

ANÁLISIS COMPARATIVO DE MODELOS DE IA OBTENCIÓN DEL CONJUNTO DE DATOS

1. Para el desarrollo de la presente investigación titulada “Modelo para la Clasificación del Alzheimer”, se utilizaron datos clínicos representados a través de imágenes de resonancia magnética (IRM) del cerebro. Estos datos fueron obtenidos de la siguiente fuente:

Jumaili, M. L. F., & Sonuç, E. (2025). *An Attention-Based CNN Framework for Alzheimer's Disease Staging with Multi-Technique XAI Visualization*. Computers, Materials & Continua, 83(2), 2947-2969.

<https://doi.org/10.32604/cmc.2025.062719>

Además, se logró crear una cuenta a través de la página oficial de figshare
<https://www.kaggle.com/>

Figura 1

Registro de usuario en Kaggle

The screenshot shows the 'Register with email' form on the Kaggle website. The form fields include: 'EMAIL' (placeholder: Enter your email address), 'PASSWORD' (placeholder: Enter password, with an eye icon for visibility), 'FULL NAME' (placeholder: Enter your full name, note: Will be displayed on your profile), and a 'reCAPTCHA' section with a checkbox for 'I'm not a robot'. Below the form are two optional checkboxes: 'Email me Kaggle news and tips' (with a note: You can opt out at any time) and 'Next' button. At the bottom left is a 'Back' button, and at the bottom right is a 'Contact Us / Support' link.

Figura 2

Inicio de Sesión en kaggle

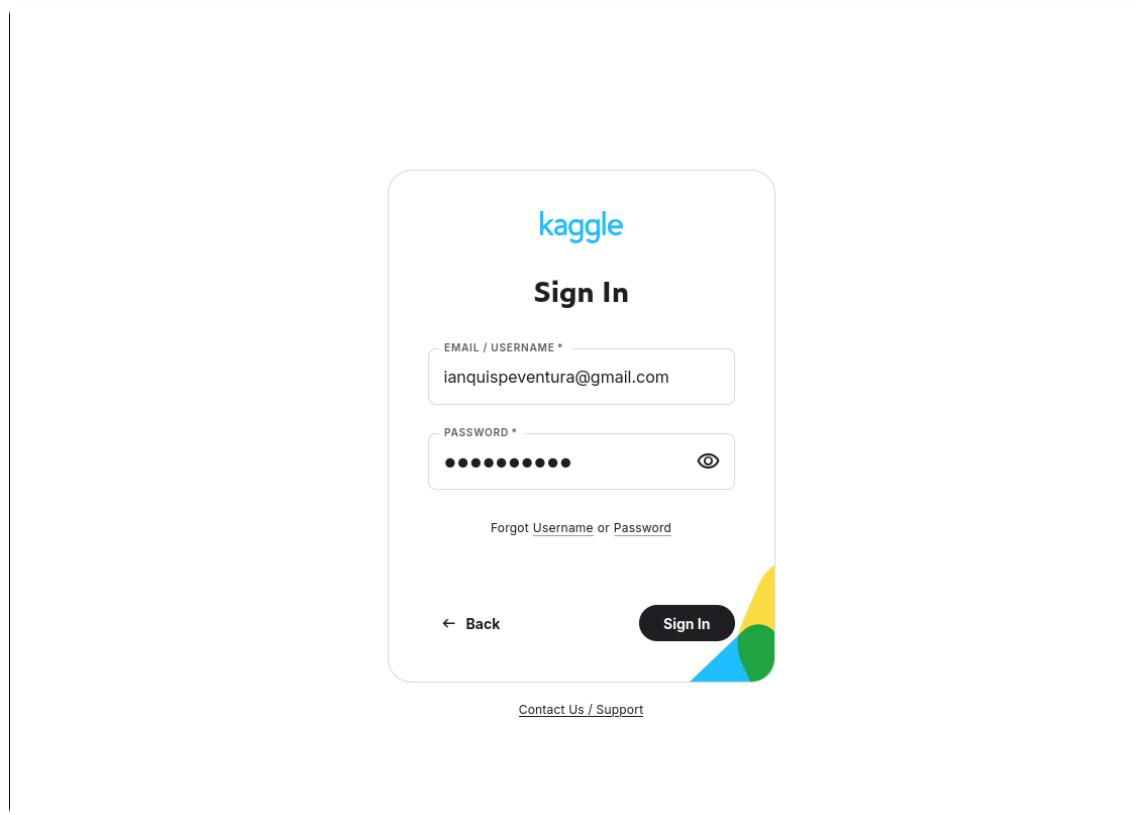


Figura 3

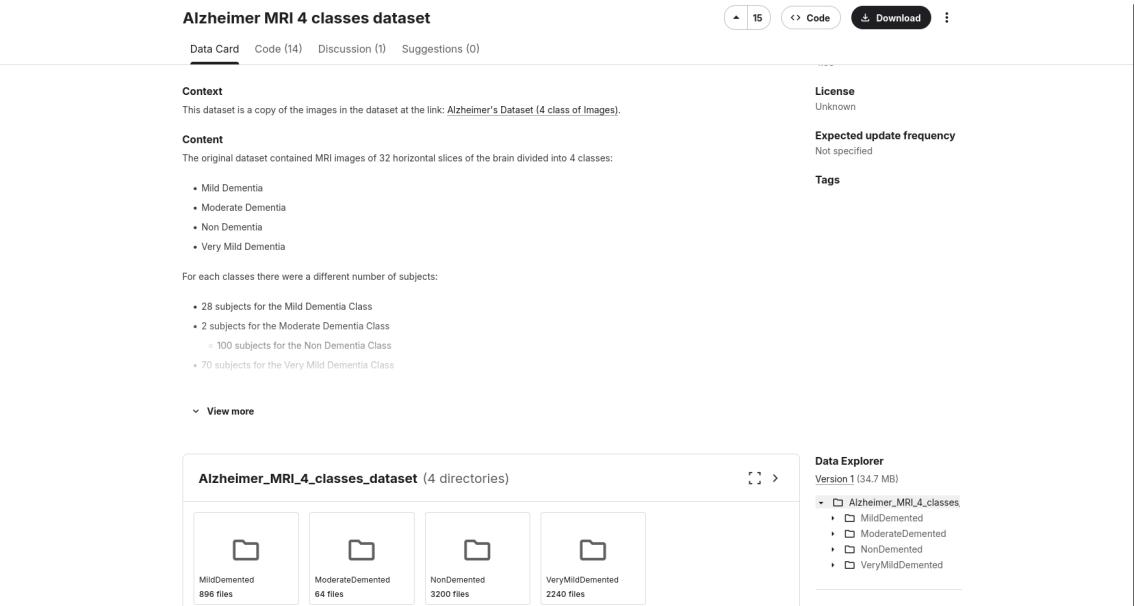
Pantalla principal en kaggle

A screenshot of the Kaggle homepage. The top navigation bar includes a menu icon, the word "kaggle", a "Create" button, and a search bar with the placeholder "Search". To the right of the search bar are three status indicators: "LOGIN STREAK" (1 day), "TIER PROGRESS" (0% to Expert), and "PUBLIC ACTIVITY" (M T W T F S S). The main content area starts with a "Welcome, pepepepe!" message and a note to "Jump back in, or start something new.". Below this are five stats boxes: "Datasets" (2 total created), "Notebooks" (13 total created), "Competitions" (2 total joined), "Courses" (1 total completed), and "Discussions" (0 total posted). A "Hide stats" link is located at the bottom right of these boxes. Further down is a section titled "How to start: Choose a focus for today" with the sub-instruction "Help us make relevant suggestions for you". This section features three cards: "Learn to compete on Kaggle" (with a "Get started" button), "Take a short course" (with a "Get started" button), and "Browse inspiring data and code" (with a "Get started" button). On the left side of the page is a sidebar with sections for "Home", "Competitions", "Datasets", "Models", "Benchmarks", "Code", "Discussions", "Learn", "More", "Your Work", "VIEWED" (listing items like "NeuroKup II Alpha", "Target Encoding", "Principal Component ..."), and "EDITED" (listing items like "Exercise: Creating Fea...", "Exercise: Mutual Infor...").

Además, con la sesión iniciada, se accedió a la siguiente ruta en la misma página:
<https://www.kaggle.com/datasets/marcopinamonti/alzheimer-mri-4-classes-dataset>

Figura 4

Alzheimer MRI 4 classes dataset



The screenshot shows the details of the 'Alzheimer MRI 4 classes dataset' on Kaggle. At the top, there are tabs for 'Data Card', 'Code (14)', 'Discussion (1)', and 'Suggestions (0)'. Below these are sections for 'Context' and 'Content'. The 'Context' section notes that this is a copy of the images from the original Alzheimer's Dataset (4 class of images). The 'Content' section describes the dataset as containing 32 horizontal slices of the brain divided into four classes: Mild Dementia, Moderate Dementia, Non Dementia, and Very Mild Dementia. It also specifies the number of subjects per class: 28 for Mild Dementia, 2 for Moderate Dementia, 100 for Non Dementia, and 70 for Very Mild Dementia. A 'View more' link is present. To the right, there are sections for 'License' (Unknown), 'Expected update frequency' (Not specified), and 'Tags'. Below this is a 'Data Explorer' section showing the structure of the dataset, which contains four directories: 'MildDemented' (898 files), 'ModerateDemented' (64 files), 'NonDemented' (3200 files), and 'VeryMildDemented' (2240 files). The 'Summary' section indicates that the dataset is at Version 1 (34.7 MB).

Este conjunto de datos incluye cuatro categorías de subconjuntos de los diferentes estados del Alzheimer: Demencia leve, Demencia moderada, Sin demencia, Demencia muy leve. A continuación, se presenta una Tabla 1 detallada con la información del conjunto de datos que contiene imágenes cerebrales obtenidas mediante imágenes de resonancia magnética.

Tabla 1

Información del conjunto de datos de la investigación

Item	Conjunto de datos	Formato de imagen	Tamaño Mega bytes	Cantidad de imágenes
1	MildDemented	png	7.1	896
2	ModerateDemented	png	0.526	64
3	NonDemented	png	26	3200
4	VeryMildDemented	png	18	2240
Total			51	6400

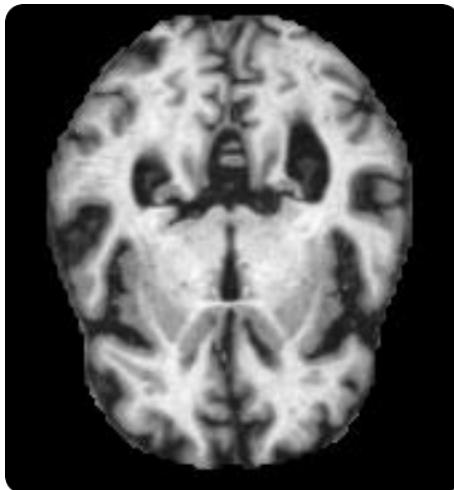
Asimismo, en la Figura 6 donde se muestra la estructura inicial del conjunto de datos:

Figura 5

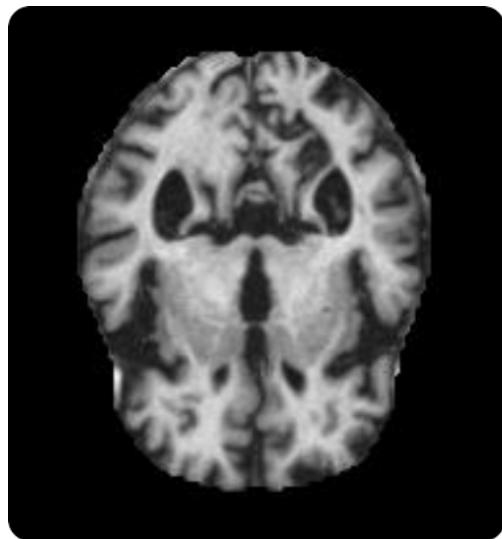
Estructura del conjunto de datos de la investigación

```
↳ Alzheimer_MRI_4_classes_dataset git:(master) tree -L 1
.
├── MildDemented
├── ModerateDemented
└── NonDemented
└── VeryMildDemented
```

