales y 3 *Fully connected* layers, sumando un total de 16 capas que da lugar a su nombre. Las capas convolucionales de esta red utilizan repetidamente un filtro 3x3 lo que característica este modelo. La Figura 2.4 muestra la arquitectura del modelo de la red VGG-16.

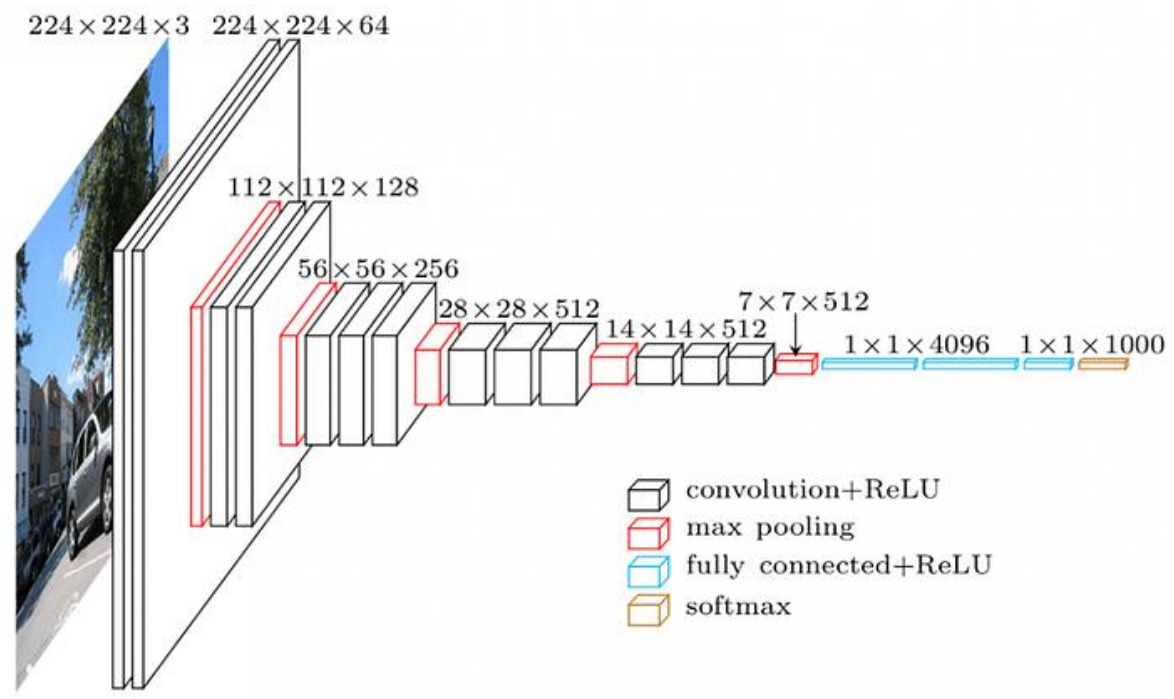


Figura 2.4 Arquitectura red neuronal convolucional VGG-16. Imagen tomada de [34]

Capítulo 3 Metodología de la solución propuesta

En este capítulo se describe el proceso de desarrollo de los algoritmos utilizados para la clasificación de los diferentes niveles cognición con las imágenes de resonancia magnética ya obtenidas y organizadas en la base de datos diseñada para este estudio. Se abordan los métodos de preparación de datos para los

diferentes algoritmos de las máquinas de aprendizaje superficiales y profundos, previos al entrenamiento y clasificación.

En este trabajo se realizó la comparación de dos tipos de modelos de aprendizaje automático: el primero fue un modelo Random Forest (RF, por sus siglas en inglés) que fue entrenado con características morfológicas extraídas de las imágenes de resonancia magnética, y el segundo método fue un modelo de red neuronal convolucional (CNN, por sus siglas en inglés). Ambos modelos fueron entrenados con la misma base de datos de entrenamiento.

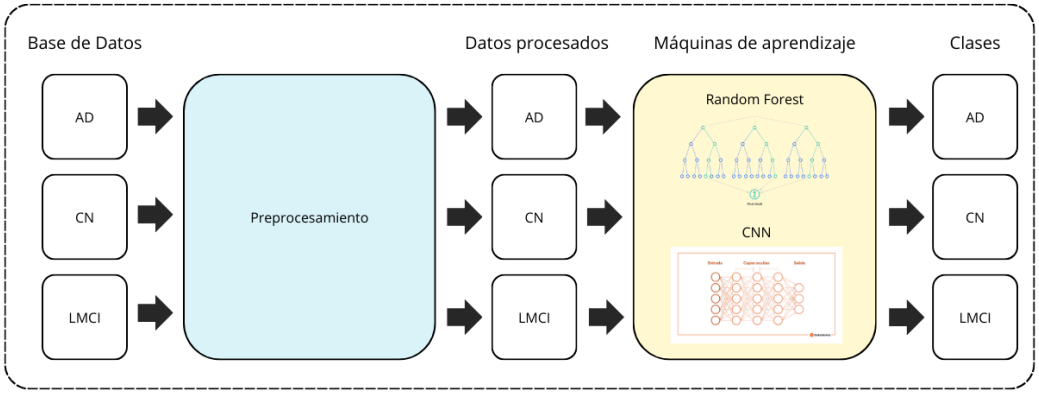


Figura 3.1 Flujo de datos general en el proceso de preparación de imágenes previos y posteriores a la etapa de clasificación. Donde AD: Enfermedad de Alzheimer, CN: cognición normal y LMCI: cognición con afectación leve. Imágenes adaptadas al diagrama de flujo de elaboración propia.

La Figura 3.1 muestra el diagrama de flujo que refleja el proceso por el que pasan los datos antes, y después de la máquina de aprendizaje. En este diagrama, se aclara que las imágenes de la base de datos también cuentan con un preprocesamiento previo, desarrollado en la sección 4.1.2, para la limpieza y reducción de ruido de las imágenes en formato médico DICOM, siendo el formato de entrada al algoritmo de clasificación en formato JPG de 256x256 pixeles (datos limpios). La Figura 4.5 muestra a detalle las fases del preprocesamiento.

3.1 Modelo Random Forest

Tras haber seleccionado el modelo Random Forest (sección 2.1.3) para el desarrollo del sistema de clasificación superficial, se realizó la investigación pertinente de los diferentes factores (sección 2.2.3) que influyen directamente con el rendimiento de dichas máquinas de aprendizaje, como lo son: la preparación de las imágenes, la extracción de características y el ajuste de los hiperparámetros del modelo. En esta sección se describen estas actividades para la máquina seleccionada.

3.1.1 Preparación de imágenes

Luego de cargar las imágenes de la base de datos en un arreglo multidimensional que describe cada imagen (*numpy en Python)*, separadas de sus etiquetas las cuales fueron organizadas en un arreglo diferente, se obtienen dos arreglos totalmente separados entre la información bruta de las imágenes y las etiquetas a cuáles pertenecen cada una. La Figura 3.2 Diagrama de flujo de los datos utilizados para la clasificación del modelo Random Forest. Se muestra los diferentes pasos de la extracción de características que son la entrada de la máquina de aprendizaje Random Forest

A continuación de desarrollan los procesos de preparación de las imágenes médicas aptas para obtener la mejor respuesta del modelo entrenado.

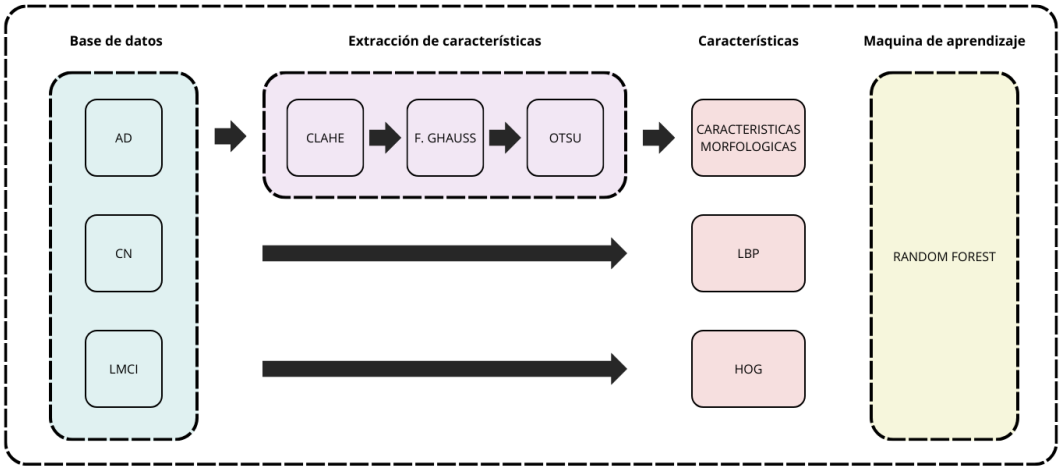


Figura 3.2 Diagrama de flujo de los datos utilizados para la clasificación del modelo Random Forest. Se muestra los diferentes pasos de la extracción de características que son la entrada de la máquina de aprendizaje Random Forest. Se muestra el orden del uso de las técnicas CLAHE, filtro de GAUSS y OTSU para la obtención de las características morfológicas, junto con los algoritmos de características LBP (patrón local binario) y HOG (Histograma de gradientes orientados).

3.1.1.1 Realce de contraste

Con base en la revisión bibliográfica y a las pruebas realizadas, se definió el método CLAHE como el indicado para el realce del contraste de las imágenes que se utilizaran para este trabajo. Se buscaba aumentar la intensidad del contraste de la región de la corteza cerebral, debido a que es la región de interés para la detección del estado de la materia gris del paciente. Se estableció como límite de recorte para el contraste un valor de 1.0 para evitar la sobreexposición, y una subsección local o *tiles* de 8x8 pixeles para el ajuste local del contraste. La Figura 3.3 son ejemplos aleatorios del conjunto de imágenes obtenidos luego de la aplicación de CLAHE.

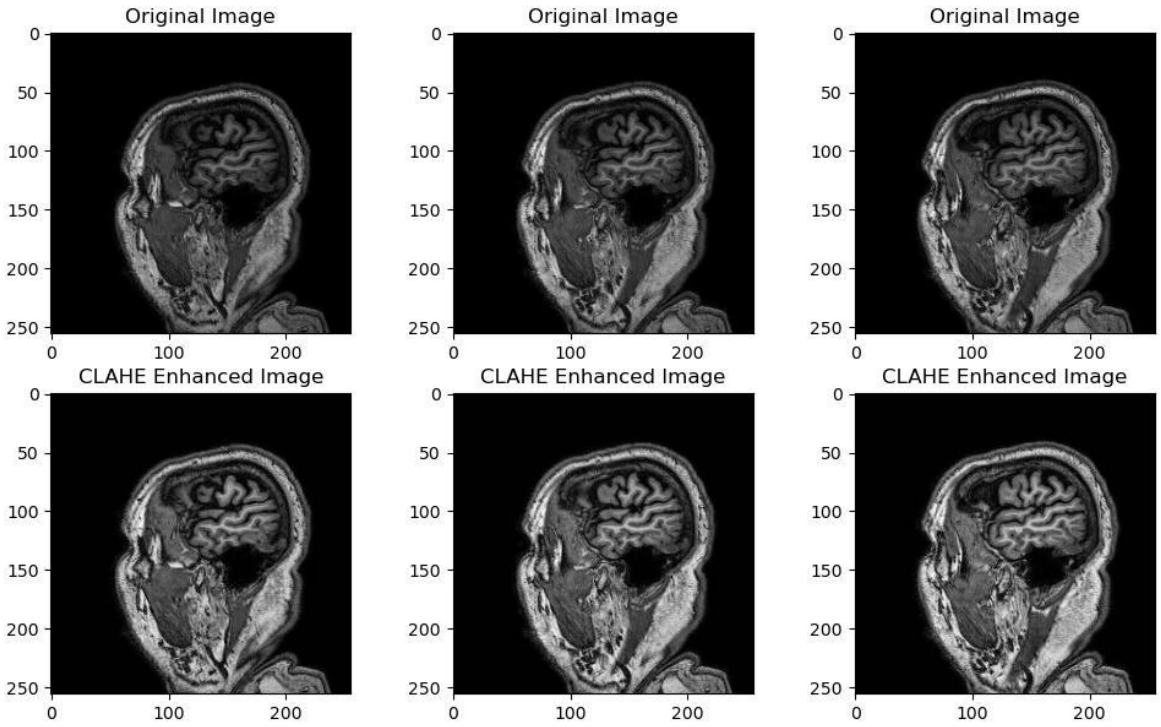


Figura 3.3 Ejemplo de tres imágenes aleatorias luego de aplicar algoritmo CLAHE para el realce del contraste, permitiendo una mayor distinción de detalles en diferentes secciones del cerebro.

3.1.1.2 Filtrado y suavizado

También se aplicó un proceso de crecimiento de regiones para generar una matriz de igual tamaño que la imagen original, pero que contenga una asignación de etiquetas por regiones. Es decir, al algoritmo se le establecen unos valores límites de contraste, para establecer punto semillas de las cuales se evaluarán una por una obteniendo un etiquetado por regiones a partir de las locaciones de las semillas seleccionadas.

Primero se aplicó un filtro de desenfoque gaussiano para remover ruido excesivo de la imagen y suavizar la imagen, preparándola para procesos posteriores de extracción de características que se puedan ver afectadas por el ruido. Este suavizado se realizó utilizando la función *GaussianBlur()* de la librería *Open CV.*

Cada píxel de la imagen original se reemplaza por un valor ponderado obtenido de los valores de intensidad de sus vecinos. En la Figura 3.4 se muestra el resultado

de la aplicación de un suavizado gaussiano agresivo, borrando detalles de la imagen, pero reduciendo el ruido al mínimo.

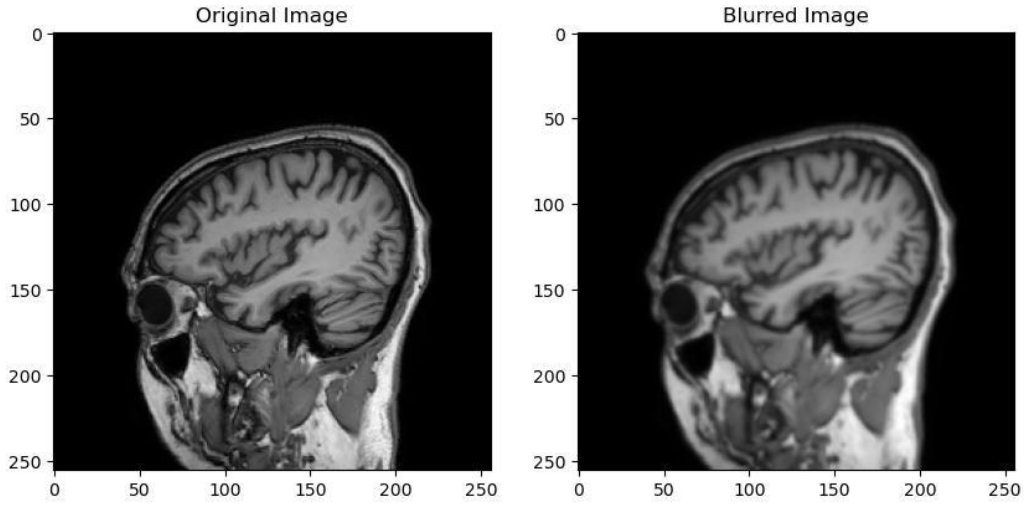


Figura 3.4 se muestra el resultado de la aplicación de un suavizado gaussiano agresivo, borrando detalles de la imagen, pero reduciendo el ruido al mínimo. A) Imagen original B) Imagen después del proceso de suavizado con parámetros x, y = 5x5 y σ = 0

Los parámetros utilizados para el suavizado gaussiano se establecen con un tamaño de matriz del Kernel de 5x5 referenciando a lo que sería el tamaño de alrededor del píxel evaluado que se utiliza para calcular el promedio ponderado; y los valores sigma se establece en 0, indicando que la función calcula la desviación estándar del filtro de forma automática con base en el tamaño del kernel.

3.1.1.3 Método Otsu

Se utilizó el método de *Otsu* para determinar el umbral óptimo de intensidad para la detección de las semillas y poder separar el fondo de la imagen y la zona de interés.

