



Pontificia Universidad
JAVERIANA
Cali

**Facultad de Ingeniería
y Ciencias**

Ingeniería Biomédica

INFORME FINAL DE TRABAJO DE GRADO

Sistema basado en deep learning para el análisis de
imágenes de cápsula endoscópica del intestino delgado.

Juan José Pabón Andrade
Ingrid Nathalia Pantoja Londoño

Director

Cristian Alejandro Torres Valencia

Codirector

Gonzalo Noreña Agudelo

18 de julio de 2025

Santiago de Cali, 18 de julio de 2025

Señores
Pontificia Universidad Javeriana – Cali
Dr. Hernán Camilo Rocha Niño
Decano
Facultad de Ingeniería y Ciencias
Ciudad

Cordial Saludo.

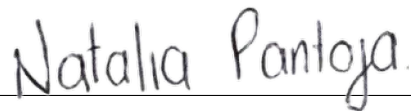
Por medio de la presente nos permitimos presentarle el Trabajo de Grado titulado “Sistema basado en deep learning para el análisis de imágenes de cápsula endoscópica del intestino delgado.”.

Esperamos que este trabajo reúna todos los requisitos académicos, cumpla el propósito para el cual fue creado y sirva de apoyo para futuros proyectos relacionados con la profesión.

Atentamente,



Juan José Pabón Andrade



Ingrid Nathalia Pantoja Londoño

Santiago de Cali, 18 de julio de 2025

Señores

Pontificia Universidad Javeriana – Cali

Dr. Hernán Camilo Rocha Niño

Decano

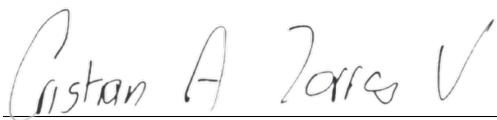
Facultad de Ingeniería y Ciencias

Ciudad

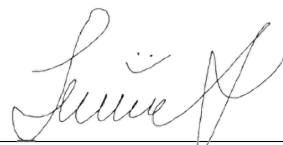
Cordial Saludo.

Certificamos que el presente Trabajo de Grado titulado “Sistema basado en deep learning para el análisis de imágenes de cápsula endoscópica del intestino delgado.”, realizado por Juan José Pabón Andrade y Ingrid Nathalia Pantoja Londoño, estudiantes de Ingeniería Biomédica, se encuentra terminado y puede ser presentado para su sustentación.

Atentamente,



Cristian Alejandro Torres Valencia
Director Trabajo de Grado



Gonzalo Noreña Agudelo
Co-Director Trabajo de Grado

Agradecimientos

Queremos expresar nuestro más profundo y sincero agradecimiento a todas las personas que, de una u otra forma, hicieron posible la realización de este proyecto. Cada palabra de aliento, cada consejo y cada gesto de apoyo se convirtieron en pilares que sostuvieron nuestro recorrido.

Agradecemos a la Pontificia Universidad Javeriana Cali y al Departamento de Ingeniería Biomédica por brindarnos los recursos académicos, técnicos y el ambiente formativo que nos permitió desarrollar este trabajo de manera adecuada.

Extendemos un agradecimiento especial a nuestros asesores académicos, Cristian Alejandro Torres Valencia y Gonzalo Noreña Agudelo, quienes, con paciencia, sabiduría y entrega, supieron guiar nuestras ideas, encauzar nuestros esfuerzos y potenciar nuestras capacidades en cada una de las etapas del proyecto.

Finalmente, agradecemos a nuestras familias por su apoyo constante, comprensión y motivación. Su respaldo fue esencial para culminar con éxito este trabajo académico.

Glosario

Símbolos

- δ_i Tiempo límite de ejecución para la i -ésima tarea [ms].
 τ_i i -ésima tarea.

Acrónimos y Abreviaturas

<i>AI</i>	Inteligencia Artificial (Artificial Intelligence)
<i>CNN</i>	Red Neuronal Convolutacional (Convolutional Neural Network)
<i>DBE</i>	Enteroscopia de Doble Balón (Double Balloon Enteroscopy)
<i>DL</i>	Aprendizaje Profundo (Deep Learning)
<i>GUI</i>	Interfaz Gráfica de Usuario (Graphical User Interface)
<i>IEEE</i>	The Institute of Electrical and Electronics Engineers
<i>IoT</i>	Internet de las Cosas (Internet of Things)
<i>VCE</i>	Endoscopia con Cápsula de Video (Video Capsule Endoscopy)
<i>AUC</i>	Área Bajo la Curva (Area Under Curve)
<i>CE</i>	Cápsula Endoscópica
<i>GPU</i>	Unidad de Procesamiento Gráfico (Graphics Processing Unit)
<i>ReLU</i>	Función de Activación Rectificada Lineal (Rectified Linear Unit)
<i>SGD</i>	Descenso Estocástico del Gradiente (Stochastic Gradient Descent)
<i>TPR</i>	Tasa de Verdaderos Positivos (True Positive Rate)
<i>FPR</i>	Tasa de Falsos Positivos (False Positive Rate)
<i>ROC</i>	Curva Característica Operativa del Receptor (Receiver Operating Characteristic)

Términos

<i>Accuracy</i>	Medida que indica qué tan frecuentemente el modelo acierta en sus predicciones.
<i>Angiectasia</i>	Dilatación anormal de vasos sanguíneos en la mucosa intestinal.
<i>Anotación</i>	Etiquetado manual o automático de datos para su uso en entrenamiento.
<i>Batch size</i>	Cantidad de muestras procesadas antes de actualizar los pesos del modelo.
<i>Cápsula endoscópica</i>	Dispositivo médico en forma de píldora que captura imágenes del tracto gastrointestinal.
<i>CE</i>	Abreviación para cápsula endoscópica.
<i>Clasificación multicategoría</i>	Proceso mediante el cual se asignan imágenes a más de dos clases posibles.
<i>DenseNet</i>	Arquitectura de red neuronal densamente conectada.
<i>Entrenamiento</i>	Proceso donde el modelo ajusta sus parámetros internos a partir de los datos.
<i>Epoca (epoch)</i>	Iteración completa sobre el conjunto de entrenamiento.
<i>Evaluación</i>	Etapas para medir el rendimiento del modelo en datos no vistos.
<i>Función de activación</i>	Función que determina la salida de una neurona dada una entrada.
<i>Función de pérdida</i>	Métrica usada para penalizar errores durante el aprendizaje.
<i>Gradiente</i>	Derivada que indica la dirección del cambio más pronunciado en la función de pérdida.
<i>Hemorragia</i>	Pérdida de sangre en el tracto gastrointestinal.
<i>Hito anatómico</i>	Punto de referencia dentro del intestino.
<i>Ileo</i>	Última porción del intestino delgado.
<i>Linfangiectasia</i>	Dilatación patológica de vasos linfáticos.
<i>Normalización</i>	Transformación de los datos para reducir la varianza y facilitar el aprendizaje.
<i>Overfitting</i>	Fenómeno donde el modelo aprende demasiado los datos de entrenamiento.
<i>Patología</i>	Condición anormal en un órgano o sistema.
<i>Píloro</i>	Esfínter entre el estómago y el duodeno.
<i>Precision</i>	Proporción de verdaderos positivos entre los casos clasificados como positivos.
<i>Preprocesamiento</i>	Conjunto de técnicas aplicadas antes del entrenamiento.
<i>Red neuronal</i>	Modelo computacional inspirado en el cerebro humano.
<i>Regularización</i>	Técnica que reduce el riesgo de sobreajuste al penalizar pesos excesivamente grandes en el modelo.
<i>Validación cruzada</i>	Técnica para evaluar el modelo dividiendo los datos.
<i>Válvula ileocecal</i>	Punto de transición entre el <i>íleon</i> y el colon.
<i>Ventana deslizante</i>	Técnica de muestreo que recorre el conjunto de datos en bloques consecutivos.

Resumen

El presente trabajo describe el diseño y desarrollo de un sistema automatizado basado en deep learning para el análisis de imágenes obtenidas mediante cápsula endoscópica del intestino delgado. El objetivo general consistió en facilitar la identificación, clasificación y localización aproximada de anomalías en dichas imágenes, buscando reducir la carga de trabajo del especialista y mejorar la precisión diagnóstica. Para ello, se seleccionó una base de datos pública que incluye imágenes correspondientes a diferentes patologías y estructuras anatómicas del intestino delgado, las cuales fueron sometidas a un proceso de preprocesamiento para estandarizar su formato y mejorar la calidad de los datos. Posteriormente, se entrenaron y validaron tres modelos de redes neuronales convolucionales, evaluando su desempeño con métricas estándar de clasificación. Los modelos entrenados lograron diferenciar de forma precisa las imágenes según su clase, obteniendo resultados satisfactorios tanto en la detección de patologías como en la identificación de hitos anatómicos, lo que facilitó aproximar la ubicación distal o proximal de dichas imágenes. Finalmente, se integraron los modelos seleccionados en una interfaz gráfica, brindando al usuario la capacidad de interactuar con el sistema de forma intuitiva. Este desarrollo constituye un aporte relevante desde la ingeniería biomédica, al proponer una herramienta de apoyo diagnóstico. Asimismo, se establecen las bases para futuras mejoras relacionadas con la ampliación de la base de datos y la implementación del sistema en escenarios reales de atención médica.

Palabras Clave: cápsula endoscópica, intestino delgado, deep learning, DenseNet, imágenes médicas, clasificación automática.

Abstract

The present work describes the design and development of an automated system based on deep learning for the analysis of images obtained through capsule endoscopy of the small intestine. The main objective was to facilitate the identification, classification, and approximate localization of anomalies in these images, aiming to reduce the specialist's workload and improve diagnostic accuracy. To achieve this, a public database was selected, which includes images corresponding to various pathologies and anatomical structures of the small intestine. These images were subjected to a preprocessing stage to standardize their format and improve data quality. Subsequently, three convolutional neural network models were trained and validated, and their performance was evaluated using standard classification metrics. The trained models successfully differentiated the images according to their class, achieving satisfactory results in both the detection of pathologies and the identification of anatomical landmarks, which allowed for the approximation of the distal or proximal location of the images. Finally, the selected models were integrated into a graphical user interface, enabling users to interact with the system intuitively. This development represents a significant contribution from biomedical engineering, by proposing a diagnostic support tool. Additionally, it lays the foundation for future improvements related to the expansion of the database and the implementation of the system in real-world healthcare settings.

Keywords: capsule endoscopy, small intestine, deep learning, DenseNet, medical imaging, automatic classification.

Índice general

1. Introducción	1
2. Planteamiento del Problema	3
3. Justificación	5
4. Objetivos	7
4.1. Objetivo General	7
4.2. Objetivos Específicos	7
5. Marco de Referencia	9
5.1. Áreas Temáticas	9
5.2. Marco Teórico	9
5.2.1. Gastroenterología	9
5.2.2. Intestino delgado	10
5.2.3. Condiciones fisiológicas y patologías comunes del intestino delgado	10
5.2.4. Panorama epidemiológico y diagnóstico de las patologías del intestino delgado	12
5.2.5. Exámenes Diagnósticos del Intestino Delgado	12
5.2.6. Ingeniería biomédica	15
5.2.7. Deep Learning	16
5.2.8. Redes Neuronales Convolucionales	17
5.3. Trabajos Relacionados	20
5.3.1. Detección automática de diversas anomalías en videos de endoscopia con cápsula mediante un sistema basado en aprendizaje profundo: un estudio multicéntrico	20
5.3.2. Detección y localización de enfermedades gastrointestinales habilitada por aprendizaje profundo utilizando imágenes de endoscopia con cápsula inalámbrica	21
5.3.3. Desarrollo y validación de un sistema de aprendizaje profundo para la detección de patologías del intestino delgado en endoscopia con cápsula: un estudio piloto en una institución de Singapur	22
5.3.4. Comparación de la utilidad clínica de sistemas basados en aprendizaje profundo para la lectura de endoscopia con cápsula del intestino delgado	22
5.3.5. Endoscopia con cápsula de video inalámbrica basada en inteligencia artificial	22
6. Materiales y Métodos	25
6.1. Criterios de Diseño	25
6.2. Equipos e Instrumentación	26