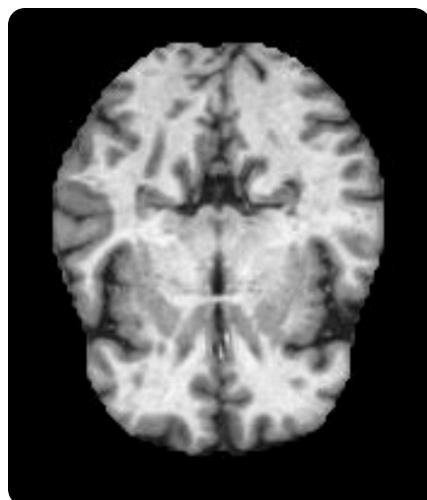
MildDemented



ModerateDemented



NonDemented



VeryMildDemented

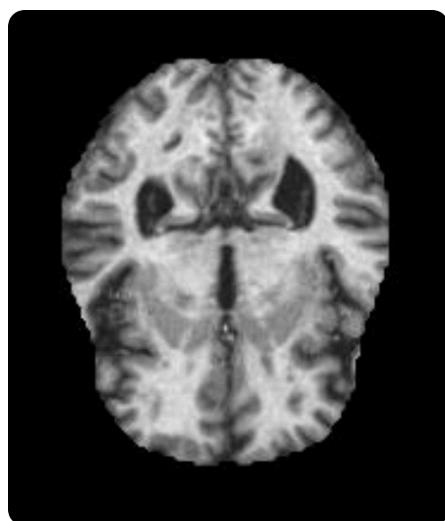
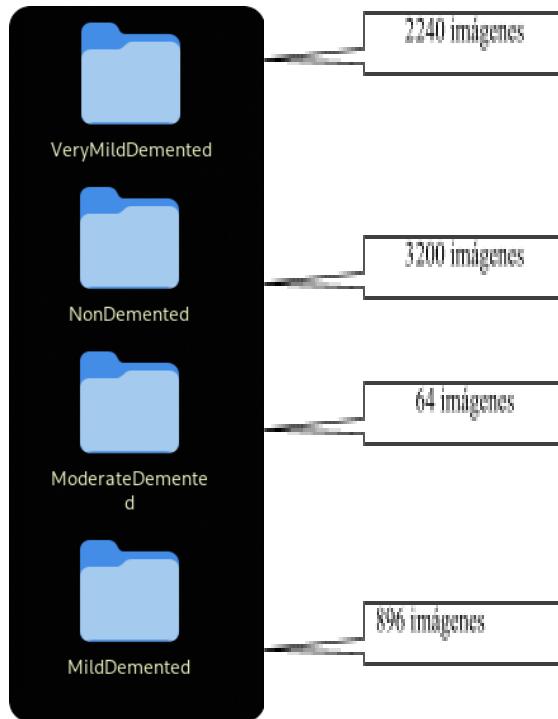


Figura 6

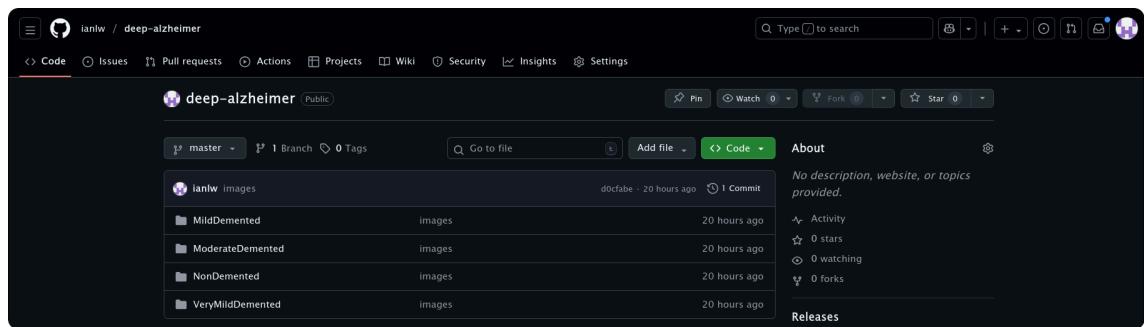
Cantidad de imágenes por carpeta en el conjunto de datos



Luego, se creó un repositorio en Github, dentro se subieron las imágenes de cada uno de los tipos.

Figura 7

Repositorio en Github



2. Posteriormente, se procedió a desarrollar una aplicación en Google Colab que permitió distribuir de forma aleatoria el conjunto de datos considerando un 80% para el conjunto de entrenamiento (*train*) y un 20% para el conjunto de validación (*val*), logrando mostrar la nueva estructura de carpetas de la presente investigación en la Figura 11.

Figura 8

Descarga de Imágenes en Colab

The screenshot shows the Google Colab interface. On the left, there's a sidebar with 'Files' containing a folder named 'deep-alzheimer' which contains subfolders 'MildDemented', 'ModerateDemented', 'NonDemented', and 'VeryMildDemented'. The main area has a code cell with the following content:

```
git clone https://github.com/ianlw/deep-alzheimer.git
Cloning into 'deep-alzheimer'...
remote: Enumerating objects: 6406, done.
remote: Counting objects: 100% (6406/6406), done.
remote: Compressing objects: 100% (6406/6406), done.
remote: Total 6406 (delta 0), reused 6406 (delta 0), pack-reused 0 (from 0)
Receiving objects: 100% (6406/6406), 32.96 MiB | 18.17 MiB/s, done.

# Forzar uso de memoria eficiente
from tensorflow.keras import mixed_precision
mixed_precision.set_global_policy('mixed_float16')
```

Tabla 2

Información del conjunto de datos de la presente investigación

Distribución 80 x 20 %	MildDemen ted	NonDement ed	ModerateDeme nted	VeryMildDe mented	Total
Conjunto de entrenamiento (Train) 80%	716 imágenes	2560 imágenes	51 imágenes	1792 imágenes	5,119
Conjunto de validación (test) 20%	180 imágenes	640 imágenes	13 imágenes	448 imágenes	1281
Total					6400

Según la Tabla 2, el conjunto de datos utilizado en esta investigación corresponde a imágenes de resonancia magnética cerebral (MRI) clasificadas según los distintos grados de la enfermedad de Alzheimer. En la Tabla 3 se muestra la cantidad de imágenes incluidas en el conjunto de datos de entrenamiento (train), que representa el 80% del total

Tabla 3

Conjunto de datos de entrenamiento (train)

Tipo de clasificación	Cantidad de imágenes
MildDemented (Demencia leve)	716
NonDemented (Sin demencia)	2,560

ModerateDemented (Demencia moderada)	51
VeryMildDemented (Demencia muy leve)	1,792
Total de imágenes	5,119

En la Tabla 4 se detalla la cantidad de imágenes correspondientes al conjunto de validación (test), el cual representa el 20% del total.

Tabla 4

Conjunto de datos de validación (test)

Tipo de clasificación	Cantidad de imágenes
MildDemented (Demencia leve)	180
NonDemented (Sin demencia)	640
ModerateDemented (Demencia moderada)	13
VeryMildDemented (Demencia muy leve)	448
Total de imágenes	1,281

Por lo tanto, el conjunto de datos total está compuesto por 6,400 imágenes, distribuidas entre entrenamiento (80%) y validación (20%). Estas imágenes fueron utilizadas para entrenar y evaluar un modelo de Deep Learning basado en Redes Neuronales Convolucionales (CNN) para la clasificación automática del grado de deterioro cognitivo en pacientes con Alzheimer.

DESARROLLO DEL MODELO DE CLASIFICACIÓN DE INTELIGENCIA ARTIFICIAL

1. Análisis de los Antecedentes de la Investigación

Para iniciar el análisis, se presenta la Tabla 5, que incluye a los autores y una descripción detallada de la investigación correspondiente.

Tabla 5

Antecedentes de investigaciones relacionadas con la detección del Alzheimer mediante IA

<i>Item</i>	<i>Referencia Principal (Título)</i>	<i>Modelo de Entrenamiento / Arquitectura</i>	<i>Banco de Datos (Fuente)</i>	<i>Descripción Breve (Objetivo y Metodología)</i>	<i>Métrica de Precisión Reportada</i>
10	"An Attention-Based CNN Framework for Alzheimer's Disease Staging with Multi-Technique XAI Visualization"	CNN con mecanismo de Atención (comparado con ResMLP, VGG16, Xception) y XAI (Grad-CAM, LIME).	Dataset primario (3714 scans MRI), ADNI.	Diseñar una CNN con atención para la estadificación (clasificación multiclas) de la EA, integrando múltiples técnicas de XAI.	<i>Accuracy: 99.18% (dataset primario), 96.64% (ADNI).</i>

2	"Automated MRI-based deep learning model for detection of Alzheimer's disease process"	<i>3D-CNN (Red Neuronal Convolucional 3D) combinada con SVM (3D-CNN-SVM)</i>	ADNI	<i>Desarrollar un modelo automatizado basado en RM y DL para la detección (clasificación binaria y ternaria) de EA, MCI y CN.</i>	Accuracy (binaria): 99.10% (NC vs AD). Accuracy (ternaria): 96.73% (AD).
15	"Predictive Modeling of the Progression of Alzheimer's Disease with Recurrent Neural Networks"	<i>Redes Neuronales Recurrentes (RNN), específicamente Long Short-Term Memory (LSTM).</i>	<i>National Alzheimer's Coordinating Center (NACC) (5432 pacientes).</i>	<i>Predecir la progresión de la EA (estado en la próxima visita) utilizando datos médicos longitudinales heterogéneos y RNN.</i>	Accuracy: > 99%
4	"An integrated deep learning model for early and multi-class diagnosis of Alzheimer's disease from MRI scans"	<i>Híbrido: IDDeepLabV3+ (segmentación), LeNet-5 (extracción), EResNext (clasificación).</i>	<i>Datos públicos (tipo ADNI)</i>	<i>Desarrollar un sistema de clasificación multiclas (4 etapas: ND, VMD, MD, MOD) usando un modelo híbrido de DL.</i>	Accuracy: 98.12% (multiclas). AUC: 0.97.