En la Figura 3.5 se muestran tres imágenes aleatorias después del proceso de segmentación. Las imágenes mostradas corresponden a la matriz de etiquetas

generadas a partir de las semillas, segmentando la región del cerebro de interés luego de aplicar el proceso de suavizado gaussiano y el método de Otsu.

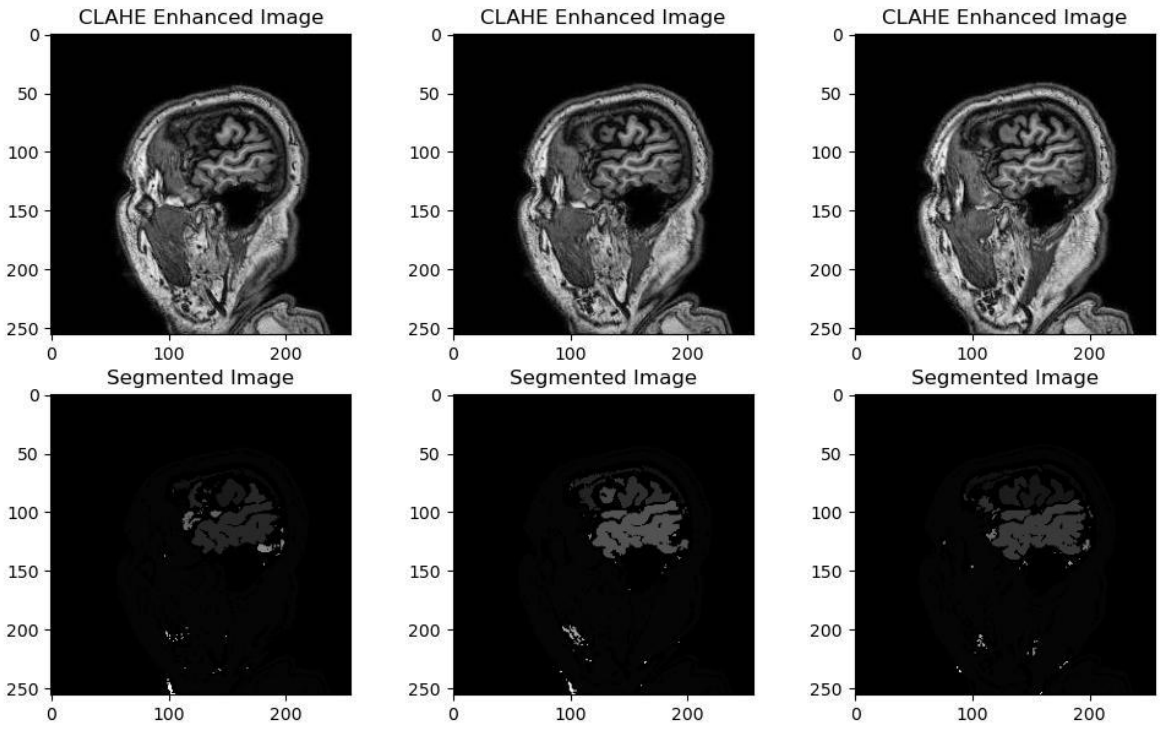


Figura 3.5 Ejemplo de tres imágenes antes y después de la segmentación utilizando el crecimiento de regiones. Se aprecia que este método permite la eliminación de información innecesaria para el estudio y dejando únicamente la zona del cerebro de cada imagen.

3.1.2 Extracción de características

3.1.2.1 Patrón Local Binario (LBP)

El proceso de realce de contraste y el crecimiento de regiones mencionados anteriormente, se realizó para la obtener un arreglo de imágenes que se utilizaron para la extracción de características en las que se basara el modelo de aprendizaje automático. Otro de los datos fundamentales que se necesitan para el proceso de extracción seleccionado, es la segmentación de las imágenes de forma que se etiquete cada píxel de la imagen tratada anteriormente y se aprecie únicamente la zona de interés bajo los criterios de contraste establecidos. Este proceso es comúnmente utilizado en imágenes médicas debido a que los diferentes tejidos que

se presentan en las MRI pueden dificultar el proceso de análisis dependiendo del estudio [21].

En la Figura 3.6 se muestra un ejemplo aleatorio del resultado de uso de este algoritmo. Se muestra tanto el histograma, la imagen original y la imagen LBP que muestra los patrones de binarización obtenidos de la imagen. Los parámetros establecidos para el algoritmo fueron: un vecindario *P=8, R=1 y Method='uniform'*

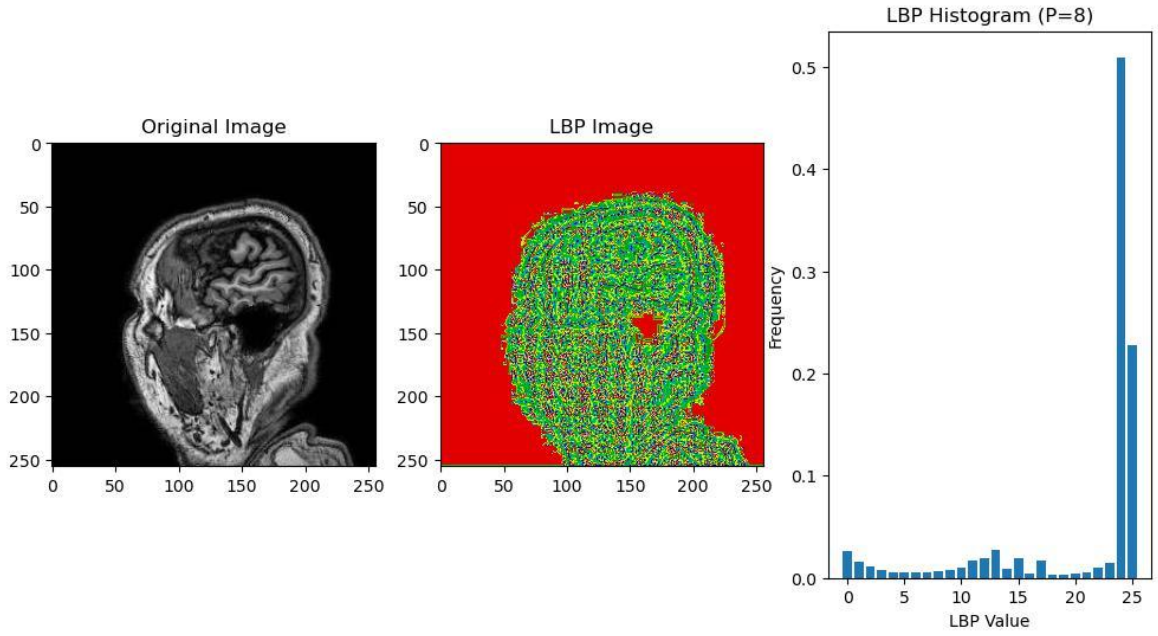


Figura 3.6 A) Imagen original. B) imagen LBP de los patrones obtenidos luego del algoritmo de comparación. C) Histograma normalizado de la frecuencia de los patrones obtenidos en la imagen LBP.

3.1.2.2 Características morfológicas

El proceso de extracción de características de cada una de las imágenes es un paso fundamental en el proceso de desarrollo de una máquina de aprendizaje para la clasificación de niveles de cognición en pacientes. Para este paso, se desarrolló una función que utiliza los algoritmos que hasta ahora se han detallado. La función definida como *Enhanced_morphologial_features* combina descriptores morfológicos, de textura e intensidad, teniendo como parámetros de entrada dos imágenes, la realzada en contraste y la segmentada.

Utilizando las imágenes con realce de contraste, las imágenes segmentadas para resaltar la zona de interés (corteza cerebral), el histograma obtenido del algoritmo LBP y las características morfológicas que se obtienen de las imágenes utilizando funciones como *regionprops, graycomatrix, solidity* entre otros, se construye un directorio con todas las características extraídas.

Para el desarrollo del algoritmo de extracción de características morfológicas, se utilizaron los paquetes *skimage.measure*, *skimage.morphology* y *skimage.feature*; las cuales brindan las herramientas y funciones que se utilizaron para esta parte del proceso. El uso de estas librerías simplificó el proceso de cálculo y obtención de información para la etapa de entrenamiento de las máquinas de aprendizaje.

Con base en la imagen segmentada, se utilizó la función *label* para asignar etiquetas a cada objeto conectado en la imagen binaria segmentada. Este nuevo etiquetado se realizó con la finalidad de analizar cada región de interés de manera individual.

Las características de textura se obtuvieron haciendo uso de la función *graycoprops()* del paquete de *features*. Se utilizó como parámetros de entrada una matriz de coocurrencia de niveles de gris (GLCM, por sus siglas en inglés), especificando el cálculo de la media de las características de interés.

```
contrast = feature.graycoprops(glcm, 'contrast').mean()
dissimilarity = feature.graycoprops(glcm, 'dissimilarity').mean()
homogeneity = feature.graycoprops(glcm, 'homogeneity').mean()
energy = feature.graycoprops(glcm, 'energy').mean()
correlation = feature.graycoprops(glcm, 'correlation').mean()
```

Figura 3.7 Fragmento de código donde se obtienen algunas características morfológicas

En las imágenes de resonancia magnética, la GLCM es utilizada ampliamente para la extracción de características texturales que brinden información sobre la composición de la imagen. Esta matriz se basa en la formación de los pixeles que contiene diferentes valores de intensidad y como están distribuidos en la imagen, midiendo la frecuencia con la que ciertos pares de valores de intensidad ocurren a una misma distancia y dirección determinadas. En [27] es utilizada para este mismo propósito, obteniendo 23 características.

3.1.2.3 Histograma de Gradientes Orientados (HOG)

Debido a que las imágenes MRI contiene bordes y texturas de gran importancia, como lo podrían ser el cerebro, masas, estructuras anatómicas anormales, entre otras, se optó por la aplicación de este algoritmo con la finalidad de que describa el posible deterioro de la materia gris en las diferentes clases o niveles de cognición. En la Figura 3.8 se muestra una imagen de la representación de gradientes que arroja el uso del algoritmo HOG, se utilizaron como parámetros celdas de 8x8 píxeles, bloques de 4 celdas (2x2) y 9 orientaciones en las que se divide el gradiente.

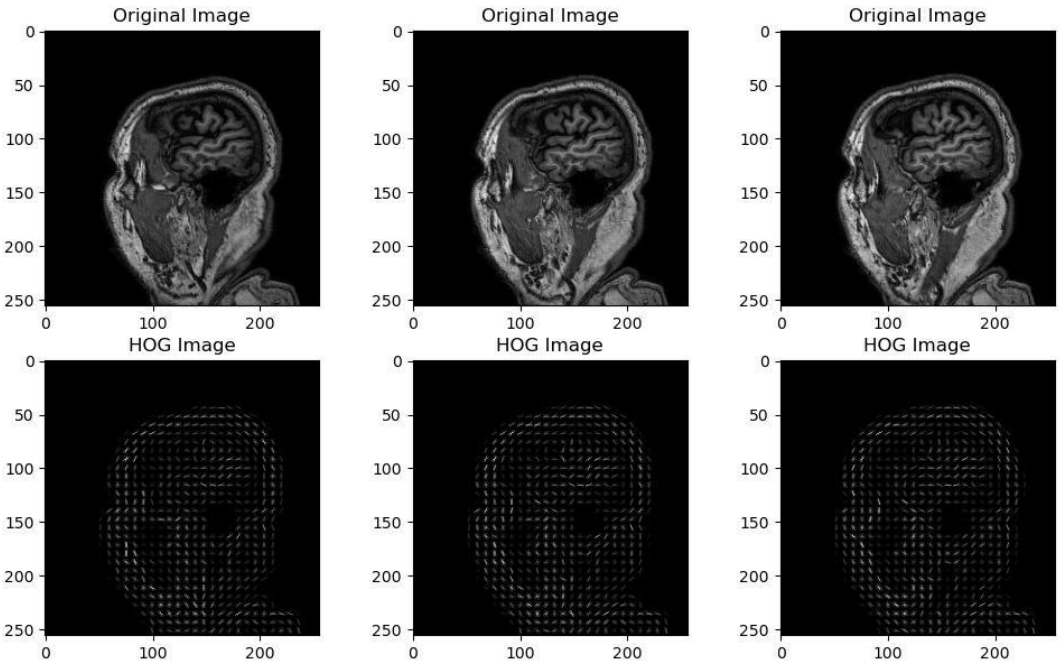


Figura 3.8 Imágenes comparativas entre las imágenes originales y las obtenidas mediante HOG que describen las orientaciones de los gradientes. Se aprecia que las imágenes resultantes enfática los bordes y formas de la imagen, eliminando detalles innecesarios como las texturas suaves.

3.1.3 Desarrollo del modelo de aprendizaje

Para el desarrollo del sistema de clasificación usando el algoritmo Random Forest, el primer pasó que se realizó fue la organización de los datos y las agrupaciones que se utilizaron para el entrenamiento y las pruebas.

Se utilizó la función *LabelEncoder* con la finalidad de transformar las etiquetas del arreglo *labels* a valores numéricos. Inicialmente, estas etiquetas se encontraban en formato de texto debido a que fueron obtenidas directamente de los nombres de las carpetas donde se encontraban las imágenes.

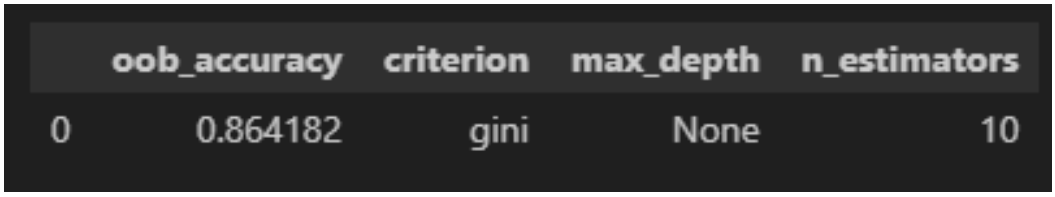
Utilizando la función *train_test_split()* se realizó la división de los conjuntos de entrenamiento y prueba de los datos. La proporción de estos subconjuntos utilizada fue de 85% de los datos serían para entrenamiento, y el 15% restante sería utilizado para las pruebas. Esta función también cuenta con un parámetro *random_state* el cual se fijó en 42 con la finalidad de que el factor de aleatorización de los datos sea siempre igual cada vez que se ejecute el código.

También se realizó una estandarización de los datos, de manera de que los valores quedaran escalados entre 0 y 1, y cada característica tenga una media calculada de 0 y una desviación estándar de 1. Este proceso de estandarización de valores ayuda a la convergencia de los modelos de aprendizaje automático, evitando que las características que poseen mayor rango de medida sean las que dominen el modelo, y el entrenamiento sea uniforme. Se utilizó la función *StandarScaler()* que utiliza los mismos valores de media y desviación estándar en los datos de test, que fueron obtenidos de los datos de entrenamiento.

Como último paso, antes de realizar el entrenamiento del modelo se realizó un balance de los datos de entrenamiento. El balanceo se realizó con la función SMOTE (Synthetic Minority Oversampling Technique) para compensar el desbalance de las clases. Debido a que cada clase tiene un número diferente de imágenes, y para evitar que alguna clase que tenga más imágenes domine el entrenamiento, se utilizó la función SMOTE para generar nuevos datos sintéticos para las clases con menor tamaño. Este balanceo se aplicó al conjunto de características estandarizado y a las etiquetas, del conjunto de entrenamiento.

Luego de aplicar las transformaciones y funciones *LabelEncoder(), StandarScaler() y SMOTE*, se utilizó la propiedad *".shape"* para obtener el tamaño del conjunto de entrenamiento escalado y compararlo con el tamaño del arreglo del mismo conjunto, y así evidenciar que el tamaño del conjunto balanceado es mayor al original.

Para el proceso de entrenamiento de este algoritmo se utilizó *RandomForestClassifier()* que se importa desde *sklearn.ensemble*, donde se definieron los hiperparámetros en que se basara el modelo clasificación. En primera instancia se realizó un entrenamiento de prueba con los hiperparámetros por defecto que tiene la función; sin embargo, la respuesta de ese modelo no fue tan pobre como se esperaba obteniendo una precisión de 86.41% con los parámetros mostrados a continuación en la Figura 3.9 Hiperparámetros que tiene por defecto el modelo RandomForestClassifier() y su prueba de precisión.



	oob_accuracy	criterion	max_depth	n_estimators
0	0.864182	gini	None	10

Figura 3.9 Hiperparámetros que tiene por defecto el modelo RandomForestClassifier() y su prueba de precisión.