7. Selección y Organización de la Base de Datos	27
7.1. Justificación Clínica de las Clases Seleccionadas	27
7.1.1. Clases de Patologías Seleccionadas	27
7.1.2. Clases de Hitos Anatómicos Seleccionados	28
7.2. Distribución por Categorías y Resolución del Conjunto de Datos	29
7.3. Patologías	30
7.4. Hitos Anatómicos	33
8. Etapa de pre-procesamiento de imagenes	35
8.1. Aumentación de Datos	35
8.2. Evaluación del conjunto aumentado	37
8.3. Conjunto Final Seleccionado	39
8.4. Ajustes Técnicos (Resolución, Color y Normalización)	39
9. Implementación modelos de deep learning	41
9.1. Visión General de las Arquitecturas DenseNet	41
9.2. Selección de Semillas y Procedimiento por Ciclo	43
9.3. Personalización de la Cabeza de Red para Cada Tarea	45
9.3.1. Clasificación de Patologías (Cinco Clases)	45
9.3.2. Detección de Hitos Anatómicos (Clasificación Binaria)	46
10. Validación del rendimiento de los modelos	49
10.1. Dinámica de Entrenamiento con Semillas	49
10.2. Resultados por Modelo y Tarea	50
10.2.1. DenseNet201 – Patologías (Cinco Ciclos)	50
10.2.2. DenseNet169 – Patologías (Cinco Ciclos)	50
10.2.3. DenseNet121 – Patologías (Cinco Ciclos)	51
10.2.4. DenseNet201 – Hitos Anatómicos (Cinco Ciclos)	51
10.2.5. DenseNet169 – Hitos Anatómicos (Cinco Ciclos)	51
10.2.6. DenseNet121 – Hitos Anatómicos (Cinco Ciclos)	52
10.2.7. Resumen Final de Métricas Promedio	52
10.3. Análisis y Justificación de Selección de Modelos	53
10.3.1. Clasificación de Patologías con DenseNet201	53
10.3.2. Detección de Hitos Anatómicos con DenseNet121	54
10.4. Análisis	56
11. Interfaz gráfica de usuario	57
11.1. Fases de diseño de la interfaz gráfica	57
11.1.1. Bocetos y planificación de flujo de navegación	57
11.2. Implementación y estructura de la interfaz gráfica	60
11.2.1. Principios de diseño y herramientas empleadas	60
11.2.2. Estructura de clases de la interfaz	60

Índice general	17
<hr/>	
11.3. Desarrollo visual y flujo de la interfaz	61
11.3.1. Pantallas principales y navegación	61
11.4. Resumen y relevancia clínica	66
12. Conclusiones	67
13. Trabajos futuros	69
14. Anexos	71
Anexos	71
Anexo 1 – Manual de Usuario	71
14.1. Instrucciones de uso	71
14.1.1. Registro de paciente y acceso al menú principal	71
14.1.2. Opciones principales del menú	72
14.2. Guía de solución de problemas	75
Anexo 2 – Mantenimiento del Sistema	77
14.2.1. DenseNet201 – Clasificación de Patologías	77
14.2.2. Resultados por ciclo	78
14.3. DenseNet201 – Detección de Hitos Anatómicos	83
14.3.1. Resumen de métricas promedio (5 ciclos)	83
14.3.2. Resultados por ciclo	83
14.4. DenseNet169 – Clasificación de Patologías	89
14.4.1. Resumen de métricas promedio (5 ciclos)	89
14.4.2. Resultados por ciclo	90
14.5. DenseNet169 – Detección de Hitos Anatómicos	95
14.5.1. Resumen de métricas promedio (5 ciclos)	95
14.5.2. Resultados por ciclo	95
14.6. DenseNet121 – Clasificación de Patologías	101
14.6.1. Resumen de métricas promedio (5 ciclos)	101
14.6.2. Resultados por ciclo	102
14.7. DenseNet121 – Detección de Hitos Anatómicos	108
14.7.1. Resumen de métricas promedio (5 ciclos)	108
14.7.2. Resultados por ciclo	108
14.8. Resumen comparativo y discusión final	114
14.8.1. Resumen Final de Métricas Promedio	114
Bibliografía	115

Índice de figuras

5.1. Olympus EC-10 endocapsule (Fuente: [1])	14
5.2. Arquitectura DenseNet (Fuente: [1])	18
7.1. Distribución original de imágenes por clase en el conjunto Kvasir-Capsule. Fuente: Thambawita et al. [1].	27
7.2. Distribución de la base de datos - patologías (Fuente: [1])	29
7.3. Distribución de la base de datos - hitos anatómicos (Fuente: [1])	30
7.4. Imagen base de datos - Mucosa Normal [1]	31
7.5. Imagen base de datos - Angiectasia [1]	31
7.6. Imagen base de datos - Ulcera [1]	32
7.7. Imagen base de datos - Hemorragia [1]	32
7.8. Imagen base de datos - Linfangiectasia [1]	33
7.9. Imagen base de datos - Píloro [1]	34
7.10. Imagen base de datos - Válvula Ilocecal[1]	34
8.1. a) Lesión ulcerosa rotada 180°: esta operación no altera los bordes característicos ni el patrón de erosión, pero introduce variabilidad en la orientación. b) Muestra de hemorragia con adición de ruido gaussiano y ajustes de brillo/contraste: simula condiciones de baja calidad de imagen para evaluar la robustez del modelo. c) Linfangiectasia sometida a una deformación elástica ligera: recrea variaciones anatómicas suaves en la forma de los vasos linfáticos.	37
8.2. Curvas de pérdida y exactitud en entrenamiento y validación (DenseNet121).	38
8.3. Matriz de confusión para el conjunto de prueba (DenseNet121).	38
9.1. Distribución porcentual de la base de datos — Patologías (70 % entrenamiento / 30 % validación).	44
9.2. Distribución porcentual de la base de datos — Hitos Anatómicos (70 % entrenamiento / 30 % validación).	44
10.1. Curvas de pérdida y exactitud (train/val) del mejor ciclo de DenseNet201 en clasificación de patologías.	53
10.2. Matriz de confusión del mejor ciclo de DenseNet201 en clasificación de patologías.	54
10.3. Curvas de pérdida y exactitud (train/val) del mejor ciclo de DenseNet121 en detección de hitos anatómicos.	55
10.4. Matriz de confusión del mejor ciclo de DenseNet121 en detección de hitos anatómicos.	55
11.1. Pantalla 1 - Boceto: ingreso de datos del paciente	58
11.2. Pantalla 2 - Boceto: selección del tipo de análisis	58

11.3. Pantalla 3 - Boceto: análisis de imagen individual	59
11.4. Pantalla 4 - Boceto: resumen de estudio completo	59
11.5. Pantalla 1 - Registro del Paciente	62
11.6. Pantalla 2 - Menú Principal	62
11.7. Pantalla 3 - Análisis de Imagen Individual	63
11.8. Pantalla 4 - Análisis de Estudio Completo	64
11.9. Galería desplegada al seleccionar la categoría “Linfangiectasia” en el análisis por lote.	65
11.10 Pantalla de historial de imágenes procesadas	65

Índice de tablas

8.1. Parámetros usados en <code>A.OneOf</code> para rotaciones discretas.	36
8.2. Parámetros usados en <code>A.Compose</code> para transformaciones adicionales.	36
8.3. Distribución de imágenes tras aumento de datos (fase inicial).	37
8.4. Métricas de rendimiento para el modelo entrenado con imágenes aumentadas.	39
8.5. Distribución final de imágenes (sin aumento de datos).	39
9.1. Comparativa de variantes DenseNet para clasificación de imágenes médicas.	43
9.2. Configuración de la cabeza de red y parámetros de entrenamiento para clasificación de patologías.	46
9.3. Configuración de la cabeza de red y parámetros de entrenamiento para detección de hitos anatómicos.	47
10.1. DenseNet201 – Clasificación de patologías (Cinco ciclos).	50
10.2. DenseNet169 – Clasificación de patologías (Cinco ciclos).	51
10.3. DenseNet121 – Clasificación de patologías (Cinco Ciclos).	51
10.4. DenseNet201 – Detección de hitos anatómicos (Cinco ciclos).	51
10.5. DenseNet169 – Detección de hitos anatómicos (Cinco ciclos).	52
10.6. DenseNet121 – Detección de hitos anatómicos (Cinco Ciclos).	52
10.7. Resumen Final – Clasificación de patologías.	52
10.8. Resumen Final – Detección de hitos anatómicos.	53

Introducción

La evolución de la medicina moderna ha estado marcada por la incorporación de tecnologías avanzadas que mejoran la precisión diagnóstica y amplían el acceso a regiones del cuerpo humano tradicionalmente difíciles de examinar. En este contexto, la cápsula endoscópica se ha consolidado como una herramienta diagnóstica no invasiva, particularmente útil para explorar el intestino delgado. Este dispositivo, del tamaño de una píldora, es capaz de capturar miles de imágenes a lo largo del tracto gastrointestinal, permitiendo una evaluación visual detallada del interior del cuerpo sin requerir procedimientos invasivos.

A pesar de sus múltiples ventajas, el análisis de las imágenes generadas por la cápsula endoscópica implica una tarea compleja y demandante para los especialistas, quienes deben revisar manualmente este volumen elevado de imágenes. Esta situación puede provocar fatiga visual, demoras en el diagnóstico y una mayor probabilidad de errores u omisiones. Estas dificultades evidencian la necesidad de implementar herramientas tecnológicas que apoyen el análisis clínico y optimicen el proceso diagnóstico.

El enfoque de este proyecto parte de la necesidad de automatizar el análisis de imágenes médicas obtenidas por cápsula endoscópica, integrando técnicas de inteligencia artificial, tales como deep learning. Este enfoque permite procesar grandes volúmenes de datos visuales y extraer patrones complejos, lo cual lo convierte en una alternativa adecuada para apoyar el trabajo de los profesionales de la salud. La ingeniería biomédica, como disciplina que conecta la tecnología con la medicina, proporciona las bases metodológicas y técnicas necesarias para el desarrollo de sistemas que mejoren la calidad de atención y el diagnóstico clínico.

El objetivo general de este proyecto consiste en desarrollar un sistema automatizado basado en deep learning que permita identificar, clasificar y localizar de forma aproximada anomalías en imágenes de cápsula endoscópica del intestino delgado. Para ello, se seleccionó un conjunto de imágenes etiquetadas de dominio público, se diseñó una etapa de preprocesamiento y aumento de datos, y se entrenaron modelos de redes neuronales convolucionales. Finalmente, se integró todo el sistema en una interfaz gráfica de usuario.

La metodología empleada se fundamenta en el desarrollo de modelos computacionales basados en aprendizaje profundo, específicamente mediante arquitecturas DenseNet. Se realizó una etapa de preparación de datos, entrenamiento y validación de los modelos, y se diseñó una interfaz gráfica que permite visualizar los resultados de forma clara e intuitiva. A lo largo del proceso, se aplicaron principios de ingeniería biomédica, procesamiento de imágenes médicas y evaluación computacional.

El presente documento se organiza en distintos capítulos. En primer lugar, se plantea el problema de investigación y se justifica la necesidad del proyecto. Luego se presentan los objetivos, el marco teórico y los antecedentes relevantes. A continuación, se expone la metodología utilizada, incluyendo

la selección de datos, el desarrollo de modelos y el diseño del sistema. Finalmente, se discuten las conclusiones del trabajo y se proponen posibles líneas de investigaciones futuras.

Planteamiento del Problema

La evolución de la ciencia ha impulsado importantes transformaciones en la medicina, especialmente en lo relacionado con la precisión y eficacia en el diagnóstico y tratamiento de enfermedades. Este progreso ha sido posible gracias al desarrollo de nuevas tecnologías que han revolucionado las prácticas clínicas. Un claro ejemplo de ello es la cápsula endoscópica, una innovación destacada en el diagnóstico de enfermedades gastrointestinales. Esta técnica permite a los médicos realizar exámenes detallados y no invasivos, capturando imágenes internas que facilitan la detección de patologías.

La cápsula endoscópica es un dispositivo del tamaño de una píldora y para cumplir su función diagnóstica es ingerida por el paciente y recorre todo el tracto gastrointestinal, capturando entre 50 000 y 60 000 imágenes en el proceso [2]. Esta característica la convierte en una herramienta fundamental para la visualización de áreas del intestino delgado que son difíciles de examinar mediante métodos tradicionales, como la endoscopia convencional o los estudios por imagen. Las imágenes capturadas por la cápsula se transmiten a un dispositivo externo para ser analizadas posteriormente por un gastroenterólogo, tarea que, aunque crucial, demanda una cantidad de tiempo considerable.

La importancia de la cápsula endoscópica radica en su capacidad para detectar patologías que podrían pasar desapercibidas con otros métodos diagnósticos. Los resultados obtenidos a partir de este examen son importantes para el diagnóstico preciso que puede conducir a tratamientos acertados de diversas enfermedades gastrointestinales, como hemorragias internas, úlceras, tumores, pólipos y enfermedades inflamatorias del intestino. Sin embargo, la efectividad de la cápsula depende en gran medida de la precisión en el análisis de las imágenes capturadas, donde la fidelidad en la interpretación de las imágenes depende de la capacidad del médico para diagnosticar correctamente las afecciones [3].

Un diagnóstico temprano y preciso mejora significativamente las tasas de supervivencia y la calidad de vida de los pacientes. No obstante, el proceso de revisión manual de estas imágenes es extremadamente laborioso y presenta para los gastroenterólogos varios desafíos que pueden llevar a errores de interpretación en las imágenes y aumentar el riesgo de diagnósticos incorrectos o tardíos, con lo que se incrementa la tasa de exámenes incompletos llegando al 20 – 30% [4]; de ahí la necesidad de buscar alternativas que garanticen precisión, eficacia y eficiencia.