14	"A predictive model using the mesoscopic architecture of the living brain to detect Alzheimer's disease"	Modelo predictivo (LASSO) basado en "rasgos mesoscópicos" de RM T1 (vector ApV).	(No especificado).	Desarrollar un modelo predictivo basado en la arquitectura mesoscópica de la RM para discriminar patologías de EA (ADrp) de las no relacionadas (nADrp).	Accuracy: 98% (interna), 81% (externa) en la discriminación de ADrp vs nADrp.
7	"A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach"	Deep Ensemble Learning (DEL) (combinando 5 modelos 2D-CNN diferentes).	Datasets públicos de EEG de Alzheimer	Desarrollar un modelo de ensamble profundo (DEL) para clasificar EA y Controles Sanos (HC) usando señales de EEG.	Average Accuracy: 97.9% (en 5-fold cross-validation).
5	"Performance Evaluation of Deep, Shallow and Ensemble Machine Learning Methods for	Comparativa: DL (CNN), ML Superficial (SVM, Random Forest), y Clasificadores Ensemble (Ensamble).	ADNI, OASIS	Comparar rigurosamente el rendimiento de diferentes paradigmas de ML (profundo, superficial, ensamble) en la clasificación de	El clasificador Ensemble propuesto alcanzó 96.52% de accuracy (una mejora del 3-5% sobre el mejor

	<i>"the Automated Classification of Alzheimer's Disease"</i>			<i>4 clases (CN, EMCI, LMCI, AD).</i>	<i>modelo individual).</i>
9	<i>"ADAMAE X—Alzheimer's disease classification via attention-enhanced autoencoders and XAI"</i>	<i>Autoencoder Convolucional con bloques de Atención (Squeeze-and-Excitation) y XAI (Grad-CAM).</i>	<i>ADNI (sMRI T2-weighted)</i>	<i>Desarrollar un modelo (ADAMAE) que combine alta precisión con interpretabilidad (XAI) para la clasificación de EA.</i>	<i>Accuracy: 96.2%</i>
8	<i>"A comprehensive hybrid model: Combining bioinspired optimization and deep learning for Alzheimer's disease identification"</i>	<i>Híbrido: Optimización Bioinspirada (GWO + HHO) y Deep Learning (DL).</i>	<i>(No especificado)</i>	<i>Combinar optimización bioinspirada (para segmentación y ajuste) con DL (para clasificación) de EA vs NC.</i>	<i>Accuracy (Segmentación): 92%. Accuracy (Clasificación): 90%.</i>

12	<i>"Predictive models of Alzheimer's disease dementia risk in older adults with mild cognitive impairment: a systematic review and critical appraisal"</i>	Revisión Sistemática (analizando modelos de ML y tradicionales).	(Revisión de 16 estudios, 9290 participantes).	Revisar sistemáticamente y evaluar críticamente los modelos existentes que predicen el riesgo de progresión de MCI a demencia por EA.	Average AUC (de 15 modelos): 0.87.
I	<i>"A deep learning model for early prediction of Alzheimer's disease dementia based on hippocampal magnetic resonance imaging data"</i>	Modelo "Time-to-Event" de Deep Learning (DL)	ADNI, AIBL	Desarrollar un modelo de DL con RM del hipocampo para predecir el tiempo de progresión de MCI a demencia por EA.	C-index: 0.762 (ADNI). (Mejora a 0.864 con datos clínicos).

11	"Development of a novel dementia risk prediction model in the general population: A large, longitudinal, population-based machine-learning study"	Machine Learning (Modelo UKB-DRP)	UK Biobank (UKB) (425,159 participantes).	Desarrollar un modelo de ML para predecir el riesgo de demencia y EA en la población general a 5, 10 y más años, usando 366 variables.	AUC: 0.862 (EA). 0.848 (demencia).
16	"Development and Validation of a Dementia Risk Prediction Model in the General Population: An Analysis of Three Longitudinal Studies"	Modelo de Predicción de Riesgo (Básico y Extendido) (estadístico, no ML puro).	Cohorte poblacional (2710 indiv.), validado en ADNI-1 y otra cohorte holandesa.	Desarrollar y validar un modelo para predecir el riesgo absoluto de demencia a 10 años en población general.	C-statistic (Extendido, con genética e imágenes): 0.86.
19	"An integrated predictive model for Alzheimer's disease progression from cognitively	Híbrido: Ensemble Transfer Learning, ViT-GAN (Generative Adversarial Network),	ADNI	Predecir la progresión de CN a AD (hasta 10 años) usando RM generadas (sintéticas) de la progresión futura y XAI.	Accuracy: 0.85 (CN a AD). F1-score: 0.86 (CN a AD).

	<i>normal subjects using generated MRI and interpretable AI"</i>	<i>3D-CNN, XAI (Grad-CAM).</i>			
18	<i>"Toward a Predictive Model of Alzheimer's Disease Progression Using Capillary Electrophoresis–Mass Spectrometry and Metabolomics"</i>	<i>Modelo de Predicción (Análisis Estadístico Multivariante) basado en metabolómica.</i>	<i>Muestras de LCR (85 sujetos).</i>	<i>Usar metabolómica de LCR (CE-MS) para desarrollar un modelo que prediga la progresión de la EA (clasificando 4 grupos: SCI, MCI-estable, MCI-progresivo, AD).</i>	<i>Accuracy (Test ciego): 83%. (97–100% en entrenamiento).</i>
6	<i>"A Multi-Classification Assessment Framework for Reproducible Evaluation of Multimodal Learning in Alzheimer's Disease"</i>	<i>Framework de evaluación multimodal (comparando métodos unimodales vs. multimodales).</i>	<i>ADNI-1</i>	<i>Proponer un marco estandarizado y reproducible para evaluar modelos multimodales (incluyendo RM y SNP) en el diagnóstico de EA.</i>	<i>Accuracy (usando SNP): 57.09%. La adición de SNP mejoró el rendimiento multimodal en 3-7%.</i>

La Tabla resume los principales estudios recientes que aplican técnicas de inteligencia artificial para el diagnóstico o predicción del Alzheimer. Se observa una fuerte predominancia del uso de imágenes de resonancia magnética (MRI) como fuente principal de datos, debido a su capacidad para reflejar cambios estructurales en regiones clave como el hipocampo y la corteza entorinal.

Los modelos convolucionales (CNN) en sus variantes 2D y 3D son los más empleados, ya que permiten extraer patrones espaciales complejos de las imágenes cerebrales.

Asimismo, los estudios basados en transfer learning (con arquitecturas como ResNet, Inception o EfficientNet) demuestran que la reutilización de pesos preentrenados mejora la precisión y reduce el tiempo de entrenamiento. Finalmente, los modelos con autoencoders, atención y técnicas de interpretabilidad (XAI) destacan como tendencias emergentes, al ofrecer explicaciones visuales sobre las decisiones del modelo, lo que contribuye a su validación médica.

Tabla 6

Arquitecturas principales basadas del CNN

Modelo / Arquitectura	Descripción	Ventajas Principales
RNN (LSTM)	Redes neuronales recurrentes que modelan secuencias temporales para predecir la progresión de la enfermedad a lo largo del tiempo.	Capturan dependencias temporales; alta precisión (>99%); adecuadas para datos clínicos longitudinales.
CNN con Mecanismo de Atención + XAI	Red convolucional con bloques de atención e interpretabilidad mediante Grad-CAM y LIME.	Alta precisión (99%); permite interpretar qué regiones del cerebro influyen más en la decisión del modelo.
3D-CNN + SVM	CNN tridimensional para extraer características espaciales de las MRI,	Muy alta precisión (\approx 99%); aprovecha tanto el aprendizaje profundo como

	combinada con SVM para la clasificación.	el poder discriminativo de SVM.
Modelo Híbrido (IDeepLabV3+ + LeNet-5 + EResNext)	Modelo compuesto: segmenta, extrae y clasifica imágenes MRI en múltiples etapas del Alzheimer.	Multiclas (4 etapas); excelente rendimiento (98%); mejora la precisión combinando tareas especializadas.
Modelo LASSO (Rasgos Mesoscópicos)	Modelo estadístico basado en la estructura mesoscópica del cerebro observada en MRI T1.	Simplicidad interpretativa; muy buena precisión (98%) sin necesidad de redes profundas.
Deep Ensemble Learning (DEL)	Conjunto de cinco CNN entrenadas sobre señales EEG para mejorar la generalización.	Robusto ante ruido; alta precisión (\approx 98%); útil cuando se dispone de EEG en lugar de MRI.
Análisis Estadístico Multivariante (Metabolómica CE-MS)	Clasifica mediante metabolitos en líquido cefalorraquídeo, diferenciando etapas del Alzheimer.	Permite diagnóstico bioquímico temprano; precisión alta (97–100%).
Autoencoder con Atención (ADAMAEX)	Autoencoder convolucional con bloques SE (Squeeze-and-Excitation) y explicabilidad XAI.	Alta precisión (96%); interpreta las regiones cerebrales clave; combina reducción de dimensionalidad y clasificación.
Ensemble Learning (DL + ML)	Combina múltiples clasificadores (CNN, SVM, RF) para mejorar la estabilidad y exactitud.	Mejora la precisión (96.5%); robusto ante variabilidad de datos; evita sobreajuste.
Transfer Learning (VGG + Selección Inteligente)	Uso de redes preentrenadas (VGG) con selección de datos basada en entropía para optimizar el entrenamiento.	Reduce drásticamente el tamaño del set de entrenamiento; mantiene o mejora la precisión.
ViT-GAN + Ensemble + XAI	Usa GANs (para generar MRI futuras) y Vision Transformers para predecir progresión CN→AD.	Predice evolución futura; combina imágenes sintéticas e interpretabilidad.

Modelo “Time-to-Event” (DL)	Predice el tiempo hasta la conversión de MCI a Alzheimer usando MRI del hipocampo.	Predicciones temporales útiles clínicamente; incorpora datos de supervivencia.
Modelo de Riesgo UKB-DRP (ML Clásico)	Modelo poblacional con variables demográficas y clínicas para estimar riesgo de demencia.	Gran escala (425k personas); útil en prevención y políticas de salud.
Modelo de Riesgo Estadístico (Extendido)	Basado en regresión con variables genéticas e imágenes (ADNI).	Alta capacidad predictiva ($C=0.86$); buena generalización entre cohortes.
Framework Multimodal (MRI + SNP)	Evalúa métodos unimodales y multimodales (genética + imagen).	Fomenta reproducibilidad; muestra mejora al combinar modalidades.
CNN (basada en biomarcadores)	Clasifica Alzheimer en base a biomarcadores como amiloide- β .	Permite diagnóstico temprano; supera métodos clásicos de ML.
RNNs (Vanilla, LSTM, GRU)	Redes recurrentes para análisis de datos longitudinales multimodales.	Modelan secuencias temporales; superan modelos estadísticos tradicionales.
Optimización Bioinspirada + DL (GWO + HHO)	Usa metaheurísticas bioinspiradas para segmentación y ajuste antes de clasificar con DL.	Mejora segmentación; integra optimización global; precisión del 90%.
Modelos Multivariantes (Metabolómica LCR)	Predicen progresión de EA a partir de metabolitos en líquido cefalorraquídeo.	Diagnóstico no invasivo; permite identificar etapas tempranas.
Revisiones Sistemáticas (ML y DL)	Analizan el estado del arte en detección y predicción del Alzheimer con IA.	Identifican vacíos, tendencias y retos futuros; no generan métricas propias.

La Tabla de arriba presenta una comparación de las arquitecturas de aprendizaje profundo y sus respectivas aplicaciones en el diagnóstico automatizado del Alzheimer. Se aprecia que las redes convolucionales (CNN) constituyen la base de la

mayoría de los modelos revisados, tanto en versiones 2D (para cortes específicos del cerebro) como 3D (para volúmenes completos).

Las estrategias de transfer learning se posicionan como una práctica estándar al permitir aprovechar representaciones preentrenadas, logrando una mejor generalización en conjuntos de datos limitados. Los autoencoders con atención y los métodos de explicabilidad (XAI) amplían la capacidad del modelo al identificar las regiones más relevantes del cerebro para la predicción, promoviendo la transparencia del sistema.

Los modelos híbridos (combinando deep learning con algoritmos clásicos o metaheurísticas bioinspiradas) muestran una tendencia a optimizar la selección de características, mientras que los RNN/LSTM y los enfoques multimodales avanzan hacia la comprensión temporal y contextual de la enfermedad. En conjunto, la literatura evidencia un tránsito hacia sistemas predictivos más interpretables, generalizables y clínicamente útiles.

Experimentación de las Arquitecturas Basadas en el Modelo de Clasificación del Alzheimer

En este apartado se presentan los resultados experimentales obtenidos al evaluar diferentes arquitecturas de redes neuronales profundas aplicadas a la tarea de clasificación automática de las etapas del Alzheimer a partir de imágenes de resonancia magnética (MRI) cerebrales.

El objetivo principal fue identificar la arquitectura con mejor capacidad de generalización, optimizando la precisión en la clasificación de las cuatro categorías establecidas en el conjunto de datos: *NonDemented*, *VeryMildDemented*, *MildDemented* y *ModerateDemented*.