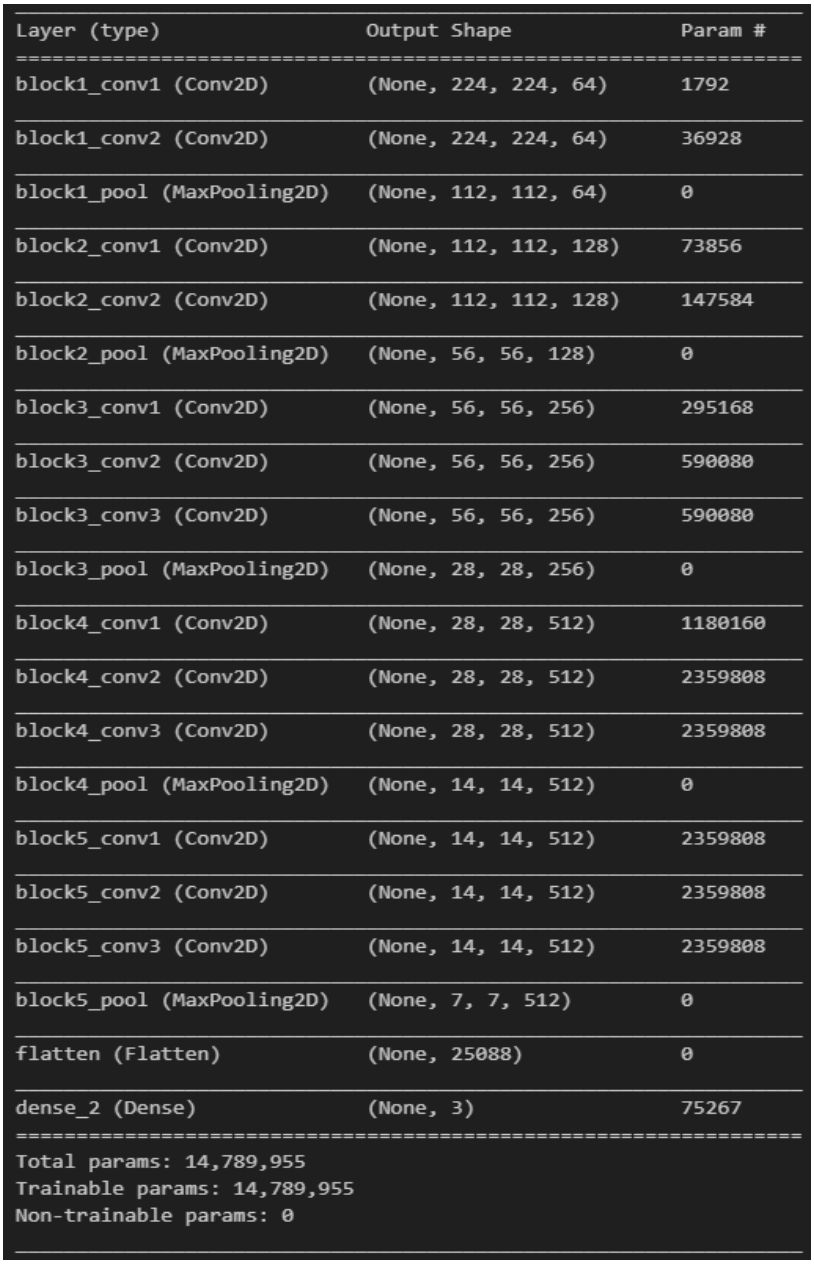
Aunque esta sección no está diseñada para plasmar los resultados de las máquinas de aprendizaje, se decidió colocar la primera prueba del este primer modelo con la finalidad de establecer las expectativas y entender que, los resultados de precisión y tiempos de ejecución varían e incluso pueden mejorar mucho con la variación de algunos hiperparámetros. Los resultados a lo largo del ajuste de la máquina de aprendizaje se mencionan en la sección 4.2

Estas variaciones y la selección final del mejor modelo obtenido, tras la variación de hiperparámetros, se desarrolló en el Capítulo 4.

3.2 Modelo Red Neuronal (VGG16)

Para el desarrollo del sistema de clasificación utilizando redes neuronales convolucionales, se debe de tener en cuenta que la complejidad de la red ayuda significativamente la clasificación de los niveles de cognición entre las imágenes; sin embargo, también va directamente relacionado con el costo computacional. Debido a lo anterior, y en busca de los mejores resultados, se decide usar Transfer Learning seleccionando una red neuronal pre-entrenada que se ha utilizado en diversos estudios de clasificación de imágenes MRI como también en la segmentación de estas ([5], [28], [29], [30], [31]).

Esta red neuronal VGG16, como se mencionó en la sección 2.2.5, cuenta con 13 capas convolucionales y 3 capas densas (fully connected layer). Se tuvo en cuenta el tamaño de entrada de las imágenes MRI para ajustarse a las capas donde se realiza la extracción de características hecha por las capas convolucionales. También se agregó una última capa para reducir las 25088 características de salida del modelo VGG16, a una salida de 3 clases correspondiente a los niveles de cognición se desean clasificar.



Layer (type)	Output Shape	Param #
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
dense_2 (Dense)	(None, 3)	75267

Total params: 14,789,955
Trainable params: 14,789,955
Non-trainable params: 0

Figura 3.10 Modelo de la Red Neuronal VGG16 entrenada.

Las imágenes de la base de datos usada tienen originalmente una dimensión de (256, 256, 1), la cual no es admitida por la red neuronal seleccionada. Por lo tanto, se hace un ajuste a las imágenes durante la carga de la base de datos para obtener una dimensión de (224, 224, 3) tal como se muestra en la estructura del modelo de la Figura 3.10.

El conjunto de imágenes que se obtiene de la base de datos, con un total de 3922 para el conjunto de entrenamiento, y un conjunto de test que consta de 840 imágenes. Además, se realiza una reorganización aleatoria de los datos con la finalidad de evitar desbalanceo en las características a clasificar durante el entrenamiento.

Previo al entrenamiento del modelo se realiza una normalización de los valores de los pixeles de las imágenes para obtener valores entre 0 y 1, lo cual es necesario para que el modelo pueda realizar una correcta clasificación. Para obtener la normalización deseada, se hizo la división entre 255 de cada píxel de los arreglos, tanto de entrenamiento como de test.

La sección 4.3 muestra las diferentes variaciones de la arquitectura de la red VGG-16 en búsqueda del mejor desempeño posible. Finalmente se encuentra que un modelo en cascada donde cada etapa utiliza un fine tunning con 4 capas entrenables presenta la mayor precisión en la clasificación de las tres clases CN, AD y LMCI.

3.3 Discusión

Es vital importancia el preprocesamiento de las imágenes previo a la etapa de clasificación, en especial en sistemas que necesitan como parámetro de entrada características extraídas manualmente como el modelo Random Forest. En este caso se realizó una extensa selección de características morfológicas, vectoriales (HOG), binarias (LPB), segmentación de zonas y realce de contraste con la finalidad de obtener la mayor cantidad de información de las imágenes posible. Cabe resaltar que para el modelo VGG-16 no fue necesario la extracción de características por medio de algoritmos debido a que la red neuronal las aprende por sí misma.

El modelo Random Forest descrito previamente, supone una buena alternativa como método de clasificación para este trabajo. Aunque la selección de características deba ser manual y, se deban utilizar algoritmos propios para la extracción de información de interés, su simplicidad de computación es un punto positivo para el acercamiento de este tipo de proyectos a la comunidad general.

La extracción de características se realizó utilizando métodos de segmentación, filtración y realce de contraste de las imágenes médicas. Además, se utilizaron descriptores texturales y métodos relativamente complejos como HOG, LBP y OTSU.

Para la red neuronal convolucional VGG-16 facilita un poco el proceso de preparación de las imágenes debido a que no se hace necesario la extracción de características manualmente, siendo la misma red neuronal la encargada identificarlas y darles su peso correspondiente. Se aprecia que esta red cuenta con 14’789.955 parámetros, de los cuales es posible modificar el total de parámetros entrenables con la finalidad de mejorar la especificidad del modelo.

Capítulo 4 Pruebas y Resultados

En este capítulo se describe todo el proceso de pruebas, variaciones a los modelos de clasificación, y finalmente validación del mejor modelo obtenido. Se presenta la selección de la base de datos y el preprocesamiento realizado utilizando filtrado Gaussiano y segmentación de zonas para obtener una base de datos lo más limpia posible.

Para cada uno de los modelos de clasificación se realizaron diversas variaciones de los hiperparámetros con la finalidad de obtener el mayor desempeño posible. Se realizaron pruebas de diferentes métodos de clasificación (cascada, *fine tunning*, *DataAugmentation*) que podrían mejorar la precisión de estos y finalmente se presenta una discusión en base a todos los métodos utilizados y los resultados obtenidos por cada uno.

4.1 Descripción base de datos

En este capítulo se describe la base de datos que utilizó para el entrenamiento y evaluación de los clasificadores seleccionados. También se describe el preprocesamiento aplicado a la base de datos para la posterior extracción de

características. Se aplicaron técnicas de segmentación de tejido, filtrado gaussiano y suavizado de imagen.

4.1.1 Selección base de datos MRI

Los archivos que se encuentran en la base de datos ADNI se pueden clasificar por los diferentes niveles de cognición, sin embargo, hay que tener en cuenta que es posible que exista un mismo sujeto de estudio en dos o más clases de la ADNI, pero en diferente momento de la toma de imágenes. Esto se da debido a que en la plataforma de ADNI se almacena la información progresivamente para estudios que podrían estar adelantando casos de estudio a lo largo del tiempo en un mismo paciente.

La base de datos que se utiliza para el entrenamiento y prueba de los modelos de aprendizaje se obtiene después de una limpieza manual de las imágenes, luego del preprocesamiento, debido a que los archivos originales se obtienen en formato DCM que incluyen diferente número de cortes (slices) para cada sujeto y/o toma del examen.

En primera instancia se realiza la descarga de los archivos en el formato DCM, desde la plataforma de ADNI. Luego se procede a la organización por clases o niveles de cognición, de forma que se obtienen tres conjuntos de imágenes que contempla 14 sujetos por cada conjunto. Estas imágenes pasan por proceso de ajuste y preprocesamiento con la finalidad de limpiar las imágenes y dejarlas en la mejor condición posible para la clasificación, y finalmente se guardan en formato jpg con la finalidad de visualizarlas fácilmente y realizar un filtro manual donde se eliminan las que no aporten información y/o afecten la clasificación en instancias posteriores. La Figura 4.1 muestra un conjunto de cinco imágenes aleatorias de la base de datos previos al preprocesamiento.

Figura 4.1 Imagen de muestra que contiene 5 imágenes aleatorias de la base de datos de MRI.

La limpieza manual de algunas imágenes de hace necesaria debido a que los archivos en formato DCM se tomaron en diferentes momentos y por diferentes máquinas y especialistas, culminando en una variedad de imágenes sin información correspondientes a secciones donde no se ve la materia gris, sino que reflejan cortes antes y después de la corteza cerebral de interés. En la Figura 4.2 se muestra cinco imágenes aleatorias de la base obtenida de la ADNI luego de la etapa de preprocesamiento y limpieza manual.

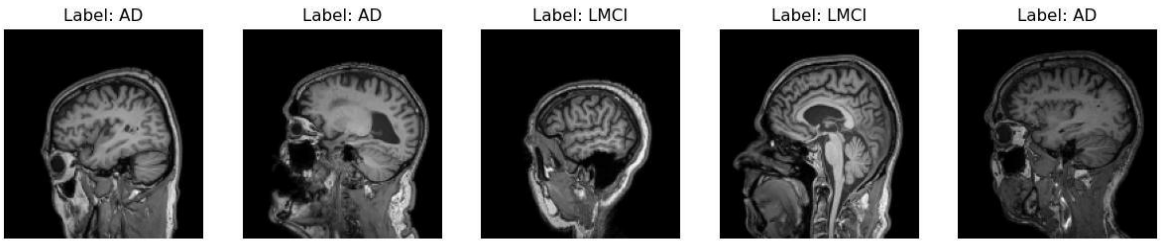


Figura 4.2 Imagen de muestra que contiene 5 imágenes aleatorias de la base de datos de la MRI luego del preprocesamiento y la limpieza manual. Con esta muestra aleatoria se evidencia que la base de datos ya no cuenta con imágenes sin información o "en negro"

Posterior al preprocesamiento de las imágenes MRI y a la limpieza manual que se realizó y se explicó anteriormente, se obtiene una base de datos para el entrenamiento con un total de 3922 imágenes, pertenecientes a 42 sujetos en total, divididas en tres conjuntos o clases. Además, también realizó el mismo procedimiento para la base de datos de test, obteniendo una base de datos de 840 imágenes en total, compuesto por 3 sujetos por clase. Cabe resaltar, que cada una de esas imágenes corresponde a un "slice", o partición, de cada MRI perteneciente a cada sujeto. El número de particiones de cada MRI varía con base en los parámetros utilizados en cada toma, variando entre 160 ~ 170 particiones por imagen. La Tabla 4.1 Tabla resumen base de datos utilizada para el entrenamiento y test de las máquinas de aprendizaje automático. (AD: Enfermedad de Alzheimer, CN: cognición normal, LMCI: afectación leve, F: femenino, M: masculino). La columna de genero indica el número de sujetos asociados a su respectivo sexo. muestra información importante sobre la base de datos utilizada.

ENTRENAMIENTO

CLASES	# SUJETOS	EDAD PROM	GENERO
AD	14	76,4	F: 5 M: 9
CN	14	78,13	F: 9 M: 5
LMCI	14	70,53	F: 7 M: 7
TEST			
AD	3	80	F: 2 M: 1
CN	3	78,33	F: 2 M: 1
LMCI	3	70	F: 2 M: 1

Tabla 4.1 Tabla resumen base de datos utilizada para el entrenamiento y test de las máquinas de aprendizaje automático. (AD: Enfermedad de Alzheimer, CN: cognición normal, LMCI: afectación leve, F: femenino, M: masculino). La columna de genero indica el número de sujetos asociados a su respectivo sexo.

4.1.2 Preprocesamiento de la base de datos

Para la remoción del ruido de la imagen se utilizó el módulo *morphology* importada desde la librería *skimage*, utilizando el método de *dilation* que se utiliza para rellenar con más pixeles en los bordes con la finalidad de poder diferenciar de forma más clara los cambios de la imagen.

Se utiliza también un ajuste en el tamaño de la imagen con la función *add_pad.* Teniendo en cuenta que todas las imágenes obtenidas no necesariamente provienen del mismo estudio, tiempo exacto o incluso especialista, estas imágenes podrían presentar diferencias en su tamaño. Se establece un estándar de 256x256 pixeles.

Se utilizaron dos algoritmos para este proceso *Z-score* y *Fuzzy C-Means*, el último método es un poco más complejo debido a que tiene unos procesos anteriores de detección de máscaras, y se importa desde *Scikit-fuzzy*. En [32] se evaluó la respuesta de la aplicación de este algoritmo a imágenes médicas DICOM para la segmentación de imágenes médicas digitales para la reconstrucción de modelos anatómicos 3D.

En la siguiente imagen se aprecia el resultado de la aplicación del algoritmo Fuzzy C-means y la extracción de la materia gris del cerebro obtenida la segmentación

anterior. Se busca particularmente la materia gris debido a que el deterioro cognitivo está relacionado con la pérdida de masa de la corteza cerebral.