Entre los principales desafíos que representa el análisis manual de las imágenes obtenidas por cápsula endoscópica está la gran cantidad de imágenes capturadas durante un solo examen, lo que implica que los especialistas deben revisar miles de cuadros en busca de posibles anomalías. Este volumen de información puede aumentar la fatiga y la posibilidad de pasar por alto lesiones pequeñas o imperceptibles a simple vista [5].

Por otra parte, la calidad de las imágenes puede verse afectada por factores comunes, como burbujas de aire, sombras fuertes, la presencia de moco o residuos alimenticios; esto puede generar

áreas oscuras o borrosas en las imágenes, ocultando detalles críticos y dificultando la visualización precisa de las lesiones [6]. Además, se pueden dar variaciones en el movimiento del intestino, lo que se convierte en una gran fuente de ruido que complica aún más la tarea del especialista en la validación e identificación correcta de las patologías.

Otro factor importante a tener en cuenta son las sombras profundas y los reflejos generados por las características internas del intestino, como su curvatura y las superficies húmedas que pueden ocultar áreas relevantes o introducir artefactos visuales que distorsionan la percepción de las lesiones. Esto resulta especialmente problemático en zonas con bajo contraste entre el tejido normal y las anomalías, donde los detalles sutiles pueden pasar desapercibidos [7].

Asimismo, la variabilidad en la apariencia de ciertas lesiones, como úlceras o pólipos, puede hacer que su detección dependa en gran medida de la experiencia del médico, lo que aumenta el riesgo de errores humanos [3]. Por ejemplo, lesiones como las angiectasias pueden ser pequeñas y de color similar al tejido circundante, lo que dificulta su identificación en imágenes de baja calidad o afectadas por ruido. En este contexto, el ruido digital inherente a las imágenes puede reducir el contraste y alterar los bordes de las lesiones [8].

Otro desafío crítico en la endoscopia con cápsula es la localización exacta de la patología dentro del intestino delgado. A diferencia de la endoscopia convencional, la cápsula no cuenta con referencias anatómicas fijas a lo largo de su recorrido, por lo que la posición de una lesión se estima generalmente a partir del tiempo de tránsito, calculado entre el paso por el píloro y la válvula ileocecal. Sin embargo, las variaciones en la motilidad intestinal y la velocidad irregular de la cápsula dificultan que este método proporcione una precisión adecuada. Además, la deformabilidad de las paredes y la falta de puntos de referencia constantes pueden generar errores significativos, estimados en varios centímetros, lo que complica la planificación del abordaje terapéutico o la orientación de procedimientos posteriores [9].

La problemática descrita subraya la necesidad de mejorar la eficiencia en el análisis de las imágenes para optimizar los resultados clínicos y obtener diagnósticos más efectivos y acertados, puesto que el impacto de un diagnóstico oportuno y correcto tiene importantes repercusiones en la pronta recuperación, en la salud y calidad de vida de los pacientes; por tanto, es esencial desarrollar métodos que reduzcan la carga de trabajo de los gastroenterólogos, minimicen el riesgo de errores y garanticen precisión en el diagnóstico. Por lo tanto, es fundamental abordar la necesidad de desarrollar un sistema para facilitar el análisis, la clasificación y la localización de anomalías en imágenes obtenidas a partir de exámenes de cápsula endoscópica del intestino delgado, optimizando el tiempo de los especialistas de manera que se avance en la eficiencia de los diagnósticos.

Justificación

La endoscopia por cápsula (VCE) es un procedimiento médico no invasivo que permite examinar el tracto gastrointestinal, especialmente las áreas del intestino delgado que son difíciles de alcanzar con la endoscopia tradicional. En Colombia, la videocápsula endoscópica ha demostrado ser útil en la detección de diversas anomalías en el tracto gastrointestinal, especialmente en el intestino delgado. Estudios indican que entre el 27 % y el 48 % de los casos evaluados con este procedimiento presentan hallazgos clínicamente relevantes, como úlceras, pólipos o cáncer. A nivel nacional, su aplicación se enfoca principalmente en pacientes con hemorragias gastrointestinales de origen oscuro o sospecha de enfermedades inflamatorias, con una incidencia significativa de detección de anomalías en estos contextos [10].

Durante el procedimiento, la cápsula se desplaza naturalmente a través del sistema digestivo y captura numerosas imágenes, las cuales son revisadas, en su mayoría de manera manual, para detectar anomalías como sangrado, inflamación, úlceras o tumores. A pesar de su efectividad como herramienta de diagnóstico de primera línea para enfermedades del intestino delgado, la revisión manual de estas imágenes es una tarea prolongada que puede tomar entre 45 y 90 minutos [8], y es propensa a errores. La precisión del análisis depende en gran medida de la experiencia del revisor, lo que ha resultado en tasas de omisión del 5,9 % para lesiones vasculares, 0,5 % para úlceras y 18,9 % para neoplasias [11].

La detección temprana y precisa de patologías gastrointestinales es crucial para la implementación de tratamientos oportunos y efectivos, lo que puede tener un impacto significativo en la supervivencia y calidad de vida de los pacientes [12]. Sin embargo, el análisis manual de miles de imágenes aumenta el riesgo de diagnósticos incorrectos o tardíos, lo que podría llevar a un tratamiento inadecuado.

De ahí que es evidente la necesidad de incorporar los avances tecnológicos de otras ciencias en el área de la salud para optimizar tanto los resultados diagnósticos como todos los procesos involucrados en el cuidado de la salud, en particular de los diagnósticos realizados a partir de imágenes de cápsula endoscópica del intestino delgado. En consecuencia, emplear tecnologías como el deep learning resulta esencial para mejorar la precisión y la eficiencia en la detección de enfermedades.

El deep learning ha emergido como una herramienta en el campo del análisis de imágenes médicas, mostrando una alta capacidad para identificar patrones complejos y detectar anomalías con una alta precisión y rapidez [13]. Estudios recientes han demostrado que los modelos de deep learning pueden alcanzar una alta precisión en la detección de diversas patologías, convirtiéndose en un enfoque prometedor para mejorar la eficiencia y eficacia del diagnóstico en la endoscopia por cápsula. Al automatizar la clasificación y localización de anomalías en las imágenes generadas por

la cápsula endoscópica, se podría mejorar la precisión diagnóstica y reducir la tasa de exámenes incompletos, optimizando así el proceso de atención al paciente.

La aplicación de deep learning ha mostrado mejoras significativas en la precisión diagnóstica; por ejemplo, el método SSAEIM alcanza una precisión del 98 % en la detección de pólipos [14]. Además, un modelo reciente que combina redes neuronales convolucionales (CNN) y redes de memoria a largo plazo (LSTM) logró una precisión, sensibilidad y especificidad superiores al 95 % para la clasificación de órganos y una estimación precisa de los tiempos de tránsito gástrico e intestinal, incluso en condiciones visuales limitadas [3].

Diseñar un sistema automatizado basado en deep learning no sólo responde a la necesidad de mejorar la precisión y eficiencia en el diagnóstico médico, sino que también tiene el potencial de transformar la práctica clínica al proporcionar una herramienta que minimiza los errores humanos y maximiza la eficacia del diagnóstico y tratamiento. Esta tecnología podría permitir a los profesionales de la salud enfocarse más en el tratamiento y menos en la tarea del análisis manual, mejorando en última instancia los resultados clínicos y la calidad de vida de los pacientes.

Un primer paso esencial para el desarrollo de este proyecto será la recopilación y preprocesamiento de un conjunto de datos amplio y robusto de imágenes del intestino delgado captadas por la cápsula endoscópica; estas servirán como la base fundamental para el entrenamiento y validación del modelo de deep learning. Este conjunto de datos debe ser exhaustivo y representativo, con las patologías correctamente etiquetadas y verificadas clínicamente. Dicho proceso garantizará que el modelo tenga una alta precisión en la identificación y clasificación de posibles anomalías situadas en el intestino delgado. La correcta preparación de los datos es crucial para asegurar que el modelo pueda aprender de manera eficiente y generalizar bien en nuevas imágenes. El rendimiento del sistema automatizado será medido utilizando métricas como precisión, sensibilidad y especificidad, parámetros clave para un análisis diagnóstico.

Por todo lo expuesto, se evidencia la importancia de este proyecto y la necesidad de su desarrollo en el contexto actual de la medicina. La implementación de un sistema automatizado basado en deep learning no sólo optimizará la detección y clasificación de anomalías en imágenes de endoscopia por cápsula, sino que también representará un aporte significativo desde el campo de la ingeniería biomédica. Al mejorar la precisión y eficiencia del diagnóstico, este enfoque contribuirá a una atención médica más efectiva.

Objetivos

4.1. Objetivo General

Desarrollar un sistema automatizado basado en deep learning para facilitar el análisis, la clasificación y localización de anomalías en imágenes obtenidas a partir de exámenes de cápsula endoscópica del intestino delgado.

4.2. Objetivos Específicos

- Seleccionar un conjunto de imágenes de un repositorio de dominio público de cápsula endoscópica relacionadas con diferentes patologías y sus respectivas etiquetas.
- Desarrollar una etapa de preprocesamiento de imágenes de cápsula endoscópica para ajustar las características propias de la imagen, usando estrategias de procesamiento en espacios de color o en escala de grises.
- Implementar modelos de deep learning para la identificación de anomalías y su ubicación aproximada, mediante el uso de imágenes de cápsula endoscópica de intestino delgado.
- Validar el rendimiento de los modelos de deep learning basados en las métricas de desempeño establecidas para los algoritmos de clasificación.
- Diseñar una interfaz gráfica de usuario que integre los modelos de deep learning, permitiendo a los especialistas en el área de gastroenterología visualizar, analizar y explorar los resultados.

Marco de Referencia

5.1. Áreas Temáticas

- Engineering Profession – Research and Development – Biomedical Engineering
- Biomedical Engineering – Clinical Engineering – Gastroenterology Applications
- Biomedical Engineering – Diagnostic Imaging – Small Intestine Pathologies
- Computing and Processing – Artificial Intelligence – Machine Learning
- Computing and Processing – Artificial Intelligence – Medical Image Classification
- Computing and Processing – Pattern Recognition – Automated Diagnosis Systems
- General Topics for Engineers – Evaluation and Testing – Validation Techniques

5.2. Marco Teórico

5.2.1. Gastroenterología

La gastroenterología es una especialidad médica enfocada en el estudio, diagnóstico y tratamiento de las enfermedades del sistema digestivo, que incluye órganos como el esófago, estómago, intestino delgado, colon, recto, páncreas, hígado y vesícula biliar. Esta disciplina atiende diversas patologías, como úlceras, enfermedades inflamatorias intestinales, trastornos hepáticos y cánceres gastrointestinales [15]. Los gastroenterólogos emplean técnicas como endoscopias y pruebas de laboratorio para identificar problemas digestivos, desarrollar planes de tratamiento y, en algunos casos, prevenir complicaciones a través de la detección temprana.

La investigación en gastroenterología es crucial para avanzar en el diagnóstico, tratamiento y prevención de enfermedades del sistema digestivo, que afectan a millones de personas en todo el mundo. Según datos recientes, los trastornos digestivos representan una de las principales causas de morbilidad, con un impacto significativo en la calidad de vida y en los sistemas de salud. Por ejemplo, se estima que las enfermedades inflamatorias intestinales afectan aproximadamente a 6 millones de personas a nivel global, y su incidencia está en aumento en países en desarrollo [16]. Este panorama resalta la necesidad de investigaciones innovadoras que desarrollen técnicas diagnósticas más precisas y terapias personalizadas para reducir la carga de estas patologías y mejorar los resultados clínicos.

La conexión entre la investigación en gastroenterología y el estudio del intestino delgado es esencial, y, a pesar de su importancia, el intestino delgado sigue siendo uno de los segmentos menos explorados del tracto digestivo, debido a su localización y extensión anatómica. Problemas como enfermedades inflamatorias, tumores y trastornos vasculares a menudo se detectan en etapas avanzadas, subrayando la necesidad de encontrar nuevos métodos y enfoques que permitan una evaluación más detallada y oportuna de este órgano.

5.2.2. Intestino delgado

En el cuerpo humano, todos los órganos desempeñan funciones vitales que aseguran el equilibrio y el correcto funcionamiento del organismo. Entre ellos, el intestino delgado es fundamental, ya que interviene directamente en los procesos de digestión y absorción de nutrientes esenciales para el mantenimiento de la salud.