Se desarrollaron dos fases experimentales:

1. Entrenamiento sin fine-tuning, utilizando los pesos preentrenados en ImageNet y ajustando únicamente las capas finales de clasificación.
2. Entrenamiento con fine-tuning, en el cual se liberaron capas intermedias de la red para permitir una mejor adaptación del modelo a las características

específicas del conjunto de datos MRI.

Las arquitecturas evaluadas fueron EfficientNetB0, ResNet50V2, MobileNetV2, DenseNet121, y un modelo CNN con mecanismo de atención (Attention-CNN) diseñado para este estudio.

La Tabla 7 muestra los resultados obtenidos en términos de *accuracy* (exactitud) para cada una de las arquitecturas evaluadas sin aplicar la técnica de fine-tuning.

Tabla 7

Resultados preliminares de los indicadores de las arquitecturas basadas en el modelo de clasificación del Alzheimer sin el uso de la técnica Fine-Tuning

Ítem	Modelo CNN	Accurac y	Adenomas				Gliomas				Meningiomas				No Tumor			
			Precisio n	Recal l	F1 Score	ROC Curve AUC	Precisio n	Recal l	F1 Score	ROC Curve AUC	Precision	Recal l	F1 Score	ROC Curve AUC	Precision	Recall	F1 Score	ROC Curve AUC
1	MobileNet	0.769	0.846	0.5 43	0.6 62	0.94 4	0.634	0.9 24	0.75 2	0.940	0.803	0.6 79	0.73 5	0.93 8	0.914	0.91 4	0.91 4	0.995
2	Xception	0.722	0.886	0.4 81	0.6 24	0.89 4	0.588	0.8 70	0.70 2	0.910	0.662	0.5 36	0.59 2	0.88 6	0.889	0.98 8	0.93 6	0.991
3	DenseNet20 1	0.716	0.780	0.3 95	0.5 25	0.85 1	0.670	0.7 72	0.71 7	0.922	0.725	0.6 90	0.70 7	0.91 2	0.730	1.00 0	0.84 4	0.995
4	ResNet101 V2	0.716	0.932	0.5 06	0.6 56	0.92 2	0.687	0.6 20	0.65 1	0.888	0.522	0.8 45	0.64 5	0.87 6	0.973	0.90 1	0.93 6	0.992
5	ResNet152 V2	0.763	0.881	0.6 42	0.7 43	0.93 1	0.676	0.7 72	0.72 1	0.905	0.678	0.7 02	0.69 0	0.89 1	0.874	0.93 8	0.93 5	0.992
6	MobileNet V2	0.778	0.759	0.7 41	0.7 50	0.91 4	0.770	0.6 20	0.68 7	0.923	0.691	0.7 98	0.74 0	0.91 3	0.898	0.97 5	0.93 5	0.991
7	DenseNet16 9	0.740	0.889	0.3 95	0.5 47	0.90 3	0.713	0.7 28	0.72 0	0.927	0.648	0.8 33	0.72 9	0.92 6	0.810	1.00 0	0.89 5	0.996

8	DenseNet121	0.725	0.762	0.395	0.520	0.898	0.7138	0.783	0.746	0.9276	0.68214	0.714	0.698	0.916	0.757	1.000	0.862	0.997
9	NASNetMobile	0.630	1.000	0.086	0.159	0.859	0.58837	0.81	0.691	0.906	0.57683	0.50	0.580	0.867	0.696	0.988	0.816	0.985
10	InceptionResNetV2	0.660	0.960	0.296	0.453	0.822	0.63563	0.69	0.649	0.888	0.50426	0.75	0.595	0.857	0.802	0.951	0.870	0.982

Nota: El proceso de entrenamiento de los modelos basados en redes neuronales profundas para la clasificación del Alzheimer se realizó siguiendo un protocolo uniforme y riguroso. Primero, se importaron todas las librerías necesarias (TensorFlow, Keras, NumPy, y Matplotlib, entre otras). Para cada modelo, se utilizó una arquitectura preentrenada (como MobileNet, ResNet, Xception, DenseNet, entre otras), cargada desde la biblioteca Keras Applications. Estas redes fueron adaptadas al problema mediante la adición de capas personalizadas, incluyendo una capa de GlobalAveragePooling2D, una capa densa con activación ReLU, y una capa de salida Softmax configurada para la clasificación multiclas de imágenes de resonancia magnética cerebral (MRI) en las cuatro categorías: *NonDemented*, *VeryMildDemented*, *MildDemented* y *ModerateDemented*. Las capas superiores del modelo base fueron congeladas para conservar los pesos preentrenados en ImageNet, permitiendo que solo las capas añadidas fueran entrenadas inicialmente. Las imágenes fueron redimensionadas a 224x224 píxeles, con un tamaño de lote de 32 y un total de 35 épocas. Para mejorar la capacidad de generalización, se aplicaron técnicas de aumento de datos (Data Augmentation) como rotación, desplazamiento horizontal/vertical y zoom aleatorio, exclusivamente sobre el conjunto de entrenamiento. Los modelos fueron compilados con el optimizador Adam, la función de pérdida categórica cruzada (Categorical Crossentropy) y se evaluaron utilizando métricas como accuracy, precision, recall, F1-score, y curvas ROC-AUC por clase. Una vez completado el entrenamiento, cada modelo fue guardado en formato “.keras” con un nombre representativo de la arquitectura utilizada. Finalmente, se realizó una evaluación comparativa exhaustiva basada en los reportes de clasificación, matrices de confusión y curvas ROC, determinando que las arquitecturas MobileNetV2 y MobileNet alcanzaron los mejores resultados iniciales en precisión global, por lo que fueron seleccionadas para la siguiente fase experimental con el uso de *Fine-Tuning*.

En la Tabla 8, se presentan los resultados obtenidos tras aplicar la técnica de *fine-tuning* a las distintas arquitecturas de redes neuronales convolucionales (CNN) evaluadas en la tarea de clasificación del Alzheimer.

La métrica accuracy (exactitud) se utilizó para medir la capacidad de cada modelo para predecir correctamente las clases de deterioro cognitivo: *NonDemented (sin demencia)*, *VeryMildDemented (demencia muy leve)*, *MildDemented (demencia leve)* y *ModerateDemented (demencia moderada)*.

Entre los modelos evaluados, MobileNet alcanzó el mayor *accuracy*, con un valor de 0.817, posicionándose como la mejor arquitectura tras la aplicación del *fine-tuning*.

Este resultado demuestra que MobileNet es altamente eficaz para la clasificación de los distintos grados de Alzheimer, logrando un aumento notable en su rendimiento respecto a los resultados obtenidos sin el uso de esta técnica.

El modelo MobileNetV2 obtuvo el segundo mejor *accuracy*, con un valor de 0.790. Al igual que MobileNet, mostró una mejora sustancial tras el ajuste fino, consolidándose como una alternativa competitiva y robusta para la clasificación automatizada de imágenes MRI de Alzheimer.

Asimismo, los modelos DenseNet201 y ResNet101V2 presentaron desempeños destacados, ambos con un *accuracy* de 0.778. Estos modelos evidenciaron una mejora clara después de aplicar *fine-tuning*, incrementando su capacidad para distinguir con precisión entre los distintos niveles de deterioro cognitivo.

Por su parte, ResNet152V2 alcanzó un *accuracy* de 0.766, mientras que DenseNet169 registró un valor de 0.757. Aunque ambos modelos se beneficiaron del proceso de ajuste fino, no lograron superar a MobileNet ni a MobileNetV2 en términos de precisión global.

En un nivel ligeramente inferior, Xception alcanzó un *accuracy* de 0.751, lo que indica una mejora con respecto a los resultados sin *fine-tuning*, aunque su rendimiento fue menor en comparación con las arquitecturas más eficientes.

Finalmente, NASNetMobile fue el modelo con el menor *accuracy*, con un valor de 0.683. A pesar de la aplicación del *fine-tuning*, su rendimiento general no resultó competitivo frente a las demás arquitecturas evaluadas.

En conclusión, la aplicación del *fine-tuning* mejoró significativamente el rendimiento de la mayoría de los modelos, siendo MobileNet y MobileNetV2 los que alcanzaron los mejores resultados.

Por lo tanto, ambas arquitecturas se consideran las principales candidatas para su implementación definitiva en el modelo propuesto de clasificación automática del Alzheimer mediante imágenes de resonancia magnética cerebral (MRI).

Tabla 8

Resultados preliminares de los indicadores de las arquitecturas basadas en el modelo CNN con el uso de la técnica Fine-Tuning

Ítem	Modelo CNN	Accuracy	Adenomas				Gliomas				Meningiomas				No Tumor			
			Precision	Recall	F1 Score	ROC Curve AUC	Precision	Recall	F1 Score	ROC Curve AUC	Precision	Recall	F1 Score	ROC Curve AUC	Precision	Recall	F1 Score	ROC Curve AUC
1	MobileNet	0.817	0.945	0.642	0.765	0.960	0.714	0.870	0.784	0.952	0.747	0.774	0.760	0.950	0.940	0.975	0.958	0.998
2	Xception	0.751	0.883	0.556	0.667	0.924	0.680	0.739	0.708	0.926	0.653	0.738	0.693	0.917	0.888	0.975	0.929	0.997
3	DenseNet201	0.778	0.975	0.481	0.645	0.939	0.645	0.909	0.758	0.932	0.753	0.726	0.739	0.930	0.889	0.988	0.936	0.997
4	ResNet101V2	0.778	0.930	0.654	0.768	0.963	0.753	0.663	0.705	0.914	0.597	0.845	0.700	0.919	0.963	0.963	0.963	0.999
5	ResNet152V2	0.766	0.912	0.642	0.754	0.956	0.675	0.815	0.739	0.923	0.625	0.655	0.640	0.894	0.939	0.951	0.945	0.995
6	MobileNetV2	0.790	0.829	0.716	0.768	0.952	0.744	0.728	0.736	0.934	0.697	0.738	0.717	0.926	0.899	0.988	0.941	0.998
7	DenseNet169	0.757	0.939	0.383	0.544	0.917	0.697	0.826	0.756	0.932	0.708	0.810	0.756	0.934	0.810	1.000	0.895	0.998

8	DenseNet121	0.728	0.837	0.50 6	0.63 1	0.92 9	0.596	0.91 3	0.72 1	0.93 5	0.719	0.48 8	0.58 2	0.92 0	0.879	0.98 8	0.93 0	0.99 8
9	NASNetMobile	0.683	0.885	0.28 4	0.43 0	0.88 6	0.561	0.90 2	0.69 2	0.90 4	0.750	0.53 6	0.62 5	0.88 2	0.769	0.98 8	0.86 5	0.98 8
10	InceptionResNetV2	0.698	0.949	0.45 7	0.61 7	0.93 9	0.558	0.89 1	0.68 6	0.91 4	0.578	0.44 0	0.50 0	0.87 0	0.909	0.98 8	0.94 7	0.99 8

Nota: El proceso de *fine-tuning* fue aplicado de manera consistente a todas las arquitecturas de Redes Neuronales Convolucionales (CNN) evaluadas. En cada modelo, se descongelaron las capas finales para permitir su ajuste a las características específicas de las imágenes de resonancia magnética cerebral (MRI) del conjunto de datos *Alzheimer MRI 4 Classes Dataset*. Posteriormente, cada modelo fue recompilado con una tasa de aprendizaje baja (1e-5) utilizando el optimizador Adam, lo que permitió optimizar las capas superiores sin perder los conocimientos generales adquiridos durante el preentrenamiento. Durante el entrenamiento, los modelos fueron sometidos a 35 épocas para ajustar los pesos de las capas descongeladas, monitoreando el desempeño mediante la métrica de *accuracy* tanto en el conjunto de entrenamiento como en el de validación. Una vez finalizado el ajuste fino, los modelos entrenados fueron guardados en formato “.keras” para su posterior evaluación y aplicación en la clasificación automática de los distintos grados de Alzheimer (*NonDemented*, *VeryMildDemented*, *MildDemented* y *ModerateDemented*). Este enfoque fue implementado de manera uniforme en todas las arquitecturas CNN, garantizando así una comparación coherente, justa y reproducible entre sus rendimientos.