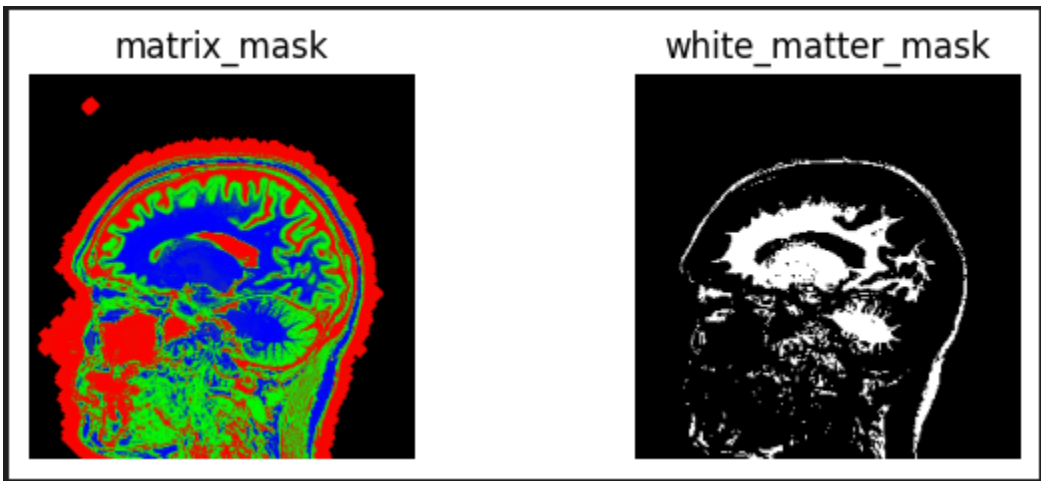


Figura 4.3 A) Imagen obtenida después de la segmentación utilizando el algoritmo Fuzzy C-means. B) Imagen de la materia gris del cerebro en escala de grises.

Se utiliza una normalización de la imagen con respecto a la intensidad de los pixeles. Este proceso se recomienda debido a que los valores del arreglo de la imagen, relacionados con la intensidad de esta, varía en función de la máquina y sus parámetros utilizados en el momento de la toma de la imagen, permitiendo la evaluación o comparación con otras imágenes, pueden ser de otros estudios u otros pacientes. Este proceso busca asegurar que las intensidades relativas de la imagen original se ajusten a las de la imagen de la materia gris, utilizando su media como referencia.

$$Imagen\ normalizada = \frac{intensidad\ imagen\ original}{intensidad\ promedio\ de\ la\ materia\ gris}$$

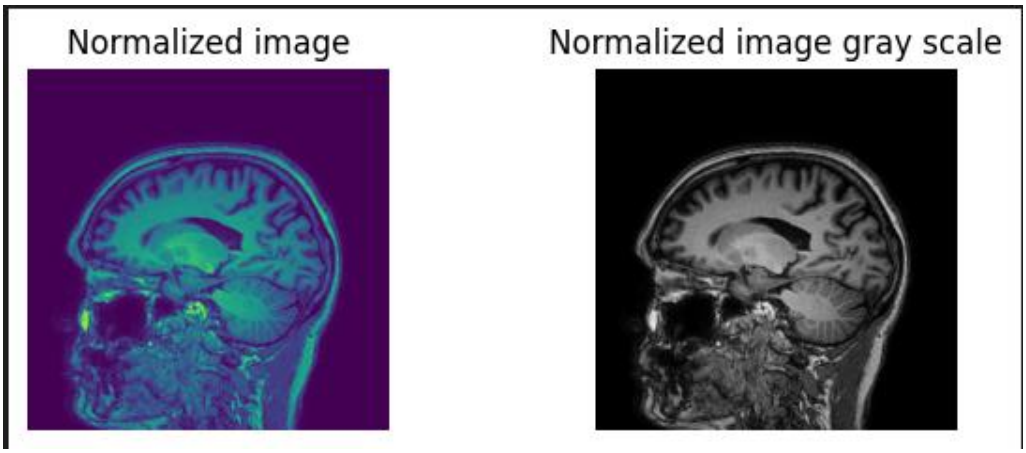


Figura 4.4 A) Imagen normalizada en función de la materia gris segmentada. B) Imagen normalizada en función de la materia gris segmentada en escala de grises.

Finalizado ese primer preprocesamiento de las imágenes y teniendo las imágenes en formato JPG, se procede a la visualización y posterior limpieza manual de las imágenes de baja calidad que no aportan información e incluso entorpecerían la fase de entrenamiento y la robustez de los clasificadores.

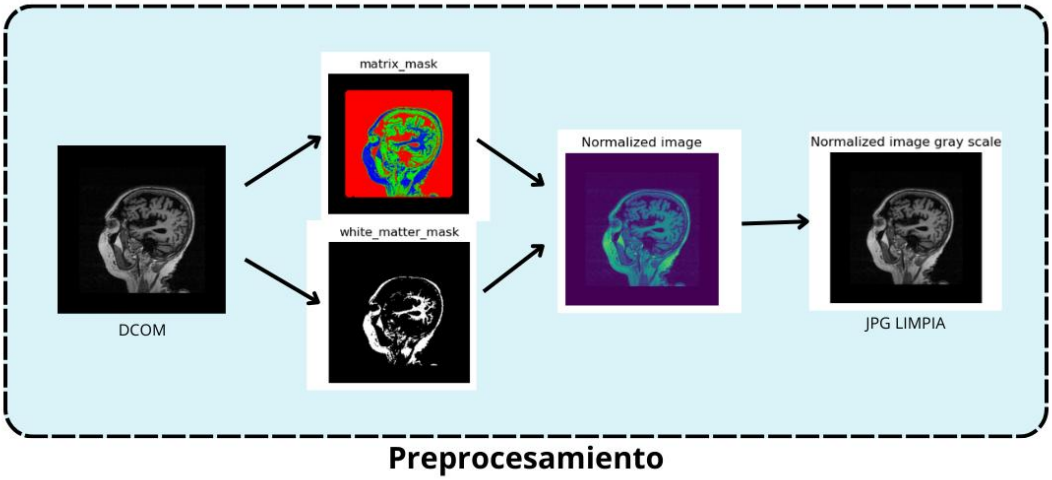


Figura 4.5 Diagrama resumido de los diferentes pasos del preprocesamiento que se hicieron las imágenes médicas originales hasta obtener la imagen en formato JPG final.

4.2 Random Forest

Como se mencionó en la sección 3.1.3, la selección de los hiperparámetros es fundamental para establecer la arquitectura del modelo clasificador, así como para obtener el mejor resultado posible.

La primera ejecución del código de entrenamiento del modelo fue realizada en la sección 3.1.3 con los parámetros mostrados en la Figura 3.9 Hiperparámetros que tiene por defecto el modelo RandomForestClassifier() y su prueba de precisión. Dicha tabla muestra los valores *n_estimators* y *max_depth* fueron establecidos en *"None"* dejando los valores en 10 y automático respectivamente. Ambos hiperparámetros están relacionado con los árboles de decisión, siendo el primero el número de árboles que tendrá el modelo y el segundo, la máxima profundidad de los árboles. El valor automático de *max_depth* es determinado por el mismo algoritmo, que detiene el árbol de decisión cuando los valores de las muestras disminuyan hasta el mínimo.

Teniendo en cuenta que este clasificador toma la decisión basándose en el voto por mayoría de los resultados obtenidos por cada árbol de decisión, se procedió con una evaluación en busca del mejor valor del hiperparámetro *n_estimators* que arrojara el mejor resultado. También se varió el hiperparametro *criterion* con "*log_loss"* y *"Gini" (*valor por defecto).

Para el ajuste del modelo con la variación de los hiperparámetros se utilizó la función *ParameterGrid()* que permite establecer arreglos de parámetros de diferentes valores, y que el modelo *RandomForestClassifier()* reconoce e itera por cada uno de esos valores, realizando un entrenamiento para cada combinación de parámetros. La información, tanto de los parámetros como de los resultados de precisión fueron organizados en orden descendente en un *DataFrame* de Pandas.

	oob_accuracy	criterion	max_depth	n_estimators
4	0.993497	log_loss	None	350
5	0.993497	log_loss	None	600
3	0.993015	log_loss	None	200
11	0.993015	gini	None	600
10	0.991811	gini	None	350

Tabla 4.2 Tabla obtenida del entrenamiento del modelo Random Forest donde se muestran los mejores cinco resultados de la evaluación por medio de la malla de parámetros, ordenados de manera descendente. La primera columna es el número de la iteración que se hizo.

En la Tabla 4.2 Se aprecia que el valor de precisión es muy similar utilizando ambos tipos de criterios. Sin embargo, como se mencionaba en [33], "*Gini"* al ser un algoritmo simple no requiere un alto costo computacional.

Con base en los resultados obtenidos por la prueba anterior, se puede evidenciar que el número de árboles decisión definitivamente afecta la precisión del clasificador, así mismo como el tiempo de ejecución. En la Tabla 4.2 se aprecia que el algoritmo de selección "*log_loss"* necesita menos árboles de decisión para obtener un mejor resultado; sin embargo, al tener un alto costo computacional, y al obtener precisiones similares al que usó "*Gini"* se plantea la reducción de las opciones para el entrenamiento del modelo.

Se decide utilizar un modelo con los siguientes criterios: *criterion="Gini" y "log_loss", n_estimators=600 y 350;* pero esta vez el parámetro *max_depth* no se ajustará en automático, sino que se hará un barrido de valores de profundidad entre 1 y 20 para cada uno de esos modelos con la finalidad de visibilizar los tiempos de ejecución y precisión.

Las Figura 4.6 y la Figura 4.7 muestran la evolución de la precisión del modelo de Random Forest durante la variación de profundidad de los árboles de decisión (*max_depth)*. La Figura 4.6 es la comparación de la precisión del modelo Con base en la profundidad de los árboles de decisión de la red utilizando únicamente los datos de entrenamiento, mientras que la Figura 4.7 es la evolución de la precisión con los datos de prueba. La finalidad de este proceso fue la validación y selección

de parámetros ideales para la máquina RF, variando la profundidad, el criterio de selección y el número de árboles de decisión.

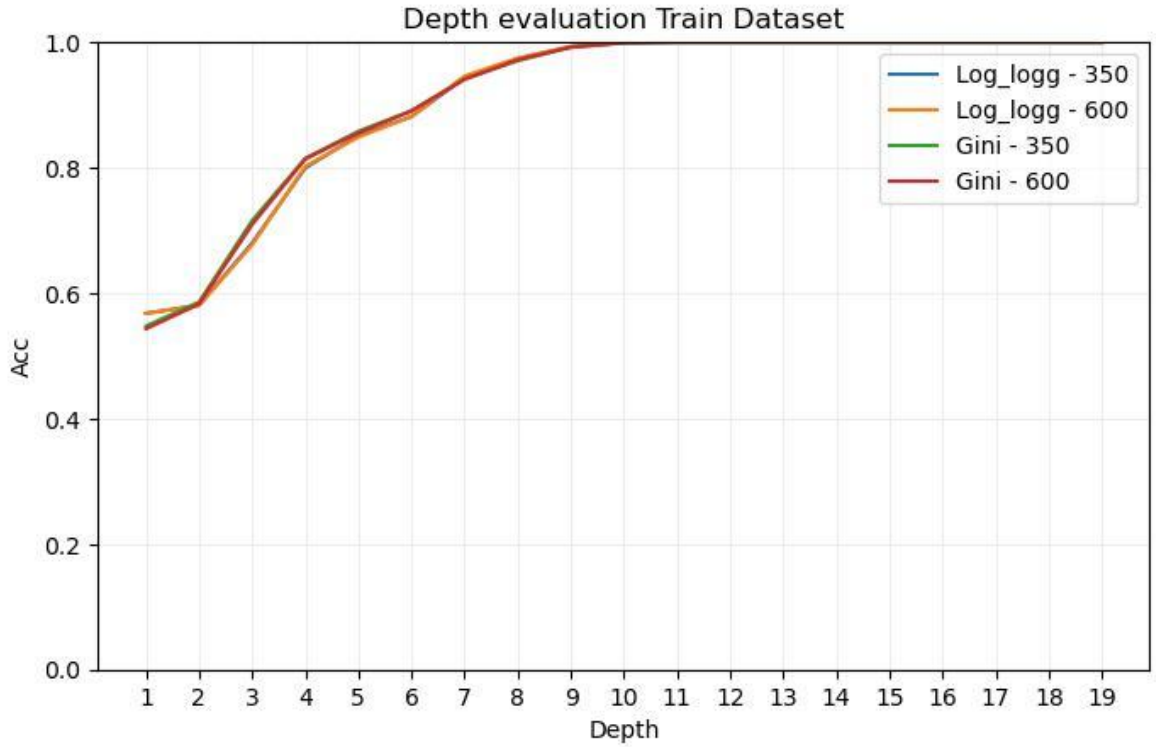


Figura 4.6 Evaluación del nivel de precisión del modelo RF Con base en la variación de la profundidad de los árboles de decisión utilizando los datos de entrenamiento.

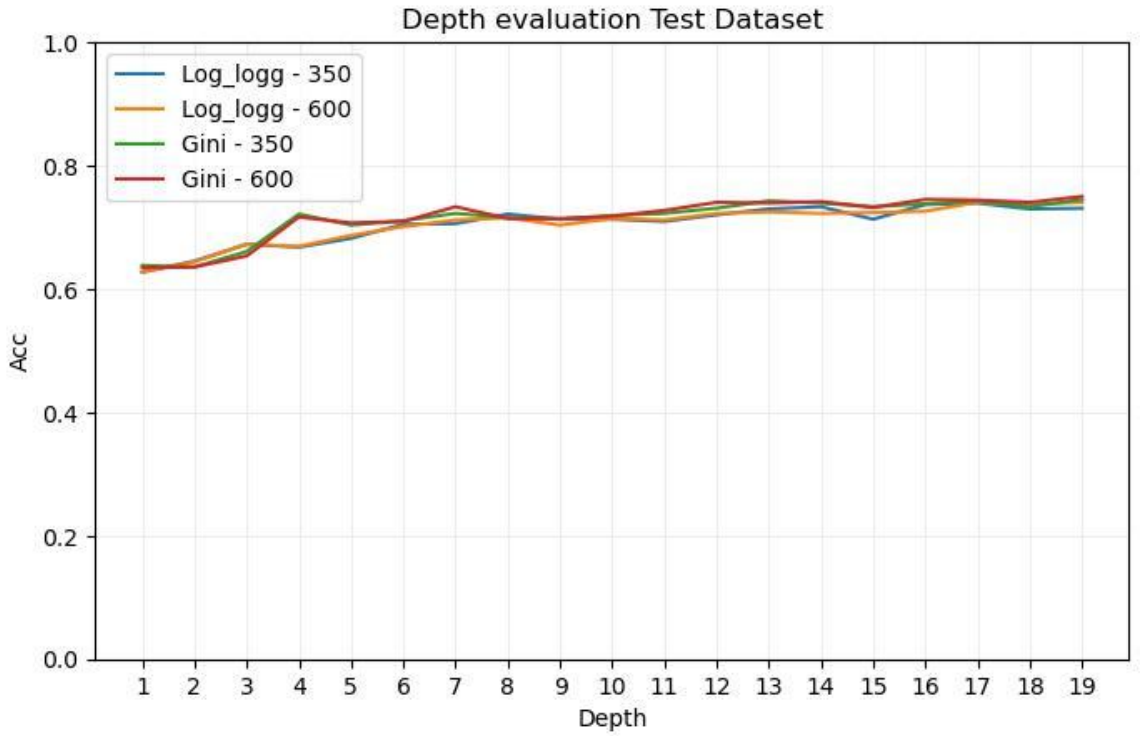


Figura 4.7 Evaluación del nivel de precisión del modelo RF Con base en la variación de la profundidad de los árboles de decisión utilizando los datos de test.

Ambos gráficos comparativos muestran la evolución de la precisión del modelo con respecto a la profundidad de los árboles del modelo RF tanto para los valores de entrenamiento como los de test. Visualmente, el comportamiento de los datos de entrenamiento es bastante similar a los de test, entre ellos tampoco se presenta una gran variación en los resultados. Los cuatro modelos de RF presentan valores similares entre los criterios *log_loss* y *Gini*, ambos evaluados con 350 y 600 árboles; razón por la cual la decisión final se tomó con base en la simplicidad computacional.

Además del número de los árboles de decisión, los gráficos anteriores ayudaron a determinar la profundidad de dichos árboles (*max_depth*), donde claramente se aprecia, para los datos de entrenamiento, que la variación entre la precisión del modelo a partir de un valor de 10 ya se acerca bastante a 100% con variaciones pequeñas. La decisión en este caso también se tomó con base en la simplicidad computacional y a los tiempos de ejecución, tomando como modelo final con los hiperparámetros consignados en la Tabla 4.3.