El intestino delgado se divide en tres secciones principales: el duodeno, que conecta con el estómago y es el sitio donde comienza la digestión química de los alimentos; el yeyuno, donde ocurre la mayor parte de la absorción de nutrientes; y el íleon, que lleva los restos no absorbidos hacia el intestino grueso. Estas tres partes trabajan de manera secuencial para descomponer y absorber nutrientes, minerales y vitaminas esenciales. Su recubrimiento interno está compuesto por pliegues y vellosidades que aumentan considerablemente su superficie, lo que facilita la absorción, pero también complica su visualización en procedimientos médicos.

El duodeno es corto y tiene una longitud aproximada de 25 a 30 cm, seguido por el yeyuno y el íleon, que juntos suman entre 5 y 6 m de longitud. Debido a esta gran extensión, su disposición serpenteante y el hecho de que esté suspendido en la cavidad abdominal por el mesenterio, el intestino delgado es difícil de examinar completamente con métodos endoscópicos tradicionales. Además, el intestino delgado no es tan rígido como el colon, lo que permite que sus bucles cambien de forma y posición, aumentando la complejidad para llegar a todas las zonas durante un procedimiento diagnóstico [17].

5.2.3. Condiciones fisiológicas y patologías comunes del intestino delgado

La mucosa del tracto gastrointestinal cumple un papel esencial en el funcionamiento del sistema digestivo, actuando como interfaz entre el medio interno y los contenidos luminales. A continuación, se describe la estructura y función de la mucosa normal en condiciones fisiológicas.

5.2.3.1. Mucosa normal

La mucosa normal del tracto gastrointestinal es la primera línea de defensa contra agresiones químicas, mecánicas y microbiológicas. Está compuesta por un epitelio especializado que varía a lo largo del tubo digestivo, con funciones de absorción, secreción y barrera. En el estómago, la mucosa contiene células parietales que secretan ácido clorhídrico y células principales productoras de pepsinógeno. En el intestino delgado, la presencia de vellosidades y microvellosidades aumenta la superficie de absorción de nutrientes. La integridad de esta mucosa es esencial para la homeostasis

digestiva y su alteración puede dar lugar a enfermedades como gastritis, enfermedad celíaca o colitis inflamatoria [18].

Por otro lado, el intestino delgado puede verse afectado por diversas patologías que alteran su funcionamiento normal. Entre las patologías más comunes encontramos:

5.2.3.2. Angiectasia

La angiectasia es una dilatación anormal y frágil de los vasos sanguíneos en la mucosa del tracto gastrointestinal, lo que puede provocar hemorragias (como se muestra en la siguiente imagen). Esta condición es más común en personas mayores y en pacientes con insuficiencia renal crónica o enfermedades cardiovasculares. Se cree que el estrés hemodinámico repetitivo contribuye a su desarrollo, especialmente en el colon derecho. Los síntomas incluyen anemia ferropénica secundaria a sangrado crónico y episodios de hemorragia oculta o manifiesta. Su diagnóstico suele realizarse mediante cápsula endoscópica [19].

5.2.3.3. Hemorragia (sangre fresca)

La hemorragia es la pérdida de sangre que ocurre dentro de esta sección del tracto gastrointestinal. Este sangrado puede ser lento, llevando a anemia, o rápido, resultando en una hemorragia significativa. Las causas comunes incluyen anomalías vasculares, úlceras, tumores y divertículos. Para diagnosticar este tipo de hemorragia, la cápsula endoscópica se ha convertido en una herramienta esencial, capturando múltiples imágenes; en la parte posterior se muestra una imagen de la patología [20].

5.2.3.4. Linfangiectasia

La linfangiectasia intestinal es un trastorno poco común caracterizado por la dilatación patológica de los vasos linfáticos en la mucosa intestinal, lo que genera una pérdida excesiva de linfa rica en proteínas y grasas. Esta afección puede ser primaria (congénita) o secundaria a obstrucciones del drenaje linfático causadas por procesos inflamatorios, infecciones o neoplasias. Los síntomas incluyen diarrea crónica, edema, ascitis, hipoproteinemia y deficiencias nutricionales debido a la malabsorción. Su diagnóstico se basa en la combinación de hallazgos clínicos, niveles bajos de albúmina y linfocitos en sangre, estudios de imágenes y biopsia intestinal [21].

5.2.3.5. Úlceras

Las úlceras en el intestino delgado son lesiones abiertas que afectan el revestimiento interno de este órgano. Estas lesiones pueden ser causadas por diversas condiciones, incluyendo la enfermedad de Crohn, el uso prolongado de antiinflamatorios no esteroideos (AINEs), infecciones y otras enfermedades inflamatorias. Los síntomas asociados pueden variar desde dolor abdominal y malestar hasta hemorragias internas, que en ocasiones pueden pasar desapercibidas [22].

5.2.4. Panorama epidemiológico y diagnóstico de las patologías del intestino delgado

Se estima que alrededor del 5 al 10 por ciento de la población general en Colombia podría sufrir alguna enfermedad gastrointestinal en algún momento de su vida, y una fracción significativa de estas condiciones afecta directamente al intestino delgado. Entre estas se incluyen angiectasia, úlceras, pólipos, hiperplasia linfoide y alteraciones en la mucosa. Estas patologías pueden manifestarse con síntomas como dolor abdominal, hemorragias o alteraciones metabólicas [23].

La angiectasia del intestino delgado es una malformación vascular frecuente en adultos mayores, con una prevalencia en el mundo que oscila del 5 al 20 por ciento, especialmente en pacientes con insuficiencia renal o enfermedades cardiovasculares. Esta condición es una de las principales causas de sangrado gastrointestinal oculto y anemia crónica. En Colombia, aunque los datos exactos son limitados, se observa un patrón similar al reportado a nivel internacional [23], [24].

Las úlceras en el intestino delgado, causadas principalmente por infecciones como *Helicobacter pylori*, enfermedades inflamatorias como Crohn o el uso prolongado de antiinflamatorios no esteroideos, tienen una incidencia más baja que en el estómago o el duodeno. A nivel nacional, se estima que aproximadamente el 10 por ciento de los pacientes con úlceras gastrointestinales presentan lesiones en el intestino delgado. Estas úlceras pueden causar perforaciones y hemorragias severas si no se tratan adecuadamente [23].

Los pólipos del intestino delgado son raros, con una prevalencia global del 1 al 2 por ciento. Los pólipos adenomatosos, en particular, representan un riesgo significativo, debido a su potencial de malignización. En Colombia, aunque la información epidemiológica es escasa, se considera que el riesgo aumenta en adultos mayores de 50 años, con factores predisponentes como antecedentes familiares de cáncer gastrointestinal [24], [25].

La sangre en el intestino delgado, ya sea visible o microscópica, es un hallazgo relevante asociado a diversas patologías como angiectasia, úlceras y tumores. Globalmente, se estima que hasta el 10 por ciento de las hemorragias digestivas tienen su origen en esta región del tracto gastrointestinal. En Colombia, las enfermedades relacionadas con sangrado gastrointestinal contribuyen significativamente a la carga hospitalaria, especialmente en adultos mayores [23].

La hiperplasia linfoide, caracterizada por un crecimiento excesivo del tejido linfático, afecta predominantemente a niños y tiene una prevalencia baja en adultos, representando menos del 1 por ciento de las patologías del intestino delgado. Esta condición generalmente está asociada a infecciones transitorias o enfermedades inflamatorias crónicas como la enfermedad celíaca. En Colombia, su diagnóstico es poco frecuente, pero se reconoce como un hallazgo incidental en estudios más amplios [25].

5.2.5. Exámenes Diagnósticos del Intestino Delgado

5.2.5.1. Enteroscopia de doble balón (DBE)

Uno de los métodos más avanzados es la enteroscopia de doble balón (DBE, por sus siglas en inglés). Este procedimiento implica la inserción de endoscopios flexibles utilizando un sistema de

dos balones que se inflan y desinflan de manera alternada, lo que permite un movimiento controlado hacia adelante y hacia atrás a través del intestino delgado. A diferencia de las técnicas tradicionales, la DBE permite explorar segmentos más profundos, brindando imágenes detalladas y facilitando la detección de lesiones como angiectasias, úlceras o tumores. Sin embargo, este procedimiento tiene limitaciones importantes: es invasivo, puede durar varias horas y requiere sedación profunda, lo que aumenta el riesgo de complicaciones respiratorias en ciertos pacientes. Además, no siempre logra alcanzar las porciones más distales del intestino delgado, especialmente en casos de anatomías complejas o intestinos particularmente largos. Además, la manipulación intensiva del intestino puede provocar perforaciones o sangrados, en pacientes con patologías preexistentes o anatomías complejas [26].

5.2.5.2. Colonoscopia

La colonoscopia sigue siendo una herramienta fundamental en la evaluación del tracto gastrointestinal, aunque está más orientada a la visualización del colon y del íleon terminal. Este procedimiento, que emplea un endoscopio flexible, permite la detección de patologías como pólipos, tumores o inflamación en la última porción del intestino delgado. Sin embargo, no es adecuada para evaluar las porciones centrales y proximales del intestino delgado, dejando áreas importantes sin explorar. Esto limita su capacidad diagnóstica para condiciones que afectan principalmente estas regiones, como ciertas úlceras, angiectasias o tumores del intestino delgado superior. Además, la preparación intestinal requerida puede ser molesta para los pacientes, y el procedimiento no está exento de riesgos, como perforaciones o sangrados [27].

5.2.5.3. Endoscopia

La endoscopia, permite examinar directamente la mucosa del tracto digestivo mediante la inserción de un tubo flexible con una cámara en su extremo. En el caso del intestino delgado, su alcance suele estar limitado al duodeno, que es la primera sección del órgano, debido a la dificultad para avanzar hacia regiones más distales. Esta técnica es particularmente útil para identificar lesiones en el duodeno, como úlceras, inflamaciones o tumores, así como para tomar biopsias dirigidas en zonas sospechosas; sin embargo, al igual que la colonoscopia, su capacidad para explorar la totalidad del intestino delgado es insuficiente, dejando áreas como el yeyuno y el íleon medio fuera del campo de visión. A pesar de estas limitaciones, la endoscopia sigue siendo una herramienta diagnóstica clave, especialmente cuando se combina con tecnologías complementarias como la enteroscopia de doble balón u otros [26].

5.2.5.4. Cápsula endoscópica



Figura 5.1: Olympus EC-10 endocapsule (Fuente: [1])

La cápsula endoscópica, es un dispositivo que surge como una alternativa frente a las dificultades que presentan los métodos de examen tradicionales, como la endoscopia y la colonoscopia. La cápsula endoscópica es un dispositivo muy pequeño que en su interior contiene una cámara de alta resolución que captura imágenes a medida que la cápsula avanza, impulsada por los movimientos naturales del sistema digestivo. La cámara suele capturar entre 2 y 6 imágenes por segundo, lo cual permite una cobertura detallada del trayecto y un análisis exhaustivo de posibles anomalías [28].

La cápsula, fue introducida en el mercado en el año 2000 por Given Imaging, conocida inicialmente como PillCam™, fue pionera en el diagnóstico no invasivo del intestino delgado [29]. Esta innovación permitió a los médicos observar áreas de difícil acceso sin la necesidad de procedimientos invasivos, revolucionando la endoscopia diagnóstica para detectar sangrados ocultos, lesiones inflamatorias y tumores en el tracto gastrointestinal [29].

Con el paso del tiempo, el uso de la cápsula endoscópica ha aumentado considerablemente debido a su efectividad en la detección de enfermedades gastrointestinales. Estudios recientes destacan que esta tecnología alcanza hasta un 77% de rendimiento diagnóstico en la identificación de afecciones en el intestino delgado [30], esta precisión en el diagnóstico, sumada a su capacidad no invasiva, ha impulsado su integración en más clínicas y hospitales a nivel mundial. Además, su utilidad en el seguimiento de enfermedades crónicas ha hecho que se consolide su rol en la práctica gastroenterológica. En Colombia, la cápsula endoscópica se ha implementado como una herramienta esencial para la evaluación de enfermedades del intestino delgado. Un estudio reciente mostró que esta tecnología se utiliza ampliamente para detectar patologías intestinales como lesiones inflamatorias y tumores, presentando una alta tasa de éxito y pocos efectos adversos [31]. Esto enfatiza su valor clínico en el país, especialmente para la investigación de problemas intestinales complejos, donde otros métodos

diagnósticos son menos efectivos.

5.2.6. Ingeniería biomédica

La medicina moderna se ha beneficiado de forma significativa de los aportes de otras ciencias, respondiendo a los múltiples desafíos que esta tiene que enfrentar permanentemente, de manera especial, se puede mencionar las contribuciones que se realiza desde la ingeniería biomédica. Esta área de la ingeniería se define como una intersección entre los principios de la ingeniería, las ciencias de la vida y la medicina, destinada al diseño, desarrollo e implementación de tecnologías que mejoren la atención médica y la calidad de vida de los pacientes. Su alcance está desde la creación de dispositivos médicos y biomateriales hasta el desarrollo de sistemas avanzados de análisis de datos clínicos, contribuyendo significativamente en la resolución de problemas críticos en el sector salud [32].