2. Propuesta de una aplicación móvil basada en un modelo de clasificación del Alzheimer

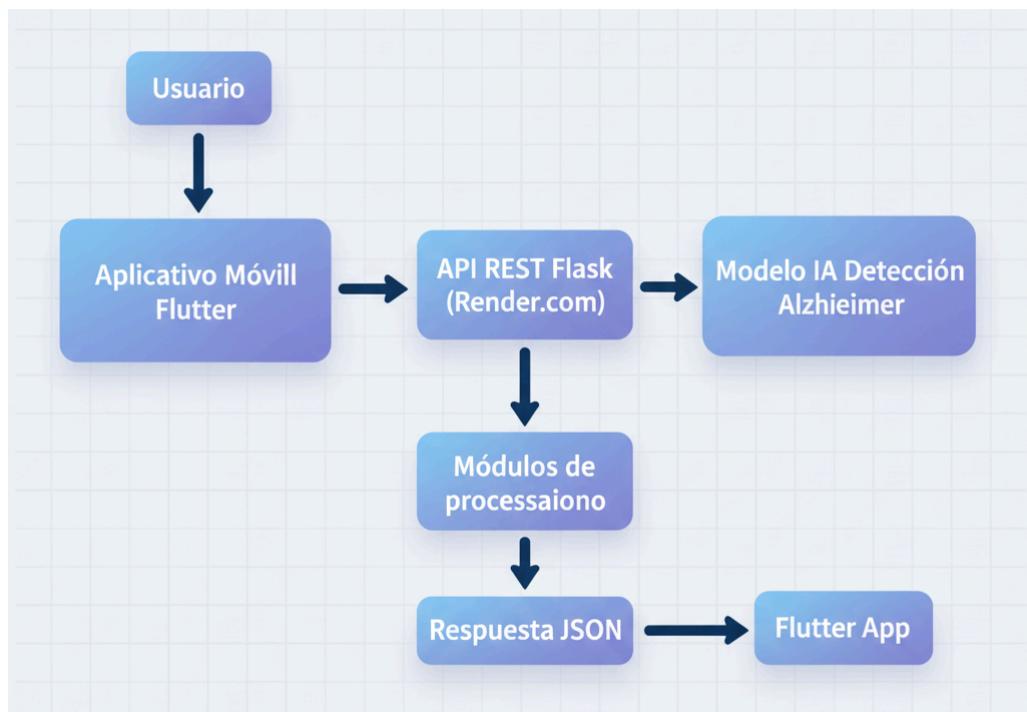
Se propone el desarrollo de una aplicación móvil inteligente que integre un modelo de clasificación entrenado mediante redes neuronales convolucionales (CNN), capaz de detectar y clasificar el grado de Alzheimer a partir de imágenes de resonancia magnética cerebral (MRI).

El modelo utilizado corresponde a una arquitectura ensamblada con las mejores configuraciones identificadas en la etapa experimental, siendo el más óptimo para la tarea de clasificación del Alzheimer según los resultados preliminares.

El diseño y la comunicación entre los componentes del sistema se sustentan en una arquitectura de software representada en la Figura 20.

Figura 20

Arquitectura de software de la aplicación móvil utilizando el modelo de clasificación del Alzheimer.



Según la Figura 20, la arquitectura del sistema se organiza bajo el patrón de diseño Modelo–Vista–VistaModelo (MVVM), que ofrece una mejor separación de responsabilidades y una interacción fluida entre los componentes.

- Vista (View): corresponde al aplicativo móvil desarrollado en Flutter, encargado de la interfaz de usuario. Desde esta vista, el usuario puede cargar imágenes MRI y enviar solicitudes al servidor para su análisis.
- VistaModelo (ViewModel): actúa como intermediario entre la vista y la lógica de negocio. Gestiona los datos recibidos desde la API, controla el estado de la aplicación y actualiza la interfaz de manera reactiva según los resultados obtenidos.
- Modelo (Model): incluye el backend desarrollado con Flask en Python, donde se aloja la API REST (Render.com). Este módulo gestiona los endpoints de comunicación, realiza el procesamiento de las imágenes, y ejecuta el modelo de IA de detección del Alzheimer.

El flujo general inicia cuando el usuario envía una imagen desde la aplicación móvil; esta solicitud es procesada por la API Flask, que realiza la inferencia mediante el modelo de IA. Luego, el resultado se devuelve en formato JSON a la aplicación Flutter, donde la VistaModelo actualiza la interfaz con la predicción correspondiente.

Esta arquitectura MVVM permite una mayor modularidad, escalabilidad y mantenimiento del código, además de optimizar la experiencia de usuario mediante una interfaz reactiva y eficiente.

En conjunto, la aplicación móvil constituye una herramienta innovadora que integra inteligencia artificial y telemedicina para el apoyo diagnóstico temprano del Alzheimer.

3. Uso de la metodología desarrollo Ágil Scrum

En la presente investigación, para el desarrollo de la aplicación móvil destinada al diagnóstico de tumores primarios, se ha adoptado la metodología Ágil SCRUM. A continuación, se presentan los aspectos más relevantes de su implementación:

3.1 Planificación del Proyecto con la Metodología SCRUM

La implementación de la metodología SCRUM en el desarrollo de esta aplicación móvil para el apoyo diagnóstico de Alzheimer permite trabajar con un enfoque flexible, iterativo y altamente adaptable. Gracias a este marco de trabajo ágil, es posible entregar incrementos funcionales del sistema de forma continua,

asegurando que cada avance sea útil, testeable y alineado con los objetivos del proyecto.

SCRUM facilita la transparencia, colaboración y comunicación constante entre los miembros del equipo, aspectos esenciales para obtener retroalimentación temprana y garantizar la calidad del software desde las primeras etapas. Asimismo, este método promueve la mejora continua mediante revisiones periódicas, retrospectivas y ajustes en respuesta a los avances técnicos o cambios en los requisitos.

3.1.1 Formación del equipo SCRUM

Para el desarrollo de la aplicación móvil enfocada en el diagnóstico asistido de Alzheimer, se conformó un equipo SCRUM con roles claramente definidos, asegurando una estructura organizativa eficiente y funcional. Cada integrante tiene responsabilidades específicas para garantizar el flujo adecuado de trabajo, la coordinación multidisciplinaria y el cumplimiento de los objetivos del proyecto.

Tabla 9

Roles y Responsabilidades del Equipo SCRUM

Rol	Responsable(s)	Función
Product Owner	Roddy Pérez Acosta	<ul style="list-style-type: none">• Define y prioriza los requisitos funcionales y no funcionales del proyecto.• Asegura que el producto final responda a las necesidades del usuario y a los objetivos del sistema de apoyo diagnóstico de Alzheimer.• Supervisa el backlog y valida los incrementos entregados.

Scrum Master	Paulo Guzmán	<ul style="list-style-type: none"> • Facilita el proceso SCRUM y asegura el cumplimiento de sus principios. • Planifica, coordina y da seguimiento a los sprints. • Elimina impedimentos y gestiona los recursos necesarios para el desarrollo fluido del proyecto.
Scrum Team (Equipo de Desarrollo)	Ian Quispe	<ul style="list-style-type: none"> • Desarrolla el backend y el frontend de la aplicación móvil. • Implementa la integración con el modelo de clasificación de Alzheimer. • Realiza pruebas técnicas y de usabilidad para garantizar la calidad y funcionalidad de cada incremento del producto.

3.2 Administración y gestión de los recursos del proyecto

3.2.1 Requisitos Funcionales (RF)

Para asegurar que la aplicación móvil cumpla con los objetivos planteados en esta investigación, se han definido una serie de Requisitos Funcionales que describen las capacidades esenciales del sistema. Estos requisitos garantizan que los usuarios puedan interactuar adecuadamente con la plataforma, desde el registro y autenticación, hasta la carga de imágenes cerebrales, evaluación mediante el modelo de IA, visualización de resultados y acceso al historial clínico.

A continuación, se presentan los requisitos funcionales clave identificados, organizados en las Tablas 10 a 17:

Tabla 10

Información del Requisito Funcional (RF-01)

RF-01	Nombre	Registrar usuario
RF-01	Descripción	Permite a los médicos crear una cuenta en la aplicación proporcionando

	información básica.
Prioridad	Alta
Entrada	Nombre, apellido, correo electrónico, contraseña.
Salida	Confirmación de registro exitoso.
Condiciones	El correo electrónico debe ser único. Los datos deben ser válidos y completos.

Tabla 11

Información del Requisito Funcional (RF-02)

RF-02	Nombre	Inicio de sesión
	Descripción	Permite que los usuarios médicos accedan al sistema mediante credenciales válidas.
	Prioridad	Alta
	Entrada	Correo electrónico y contraseña del usuario.
	Salida	Acceso al panel principal de la aplicación.
	Condiciones	Las credenciales deben coincidir con un usuario registrado.

Tabla 12

Información del Requisito Funcional (RF-03)

RF-03	Nombre	Cierre de sesión
	Descripción	Permite al usuario salir del sistema de forma segura.
	Prioridad	Media
	Entrada	Acción manual del usuario por parte del usuario.
	Salida	Confirmación de cierre de sesión.
	Condiciones	El usuario debe estar previamente autenticado. Tabla 13.

Tabla 13

Información del Requisito Funcional (RF-04)

RF-04	Nombre	Carga de imágenes médicas
	Descripción	Permite subir imágenes MRI para análisis automático.
	Prioridad	Alta
	Entrada	Imagen en formato compatible (PNG).
	Salida	Confirmación de carga exitosa de la imagen.
	Condiciones	La imagen debe cumplir con los formatos y tamaño permitido.

Tabla 14

Información del Requisito Funcional (RF-05)

RF-05	Nombre	Procesamiento de imágenes
	Descripción	La aplicación analiza las imágenes cerebrales cargadas mediante el modelo de clasificación Alzheimer.
	Prioridad	Alta
	Entrada	Imágenes MRI cargadas.
	Salida	Resultados automáticos del análisis.
	Condiciones	El análisis debe ser rápido, preciso y ejecutarse en el servidor IA.

Tabla 15*Información del Requisito Funcional (RF-06)*

RF-06	Nombre	Visualización de resultados
	Descripción	Muestra el resultado del análisis de IA al usuario médico.
	Prioridad	Alta
	Entrada	Resultados generados por el modelo IA.
	Salida	Clasificación del caso (normal, leve, moderado, severo), mapas de calor y lectura interpretativa.
	Condiciones	Debe presentarse con claridad y con interpretación médico–orientada.

Tabla 16*Información del Requisito Funcional (RF-07)*

RF-07	Nombre	Historial de diagnósticos
	Descripción	Permite consultar los análisis previos asociados al paciente.
	Prioridad	Media
	Entrada	Registros almacenados de imágenes y diagnósticos.
	Salida	Línea de tiempo con diagnósticos previos y evolución.
	Condiciones	Los datos deben almacenarse de forma segura y accesible.

Tabla 17*Información del Requisito Funcional (RF-08)*

RF-08	Nombre	Información sobre Alzheimer
	Descripción	Muestra contenido informativo relacionado con el Alzheimer.
	Prioridad	Media
	Entrada	Consulta del usuario.
	Salida	Información educativa sobre síntomas, causas, fases, cuidados, tratamientos, etc.
	Condiciones	El contenido debe ser veraz, actualizado y comprensible.