Random Forest	
Hiperparámetros	**Valor**
N_estimators	600
Max_depth	12
Criterion	Gini
Ranodm State	42
N_jobs	-1

Tabla 4.3 Muestra los hiperparámetros que se utilizaron para el entrenamiento del modelo final de RF.

Con base en el modelo de aprendizaje entrenado con los hiperparámetros consignados en la Tabla 4.3 Muestra los hiperparámetros que se utilizaron para el entrenamiento del modelo final de RF., se obtiene un gráfico de barras que muestra el peso de las características que se utilizaron para el entrenamiento del RF. Debido al alto número de características extraídas de las imágenes, además de la forma en la que se organizaron (un único vector por imagen), se realiza una limitación y en la Figura 4.9 únicamente se muestran las 35 características con más peso o importancia del modelo de clasificación RF.

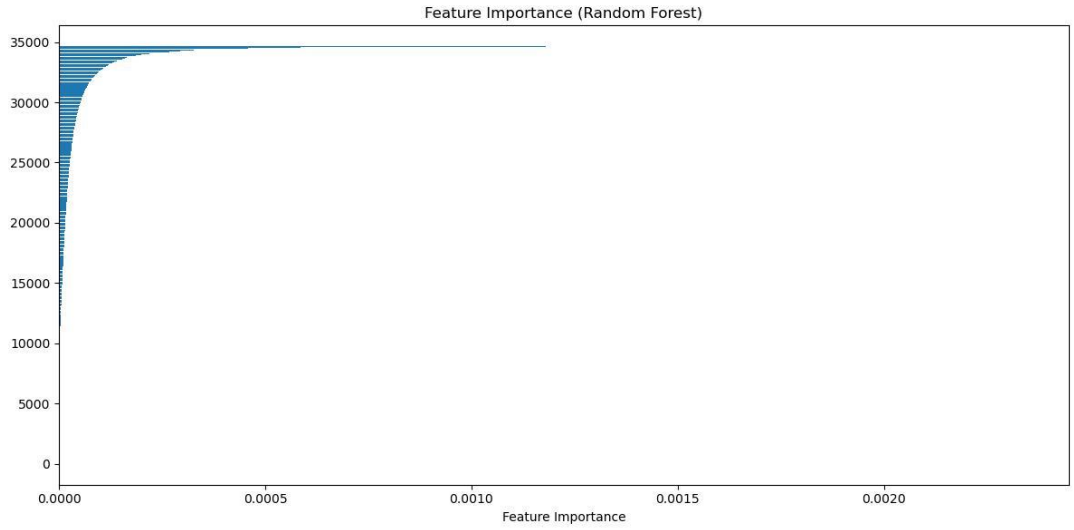


Figura 4.8 Gráfico de barras sobre las características más importantes del modelo de clasificación RF.

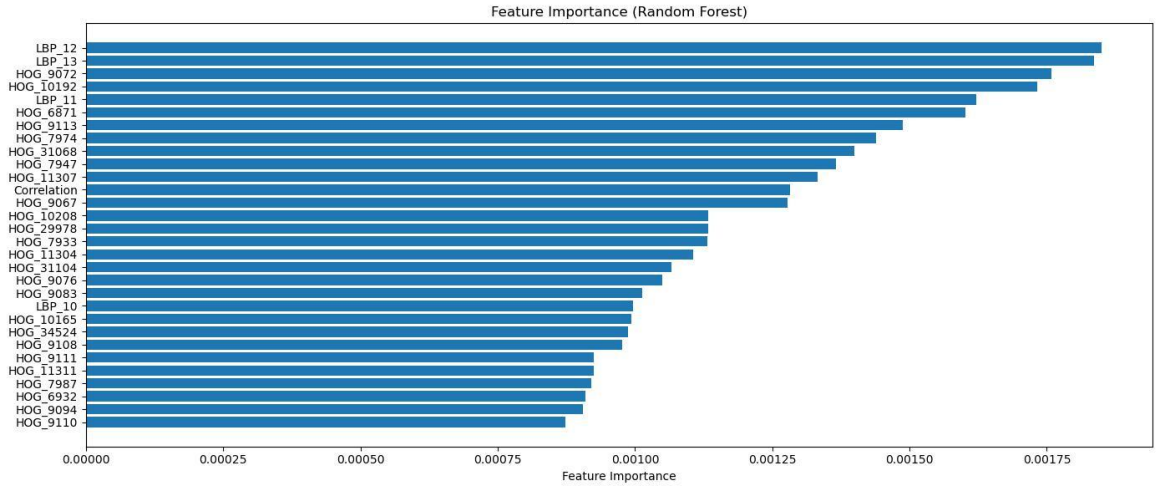


Figura 4.9 Gráfico de barra con las 35 características más importantes del modelo de clasificación RF.

Con la finalidad de mostrar los resultados del modelo de aprendizaje RF, se genera la matriz de confusión (Figura 4.10), que indica un comportamiento bastante bueno para la clasificación de pacientes con Alzheimer; sin embargo, se aprecia la gran dificultad que posee el modelo para la clasificación entre LMCI y CN.

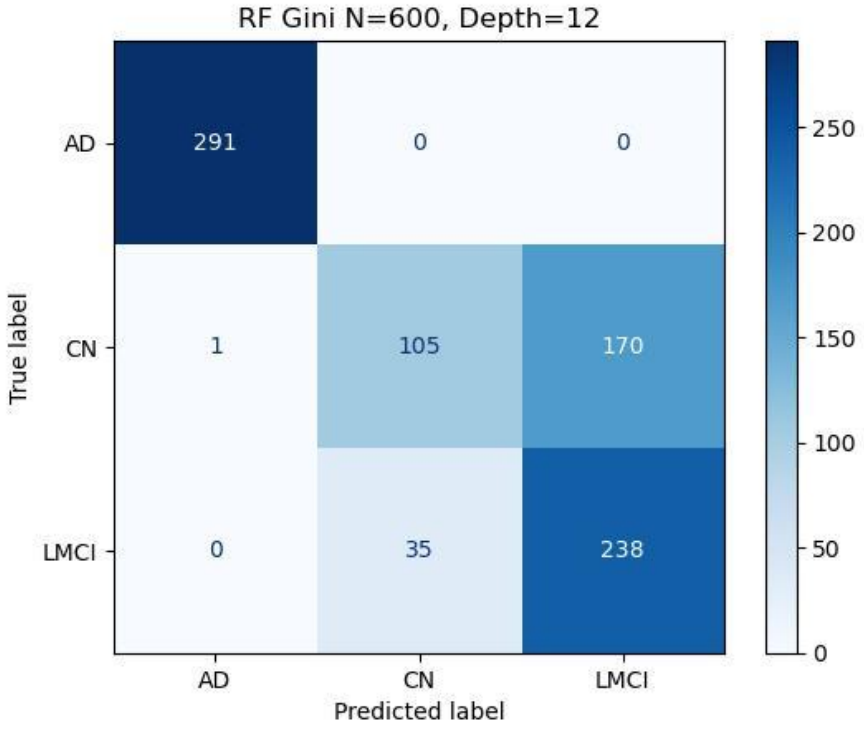


Figura 4.10 Matriz de confusión del modelo RF.

En la Figura 4.9, se evidencia que hay diversas características de las que fueron extraídas en la sección 3.1.2 permitiendo analizar cual tipo de característica podría tener un peso fundamental. En primera instancia se aprecia que el HOG presenta un claro dominio dentro de las características de más peso del modelo.

Debido a que el proceso de extracción de características utilizado inicialmente posee un gran costo computacional y un tiempo de ejecución relativamente largo (72m y 34.6s), se planteó la posibilidad de un modelo con un número limitado de características que permitieran reducir el costo computacional sin perder la precisión del clasificador.

Con base en lo anterior, se realizó el entrenamiento de un modelo RF, con los mismos hiperparámetros de la Tabla 4.3 pero esta vez utilizando únicamente el vector de características del histograma de gradientes orientados (HOG). El resultado de esta prueba fue regular, obteniendo una precisión del 74,1% y una reducción de tiempo de ejecución de más de 10 veces en el proceso de extracción de características, demorándose únicamente 5m 36s. La Figura 4.11 muestra las características más importantes dentro del vector de HOG.

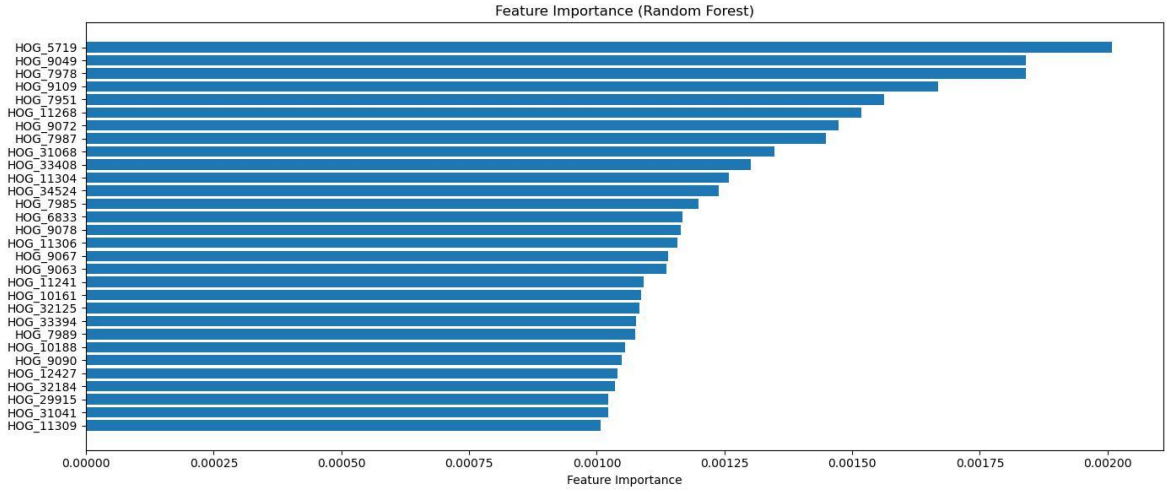


Figura 4.11 Gráfico de barras con las 35 características más importantes del modelo de clasificación.

Para mostrar el comportamiento del modelo de clasificación RF, se muestra en la Figura 4.12 la matriz de confusión que evidencia una clara deficiencia para la identificación entre los niveles de cognición con afectación leve y condición normal;

sin embargo, el modelo presenta una alta precisión en cuanto a la detección de Alzheimer.

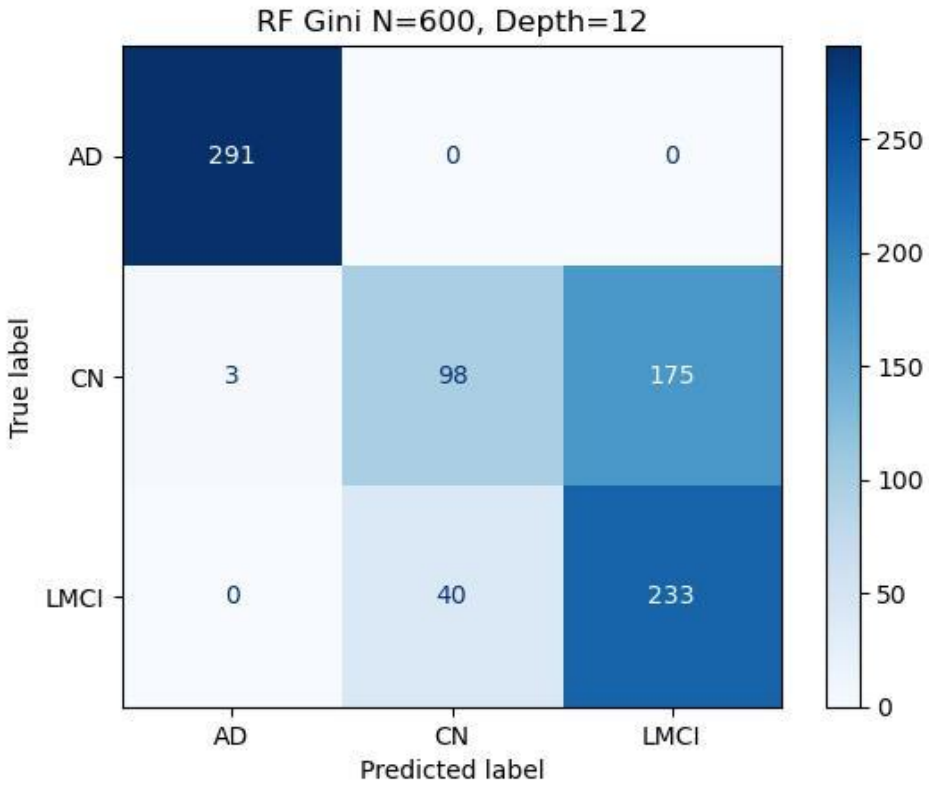


Figura 4.12 Matriz de confusión del modelo RF, únicamente con las características HOG

Con la finalidad de mejorar el desempeño del modelo Random Forest, se planteó modificar el sistema de clasificación y utilizar un modelo en cascada. Este tipo de modelos se utiliza cuando dos o más clases presenten claras dificultades para la identificación entre ellas. En este caso se aprecia la clara dificultad para clasificar entre CN y LMCI, mientras que la enfermedad de Alzheimer presenta una alta precisión de identificación.

Inicialmente se planteó realizar una categorización de “super clase” que agrupara las clases CN y LMCI debido a que la clase AD era las precisa. Se estableció un modelo de clasificación que identificara entre dos clases AD y Not_AD. Este enfoque se toma con base a que se esperaría que dos clasificadores binarios robustos, fueran mejores de un solo clasificador de tres clases.

Para evitar el desbalanceo de las clases, se tomó toda la base de datos clasificada con AD, y solo el 50% de la base de datos de cada una de las clases CN y LMCI. El proceso de clasificación del modelo completo se basa en evaluar un primer

modelo para la identificación entre AD y Not_AD, que dependiendo del resultado establece la clase o de lo contrario, pasa a la evaluación con el siguiente modelo que termina de diferenciar entre CN y LMCI, y de esta manera obtener una clasificación más precisa del modelo de identificación de las tres clases de cognición.

La primera etapa del modelo, encargada de la clasificación entre AD vs Not_AD, tuvo una precisión de 34% y de un 56% para la segunda etapa (CN vs LMCI). Este modelo de clasificación obtuvo un desempeño global de 35%.

Debido al bajo desempeño del modelo planteado, se considera volver a aplicar el método de clasificar en cascada; sin embargo, esta vez se definieron los grupos de clasificación como: CN vs Not_CN y AD vs LMCI. Esta decisión se tomó debido a que es coherente pensar que las clases AD y LMCI podrían diferenciarse más de la clase CN. La primera etapa de este nuevo modelo arroja una precisión del 77% (CN vs Not_CN) y del 98% para las clases AD vs LMCI, obteniendo un desempeño global de 77%.

4.3 VGG-16

Inicialmente el entrenamiento se realizó por defecto, utilizando únicamente la CPU con un tamaño de batch de 32 y utilizando 30 épocas correspondientes a la repetición del entrenamiento. Así mismo aclarar, que en la construcción del modelo se estableció todas las capas entrenables Con base en la información de la iteración anterior, permitiendo una mayor velocidad de convergencia y de aprendizaje de la red neuronal.

A este primer modelo de la red neuronal VGG-16 se le realizó una interrupción abrupta debido al largo tiempo de la fase de entrenamiento. Se estimaba una duración de alrededor de seis horas. Para solventar el inconveniente, se creó un entorno de ejecución que utilizara la GPU del computador para realizar el procesamiento de la información. Se utilizó la versión 2.0 de Tensorflow y una GPU GTX 1660 super 6gb.

Este primer modelo, tuvo un resultado no tan satisfactorio de 53,09% de precisión y una pérdida de 4,8396 con los datos de test. En las Figura 4.13 y Figura 4.14 se aprecia el comportamiento de los valores de precisión y pérdida de los datos de test durante el entrenamiento.