Para el objeto de esta investigación ante la problemática identificada se requiere la creación de un sistema automatizado para clasificar imágenes sospechosas en los resultados de la cápsula endoscópica, lo cual representa un reto significativo en el diagnóstico. Sin una herramienta que priorice áreas potencialmente anómalas, los especialistas deben revisar cada imagen individualmente, lo que aumenta la carga de trabajo y eleva el riesgo de errores humanos, especialmente en evaluaciones largas. La automatización ayudaría a reducir estos errores al permitir que el gastroenterólogo se concentre en las áreas críticas, mejorando la precisión y efectividad del diagnóstico.

Por tanto, es de vital relevancia valerse de los conocimientos de la ingeniería biomédica aplicados a este campo, puesto que esta disciplina aporta herramientas y técnicas, como algoritmos de inteligencia artificial y procesamiento de imágenes, que permiten automatizar la clasificación de las anomalías. La implementación de un sistema basado en ingeniería biomédica ayudaría a reducir errores al priorizar imágenes con características sospechosas y mejorar tanto la precisión como la efectividad del diagnóstico.

5.2.6.1. Análisis de Imágenes Médicas

El análisis de imágenes médicas es un campo interdisciplinario que combina técnicas avanzadas de procesamiento digital de imágenes con herramientas computacionales para interpretar datos visuales generados por dispositivos médicos. Estas imágenes incluyen radiografías, resonancias magnéticas, tomografías computarizadas y, más recientemente, imágenes obtenidas mediante cápsulas endoscópicas; este proceso es esencial, considerando que más del 80 por ciento de los datos médicos provienen de imágenes diagnósticas, y su interpretación precisa puede mejorar significativamente la detección y tratamiento de enfermedades [13]. Este análisis tiene como propósito principal mejorar la detección, diagnóstico y monitoreo de diversas patologías, optimizando la precisión y eficacia de los procedimientos clínicos.

La ingeniería biomédica desempeña un papel crucial en este ámbito al desarrollar tecnologías y sistemas que integran hardware médico con software de análisis avanzado. Los ingenieros biomédicos son responsables de diseñar dispositivos como las cápsulas endoscópicas, asegurando que cumplan con los estándares de seguridad, miniaturización y eficacia clínica. Además, trabajan en el desarrollo

de algoritmos que procesan y analizan los datos generados por estos dispositivos, incorporando técnicas de inteligencia artificial como el deep learning para automatizar tareas complejas y reducir errores humanos [13].

El análisis de imágenes médicas también abarca el preprocesamiento de datos para mejorar la calidad de las imágenes y resaltar características relevantes. Este proceso incluye técnicas enfocadas en la reducción de ruido mediante filtros y la normalización de imágenes para asegurar consistencia en los análisis. Estas etapas, diseñadas e implementadas por ingenieros biomédicos, son fundamentales para garantizar que los modelos computacionales puedan interpretar de manera efectiva los datos visuales, mejorando tanto la precisión del diagnóstico como la confianza del médico en los resultados generados por sistemas automatizados [2].

Además de las aplicaciones clínicas, el análisis de imágenes médicas tiene un impacto en la investigación biomédica. Permite explorar nuevas metodologías para identificar patrones asociados con enfermedades emergentes, validar tratamientos y diseñar dispositivos médicos más avanzados.

5.2.6.2. Procesamiento de imágenes endoscópicas

El procesamiento de imágenes endoscópicas es un campo especializado dentro del análisis de imágenes médicas que se enfoca en el tratamiento y análisis de datos visuales capturados por dispositivos endoscópicos. Este procesamiento tiene como objetivo mejorar la calidad de las imágenes, resaltar características relevantes y facilitar la identificación de anomalías en estructuras internas del cuerpo. Los dispositivos endoscópicos, incluidos los endoscopios tradicionales y las cápsulas endoscópicas, generan imágenes de alta resolución que requieren técnicas avanzadas para su interpretación eficiente [11].

Entre los pasos fundamentales en el procesamiento de imágenes endoscópicas se encuentran el preprocesamiento y la mejora de las imágenes, que son cruciales para obtener resultados de calidad que permitan un diagnóstico fiable. En esta etapa, se aplica la normalización del brillo y el contraste para ajustar las condiciones de iluminación heterogénea dentro del tracto digestivo, asegurando una visualización uniforme de los tejidos. Además, la reducción de ruido es esencial, ya que las imágenes capturadas por los dispositivos endoscópicos suelen presentar interferencias causadas por el movimiento del intestino o la presencia de burbujas de aire. Para abordar este desafío, se utilizan filtros como el gaussiano o el mediano, que permiten preservar los bordes y detalles anatómicos mientras eliminan el ruido [13].

Otra corrección importante es la eliminación de artefactos visuales, como sombras fuertes, áreas desenfocadas o residuos alimenticios adheridos a las paredes intestinales. Estos elementos pueden ocultar lesiones críticas, dificultando su detección. También se emplean algoritmos para ajustar las variaciones de color en las imágenes, lo que mejora la percepción de las estructuras internas y facilita la comparación entre diferentes áreas anatómicas [33].

5.2.7. Deep Learning

El deep learning se perfila como una solución prometedora para mejorar la eficiencia en el análisis de imágenes de cápsulas endoscópicas. Gracias a su capacidad para procesar y clasificar grandes

volúmenes de datos visuales, un modelo de deep learning puede identificar patrones asociados con anomalías y priorizar las imágenes que requieren mayor atención. Esto no solo reduciría significativamente el tiempo que el especialista debe dedicar a la revisión, sino que también minimizaría los errores humanos, logrando un diagnóstico más rápido y preciso al automatizar la detección de áreas críticas en el intestino delgado.

El deep learning es una subcategoría del aprendizaje automático que utiliza redes neuronales artificiales con múltiples capas para aprender de grandes cantidades de datos. En el contexto del análisis de imágenes, el deep learning permite que se reconozcan patrones y características complejas en las imágenes sin necesidad de intervención humana directa.

El funcionamiento del deep learning en el análisis de imágenes se basa en el uso de redes neuronales convolucionales (CNN), que son especialmente diseñadas para procesar datos con una estructura en cuadrícula, como las imágenes. Estas redes utilizan capas convolucionales que aplican filtros a las imágenes para extraer características locales, como bordes, texturas y formas. A medida que las imágenes pasan a través de las distintas capas de la red, el modelo aprende a reconocer patrones cada vez más abstractos y complejos [34].

Un kernel puede ser un pequeño conjunto de valores que se superpone a una porción de la imagen y realiza operaciones matemáticas (como la convolución) para resaltar ciertas características, lo que permite que la red aprenda sobre las estructuras presentes en los datos de entrada [35].

Por ejemplo, las primeras capas pueden detectar bordes simples, mientras que las capas más profundas pueden identificar objetos completos o incluso características más complejas relacionadas con diferentes patologías [36]. Este enfoque jerárquico es lo que permite a las CNN ser tan efectivas en el análisis de imágenes médicas, ya que pueden captar desde las características más básicas hasta las relaciones más complejas entre ellas.

Este enfoque ha demostrado ser altamente eficaz en aplicaciones médicas, donde el reconocimiento preciso de características en imágenes, como radiografías, tomografías computarizadas y resonancias magnéticas, es crucial para el diagnóstico y tratamiento de enfermedades. A través del entrenamiento con grandes volúmenes de imágenes etiquetadas, las CNN pueden generalizar su aprendizaje y mejorar continuamente su capacidad para clasificar y detectar diferentes condiciones [37].

5.2.8. Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNNs) son un tipo de arquitectura de deep learning diseñada específicamente para el procesamiento de datos en forma de cuadrícula, como imágenes y videos. A diferencia de las redes neuronales artificiales convencionales (ANNs), las CNNs utilizan capas convolucionales para extraer características espaciales de manera jerárquica, lo que las hace eficientes para tareas de visión artificial [38].

Uno de los principios fundamentales de las CNNs es la operación de convolución, que consiste en aplicar un conjunto de filtros (kernels) sobre la imagen de entrada. Cada filtro es capaz de detectar patrones específicos, como bordes, texturas y formas geométricas, que luego se combinan en niveles superiores para formar representaciones más abstractas. Esta propiedad permite que las

CNNs identifiquen objetos en imágenes con gran precisión [39].

Las CNNs organizan la información en distintos niveles de abstracción. En las primeras capas se identifican características de bajo nivel, como contornos y esquinas. En las capas intermedias se extraen patrones más complejos, como texturas y formas básicas, y finalmente en las capas profundas se reconoce la estructura completa de los objetos, lo que permite clasificaciones avanzadas, como la identificación de rostros o anomalías médicas en imágenes [40].

Una de las principales ventajas de las CNNs es la reducción de la complejidad computacional mediante el uso de parámetros compartidos. En lugar de conectar cada neurona con todas las neuronas de la capa anterior, como ocurre en las ANNs tradicionales, las CNNs reutilizan los mismos filtros en diferentes regiones de la imagen. Esto disminuye la cantidad de parámetros a entrenar y mejora la eficiencia del modelo, haciéndolo más adecuado para aplicaciones en entornos con grandes volúmenes de datos [41].

Otro aspecto clave es la invariancia espacial lograda mediante técnicas de pooling, que reducen la dimensionalidad de la información y hacen que la red sea menos sensible a pequeñas variaciones en la posición de los objetos. El max pooling, por ejemplo, selecciona el valor máximo dentro de una ventana determinada, conservando las características más relevantes y reduciendo la cantidad de información redundante en la red [42].

Las CNNs suelen entrenarse en entornos supervisados utilizando algoritmos de optimización como backpropagation y gradient descent. Sin embargo, el entrenamiento desde cero requiere grandes volúmenes de datos y altos costos computacionales. Para superar esta limitación, se emplea el transfer learning, que permite reutilizar modelos preentrenados, como VGG, ResNet y DenseNet, en nuevas tareas. Esto ha sido útil en el campo del diagnóstico médico, donde modelos entrenados en grandes bases de datos pueden adaptarse para identificar enfermedades en imágenes de rayos X o resonancias magnéticas con precisión [43].

Gracias a estas características, las CNNs han revolucionado diversas áreas, incluyendo la visión por computadora, el análisis biomédico y la conducción autónoma. Su capacidad para extraer y aprender representaciones jerárquicas ha permitido grandes avances en inteligencia artificial, consolidándose como una de las arquitecturas más efectivas en la actualidad.

5.2.8.1. DenseNet: Redes Densamente Conectadas

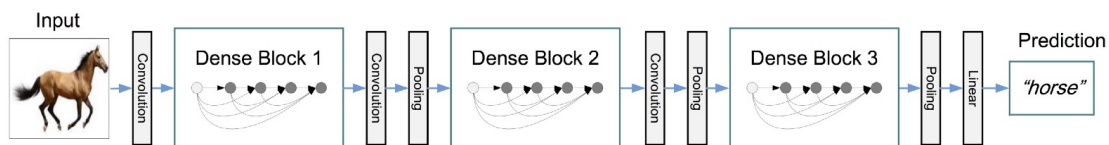


Figura 5.2: Arquitectura DenseNet (Fuente: [1])

DenseNet (Densely Connected Convolutional Network) fue introducida por Huang et al. en 2017, con el objetivo de mejorar la reutilización de características y reducir el número de parámetros en

redes profundas. A diferencia de ResNet, donde las conexiones de atajo suman la entrada a la salida, en DenseNet cada capa está conectada con todas las capas anteriores, lo que se expresa como:

$$\mathbf{x}_l = H_l([x_0, x_1, \dots, x_{l-1}])$$

donde H_l representa la transformación de la capa l y $[x_0, x_1, \dots, x_{l-1}]$ denota la concatenación de todas las características aprendidas en capas anteriores. Esta estructura permite un mejor flujo de información y reutilización de características, lo que reduce la redundancia y mejora la eficiencia del modelo [44].

Una de las ventajas más destacadas de DenseNet es la reducción de parámetros en comparación con arquitecturas convencionales. Al compartir características extraídas en capas anteriores, evita la redundancia de cálculos, lo que permite alcanzar un rendimiento competitivo con menos parámetros y menor consumo de memoria [45]. Esta optimización estructural no solo acelera el entrenamiento del modelo, sino que también lo hace más robusto frente a problemas como el oversize y el sobreajuste en conjuntos de datos de menor tamaño [46].

Además, DenseNet mejora la propagación del gradiente, facilitando el aprendizaje en redes profundas. En arquitecturas tradicionales, los gradientes pueden debilitarse a medida que avanzan por la red, dificultando la actualización efectiva de los pesos. En DenseNet, la conexión directa entre capas permite que los gradientes fluyan con mayor facilidad, lo que resulta en un entrenamiento más estable y eficiente [47].