3.2.2 Requisitos no Funcionales (RNF)

Además de los Requisitos Funcionales, es indispensable considerar una serie de Requisitos No Funcionales que aseguren el correcto rendimiento, seguridad, usabilidad y calidad de la aplicación móvil. Estos requisitos permiten que el sistema no solo cumpla sus objetivos clínicos y tecnológicos, sino que lo haga de manera eficiente, confiable y alineada a los estándares del sector salud.

En la Tabla 18 se presentan los Requisitos No Funcionales clave definidos para este proyecto.

Tabla 18*Información de los Requisitos no Funcionales*

Requisito No Funcional	Descripción
RNF-01 Seguridad y Privacidad	Cumplimiento normativo: La aplicación debe cumplir con normativas de protección de datos médicos como HIPAA, GDPR o equivalentes nacionales. Cifrado de datos: Las imágenes cerebrales, diagnósticos y datos clínicos deben estar cifrados tanto en tránsito (HTTPS/TLS) como en reposo.
RNF-02 Rendimiento	Tiempo de respuesta: El procesamiento de imágenes y generación de resultados debe realizarse en tiempos aceptables para el usuario. Escalabilidad: La aplicación debe soportar un incremento de usuarios, imágenes procesadas y tamaño de los datos sin degradar su desempeño.
RNF-03 Usabilidad	Interfaz intuitiva: Debe diseñarse una interfaz clara orientada a profesionales de salud, cuidadores y familiares. Accesibilidad: La aplicación debe cumplir los estándares WCAG, facilitando su uso a personas con limitaciones visuales, motoras o cognitivas.
RNF-04 Fiabilidad	Alta disponibilidad: El sistema debe estar operando con mínimo tiempo fuera de servicio. Tolerancia a fallos: Debe contar con mecanismos de recuperación ante errores, pérdida de conexión o fallos en el servidor.
RNF-05	Código sostenible: El sistema debe implementarse con buena documentación, modularidad y

Mantenibilidad	estandarización. Actualizaciones continuas: Debe permitir aplicar actualizaciones para mejora del modelo IA, de la seguridad o funcionalidad sin afectar la operación.
RNF-06 Integración	Interoperabilidad médica: El sistema debe ser compatible con formatos y protocolos estándar en imágenes médicas (DICOM, NIfTI). APIs documentadas: Se deben proveer APIs seguras y estables para la integración con sistemas médicos externos, repositorios o servicios hospitalarios.

3.2.3 Historias de Usuario (HU)

Las Historias de Usuario permiten capturar las necesidades reales de los usuarios de forma clara, simple y orientada al valor. Cada historia describe una funcionalidad desde la perspectiva del usuario, facilitando el trabajo incremental dentro de cada sprint del marco ágil SCRUM.

En la Tabla 19 se presentan las Historias de Usuario priorizadas para el desarrollo de la aplicación móvil de apoyo al diagnóstico de la enfermedad de Alzheimer mediante imágenes médicas e inteligencia artificial.

Tabla 19

Historias de Usuario para el desarrollo de la aplicación móvil

Identificador (ID)	Enunciado de la Historia	Alias	Estado	Dimensión / Esfuerzo	Iteración (Sprint)	Prioridad
HU-001	Como médico especialista, quiero registrar una cuenta en la aplicación proporcionando mis datos, para acceder al sistema y	Registro de usuario	Hecho	12 horas	Sprint 1	Alta

	gestionar mis pacientes.					
HU-002	Como usuario registrado, quiero iniciar sesión con mis credenciales, para acceder de forma segura a mis datos y herramientas de diagnóstico.	Inicio de sesión	Hecho	12 horas	Sprint 1	Alta
HU-003	Como usuario autenticado, quiero cerrar sesión cuando termine, para garantizar la privacidad y seguridad de mis datos y los de mis pacientes.	Cerrar sesión	Hecho	8 horas	Sprint 1	Media
HU-004	Como profesional de salud, quiero cargar imágenes de resonancia magnética cerebral, para que el sistema las procese con el modelo de clasificación.	Carga de imágenes	Hecho	14 horas	Sprint 2	Alta
HU-005	Como médico, quiero que el sistema procese las imágenes cargadas mediante IA, para obtener una clasificación del estado potencial de Alzheimer.	Procesamiento IA	Hecho	16 horas	Sprint 3	Alta

HU-006	Como usuario, quiero visualizar los resultados del análisis de manera clara y estructurada, para interpretar el nivel de deterioro cognitivo asociado.	Visualizar resultados	Hecho	12 horas	Sprint 3	Alta
HU-007	Como médico, quiero acceder al historial clínico de cada paciente, para evaluar su progreso y comparar evaluaciones anteriores.	Historial de diagnósticos	Hecho	14 horas	Sprint 4	Media
HU-008	Como usuario, quiero acceder a material educativo sobre Alzheimer, para comprender mejor la enfermedad, su progresión y manejo.	Información educativa	Hecho	10 horas	Sprint 4	Media

3.2.4 Sprints

El desarrollo de la aplicación móvil se organizó mediante una planificación incremental basada en sprints, permitiendo al equipo enfocarse en funcionalidades específicas en cada iteración. Este enfoque facilitó la validación continua, la retroalimentación temprana y la construcción gradual del sistema de clasificación de Alzheimer basado en imágenes de resonancia magnética.

3.2.4.1 Sprint 1

El objetivo principal del Sprint 1 fue implementar las funcionalidades base relacionadas con el acceso del usuario al sistema, asegurando el control de autenticación y la interfaz inicial.

Se completaron las Historias de Usuario HU-001, HU-002 y HU-003, además del diseño preliminar de la interfaz y la ejecución de pruebas unitarias para garantizar estabilidad.

Tabla 20

Información del Sprint 1

Actividad	Prioridad
Desarrollo del módulo de registro, inicio y cierre de sesión (HU-001, HU-002, HU-003)	Alta
Diseño inicial de interfaces de usuario	Alta
Ejecución de pruebas unitarias iniciales	Alta
Revisión técnica del avance del sprint	Media

3.2.4.2 Sprint 2

El Sprint 2 se enfocó en establecer la arquitectura del sistema y la estructura de datos necesaria para almacenar imágenes médicas y resultados diagnósticos.

Se definieron los modelos y diagramas ER, así como la creación de las tablas para el manejo de usuarios, imágenes y clasificaciones. También se realizó la maquetación inicial del frontend.

Tabla 21

Información del Sprint 2

Actividad	Prioridad
Diseño de interfaces y pantallas principales (Frontend)	Alta
Elaboración del diagrama ER para la base de datos	Alta
Creación de tablas y relaciones en la base de datos	Alta

Maquetación general del diseño de la aplicación	Media
---	-------

3.2.4.3 Sprint 3

En el **Sprint 3** se comenzó la implementación del backend y la estructura del modelo IA para la clasificación del Alzheimer. Además, se configuraron los endpoints y el despliegue inicial de la API.

Tabla 22

Información del Sprint 3

Actividad	Prioridad
Implementación de endpoints REST para conexión con el frontend	Media
Integración completa del flujo de autenticación en backend	Alta
Despliegue de la API en servidor (Render u otro servicio cloud)	Alta
Obtención, preparación y organización del dataset de resonancias MRI	Media

3.2.4.4 Sprint 4

El Sprint 4 estuvo orientado a la conexión entre el frontend, backend y el modelo IA, así como la validación clínica y funcional del sistema. Se entrenó la IA, se integró el historial de resultados y se realizaron pruebas de aceptación.

Tabla 23

Información del Sprint 4

Actividad	Prioridad
Implementación del historial de diagnósticos en la interfaz	Media
Ejecución de pruebas unitarias y de integración	Alta
Pruebas de aceptación con usuarios (médicos o testers)	Alta
Entrenamiento, validación y comparación entre modelos IA	Alta
Preparación de la documentación técnica para despliegue final	Media

3.3 Diseño de la Aplicación

3.3.1 Frontend

Para el diseño del frontend de la aplicación, se utilizó la herramienta Figma con el objetivo de crear prototipos visuales funcionales y centrados en la experiencia del usuario. Los prototipos reflejan las interfaces clave que guían al usuario a través del proceso de diagnóstico asistido por IA para Alzheimer a partir de imágenes médicas.

En la Figura 21, se muestra la pantalla de inicio de sesión, donde los usuarios pueden ingresar sus credenciales para acceder de forma segura a la plataforma.

La Figura 22 presenta la pantalla de registro, diseñada para que nuevos usuarios puedan crear una cuenta rápidamente y comenzar a utilizar la aplicación.

Una vez iniciada la sesión, el usuario es dirigido a la pantalla principal, mostrada en la Figura 23, donde puede visualizar sus últimas imágenes de resonancia magnética cargadas, así como consultar los resultados anteriores generados por el modelo de IA.

En la Figura 24, se ilustra la pantalla para subir una nueva imagen médica, que permite al usuario cargar imágenes de resonancias cerebrales desde la galería o directamente desde la cámara del dispositivo móvil.

Después de una subida exitosa, la aplicación muestra una confirmación visual, como se aprecia en la Figura 25, garantizando al usuario que la imagen ha sido enviada correctamente para su análisis.

Finalmente, la Figura 26 muestra la pantalla de resultados del diagnóstico, donde se observa de manera clara el resultado del análisis realizado por la IA, indicando la probabilidad de Alzheimer y permitiendo al usuario interpretar fácilmente los datos de salud proporcionados por el sistema.

Figura 21

Pantalla de inicio de sesión



Figura 22

Pantalla de registro de usuario

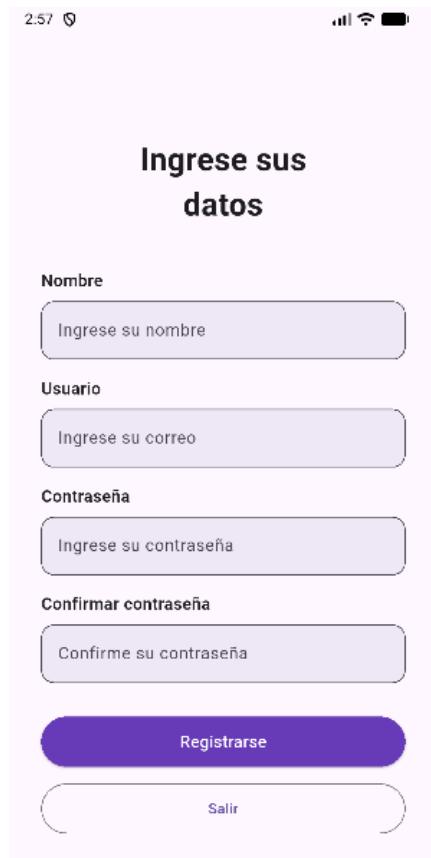


Figura 23

Pantalla principal del aplicativo móvil

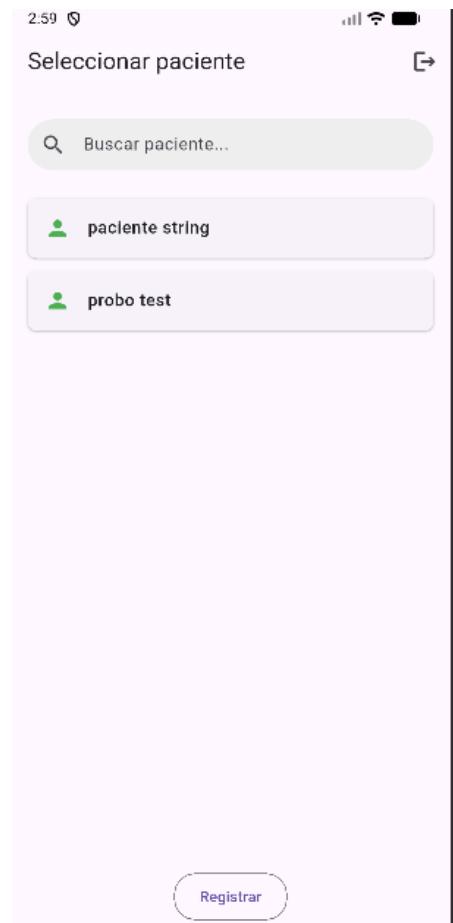


Figura 24

Pantalla de subir nueva muestra de imágenes

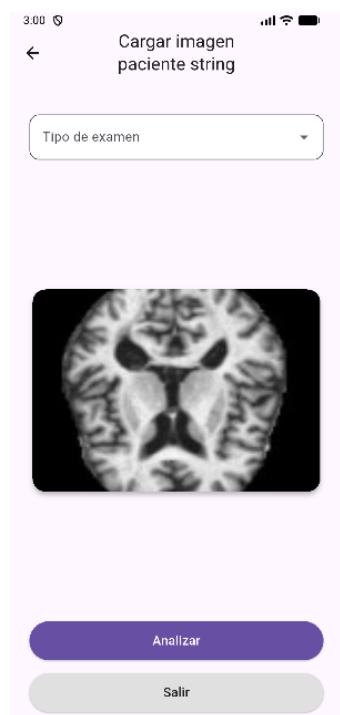


Figura 25

Pantalla de mensaje de carga exitosa



Figura 26

Pantalla de resultado



3.3.2 Backend

En el lado del backend, se utilizó el microframework Flask para el desarrollo de la aplicación. Este framework permitió administrar la base de datos, que está organizada en tres tablas principales:

- **Usuarios:** Contiene la información necesaria para el inicio de sesión de los médicos.
- **Pacientes:** Almacena la información relevante de los usuarios, así como la relación con sus respectivos médicos.
- **Análisis:** Guarda los resultados de las imágenes cerebrales subidas por los usuarios, incluyendo la clasificación de las etapas del Alzheimer.