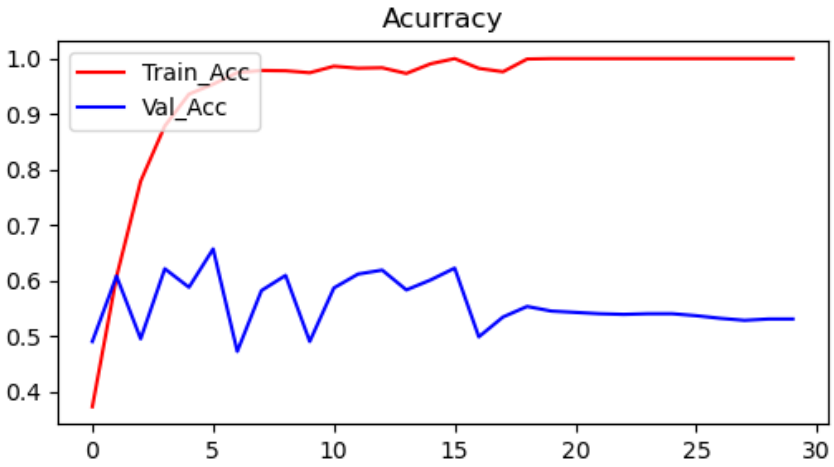


Figura 4.13 Gráfica lineal del comportamiento de la precisión durante el entrenamiento del primer modelo de red neuronal VGG16.

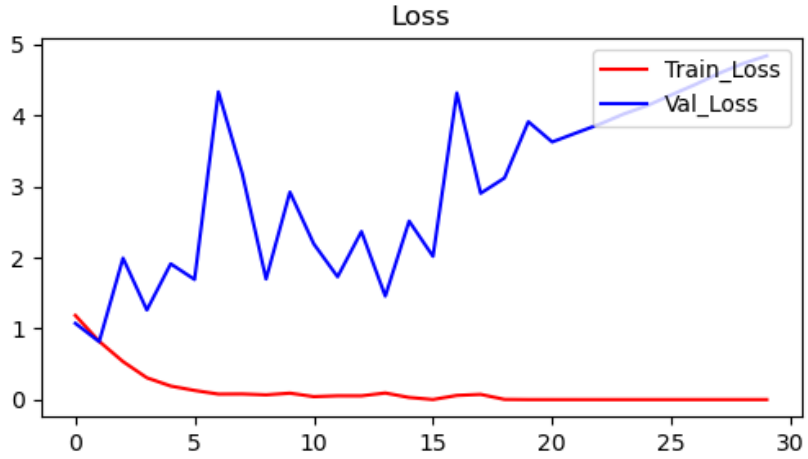


Figura 4.14 Gráfica lineal del comportamiento de la pérdida durante el entrenamiento del primer modelo de red neuronal VGG16.

Con la finalidad de mejorar el rendimiento de la red neuronal, se propone añadir una capa de dropout con un valor de 0.5 justo después de la salida del modelo pre entrenado VGG16. Además de eso, se desactiva la característica de “entrenable” de gran parte de la red VGG16, utilizando únicamente los valores base de la red VGG-16, y dejando 1539 parámetros entrenables. Este modelo arrojó un valor de precisión de 50,23% y una pérdida de 1,0394. La Figura 4.15 muestra el

comportamiento de los valores de precisión y pérdida de este segundo modelo a lo largo de las épocas de entrenamiento.

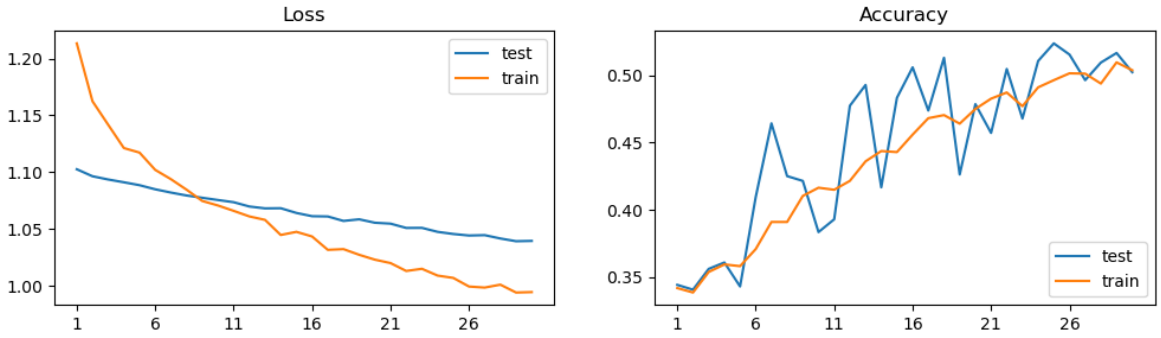


Figura 4.15 Gráfica lineal del comportamiento de la precisión y pérdida durante el entrenamiento del segundo modelo de red neuronal con los datos de test.

Los entrenamientos previos se realizaron utilizando todas las capas y sus valores pre entrenados del modelo base de la red neuronal VGG16, si bien es cierto que se agregó una capa de dropout para mejorar el rendimiento, los hiperparámetros de las capas pueden no ser las ideales para este preciso caso de estudio. Buscando una mejoría en el rendimiento y la precisión de esta red neuronal, se realizaron variaciones en las capas del modelo.

Se realizaron diversas pruebas con la variación del número de capas entrenables para buscar más especificidad del caso y el número de épocas ideales de la red. Las variaciones previamente mencionadas se muestran en la Tabla 4.4 de manera resumida. La primera prueba utilizando método “fine tuning” se dejó únicamente activas las últimas cuatro capas de la red neuronal activas, obteniendo 7’080,963 parámetros entrenables; una gran diferencia con la red anterior. En la Figura 4.16 se aprecia el comportamiento de los valores de pérdida y precisión de los datos de test, obteniendo un valor de precisión final de 65,96% y de pérdida de 1,2523.

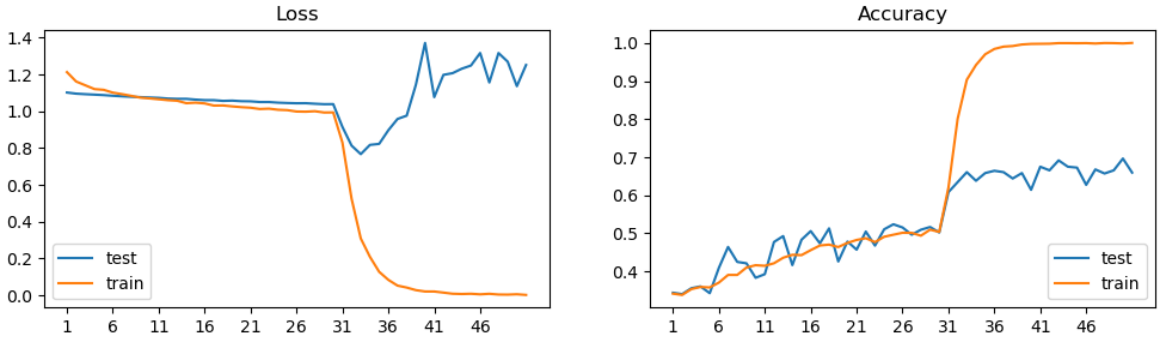


Figura 4.16 Gráfica lineal del comportamiento de los valores de precisión y pérdida durante la fase de entrenamiento y test luego de dejar activas únicamente cuatro capas de la red pre entrenada.

1	PARAMS	VALUES	VAL_ACC	VAL_LOSS
	train_params	14789955	0,5309	4,8396
	nontr_params	0		
	epochs	30		
	layers	19		
2		VALUES	VAL_ACC	VAL_LOSS
	train_params	1539	0,5023	1,0394
	nontr_params	14714688		
	epochs	30		
	layers	0		
3		VALUES	VAL_ACC	VAL_LOSS
	train_params	7080963	0,6595	1,2523
	nontr_params	7635264		
	epochs	50		
	layers	4		
4		VALUES	VAL_ACC	VAL_LOSS
	train_params	7080963	0,6797	1,5934
	nontr_params	7635264		
	epochs	40		
	layers	4		
5		VALUES	VAL_ACC	VAL_LOSS
	train_params	4721155	0,6154	1,4883
	nontr_params	9995072		
	epochs	50		
	layers	3		
6		VALUES	VAL_ACC	VAL_LOSS
	train_params	9440771	0,6047	1,8717

	nontr_params	5275456		
	epochs	30		
	layers	5		
7		VALUES	VAL_ACC	VAL_LOSS
	train_params	7080963	0,6773	1,2105
	nontr_params	7635264		
	epochs	20		
	layers	4		

Tabla 4.4 Tabla resumen donde se muestran las diferentes variaciones que se realizaron a la etapa de fine tuning de la red neuronal VGG-16 para el entrenamiento. Se aprecia las variaciones de capas entrenables, épocas y las características que son entrenables.

Cabe resaltar que el entrenamiento del modelo luego de apagar casi todas las capas del modelo VGG16 se realizó en continuación al segundo modelo probado. Se aprecia en la Figura 4.16 que las primeras 30 épocas es la misma gráfica de la Figura 4.15 y a partir de ahí, se continuó con otras 20 iteraciones del entrenamiento con el ajuste de las capas mencionado; dando un total de 50 épocas o iteraciones de entrenamiento que se realizó en este nuevo modelo.

Debido a que en las gráficas anteriores del comportamiento de precisión y pérdida se aprecia que hubo una clara mejora en el rendimiento del modelo luego del ajuste del número de capas entrenables, se plantea un nuevo entrenamiento con el mismo número de capas, pero esta vez con menos iteraciones. El punto de realizar menos iteraciones en el modelo se basa en que la mejoría de la precisión no es significativa luego de ciertas épocas.

En la Figura 4.17 se muestra las gráficas de los valores de precisión y pérdida del modelo, con un valor de épocas de 40 y dejando las mismas cuatro capas entrenables del modelo base, además con la claridad de que se hará un entrenamiento desde el inicio en estas condiciones. Se obtiene un valor de precisión de 67,97% y una pérdida de 1,5934.

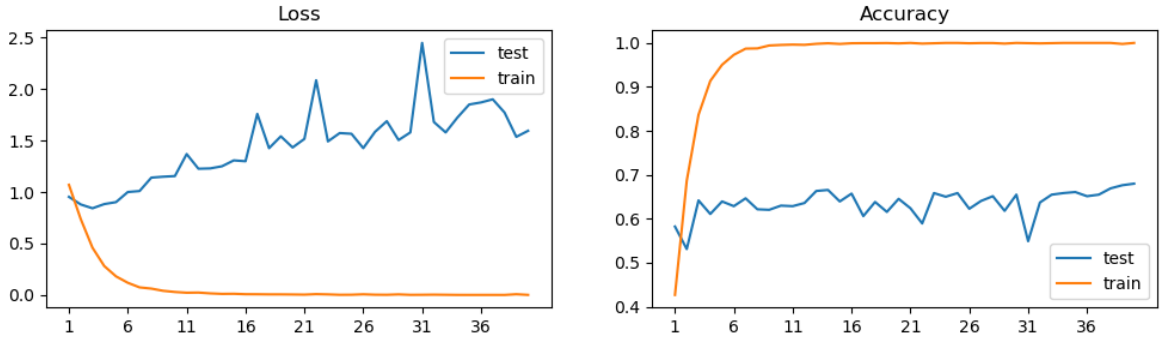


Figura 4.17 Gráfica lineal de los valores de precisión y pérdida durante el entrenamiento del modelo con cuatro capas entrenables y 40 iteraciones.

En las gráficas anteriores, se puede apreciar que a medida que se va realizando el entrenamiento, los valores de las pérdidas del modelo van aumentando con algunos picos de subida y bajada. Con base en las observaciones anteriores, se plantea la reducción de las épocas de entrenamiento con la finalidad de que permita reducir ese valor de error sin sacrificar mucho los valores de precisión en la clasificación de los niveles de cognición.

Teniendo en cuenta lo anterior, se realizó una última variación al modelo de aprendizaje con la reducción de las iteraciones a 20 épocas. Con este ajuste, el modelo finaliza con una precisión de 67,73% y una pérdida de 1,2105. Se aprecia que la variación de la precisión obtuvo una diferencia de 0,24% comparado con el modelo anterior; sin embargo, la reducción del valor de pérdida es significativa con una diferencia de 0,3829.

La Figura 4.18 muestra la matriz de confusión de este modelo final, donde se identifica una clara dificultad del clasificador para la correcta identificación entre las imágenes que representan una afectación cognitiva leve y las de control o cognición normal.

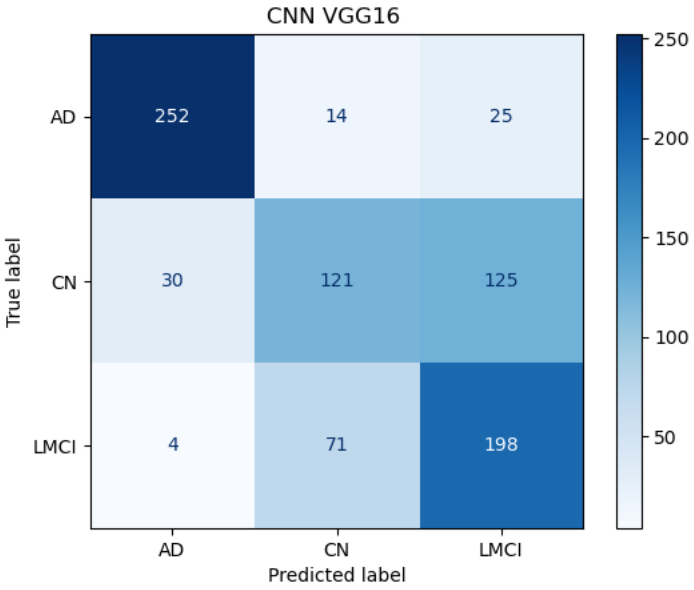


Figura 4.18 Matriz de confusión del modelo de red neuronal basado en VGG16 aplicando una técnica de fine tuning permitiendo la especificidad del entrenamiento con las ultimas 4 capas del modelo base.

Buscando la manera de mejorar el rendimiento y la precisión de la red neuronal, se planteó el uso de *Data Augmentation* para usar la variabilidad de diferentes parámetros de cada imagen durante en el entrenamiento. La finalidad de implementar esta técnica fue buscar mayor robustez en el modelo de clasificación VGG-16.

Para aplicar la técnica "*Data Augmentation"* se utilizó la función *ImageDataGenerator()* de la librería Keras. Esta función permite la variación en rotación, ancho y largo de la imagen, zoom entre otras características. Se seleccionaron estas variables debido a que la naturaleza de las MRI no es aconsejable la variación de intensidades y/o de colores de las imágenes, al igual que las variaciones deben ser moderadas. Esta función solo se utilizó con los datos de entrenamiento. Se realizaron cuatro ejecuciones de entrenamiento, cada una variando los diferentes parámetros mencionados anteriormente.

Inicialmente se establecieron los parámetros modificando únicamente su ubicación espacial; se establecieron valores de rotación de máximo 40°, acercamiento aleatorio de máximo 10%, variación en el ancho y altura de máximo 20%. Con la configuración mencionada anteriormente, se obtuvo una precisión de 47,61% para los datos de test.

Debido a que la precisión desmejoró, se optó por realizar cambios aún más limitados y además añadir un filtro gaussiano para agregar ruido a la imagen y permita una mayor variabilidad. Para la segunda ejecución del modelo se estableció una variación de rotación de máximo 5°, variación de ancho y largo de máximo 5% y se mantiene el acercamiento de 10%. Se obtuvo una precisión de 43.21% para los datos de test.

Finalmente, se decidió realizar el entrenamiento de la red neuronal únicamente con la variabilidad de ruido en las imágenes, obteniendo un desempeño de 45.83% usando un filtro Gaussiano de 5x5, y de 49,64% con el mismo filtro, pero con dimensión 3x3.

Los resultados que se obtuvieron con este nuevo modelo no fue lo esperado, reduciendo su precisión se manera significativa para los datos de test. Los demás parámetros fueron los mismos establecidos en la Tabla 4.4. Teniendo en cuenta que ya se tiene establecido el número de épocas ideal para el entrenamiento, se plantea la variación del tamaño del batch. Se realizaron evaluaciones del modelo VGG-16 variando entre 16, 32 y 64. Estos resultados se muestra en la Tabla 4.5 donde se consignaron las precisiones de las diferentes fases del entrenamiento del modelo con la finalidad de apreciar las mejores y/o desmejoras.