En términos de aplicaciones, DenseNet ha demostrado un rendimiento sobresaliente en tareas de visión artificial y análisis biomédico. En imágenes médicas, por ejemplo, se ha empleado con éxito en la clasificación de tumores, detección de neumonía en radiografías y diagnóstico de enfermedades oftalmológicas mediante retinografías [48]. Su capacidad para capturar detalles finos en imágenes ha permitido mejorar la precisión en la identificación de patrones anómalos, superando en algunos casos a radiólogos expertos en la detección de patologías específicas [49].

Otra característica clave de DenseNet es su eficiencia en la segmentación de imágenes, particularmente en aplicaciones de segmentación semántica. Modelos como U-Net y variantes de DeepLab han integrado estructuras inspiradas en DenseNet para mejorar la identificación precisa de regiones en imágenes médicas, optimizando el diagnóstico asistido por IA [50].

Gracias a estas ventajas, DenseNet sigue siendo una de las arquitecturas más utilizadas en deep learning, con aplicaciones en reconocimiento de imágenes, visión por computadora y detección de anomalías en imágenes médicas. Su capacidad para maximizar la reutilización de características y optimizar la eficiencia del modelo la convierten en una opción preferida en entornos con limitaciones computacionales o conjuntos de datos reducidos [51].

- **DenseNet-121:** Es la versión más compacta de la familia, con 121 capas distribuidas entre convoluciones y capas completamente conectadas. Debido a su menor número de parámetros, ofrece una solución eficiente para dispositivos con recursos computacionales limitados, como sistemas médicos portátiles o entornos clínicos con hardware estándar. A pesar de su tamaño reducido, mantiene un desempeño competitivo en tareas de clasificación y segmentación de imágenes médicas [52].

- **DenseNet-169:** Posee 169 capas, lo que le permite extraer características más complejas y capturar detalles más finos en las imágenes endoscópicas. Su profundidad adicional contribuye a una mejor generalización en el reconocimiento de anomalías, siendo una opción intermedia entre eficiencia computacional y precisión diagnóstica. Este modelo es ideal para aplicaciones donde se requiere mayor capacidad de análisis sin incrementar excesivamente los tiempos de entrenamiento e inferencia [52].
- **DenseNet-201:** Es la versión más profunda evaluada en este proyecto, con 201 capas, lo que le permite una capacidad superior de extracción de características y detección de patrones complejos en imágenes médicas. Su profundidad hace que sea más adecuado para conjuntos de datos grandes y tareas de diagnóstico donde la precisión es crítica. Sin embargo, su mayor cantidad de parámetros implica un mayor consumo computacional y tiempos de entrenamiento más prolongados, por lo que se evaluará su viabilidad en términos de eficiencia y exactitud [44].

5.3. Trabajos Relacionados

El intestino delgado, por su longitud y anatomía compleja, ha sido una de las regiones más difíciles de explorar con técnicas endoscópicas convencionales, como la colonoscopia o la enteroscopia de doble balón, las cuales presentan limitaciones para alcanzar sus segmentos más distales. Para superar estas barreras, se desarrolló la cápsula endoscópica, un dispositivo autónomo que recorre el tracto digestivo y permite una visualización completa del intestino delgado, facilitando el diagnóstico de patologías antes inaccesibles [53].

En mayo de 2000, la cápsula endoscópica (CE) fue introducida por Given Imaging y aprobada por la FDA, marcando un avance significativo en la exploración no invasiva del intestino delgado [35]. Desde su aparición, ha experimentado una evolución constante, centrada en mejorar la calidad de imagen, la duración de la batería y las capacidades de registro, consolidándose como una herramienta clave en la evaluación de enfermedades del intestino delgado.

Los avances recientes en redes neuronales profundas han optimizado la detección y clasificación automática de anomalías en grandes volúmenes de datos médicos, transformando el análisis en cápsula endoscópica. A continuación, se presentan los desarrollos más relevantes aplicados en este proyecto:

5.3.1. Detección automática de diversas anomalías en videos de endoscopia con cápsula mediante un sistema basado en aprendizaje profundo: un estudio multicéntrico

La introducción de redes neuronales convolucionales (CNN) ha permitido la automatización de la detección y segmentación de anomalías en las imágenes de cápsula endoscópica (CE). Se desarrolló un sistema basado en CNN para la detección de diversas anomalías en las imágenes obtenidas por CE, entrenado con un conjunto de datos de 66 028 imágenes, de las cuales 44 684 correspondían

a anomalías y 21 344 a imágenes normales. Este enfoque fue validado utilizando vídeos de CE de múltiples centros, lo que asegura su aplicabilidad en diferentes contextos clínicos.

El sistema mostró una tasa de detección del 99 %, significativamente superior al 89 % alcanzado por el modo QuickView, una herramienta existente de selección de imágenes. Además, el sistema basado en CNN presentó tasas de detección del 100 % para rupturas de la mucosa y el contenido sanguíneo, del 97 % para angiectasia y del 99 % para lesiones protruyentes, lo que demuestra una alta precisión en la identificación de distintas anomalías gastrointestinales.

Una de las principales ventajas de este sistema es su capacidad para procesar imágenes a alta velocidad, con un tiempo promedio de 0,09 segundos por imagen, lo que permite realizar análisis mucho más rápidos que los métodos manuales. Esta velocidad, junto con la alta precisión en la detección, no solo mejora la eficiencia del diagnóstico, sino que también reduce la carga de trabajo de los gastroenterólogos, minimizando los errores humanos y optimizando el tiempo de lectura. De esta manera, este modelo basado en CNN se presenta como una herramienta prometedora para la detección temprana de patologías gastrointestinales, con un gran potencial para mejorar la práctica clínica [36].

5.3.2. Detección y localización de enfermedades gastrointestinales habilitada por aprendizaje profundo utilizando imágenes de endoscopia con cápsula inalámbrica

Un avance significativo en el campo de la endoscopia por cápsula (CE) ha sido el desarrollo del modelo GINet, una red neuronal convolucional de 13 capas, que ha alcanzado una precisión de clasificación del 99,0 % en imágenes de CE. Este modelo ha sido entrenado y validado utilizando un conjunto de datos de 3 658 imágenes de endoscopia por cápsula, logrando una sensibilidad del 99,6 % y una especificidad del 99,86 %. El GINet ha sido diseñado para detectar enfermedades gastrointestinales (GI) como angiectasia, linfangiectasia, sangrado gastrointestinal y úlceras, demostrando su capacidad para realizar diagnósticos altamente precisos.

Una de las características más destacadas de GINet es su uso de técnicas avanzadas de visualización como GradCAM y Guided-GradCAM. Estas herramientas permiten a los gastroenterólogos interpretar las predicciones del modelo, proporcionando explicaciones visuales de las áreas específicas en las que el modelo se enfoca para tomar sus decisiones. Esta capacidad de interpretación aumenta la confianza de los especialistas en el modelo y mejora la colaboración entre humanos y máquinas en el diagnóstico. La capacidad de localizar lesiones de manera precisa es crucial para que los médicos puedan confirmar el diagnóstico y tomar decisiones de tratamiento con mayor rapidez y precisión [2].

El modelo GINet también ha resuelto un desafío importante relacionado con el análisis de imágenes provenientes de un mismo paciente. A través de técnicas de adaptación de dominio, el modelo mitiga los sesgos que podrían surgir al combinar imágenes de diferentes fuentes, garantizando un diagnóstico sin distorsiones. Esto es esencial en el entorno clínico, donde los exámenes pueden generar grandes volúmenes de datos. GINet no solo mejora la eficiencia en el análisis de imágenes de endoscopia por cápsula, sino que también reduce significativamente el tiempo necesario para realizar

un diagnóstico, lo que a su vez disminuye los costos y mejora los resultados para los pacientes [2].

5.3.3. Desarrollo y validación de un sistema de aprendizaje profundo para la detección de patologías del intestino delgado en endoscopia con cápsula: un estudio piloto en una institución de Singapur

La implementación de modelos que evalúan la calidad de la imagen y detectan anomalías ha mostrado mejoras notables en la eficiencia del análisis. Por ejemplo, un estudio desarrolló un modelo para evaluar la calidad de las imágenes y otro para la detección de anomalías, logrando procesar entre 200 y 250 imágenes por segundo. Estos avances no solo ayudan a reducir el tiempo de diagnóstico, sino que también mejoran la precisión general del proceso. El modelo de evaluación de la calidad de la preparación intestinal, basado en una arquitectura ResNet50 preentrenada, alcanzó un área bajo las curvas ROC y PR de $0,969 \pm 0,008$ y $0,843 \pm 0,041$, respectivamente.

Además, el modelo de detección de anomalías, también basado en ResNet50, mostró una alta precisión con una exactitud top-1 de $84,03 \pm 0,051$ y una exactitud top-2 de $94,78 \pm 0,028$, lo que refuerza la efectividad del sistema en la identificación de anomalías críticas como el sangrado. Estos modelos demuestran el potencial de los enfoques de deep learning para mejorar la eficiencia y la precisión en los flujos de trabajo de endoscopia por cápsula, optimizando los tiempos de diagnóstico y apoyando la toma de decisiones clínicas de manera más efectiva [54].

5.3.4. Comparación de la utilidad clínica de sistemas basados en aprendizaje profundo para la lectura de endoscopia con cápsula del intestino delgado

La comparación de sistemas de CNN existentes con nuevos modelos ha demostrado que los sistemas avanzados no solo mantienen una alta tasa de detección de anomalías, sino que también reducen significativamente el tiempo de lectura y el estrés psicológico para los endoscopistas. En un estudio, un nuevo sistema de CNN logró reducir el tiempo de lectura a 10,1 minutos por paciente, en comparación con 33,6 minutos utilizando métodos tradicionales. Estos avances no solo mejoran la eficiencia en el diagnóstico, sino que también optimizan el proceso, permitiendo que los endoscopistas realicen una lectura más rápida y efectiva.

Este sistema novel también resultó en una reducción significativa del estrés psicológico durante el proceso de lectura. La puntuación promedio de estrés psicológico en una escala de 1 a 5 fue de 1,8 para el nuevo sistema, mientras que en los otros procesos fue de 2,8 y 3,2, respectivamente. Esto demuestra que, además de aumentar la eficiencia, los sistemas de CNN avanzados tienen un impacto positivo en la carga mental de los endoscopistas, mejorando tanto el rendimiento como el bienestar en el contexto clínico [55].

5.3.5. Endoscopia con cápsula de video inalámbrica basada en inteligencia artificial

La introducción de cápsulas endoscópicas con IA avanzada ha marcado un hito en la detección temprana de enfermedades gastrointestinales, como el cáncer colorrectal. Estas cápsulas incorporan

procesamiento de imágenes en tiempo real, técnicas de imagen dual y programación para mejorar la precisión y aceptación del paciente. Las pruebas en modelos in vivo han demostrado la efectividad de estas innovaciones. La integración de algoritmos avanzados permite mejorar la calidad de las imágenes y la capacidad de detectar anomalías, lo que contribuye a una evaluación más precisa y menos invasiva de las condiciones gastrointestinales.

Además, el uso de algoritmos como el SAVE (Spectrum Aided Visual Enhancer), que convierte imágenes de endoscopia convencional en imágenes mejoradas, ha mostrado gran potencial para la visualización de anomalías, como rupturas mucosas y cáncer esofágico temprano. Este algoritmo mejora el contraste y la resolución de las imágenes, facilitando la identificación de áreas anormales y mejorando la capacidad de diagnóstico. Estos avances no solo optimizan la precisión diagnóstica, sino que también proporcionan una experiencia más cómoda para los pacientes al reducir la necesidad de procedimientos invasivos, lo que podría llevar a una mejora significativa en la tasa de detección y tratamiento de patologías [56].

Como se evidencia, diversos trabajos relacionados han demostrado que es posible desarrollar sistemas basados en deep learning capaces de clasificar de manera automática diferentes patologías a partir de imágenes obtenidas mediante cápsula endoscópica. Estos estudios muestran avances significativos en la identificación computacional de lesiones gastrointestinales, contribuyendo a optimizar el proceso diagnóstico y a reducir la carga de trabajo de los especialistas. No obstante, la mayoría de estas soluciones se centran únicamente en la clasificación de las patologías, sin considerar la estimación confiable de su ubicación anatómica dentro del intestino delgado, lo que limita su aplicabilidad clínica de forma integral.

Por esta razón, se reconoce la necesidad de contar con un sistema automatizado que integre dos capacidades fundamentales: por un lado, la clasificación precisa de las distintas patologías que pueden presentarse en las imágenes obtenidas mediante cápsula endoscópica, y por otro, la estimación confiable de su ubicación anatómica dentro del intestino delgado. Esta necesidad surge de la limitación que implica trabajar únicamente con imágenes estáticas sin información de secuencias de video ni referencias espaciales claras, lo que dificulta para los especialistas determinar con exactitud la localización de una lesión identificada. Un sistema que combine ambas funcionalidades no solo permitiría optimizar el proceso diagnóstico al reducir el tiempo y la carga de trabajo del profesional, sino que también aportaría información valiosa para la planificación de intervenciones terapéuticas posteriores. Además, ofrecer una localización aproximada basada en hitos anatómicos clave dentro del intestino delgado puede mejorar significativamente la interpretación clínica, fortaleciendo la confianza del especialista en los resultados generados por herramientas basadas en deep learning. Por ello, el desarrollo de soluciones que articulen la detección de patologías con su ubicación anatómica representa un paso clave hacia la automatización integral del análisis de estudios endoscópicos y la mejora de la calidad de la atención médica en gastroenterología.