En la Figura 27 se presenta el modelo de la base de datos utilizado en la aplicación.

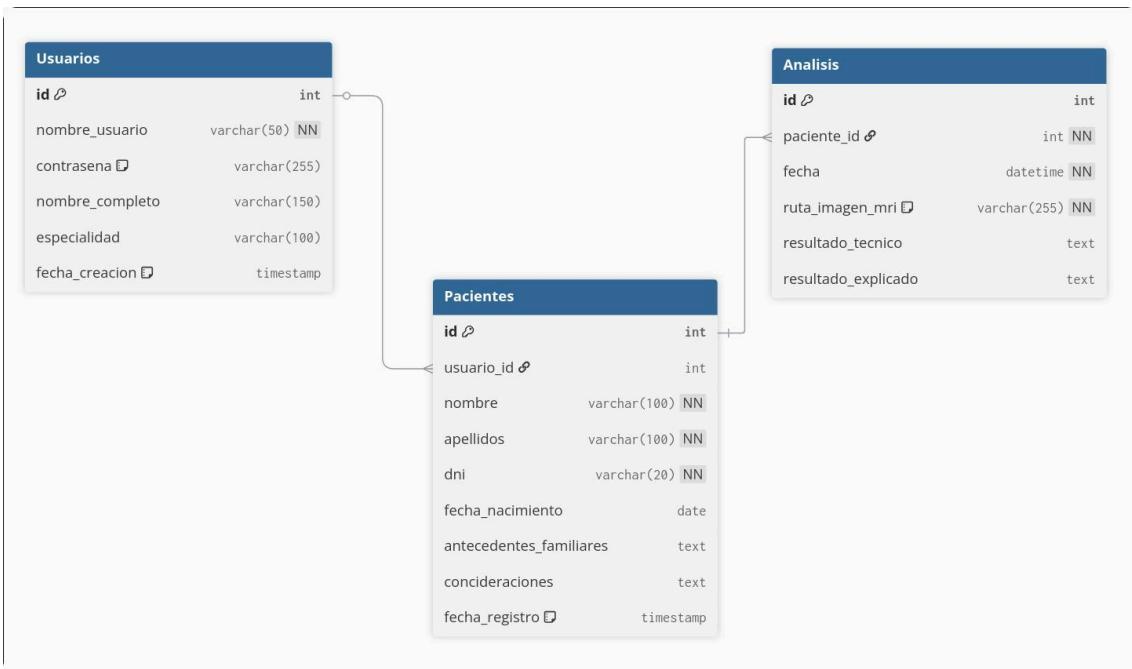


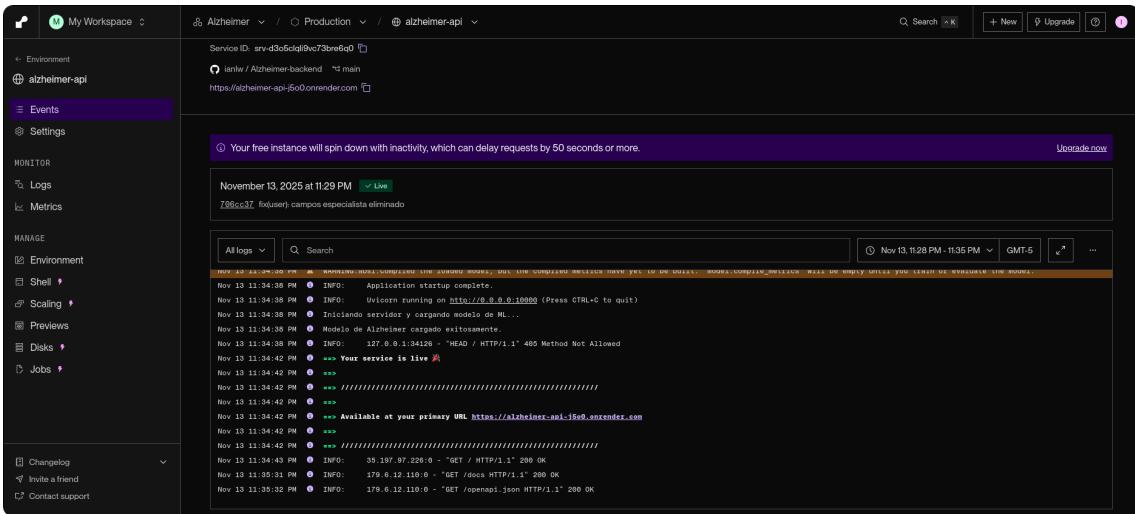
Figura 27

Modelo de la Base de Datos

Además, el código Python, soportado por Flask, permitió la generación y prueba de los endpoints necesarios para el correcto funcionamiento de la aplicación. En la Figura 28, se visualizan las pruebas realizadas a estos endpoints, verificando su capacidad para responder adecuadamente a las solicitudes y asegurar el buen desempeño del sistema.

Figura 29

Backend desplegado y disponible en Render



Finalmente, se realizaron diversas peticiones, como se muestra en la Figura 9, para comprobar la funcionalidad de los endpoints, confirmando que el backend opera de manera óptima conforme a los requisitos establecidos.

Figura 29

Endpoints disponibles en /docs proporcionado por Flask

API para Detección de Alzheimer 1.0.0 OAS 3.1

/openapi.json

Backend para gestionar médicos, pacientes y análisis de MRI.

Authorize

Usuarios y Autenticación

Pacientes y Análisis

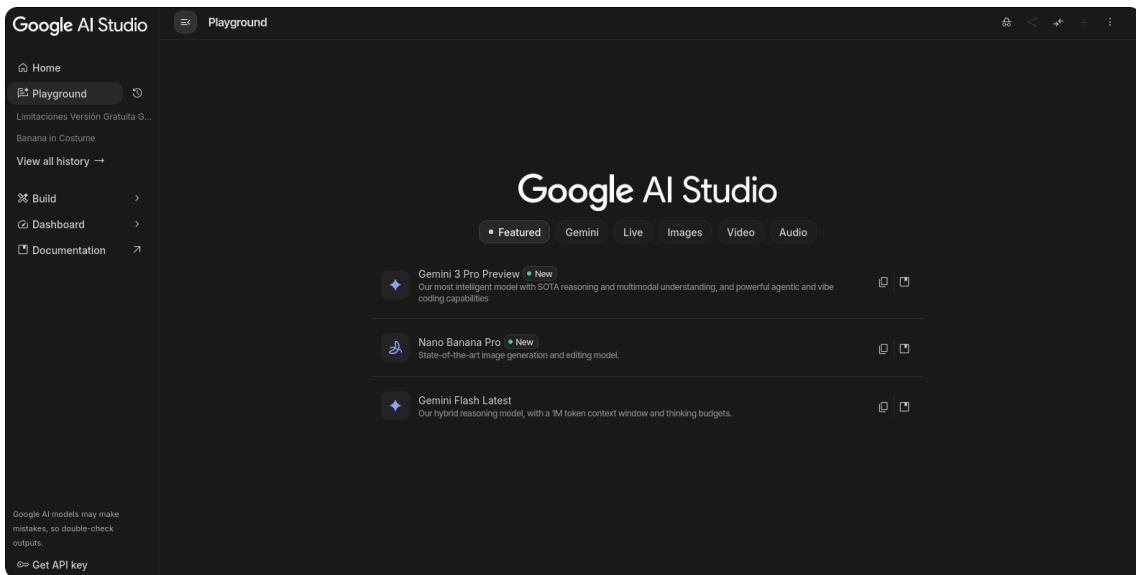
- POST** /pacientes/ Create Paciente
- GET** /pacientes/ Read Pacientes
- POST** /pacientes/{paciente_id}/analisis/ Create Analysis
- GET** /pacientes/{paciente_id}/analisis/ Read Analysis For Paciente

default

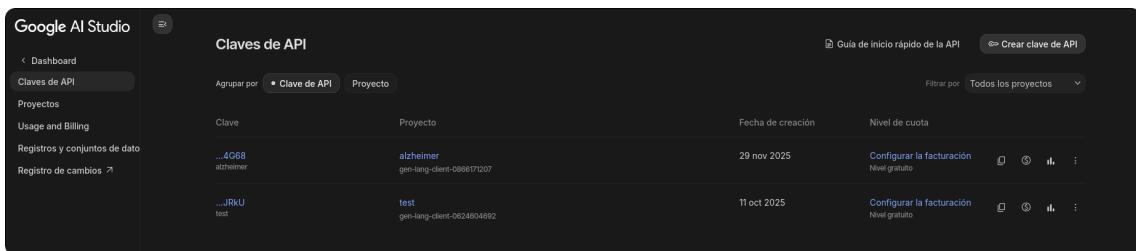
- GET** / Read Root

3.3.3 XAI

Para que el resultado encontrado sea más entendible y mejor entendido, usamos Google Gemini para poder explicar este resultado.



Para poder acceder a Google Gemini usaremos una api y la configuraremos con el modelo gemini-2.5-flash



3.4 Proceso de Prueba de Usabilidad

3.4.1 Objetivos

El objetivo de las pruebas de usabilidad fue evaluar la facilidad de uso, fluidez y comprensión de las funcionalidades principales de la aplicación. En particular, se buscó:

- Verificar si los usuarios pueden realizar tareas esenciales como el registro, inicio de sesión y carga de imágenes médicas.
- Identificar dificultades en la navegación entre módulos, pantallas y acciones.
- Detectar problemas de diseño o interacción que afecten la experiencia del usuario médico o familiar.
- Evaluar la comprensión de los resultados generados por la IA y la lectura del historial clínico.

3.4.2 Tareas y Usuarios Participantes

Las pruebas se realizaron con usuarios simulando el rol de médicos o familiares responsables. A cada participante se le asignaron tareas específicas cubriendo las funcionalidades principales de la aplicación.