Batch size = 16			**Batch size = 32**			**Batch size = 64**		
VGG-16	**fine tuning**	**Augmentation**	**VGG-16**	**fine tuning**	**Augmentation**	**VGG-16**	**fine tuning**	**Augmentation**
AD-CN-LMCI								
33,33%	61,79%	47,26%	38,33%	65,95%	46,67%	24,64%	67,85%	46,67%

Tabla 4.5 Tabla resumen de los valores precisión obtenidos a lo largo de las diferentes etapas de entrenamiento

La Figura 4.19 muestra las gráficas de precisión y pérdida durante el entrenamiento. Cada gráfica muestra tanto los datos de entrenamiento como los de test. Se aprecia una mejora constante con los datos de entrenamiento, tanto en la disminución de la pérdida como en el aumento de la precisión. Por otro lado, los resultados con los datos de test presentan una fluctuación con una convergencia mínima.

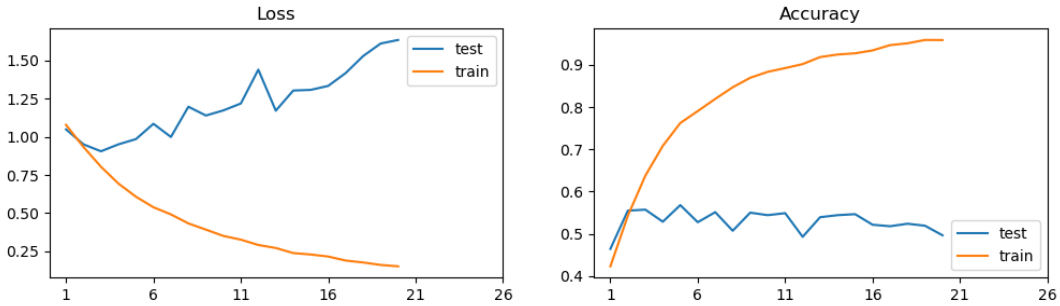


Figura 4.19 Evolución de los valores de pérdida y precisión durante el proceso de entrenamiento para los datos de entrenamiento y test.

La Figura 4.20 muestra la matriz de confusión del modelo donde se aprecia una gran complicación del modelo para distinguir las clases de CN y LMCI. El modelo de clasificación identifica la mayoría de los datos como si tuviera la enfermedad de Alzheimer.

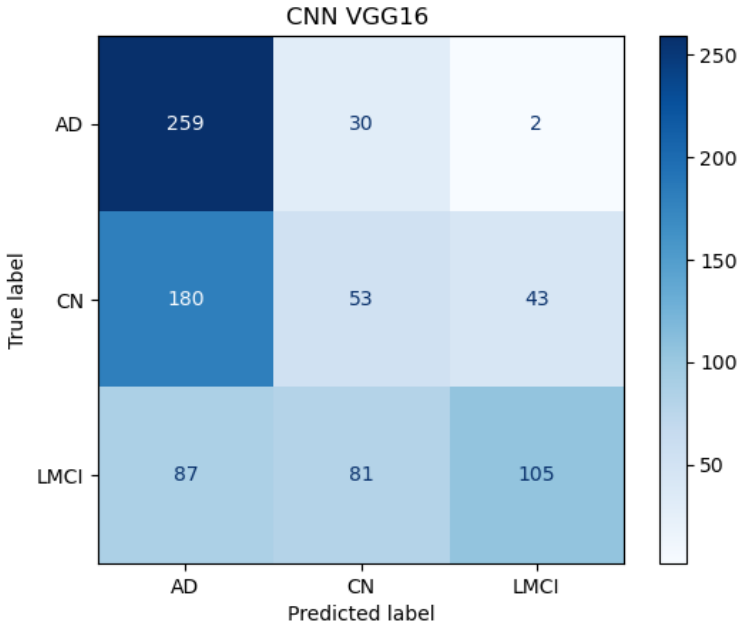


Figura 4.20 Matriz de confusión del modelo de clasificación VGG-16 aplicando Data Augmentation a los datos de entrenamiento.

Con base a los resultados mostrados previamente, y a la Tabla 4.5 se plantea aplicar el mismo modelo en cascada desarrollado en la sección 4.2 para el

clasificador Random Forest. Inicialmente se realizó la agrupación de las clases CN y LMCI, ambos conjuntos con el 50% de las imágenes con la finalidad de no afectar el balanceo de la base de datos. La etapa uno del modelo en cascada consta de la clasificación entre las clases AD y Not_AD, obteniendo un desempeño del 92% de precisión, mientras que para le etapa dos se obtuvo una precisión de 56% (CN vs LMCI). Para el desempeño final de este clasificador en cascada, se obtuvo una precisión final del 64% mostrando claras dificultades la distinción entre CN y LMCI.

Al igual que con el modelo Random Forest, se modificó la agrupación de las clases para el clasificador en cascada para obtener un mejor resultado final. Para la etapa uno se decidió realizar la clasificación de las clases CN y Not_CN. Para esta primera etapa se obtuvo un desempeño del 72%, mientras que para la segunda etapa se tuvo una precisión del 99% (AD vs LMCI), suponiendo una mejora en ambos clasificadores. Finalmente, al momento de evaluar la lógica de decisión del modelo en cascada, se obtuvo una precisión final del 71% al diferenciar entre los tres niveles de cognición evaluados.

4.4 Discusión

Tras realizarse diversos ajustes de hiperpárametros del modelo Random Forest, finalmente se establecieron las consignadas en la Tabla 4.3. Esta configuración permitió la mayor precisión del modelo con un 75,47%. Inicialmente se realizó el entrenamiento con diversas características morfológicas, segmentación de imágenes, descriptores de gradientes (sección 3.1); sin embargo, al variar las características extraídas se observó que utilizando únicamente el histograma del gradiente orientado se obtuvo el menor costo computacional y una precisión del 74%, pudiendo concluir que es la característica de mayor peso.

El relativo bajo desempeño del clasificador Random Forest podría deberse a las etapas de preprocesamiento de imágenes realizadas en este trabajo, debido a que se utilizaran técnicas relativamente sencillas (Filtrado gaussiano, método OTSU, CLAHE). Además, se debe tener en cuenta la complejidad del sistema clasificación planteado donde las imágenes de resonancia magnética presentan diferencias sutiles entre las clases de interés AD, CN y LMCI, que junto con la variación de imágenes provenientes de diferentes equipos de medición dificultan la definición de las características de cada clase.

Finalmente, el modelo en cascada que se planteó para la distinción entre las clases CN vs Not_CN y AD vs LMCI parece ser una solución interesante para este tipo de

sistemas donde existe una clara confusión o una gran similitud entre dos clases en particular (CN vs LMCI), mejorando un 4% la precisión final del sistema, alcanzando un desempeño global del 77%.

Cabe resaltar que se utilizó validación cruzada utilizando un conjunto de datos para el entrenamiento, y uno completamente diferente para el test. Estos conjuntos cuentan con 3922 imágenes para el entrenamiento y 840 imágenes para el test. El modelo presenta una clara dificultad para la clasificación de las clases CN y LMCI, con una precisión de 38% y 70% para verdaderos positivos respectivamente; mientras que la clase AD obtuvo un 100% para verdaderos positivos.

La máquina de Random Forest obtuvo un menor rendimiento que en [2] (85,4%) que clasifica los diferentes tejidos del cerebro a partir de MRI; sin embargo, se debe tener en cuenta que en ese trabajo se clasificó entre tejido blando, hueso y aire presente en el cerebro, mientras que en este trabajo se buscaba identificar diferencias entre el mismo tejido blando del cerebro para buscar anomalías.

Para la red neuronal VGG-16 se realizaron diferentes variaciones en los hiperparámetros para encontrar la arquitectura que permitiera la mayor precisión posible. Se hizo variación entre las capas de entrenamiento, épocas y características entrenables del modelo, y se consignaron en la Tabla 4.4. También se realizaron variaciones del tamaño del batch con las diferentes técnicas aplicadas *fine tunning* y *Data augmentation*; estas quedaron consignadas en Tabla 4.5.

Con base a los resultados obtenidos, donde se aprecia una clara dificultad para la identificación de las clases CN y LMCI, con una precisión de 50% y 65% respectivamente; se realizó la implementación de un clasificador en cascada que permitiera una clasificación binaria esperando mayor robustez del modelo. Inicialmente, se realizó la clasificación de las clases de la siguiente manera: etapa 1, AD vs Not_AD, y etapa 2 clasifica CN vs LMCI, obteniendo una precisión del 35% para el modelo Random Forest y del 64% para la red VGG-16.

También se decidió variar la agrupación de las clases se la siguiente manera: etapa 1 clasifica entre CN vs Not_CN, y etapa 2 clasifica AD vs LMCI. Finalmente, con estos cambios se obtuvo una precisión final del modelo de clasificación en cascada del 77% utilizando la arquitectura Random Forest, y del 71% utilizando la red neuronal VGG-16.

Al igual que el clasificador Random Forest, la red neuronal VGG-16 podría estar siendo afectada por la etapa de preprocesamiento de las imágenes, donde los

métodos de suavizado y reducción de ruido utilizados, podrían estar eliminando información importante de cada imagen. Teniendo en cuenta que la red VGG-16 tiene como parámetro de entrada las imágenes MRI, y la misma red se encarga de la identificación de características y asignación de pesos; el tamaño de la base de datos impacta el rendimiento al no tener la suficiente variabilidad anatómica entre las clases, en particular las clases CN y LMCI que presentan cambios leves que son difíciles de distinguir por el modelo.

Comparando los resultados obtenidos al probar los clasificadores con los datos de test, se observa que el modelo del clasificador Random Forest presenta una mayor precisión, alcanzando un 74% con respecto al 67,85% de la red neuronal VGG-16. Teniendo en cuenta que el modelo Random Forest tiene un método de aprendizaje más simple que la red convolucional, se obtiene un mayor desempeño que puede estar relacionado a que la red VGG-16 necesita una mayor cantidad de datos para el entrenamiento, factor que es afectado en este trabajo debido al reducido tamaño de la base de datos utilizada comparado con la literatura investigada.

En cuanto a la red neuronal convolucional VGG-16, que se obtuvo una precisión de 67,85% presentó un menor desempeño comparado a [6] con 97,12%, [14] con 99,31% y [4] con 93,48% que utilizan el mismo modelo precargado. Cabe resaltar que el tamaño de la base de datos imágenes utilizado en este trabajo es significativamente menor a los utilizado en la literatura mencionada anteriormente, lo cual afecta directamente en el entrenamiento de la red.

Por último, se puede observar que, en ambos modelos de clasificación, se hicieron pruebas de diferentes técnicas para mejorar la robustez, implementando algoritmos que extraen múltiples características en el caso del Random Forest (características morfológicas, LBP, HOG); y técnicas de variabilidad de datos en la red convolucional (*Fine tuning* y *DataAugmentation*). Estas variaciones y pruebas que se realizaron con la finalidad de obtener el mejor desempeño de cada una de las máquinas, que finalmente se redujeron al uso de del histograma del gradiente orientado (HOG) para el Random Forest, y *Fine tuning* para la red VGG-16. Cabe resaltar que los algoritmos desarrollados no están diseñados ni capacitados para brindar resultados altamente confiables en áreas médicas certificadas. Este se debe a que no se realizaron pruebas en entorno médicos reales con personal de la salud, además de la limitante de los algoritmos de preprocesamiento que únicamente utilizan imágenes médicas en formato DICOM.

4.1 Alcances

En este trabajo de grado se desarrollaron dos máquinas de aprendizaje automático para la clasificación de tres niveles de cognición presentes en el proceso de evolución de la enfermedad de Alzheimer (Cognición normal, Afectación leve y enfermedad de Alzheimer) a partir de imágenes de resonancia magnética del plano sagital del cerebro en secuencia T2. Previos a la clasificación, el algoritmo realiza una limpieza de ruido por medio de filtros gaussianos segmentación para el realce de las zonas de interés.

La primera máquina de clasificación utiliza un algoritmo de Random Forest que logra identificar los tres niveles de cognición de interés con una precisión del 74%. Durante las primeras pruebas de validación del sistema se utilizaron tres tipos de características fundamentales que fueron: morfológicas, patrones locales binarios (LBP y el histograma de gradientes orientados (HOG). Esta precisión se obtiene con la reducción de características utilizadas a las obtenidas por medio del algoritmo HOG, que es una representación vectorial de las imágenes, debido a que estas últimas características fueron las de mayor peso en la clasificación.

Por otro lado, la segunda máquina de clasificación es una red neuronal convolucional que implementa un modelo pre entrenado VGG-16, que con los datos de test se obtiene una precisión de 67,73% aplicando técnicas de fine tuning para mejorar el rendimiento. También se aplicó, en conjunto, la técnica *DataAugmentation()* para mejorar la robustez del modelo. Esta última técnica se utilizó con un filtro Gaussiano de 3x3 y una variación máxima de 10 grados en la rotación de las imágenes, y obtuvo un desempeño de 49,64% reduciendo el rendimiento del modelo. Además de la clasificación de las tres clases, se realizó la clasificación utilizando un modelo en cascada (CN vs AD + LMCI y AD vs LMCI) con el fin de obtener un mejor desempeño. Se obtuvo una precisión de 77% para la arquitectura Random Forest, y de 71% utilizando la red VG-16.

Se desarrolló una interfaz gráfica sencilla, con la finalidad de presentar un primer acercamiento a una aplicación que permitiese la evaluación de diferentes máquinas de aprendizaje y la evaluación de las imágenes de un paciente que muestra la predicción del nivel de cognición en que se encuentra éste.

4.2 Limitaciones

La limitación principal del trabajo de grado realizado sería la cantidad de imágenes de resonancia magnética para el entrenamiento del modelo de clasificación. La natural diferencia entre la estructura del cerebro de las diferentes personas representa una gran dificultad al identificar características propias de las diferentes fases de cognición de interés en este trabajo.

El no poder proporcionar la suficiente variabilidad anatómica en las imágenes al modelo de clasificación, este se ve directamente afectado en la incapacidad de discernir correctamente entre las diferentes clases, mostrando especial dificultad con las clases CN y LMCI.

Aunque el modelo basado en Random Forest tuvo un desempeño del 74,57%, existe la posibilidad de mejorar la precisión de clasificación haciendo uso de técnicas más complejas para la extracción de características de las MRI, como las incluidas en el paquete de Python *Py-radiomics.* Este paquete pertenece a un paquete de código abierto de Python que fue diseñado para la extracción de características cuantitativas del cerebro.

Por otro lado, la red neuronal convolucional VGG-16 presenta una gran dificultad en la distinción de las tres clases abordadas. Esta limitación puede darse debido al tamaño reducido de la base de datos obtenida, y a la capacidad de recursos computacionales disponibles. Cabe resaltar que ambos modelos, tanto el Random Forest como la red VGG-16 presentan unas claras dificultades en la distinción entre los niveles de cognición normal y de afectación leve. Además de la clasificación de las tres clases, se realizó la clasificación utilizando un modelo en cascada (CN vs AD + LMCI y AD vs LMCI) con el fin de obtener un mejor desempeño. Se obtuvo una precisión de 77% para la arquitectura Random Forest, y de 71% utilizando la red VGG-16.

La interfaz gráfica desarrollada, aunque muestra una representación gráfica rápida y simple del modelo y su precisión, no podría ser considerada lo suficientemente confiable y robusta para el uso en ámbitos médicos reales debido al nivel de desempeño obtenido en la sección de pruebas.

Una limitante fundamental al usar los modelos de clasificación implementados sería la necesidad de usar imágenes médicas con el formato DICOM, debido a que los exámenes de resonancia magnética para la obtención de las imágenes van a variar en sus parámetros de calibración y características físicas de los equipos médicos;

brindando imágenes en diferentes formatos como NIfTI (Neuroimaging Informatics Technology Initiative).

Capítulo 5 Interfaz de visualización

Para facilitar el uso del sistema de reconocimiento desarrollado, por personas con y sin conocimiento especializado sobre máquinas de aprendizaje, se implementa una interfaz gráfica que permite seleccionar el modelo de clasificación y el conjunto de imágenes MRI del cerebro y generar una estimación del grado de cognición que presente la persona.

La aplicación desarrollada permite la selección entre los dos diferentes modelos de clasificación implementados, Random Forest o VGG-16. La Figura 5.1 muestra la interfaz simple y las diferentes opciones que tiene.