Materiales y Métodos

6.1. Criterios de Diseño

El diseño del sistema automatizado para la clasificación de imágenes de cápsula endoscópica del intestino delgado se basa en la implementación de redes neuronales convolucionales (CNNs) con el objetivo de identificar anomalías de manera precisa y eficiente. Para ello, se han establecido criterios específicos que guían el desarrollo del proyecto, garantizando su efectividad y, en futuros trabajos, la aplicabilidad en el ámbito clínico.

El sistema trabaja exclusivamente con imágenes del intestino delgado obtenidas mediante cápsula endoscópica. Se excluirán imágenes de otros segmentos del tracto gastrointestinal, como el esófago, el estómago o el colon, con el fin de optimizar el modelo para la detección de patologías específicas de esta región. Entre las anomalías a identificar se incluyen pólipos, úlceras, angiectasias y la presencia de sangre, las cuales representan condiciones de gran relevancia en el diagnóstico de enfermedades gastrointestinales. Para asegurar la disponibilidad de datos adecuados, se utilizaron imágenes provenientes de bases de datos médicas especializadas, en este caso, Kvasir y Endoscopy Video Database [1], seleccionando únicamente aquellas que cumplan con criterios de calidad y etiquetado preciso.

El desarrollo y entrenamiento del modelo se lleva a cabo en Google Colab, una plataforma en la nube que ofrece acceso a unidades de procesamiento gráfico (GPU) y unidades de procesamiento tensorial (TPU), optimizando el procesamiento de datos y reduciendo el tiempo de entrenamiento. La elección de este entorno se debe a su capacidad para integrar bases de datos externas, ejecutar modelos de deep learning sin depender de recursos computacionales locales y almacenar grandes volúmenes de imágenes médicas en la nube. Esto garantiza la seguridad, disponibilidad y trazabilidad de los datos, facilitando su gestión y evitando la sobrecarga en sistemas locales.

Para evaluar el rendimiento del modelo y garantizar que cumpla con los estándares requeridos en aplicaciones médicas, se emplean métricas de evaluación utilizadas en la clasificación de imágenes. La precisión (accuracy) permitirá conocer el porcentaje total de imágenes correctamente clasificadas, mientras que la sensibilidad (recall) medirá la capacidad del modelo para identificar correctamente las imágenes con anomalías, reduciendo el riesgo de falsos negativos. La especificidad (precision) evalúa la proporción de imágenes diagnosticadas como anormales que realmente contienen una lesión, minimizando falsos positivos.

El diseño del sistema se basa exclusivamente en la arquitectura DenseNet (Densely Connected Convolutional Networks) debido a su capacidad para reutilizar características entre capas y optimizar la propagación del gradiente, lo que mejora su eficiencia sin aumentar excesivamente el número de parámetros. Se evaluaron tres versiones de DenseNet: DenseNet-121, que se caracteriza por su

menor cantidad de parámetros y su rapidez en la inferencia; DenseNet-169, que ofrece un balance entre profundidad de la red y rendimiento en la clasificación; y DenseNet-201, que es la versión más profunda y con mayor capacidad para la extracción de características complejas, aunque con un mayor costo computacional. Cada uno de estos modelos será entrenado utilizando las mismas imágenes y evaluado bajo los mismos criterios de desempeño. Se seleccionará la versión que logre los mejores resultados en la clasificación de anomalías del intestino delgado, asegurando su integración en el sistema final.

6.2. Equipos e Instrumentación

- Computador portátil: Acer Nitro 5 AN515-55-76M1 (Intel Core i5-10750H).
- Plataforma de desarrollo: Google Colab Pro y Visual Studio.
- Base de datos utilizada: Kvasir-Capsule (base de datos de dominio público para imágenes de cápsula endoscópica).

Selección y Organización de la Base de Datos

Como primer paso en el desarrollo del sistema automatizado, se llevó a cabo la selección de un conjunto de datos clínicos provenientes de repositorios públicos especializados. A continuación, se describen el origen, la estructura y las características técnicas de las imágenes utilizadas, así como los criterios de clasificación aplicados a las categorías patológicas y a los hitos anatómicos.

Para el desarrollo del sistema automatizado, se empleó un subconjunto de la base de datos Kvasir-Capsule, la cual fue recopilada en varios hospitales de Noruega mediante cápsula endoscópica (VCE). Esta base original constó de 117 vídeos, de los cuales pudieron extraerse hasta 4 741 504 fotogramas; de ellos, 47 238 fueron previamente etiquetados y validados por expertos médicos, con cajas delimitadoras que abarcan catorce categorías de hallazgos clínicos [1].

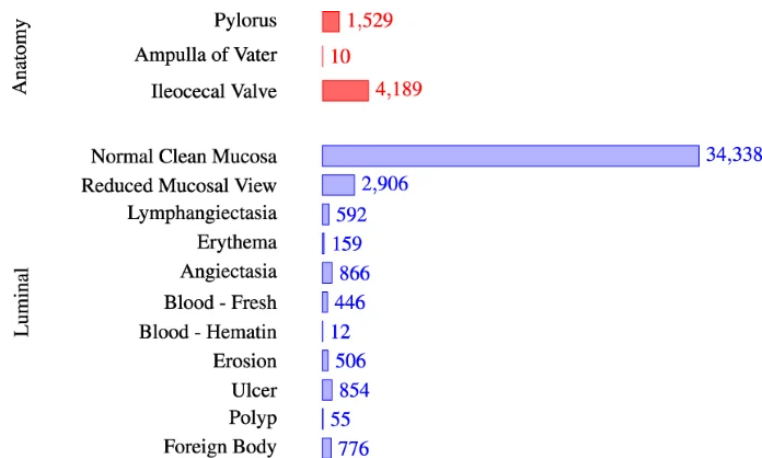


Figura 7.1: Distribución original de imágenes por clase en el conjunto Kvasir-Capsule. Fuente: Tham-bawita et al. [1].

7.1. Justificación Clínica de las Clases Seleccionadas

7.1.1. Clases de Patologías Seleccionadas

De este amplio repositorio se seleccionaron cinco clases de interés clínico para la evaluación del intestino delgado: Mucosa Normal, Angiectasia, Úlcera, Hemorragia y Linfangiectasia. Esta selección

tuvo como propósito entrenar un modelo capaz de distinguir de manera precisa entre tejido sano y distintas formas de lesiones, considerando la variabilidad clínica y técnica (iluminación, ángulo de captura, presencia de artefactos) observada en estudios reales. Además, se tomó en cuenta la frecuencia relativa de cada hallazgo en la población sometida a cápsula endoscópica, así como su importancia diagnóstica para la toma de decisiones terapéuticas.

La inclusión de la clase Mucosa Normal resultó esencial para proporcionar al modelo una referencia clara del aspecto fisiológico del intestino delgado. Esta categoría permitió discriminar entre tejido patológico y sano, y actuó como una clase base que mejoró la especificidad del sistema. En la práctica clínica, la capacidad de identificar correctamente segmentos sin alteraciones es tan importante como detectar lesiones, ya que evita falsos positivos y permite enfocar la atención del especialista en áreas sospechosas. Asimismo, incluir imágenes de Mucosa Normal en el entrenamiento fortaleció la capacidad del modelo para generalizar frente a la gran heterogeneidad del intestino delgado, permitiéndole diferenciar hallazgos reales de artefactos visuales o variaciones anatómicas benignas.

La elección de estas cinco clases respondió también a su prevalencia y valor diagnóstico en la práctica clínica habitual. La clase Mucosa Normal no solo representó el estado fisiológico base, sino que funcionó como contraste esencial frente a cualquier anomalía. Las otras cuatro patologías seleccionadas —Angiectasia, Úlcera, Hemorragia y Linfangiectasia— fueron escogidas por su impacto clínico y por representar un espectro diverso de hallazgos relevantes en la evaluación endoscópica del intestino delgado. La angiectasia es una causa común de sangrado digestivo oculto, especialmente en pacientes de edad avanzada. Las úlceras, por su parte, pueden estar asociadas a enfermedades inflamatorias intestinales, uso de medicamentos o infecciones, y su detección es fundamental para guiar el tratamiento. La hemorragia activa o reciente es uno de los principales motivos de indicación de cápsula endoscópica y debe identificarse con precisión para localizar la fuente de sangrado. Finalmente, la linfangiectasia puede ser un hallazgo de enfermedades sistémicas o procesos locales obstructivos, y su reconocimiento tiene implicancias terapéuticas y pronósticas.

7.1.2. Clases de Hitos Anatómicos Seleccionados

En contextos en los que únicamente se cuenta con imágenes estáticas del intestino delgado, sin información secuencial ni marcas temporales, resulta fundamental identificar claramente los límites proximal y distal a través de hitos anatómicos específicos. Esta estrategia se sustenta en que cualquier imagen etiquetada como *píloro* se asocia con el inicio proximal del intestino delgado (duodeno), mientras que aquella identificada como *válvula ileocecal* representa la delimitación distal del íleon. Esta elección, si bien no exenta de limitaciones, está validada por características anatómicas visualmente reconocibles y consistentemente reportadas en la literatura clínica.

Desde un punto de vista anatómico y clínico, el píloro se caracteriza por su estructura muscular prominente, formando un anillo fácilmente identificable acompañado por pliegues gástricos que desaparecen abruptamente al inicio del duodeno. Por otro lado, la válvula ileocecal presenta una clara transición mucosa, manifestada visualmente como líneas convergentes que marcan el límite entre el íleon y el colon. Estos hitos anatómicos han sido extensamente documentados en protocolos de en-

doscopya gastrointestinal, siendo regularmente utilizados para segmentar espacialmente el intestino delgado incluso cuando se carece de información temporal continua.

Adicionalmente, existen diferencias macroscópicas bien establecidas entre los extremos proximal y distal del intestino delgado que respaldan la validez de esta aproximación. El duodeno y el yeyuno proximal presentan una mucosa con pliegues más pronunciados y una vascularización prominente, en contraste con la mucosa del íleon distal, donde predominan las placas linfoides de Peyer, una vascularización menos evidente y una tonalidad más pálida. Estas características diferenciales permiten inferir con relativa precisión la ubicación aproximada de una imagen aislada dentro del intestino delgado, facilitando así la localización clínica de patologías detectadas.

Es importante asumir y reconocer ciertos sesgos inherentes a esta aproximación, especialmente relacionados con variaciones anatómicas individuales y la presencia de artefactos lumbales como burbujas, contenido intestinal residual o secreciones que pueden afectar la identificación precisa de los hitos. No obstante, considerando bases de datos retrospectivas o escenarios clínicos como la telemedicina, donde la información temporal no está disponible, utilizar estas referencias anatómicas visuales constituye una estrategia válida y ampliamente aceptada en la práctica clínica. Por tanto, a pesar de sus limitaciones, esta metodología representa una alternativa razonable y operativamente efectiva para orientar espacialmente el análisis de imágenes aisladas del intestino delgado.

7.2. Distribución por Categorías y Resolución del Conjunto de Datos

Todas las imágenes utilizadas, tanto las correspondientes a patologías (Angiectasia, Mucosa Normal, Úlcera, Hemorragia y Linfangiectasia) como a hitos anatómicos (el píloro y la válvula ileocecal), presentan una resolución uniforme de 336×336 píxeles, están compuestas por tres canales de color (RGB) y almacenadas como matrices de tipo de dato entero sin signo de 8 bits (uint8). A continuación, se presenta la distribución de imágenes por categoría, diferenciando entre las cinco clases patológicas seleccionadas y los dos grupos correspondientes a la localización anatómica.

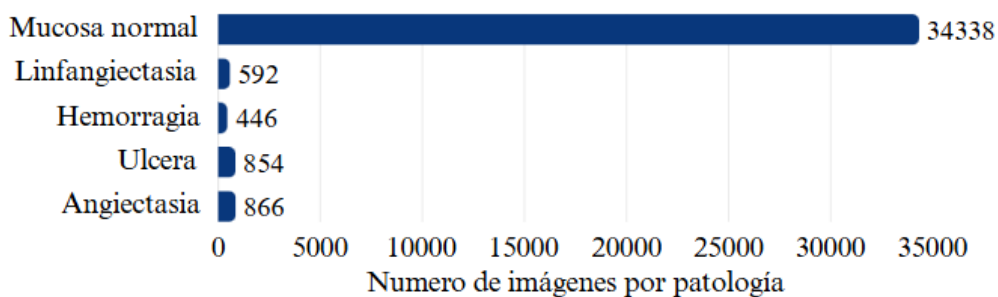


Figura 7.2: Distribución de la base de datos - patologías (Fuente: [1])

Como se muestra en la Figura 1, la clase predominante en el conjunto de datos es Mucosa Normal, con un total de 34.138 imágenes, seguida de Linfangiectasia con 592 imágenes, Úlcera con

854 imágenes, Angiectasia con 866 imágenes y Hemorragia con 446 imágenes.