La siguiente tabla presenta las tareas asignadas:

Tabla 23

Información del Sprint 4

Tarea	Usuario
Registrar un nuevo médico	Participante 1 (anónimo)
Iniciar sesión en la aplicación	Participante 2 (anónimo)
Cerrar sesión	Participante 2 (anónimo)
Registrar un nuevo paciente	Participante 3 (anónimo)
Cargar una imagen cerebral (MRI)	Participante 4 (anónimo)
Visualizar los resultados del análisis realizado por IA	Participante 5 (anónimo)
Consultar el historial clínico de un paciente	Participante 6 (anónimo)
Total de participantes	6

3.4.3 Metodología

Figura 30

Metodología de Prueba de Usabilidad aplicada en la aplicación móvil



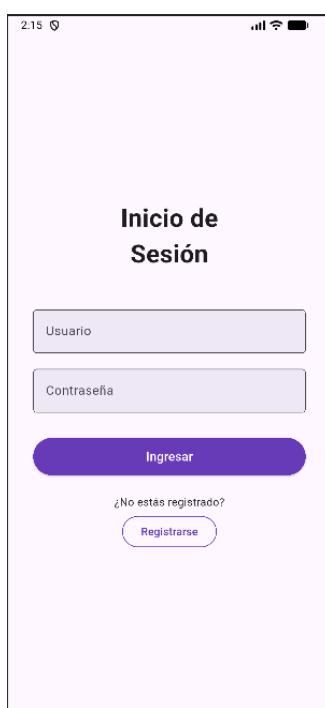
3.4.4 Entorno de Prueba

a) Preparación del Entorno

Una vez completado el desarrollo de las interfaces basadas en los prototipos, se procedió a la preparación del entorno de prueba utilizando un emulador para ejecutar la aplicación. El proceso de prueba comenzó con la pantalla de Inicio de Sesión, como se muestra en la Figura 30.

Figura 30

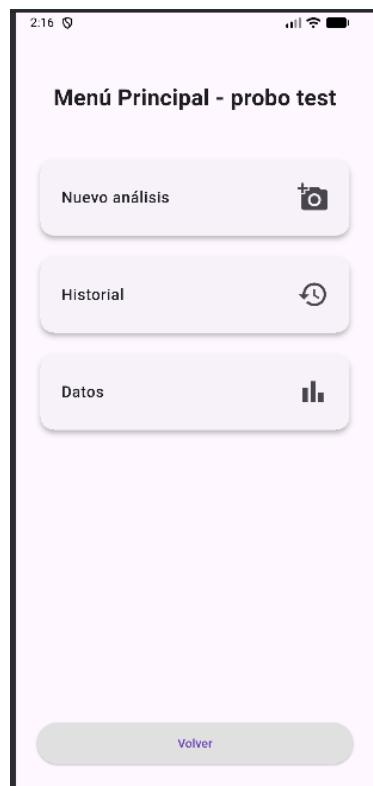
Interfaz de Inicio de sesión



Después de que el usuario haya ingresado sus credenciales, se redirige a la Página Principal, donde puede acceder a las imágenes médicas subidas y a los resultados de sus diagnósticos, tal como se muestra en la Figura 31.

Figura 31

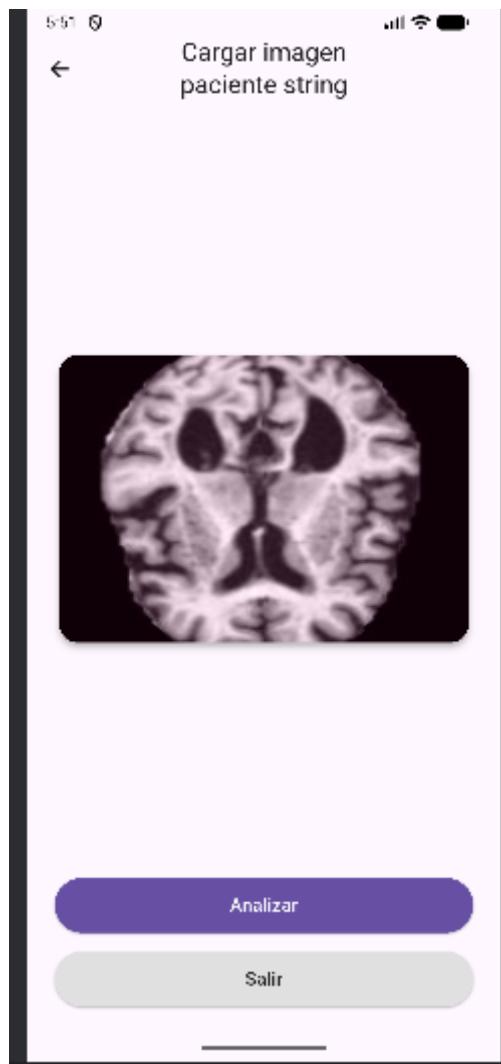
Interfaz Página Principal



A continuación, se procede a la carga de una nueva muestra para la detección de tumores primarios, ilustrado en la Figura 32.

Figura 32

Interfaz de subir nueva muestra de imágenes



Una vez realizada la carga, aparece un mensaje de confirmación, como se observa en la Figura 33.

Figura 33

Interfaz de mensaje de carga exitosa



Finalmente, el usuario podrá revisar el resultado del análisis, los cuales se muestran en la Figura 34. En esta pantalla, se especifica el tipo de tumor identificado en la imagen subida, así como el porcentaje de exactitud del diagnóstico. Si el usuario desea consultar el historial de sus últimos resultados, puede acceder a la pantalla principal, tal como se ilustra en la Figura 31, donde se presenta la fecha y hora de cada análisis para facilitar un mejor seguimiento.

Figura 34

Interfaz de resultado



b) Escenarios de prueba

- **Funcionalidades principales:** Registro e Inicio de Sesión, Carga y análisis de imágenes, Visualización de resultados.
- **Navegación y usabilidad:** Asegurar que los usuarios encuentren fácilmente las funciones y la información que necesitan.
- **Rendimiento y seguridad:** Medir el tiempo que tarda la aplicación en procesar y analizar las imágenes, monitorear el uso de recursos y asegurar que los datos del usuario se manejen de manera confidencial.

c) Pruebas de usabilidad

- Crear escenarios que simulen situaciones reales.
- Registrar métricas como el tiempo de carga, tiempo de análisis.
- Realizar un feedback en tiempo real.

RESULTADOS Y DISCUSIÓN

4. Resultados

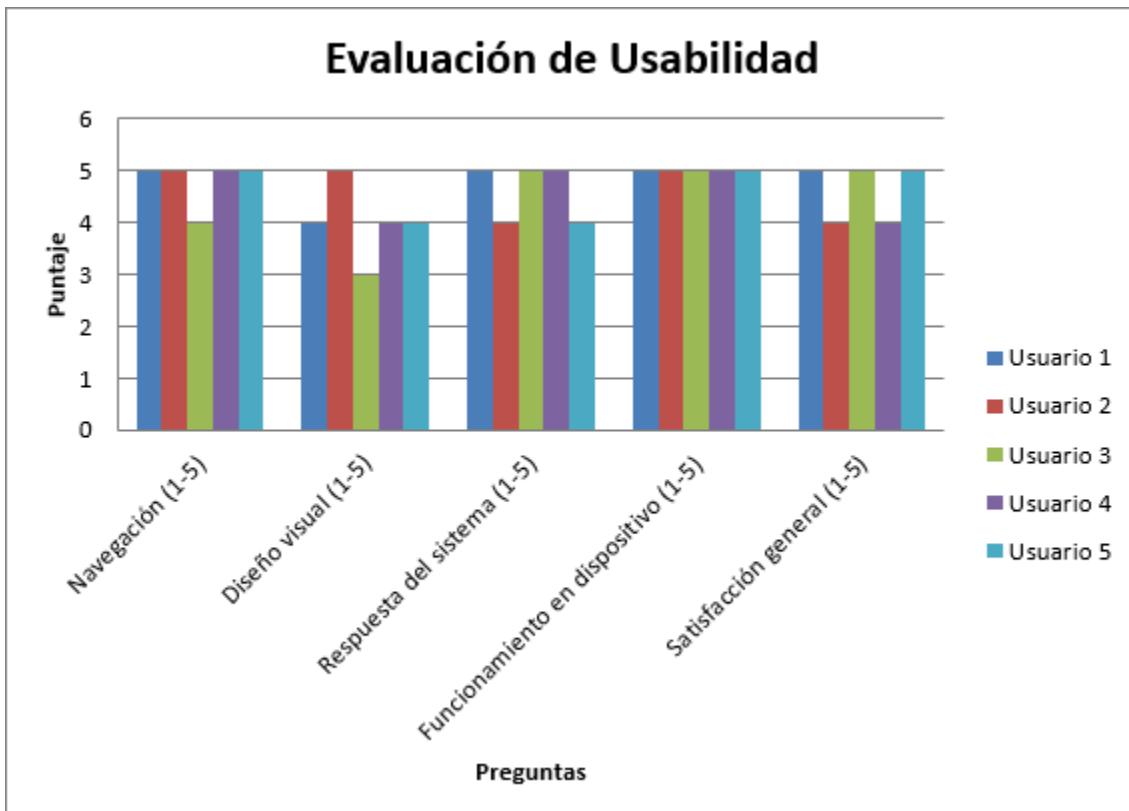
Para evaluar la usabilidad de la aplicación móvil destinada al apoyo en el diagnóstico de Alzheimer mediante imágenes cerebrales (MRI), se aplicó una encuesta a los usuarios participantes durante el proceso de pruebas. La evaluación se centró en aspectos clave como:

- Facilidad de navegación
- Diseño visual
- Respuesta del sistema durante la carga y análisis de imágenes
- Funcionamiento en dispositivos móviles
- Satisfacción general con el uso de la aplicación

Los resultados obtenidos se resumen en la Figura 35, correspondiente a la Evaluación de Usabilidad.

Figura 35

Evaluación de Usabilidad



Para cada integrante de los usuarios participantes se realizó una encuesta para evaluar la usabilidad de la aplicación con las siguientes preguntas.

- **En una escala del 1 al 5, ¿Qué tan fácil fue navegar por la interfaz de la aplicación?**

Todos los participantes calificaron positivamente la navegación, destacando que las pantallas son claras, ordenadas y permiten realizar las tareas sin dificultades.

- **En una escala del 1 al 5, ¿Qué tan atractivo te pareció el diseño visual de la aplicación?**

Los usuarios coincidieron en que la aplicación tiene un diseño visual agradable y minimalista. Se menciona que, aunque cumple con lo necesario, existen oportunidades para incorporar más elementos visuales y estéticos dirigidos a profesionales médicos.

- **En una escala del 1 al 5, ¿Cómo calificarías la respuesta de la aplicación a tus acciones (por ejemplo, al cargar imágenes o al procesar datos)?**

Los participantes no reportaron errores ni demoras importantes. La carga de MRI y el análisis automático mediante IA fueron evaluados como eficientes y satisfactorios.

- **En una escala del 1 al 5, ¿La aplicación funcionó correctamente en tu dispositivo móvil?**

La totalidad de los participantes reportó un funcionamiento óptimo en sus dispositivos, sin cierres inesperados ni problemas de compatibilidad.

- **En una escala del 1 al 5, ¿cuál es tu nivel de satisfacción general con la aplicación? ¿La recomendarías?**

La mayoría expresó satisfacción general con el sistema y aseguró que recomendaría la aplicación tanto por su utilidad clínica como por su facilidad de uso.

Discusión

Los resultados reflejan que la aplicación móvil cumple con los objetivos planteados en términos de: usabilidad, accesibilidad, navegación simple e intuitiva y buen rendimiento en dispositivos móviles

Además, se comprobó que el modelo de IA integrado clasificó correctamente las imágenes cerebrales durante la prueba, demostrando que el entrenamiento y validación previa fueron adecuados.

La precisión obtenida en el modelo permitió identificar correctamente casos de Alzheimer en diferentes grados, lo que se ve reflejado en la satisfacción de los usuarios durante la prueba práctica.

Estos resultados confirman que la aplicación no solo es funcional desde el punto de vista técnico, sino también útil y aceptada desde la experiencia del usuario final.