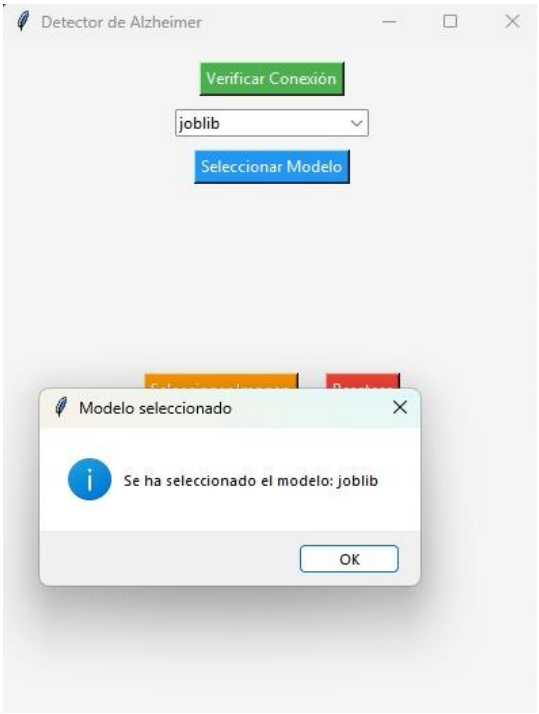


Figura 5.1 Interfaz gráfica simple que muestra la selección del modelo de clasificación.

La interfaz se desarrolló en el entorno virtual de Python para que facilitara la interacción con el mismo lenguaje de las máquinas de clasificación. La aplicación

cuenta con un botón verificar la conexión entre la interfaz y el entorno para asegurarse de la carga de los documentos no va a generar ningún tipo de conflicto.

También la aplicación cuenta con una barra desplegable para la selección del modelo de clasificación que se desea evaluar y finalmente con el botón “seleccionar modelo” se carga el modelo de clasificación. Esta aplicación cuenta con dos opciones, el modelo Random Forest y el modelo VGG-16.

Finalmente, la aplicación cuenta con el botón para la seleccionar la imagen a evaluar y/o el paciente. La interfaz muestra una imagen previa seleccionada y justo al lado se tiene el botón de “evaluar” que se encarga de obtener la predicción del modelo previamente cargado. La Figura 5.2 muestra un ejemplo de este último proceso.

Figura 5.2 Interfaz gráfica que muestra la previsualización de la imagen seleccionada para la evaluación del modelo.

Capítulo 6 Conclusiones

Se desarrolló un sistema de reconocimiento de patrones con dos máquinas de aprendizaje (Random Forest y VGG-16), con una precisión de 74,57% y 67,85% respectivamente, que permiten la clasificación entre tres niveles de cognición cerebral para la detección temprana de la enfermedad de Alzheimer utilizando imágenes de resonancia magnética (MRI). Los niveles de cognición son: Cognición normal (CN), Cognición con alteración leve (LMCI) y enfermedad de Alzheimer. La base de datos utilizada fue obtenida de la ADNI, que debido a su alta acogida en el área científica se convierte en una fuente de datos confiable, dando así cumplimiento al primer objetivo específico.

Para el preprocesamiento de las imágenes de la base de datos seleccionada se realizaron filtrado, suavizado e incluso distinción de máscaras de tejido que permitieron una segmentación de la zona del cerebro, ignorando información extra como lo son el cráneo, los tejidos acuosos y aire. Aunque los resultados obtenidos con esas técnicas fueron decentes, las técnicas que se utilizaron no son las más poderosas ni las más complejas, dejando un amplio margen de mejora en esta fase, además de la gran cantidad de técnicas de preprocesamiento que existen para las imágenes MRI.

La primera máquina de aprendizaje se basa en un modelo de Random Forest la cual necesitó de una extracción de características manuales mediante técnicas de filtrado y segmentación de zonas de interés de las imágenes de resonancia magnéticas del plano sagital. Durante las pruebas de validación del sistema de clasificación se realizó la extracción de características morfológicas utilizando las técnicas *CLAHE* para el realce de contraste, un filtrado gaussiano para suavizar las imágenes y el método *OTSU* para la segmentación del cerebro; también se utilizaron los patrones locales binarios (LBP) y el histograma de gradientes orientados (HOG). Se identificó que las características más relevantes para la detección de Alzheimer se obtienen mediante el histograma de gradientes orientados (HOG) obteniendo una precisión de 74,57% con los datos de test, y permitiendo una reducción de 10 veces en el tiempo de extracción de características, comparado al sistema que contempla las características morfológicas y el LPB.

La segunda máquina de aprendizaje es una red neuronal convolucional que utiliza una arquitectura precargada VGG-16 que, debido a su naturaleza, la etapa de extracción de características no es necesaria y solo hace uso de las imágenes de

la base de datos de entrenamiento. Esta red se ajustó para obtener una precisión en la clasificación de los niveles de cognición de 67,26%. %. En esta red neuronal se aplicaron dos técnicas para buscar la especificidad y mejorar el desempeño del modelo. Las técnicas utilizadas fueron *Fine tuning* y *Data Augmentation* las cuales se implementaron por separado y en conjunto; sin embargo, la técnica *Data Augmentation* falla debido a que los leves cambios no representan el comportamiento real y variación del tejido blando entre personas. Esta última representó una desmejora en el desempeño de la red neuronal obteniendo una precisión de 46,67% utilizando un filtro Gaussiano 5x5, y de 49,64% con el mismo filtro de dimensión 3x3. La decisión de reducir el tamaño del filtro se planteó inicialmente para evitar perder información importante de la imagen en texturas finas y bordes del cerebro; sin embargo, aunque mejoró el desempeño sigue siendo menor a la red que únicamente utilizó *Fine Tuning*.

Esto puede darse debido a que modificar las imágenes de resonancia magnética que proceden de diferentes pruebas y diferentes resonadores, se estaría afectando la capacidad de discernir entre las anomalías del tejido blando del cerebro naturales. Esto último es de vital importancia debido a que el deterioro del tejido cerebral varía de persona a persona, por factores tales como niveles de escolaridad, género y hábitos deportivos y alimenticios.

Aunque se aprecia una diferencia notable en el desempeño de ambas máquinas de aprendizaje, 77% en el modelo Random Forest y 71% con el modelo de la red neuronal VGG-16, en ambos modelos es evidente la dificultad que tienen para la distinción entre las imágenes que representan una condición de cognición normal y las que presentan una alteración leve. La clase que representa a las personas con enfermedad de Alzheimer diagnosticada, en ambos modelos, fue la que más precisión obtuvo con 98% para el modelo Random Forest y, 99% para la red VGG-16.

Con base a las diversas herramientas y métodos implementados, se concluye que un modelo de clasificadores en cascada es la mejor opción para este caso de estudio debido a las grandes similitudes que hay entre las clases CN y LMCI, que afectan bastante el desempeño global del clasificador.

Capítulo 7 Trabajos futuros

Basado en el trabajo realizado y los resultados obtenidos en ambas máquinas de aprendizaje, se plantea como trabajos futuros las siguientes propuestas:

En primer lugar, se propone realizar una investigación más exhaustiva de las técnicas de extracción de características y/o segmentación de las zonas del cerebro más relevantes para el diagnóstico de la enfermedad de Alzheimer. En este trabajo se utilizaron técnicas de segmentación y detección de tejido blando simples; sin embargo, existen algoritmos e incluso redes neuronales más complejas que permiten la extracción de características claves de manera más limpia, pero con un costo computacional más alto.

Se propone además utilizar la técnica del histograma de gradientes orientados en una máquina de soporte vectorial que podría brindar un mejor resultado y mayores variaciones en sus hiperparámetros, comparándose con el modelo de Random Forest.

Otro punto para tener en cuenta a futuro sería que las redes neuronales, a medida que pasa el tiempo, se vuelven más complejas y robustas, permitiendo una clasificación más precisa y el uso en condiciones más específicas. Se propone realizar un sistema de clasificación que se base en la red neuronal convolucional VGG-19, qué es la versión mejorada de la utilizada en este trabajo, al que se le agregan tres nuevas capas convolucionales. Este modelo se evaluó en [3] se obtuvo una precisión de 98,47%.

En cuanto a la interfaz gráfica, en este trabajo se desarrolló una aplicación sencilla que muestra el resultado de los dos sistemas de clasificación implementados; sin embargo, sería recomendable mejorar y/o desarrollar una interfaz más dinámica y compleja que permitiera el uso de diferentes modelos de clasificación y el uso de diferentes tipos de formatos de imágenes médicas que podrían brincar otro tipo de información y características relevantes para la detección temprana de la enfermedad del Alzheimer.

Capítulo 8 Bibliografía

[1] “Alzheimer’s Disease Neuroimaging Initiative.”

[2] Y. Lei *et al.*, “MRI classification using semantic random forest with auto-context model,” *Quant Imaging Med Surg*, vol. 11, no. 12, pp. 4753–4766, Dec. 2021, doi: 10.21037/qims-20-1114.

[3] R. Khan *et al.*, “A transfer learning approach for multiclass classification of Alzheimer’s disease using MRI images,” *Front Neurosci*, vol. 16, 2023, doi: 10.3389/fnins.2022.1050777.

[4] Paniza Valentina, “Deep Learning en la detección de Alzheimer utilizando imágenes de resonancia magnética,” *Instituto Tecnologico de Buenos Aires*, Sep. 2020.

[5] W. Feng *et al.*, “Automated MRI-Based Deep Learning Model for Detection of Alzheimer’s Disease Process,” *Int J Neural Syst*, vol. 30, no. 6, pp. 1–14, 2020, doi: 10.1142/S012906572050032X.

[6] R. Khan *et al.*, “A transfer learning approach for multiclass classification of Alzheimer’s disease using MRI images,” *Front Neurosci*, vol. 16, 2023, doi: 10.3389/fnins.2022.1050777.

[7] A. Abad Martín, “Aplicación de técnicas de deep learning en la ayuda al diagnóstico de la enfermedad de Alzheimer,” pp. 1–130, 2020, [Online]. Available: https://uvadoc.uva.es/handle/10324/43256

[8] G. P. Tutor, C. S. Gotarredona, and A. Pi, “Proyecto Fin de Carrera Ingeniería de Telecomunicación Análisis de imágenes de Resonancia Magnética para la detección del Alzheimer,” 2018.

[9] J. H. Noh, J. H. Kim, and H. D. Yang, “Classification of Alzheimer’s Progression Using fMRI Data,” *Sensors*, vol. 23, no. 14, 2023, doi: 10.3390/s23146330.

[10] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, “Multimodal Neuroimaging Feature Learning with Multimodal Stacked Deep Polynomial Networks for Diagnosis of Alzheimer’s Disease,” *IEEE J Biomed Health Inform*, vol. 22, no. 1, pp. 173–183, 2018, doi: 10.1109/JBHI.2017.2655720.

[11] G. A. Peláez Brioso, “Métodos para la Clasificación Automática de Imágenes de Resonancia Magnética del Cerebro,” p. 70, 2014, [Online]. Available: https://repositorio.uam.es/bitstream/handle/10486/670427/Pelaez_Briosio_Geraro_Arturo_tfm.pdf?sequence=1&isAllowed=y

[12] Fredy Hernán Sánchez Montaña, “Multiclasificador Deep Learning para la enfermedad de Alzheimer Usando Imágenes FDG-PET,” 2020.

[13] Y. Lei *et al.*, “MRI classification using semantic random forest with auto-context model,” *Quant Imaging Med Surg*, vol. 11, no. 12, pp. 4753–4766, Dec. 2021, doi: 10.21037/qims-20-1114.

[14] A. Nawaz, S. M. Anwar, R. Liaqat, J. Iqbal, U. Bagci, and M. Majid, “Deep Convolutional Neural Network based Classification of Alzheimer’s Disease using MRI Data,” *Proceedings - 2020 23rd IEEE International Multi-Topic Conference, INMIC 2020*, 2020, doi: 10.1109/INMIC50486.2020.9318172.

[15] M. F. Romano, M. D. Nissen, N. Maria, D. Huerto, and C. A. Parquet, “Enfermedad de alzhei mer,” pp. 9–12, 2007.

[16] C. D. E. Revisi, “Resonancia magnética cerebral : secuencias básicas e interpretación,” pp. 292–306, 2011.

[17] L. Bayliss, R. G. Ruiz-garcía, J. Chacón-gonzález, and L. Bayliss, “Neuropsiquiatría del síndrome de Susac : a propósito de un caso,” *Rev Colomb Psiquiatr*, no. February, 2020, doi: 10.1016/j.rcp.2019.10.007.

[18] López Chiret Alvaro, “López - Clasificación de imágenes de resonancia magnética cerebral mediante redes neuronales para...,” *Universidad Politecnica de Valencia*, 2019.

[19] I. tecnológico superior de S. L. P. Ramirez Rangel, Cuevas-Tello JC, *Y SISTEMAS : Volumen I*, no. May 2021. Universidad Autónoma San Luis Potosí, 2019.

[20] A. Bereciartua Pérez, “Desarrollo de algoritmos de procesamiento de imagen avanzado para interpretación de imágenes médicas. Aplicación a segmentación de hígado sobre imágenes de Resonancia Magnética multisecuencia,” Bilbao, 2016.

[21] M. R. Cárdenas, S. Doctorando, and G. Mato, “SEGMENTACION DE IMAGENES MEDICASMEDIANTE INFERENCIA BAYESIANA,” 2023.

[22] I. Tecnológico De León *et al.*, “ANALISIS DE IMAGENES MEDICAS PARA LA EXTRACCION DE BIOMARCADORES MEDIANTE EL ALGORITMO LOCAL BINARY PATTERNS,” 2021.

[23] Villa Jonatan, Hernandez Mario, Morales Severino, and Hernandez Jose, “Detección de objetos basados en HOG/SVM en una imágen,” *InnovaIngeniería*, May 2020.

[24] J. Kalyani and M. Chakraborty, “Contrast Enhancement of MRI Images using Histogram Equalization Techniques,” 2020.

[25] R. Días, “Ranodm Forest - Bagging y árboles de decisión,” The Machine Learners.

[26] Amat Joaquín Rodrigo, “Random Forest con Python,” Ciencia De Datos - https://cienciadedatos.net/documentos/py08_random_forest_python.

[27] N. Gutierrez, “Extracción de Características Texturales de Imágenes de Resonancia Magnética del Cerebro,” UniversidadPolitecnicaMadrid, Madrid, 2019.

[28] A. W. Salehi, P. Baglat, B. B. Sharma, G. Gupta, and A. Upadhya, “A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI,” *Proceedings - International Conference on Smart Electronics and Communication, ICOSEC 2020*, no. Icosec, pp. 156–161, 2020, doi: 10.1109/ICOSEC49089.2020.9215402.

[29] S. Yalamanchili *et al.*, “MRI Brain Tumor Analysis on Improved VGG-16 and Efficient NetB7 Models,” *Journal of Image and Graphics(United Kingdom)*, vol. 12, no. 1, pp. 103–116, 2024, doi: 10.18178/joig.12.1.103-116.

[30] A. A. Pravitasari *et al.*, “UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation,” *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1310–1318, 2020, doi: 10.12928/TELKOMNIKA.v18i3.14753.

[31] S. Raghuvanshi and S. Dhariwal, “The VGG16 Method Is a Powerful Tool for Detecting Brain Tumors Using Deep Learning Techniques †,” *Engineering Proceedings*, vol. 59, no. 1, 2023, doi: 10.3390/engproc2023059046.

[32] G. Lorca, J. Arzola, and O. Pereira, “Segmentación de Imágenes Médicas Digitales mediante Técnicas de Clustering Digital Medical Image Segmentation with Clustering Techniques Artículo original,” 2010.

[33] P. Aznar, “DEcisión Trees: Gini cs Entropy,” QuantDare:https://quantdare.com/decision-trees-gini-vs-entropy/.

[34] Loukadakis Manolis, Cano Jose, O’Boyle Michael, “Accelerating Deep Neural Networks on Low Power Heterogeneous Architectures”, 2018,

Capítulo 9 ANEXOS

La consignación de los programas creados en el entorno Python, con las diferentes fases de las máquinas de aprendizaje implementadas se encuentran en un repositorio personal de Github: *https://github.com/JaimeVG98/Classification-of-AD-CN-LMCI-using-Random-Forest-and-VGG-16-MRI-/tree/main*