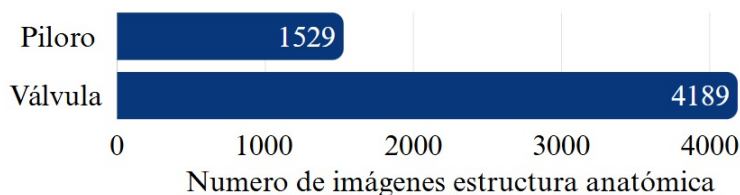


Figura 7.3: Distribución de la base de datos - hitos anatómicos (Fuente: [1])

En lo que respecta a los hitos anatómicos, ilustrados en la Figura 2, se cuenta con 1.529 imágenes del píloro y 4.189 imágenes de la válvula ileocecal, utilizadas específicamente para la tarea de localización anatómica.

A continuación, se presentan ejemplos representativos de cada una de las cinco patologías incluidas en el conjunto, así como de los dos hitos anatómicos considerados. Estas imágenes ilustran la diversidad visual entre clases y justifican la necesidad de abordar de forma diferenciada tanto la clasificación de lesiones como la estimación de su ubicación en el tracto intestinal.

7.3. Patologías

Antes de abordar las distintas patologías del intestino delgado, resulta pertinente incluir una descripción de la mucosa normal. Aunque no se trata de una condición patológica, su caracterización permite establecer una base de comparación que facilita la identificación de anomalías en las imágenes clínicas.

- **Mucosa Normal:** se observa como una superficie delgada y brillante, de color rosa pálido o salmón, sin interrupciones en su continuidad. Los pliegues están bien definidos, dispuestos de forma ordenada, con un patrón repetitivo que denota peristalsis moderada. No hay áreas de engrosamiento, friabilidad ni secreciones anormales; la luz refleja de forma uniforme, permitiendo distinguir claramente las vellosidades y los pequeños vasos submucosos. En ocasiones, se aprecian pequeñas sombras entre los pliegues debido a la inclinación de la cápsula, pero sin distorsionar la estructura general de la mucosa [57].

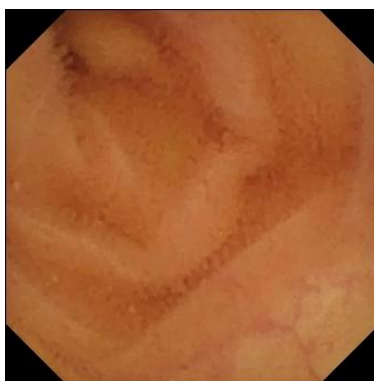


Figura 7.4: Imagen base de datos - Mucosa Normal [1]

A continuación, se encuentra la descripción de las principales patologías que pueden afectar al intestino delgado, destacando sus características clínicas en el contexto del análisis de imágenes endoscópicas.

- **Angiectasia:** se caracteriza por manchas rojas pequeñas o disposiciones lineales de vasos dilatados sobre el fondo de la mucosa, con una tonalidad rojiza intensa que a menudo contrasta sutilmente con el tejido circundante. Estas lesiones pueden presentar bordes difusos y tamaño variable, desde puntos minúsculos hasta agrupaciones ligeramente más extensas. En condiciones de iluminación óptima se aprecia que los vasos sobresalen de la mucosa, mientras que bajo iluminación tenue pueden confundirse con artefactos de reflejo o burbujas. La textura de la zona afectada suele ser más lisa que en otras patologías hemorrágicas, y la coloración uniforme de los puntos o líneas vasculares es un indicio clave [58].



Figura 7.5: Imagen base de datos - Angiectasia [1]

- **Úlcera:** Se presenta como una depresión o cavidad en la mucosa con bordes elevados y de contornos irregulares, cuyo centro suele mostrar un tono grisáceo o amarillento debido a la

presencia de fibrina o tejido necrótico. El área circundante se encuentra enrojecida por el eritema inflamatorio, que se extiende en un halo difuso más allá del borde de la lesión. Según la severidad, la úlcera puede variar desde una pequeña erosión superficial hasta una lesión profunda que expone el submucoso, lo cual se aprecia como una zona más oscura o de color marrón [59] .



Figura 7.6: Imagen base de datos - Úlcera [1]

- **Hemorragia:** se identifica por áreas con acumulaciones de sangre de tonalidad rojo oscuro o brillante, que a veces aparecen dispersas como manchas difusas sobre la mucosa. En imágenes dinámicas, puede percibirse movimiento líquido, donde el sangrado activo forma estelas o se desplaza con la peristalsis. Cuando hay coágulos, estos se adhieren a la pared intestinal y presentan bordes irregulares y textura granulada, de color rojo oscuro a marrón. La presencia de sangre fresca suele opacar parcialmente la vista de la mucosa subyacente, generando un contraste marcado con las zonas circundantes. En casos de sangrado menor, la mucosa puede mostrar pequeños puntos o líneas de sangre, casi imperceptibles a simple vista [60].

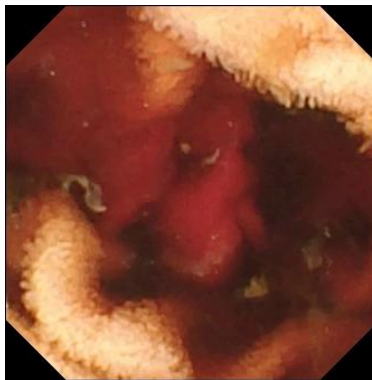


Figura 7.7: Imagen base de datos - Hemorragia [1]

- **Linfangiectasia:** se caracteriza por parches o quistes translúcidos de color blanquecino o amarillento claro, delimitados por una mucosa circundante ligeramente enrojecida debido al edema local. Estas estructuras elevadas pueden variar en tamaño, desde pequeñas vesículas hasta áreas con aspecto de racimo de uvas, y suelen reflejar luz de forma uniforme, dando una apariencia vidriosa. Las regiones adyacentes muestran pliegues suavemente engrosados y, en ocasiones, se observan finas líneas de linfa transparente. Dependiendo del grado de dilatación, la mucosa puede lucir distendida, con pliegues menos marcados y aspecto más liso, lo que en conjunto facilita la identificación de estos quistes como elementos sobreelevados sobre la superficie intestinal [57].

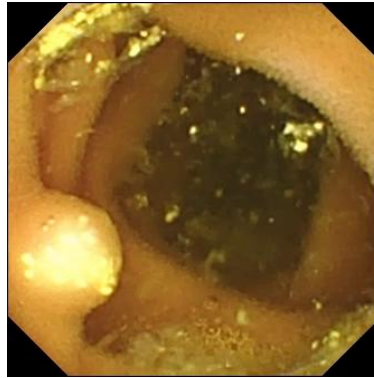


Figura 7.8: Imagen base de datos - Linfangiectasia [1]

7.4. Hitos Anatómicos

- **Píloro:** anillo muscular con luz estrechada, identificado por un pliegue circular pronunciado y sombras marcadas que definen el orificio de salida del estómago. Funciona como referencia proximal en el recorrido de la cápsula, de modo que cualquier anomalía detectada antes o después de este punto se sitúa en la región más próxima al estómago dentro del intestino delgado [61].



Figura 7.9: Imagen base de datos - Píloro [1]

- **Válvula Ileocecal:** unión entre íleon terminal y ciego, reconocible por una transición abrupta en la textura de la mucosa y la presencia de pliegues más anchos y convexos. Sirve como marcador distal, permitiendo determinar si una lesión está ubicada hacia la porción final del intestino delgado, cerca de la entrada al colon [62].



Figura 7.10: Imagen base de datos - Válvula Ileocecal[1]

La organización y caracterización de las imágenes clínicas, tanto patológicas como anatómicas, sentó las bases para el diseño del sistema. Sin embargo, las diferencias marcadas en la cantidad de instancias por clase y la variabilidad visual inherente a las imágenes endoscópicas plantearon desafíos importantes para el entrenamiento de modelos de aprendizaje profundo. Por ello, antes de iniciar la fase de modelado, fue necesario aplicar un proceso cuidadoso de preprocesamiento, orientado a optimizar la calidad, balance y representatividad del conjunto de datos. Esta etapa se describe en detalle en el siguiente capítulo.

Etapa de pre-procesamiento de imágenes

En el proceso de desarrollo de un sistema de clasificación basado en aprendizaje profundo, la etapa de preprocesamiento de las imágenes resulta fundamental para garantizar la calidad y representatividad de los datos de entrada. En particular, cuando se trabaja con conjuntos de datos médicos que presentan un marcado desbalance entre clases—como ocurre en nuestro caso con una gran cantidad de ejemplos en la categoría de Mucosa Normal y pocas instancias en las clases patológicas—es necesario adoptar estrategias que permitan equilibrar la distribución de muestras sin comprometer la información crítica. Por ello, en la fase inicial de preprocesamiento se implementó un esquema de aumento de datos focalizado en las clases patológicas, con el objetivo de proporcionar al modelo una exposición adecuada a cada tipo de lesión, a la vez que se conserva íntegra la representación de los hitos anatómicos utilizados como referencia para la estimación de ubicación.

En la fase inicial de preprocesamiento, se implementó una estrategia de aumento de datos dirigida específicamente a las clases patológicas, con el fin de mitigar el marcado desequilibrio observado en el conjunto de imágenes. Este desbalance se hizo evidente al comparar la clase Mucosa Normal, con más de 34 000 ejemplos, frente a las categorías patológicas, cada una con menos de 900 instancias. Para nivelar la representación de Sangre, Linfangiectasia, Angiectasia y Úlcera, se generaron imágenes sintéticas únicamente en estas clases, preservando la distribución original de los hitos anatómicos (píloro y válvula ileocecal), que quedaron excluidos de este proceso. Esta decisión respondió a la necesidad de exponer al modelo a un número más equilibrado de manifestaciones patológicas, evitando a la vez introducir artefactos en la referencia anatómica, cuya función era servir como punto de anclaje para la estimación de ubicación.

8.1. Aumentación de Datos

El aumento de datos se llevó a cabo empleando la biblioteca **Albumentations**, elegida por su versatilidad y capacidad para aplicar transformaciones visuales complejas de manera eficiente. Esta librería permite definir pipelines modularizados y combinaciones de operaciones que preservan las características clínicas esenciales de las lesiones, al tiempo que generan variaciones suficientes para robustecer el entrenamiento. Dado el marcado desequilibrio entre la clase *Mucosa normal*—con más de 34 000 ejemplos—y las categorías patológicas (cada una con menos de 900 instancias), resultó necesario centrar las técnicas de aumentación en estas últimas, de modo que el modelo recibiera un número de muestras comparable para cada tipo de lesión.

Para ello, se diseñaron dos bloques principales de transformaciones aplicadas exclusivamente a las clases minoritarias (*Sangre*, *Linfangiectasia*, *Angiectasia* y *Úlcera*). En primer lugar, se utilizaron

rotaciones discretas de 90° , 180° y 270° , con probabilidades equilibradas para cada ángulo. Con esta estrategia se busca simular distintos puntos de vista de la misma lesión sin alterar la información diagnóstica: un borde ulceroso, un patrón de sangrado o las características de un vaso linfático dilatado se mantienen legibles independientemente de la orientación. Los parámetros empleados en la función `A.OneOf(...)` de `Albumentations` se resumen en la tabla siguiente:

Transformación	Ángulo	Probabilidad (p)
Rotación Discreta	90°	0.33
Rotación Discreta	180°	0.33
Rotación Discreta	270°	0.34

Cuadro 8.1: Parámetros usados en `A.OneOf` para rotaciones discretas.

A continuación, se definió un segundo bloque de transformaciones aleatorias, con el objetivo de emular condiciones adversas de iluminación, ruido o ligeras variaciones anatómicas que suelen presentarse en la práctica endoscópica. Estas operaciones incluyen ajustes de brillo y contraste para reproducir entornos con iluminación débil o excesivamente brillante; la adición de ruido gaussiano para simular artefactos de movimiento o baja calidad de imagen; deformaciones elásticas suaves para recrear sutiles cambios en la forma del tracto intestinal; y cambios en escala y recorte para presentar fragmentos de la lesión centrados en distintas proporciones del campo visual. Los parámetros usados en esta segunda fase—implementada con `A.Compose(...)`—se detallan en la siguiente tabla:

Transformación	Parámetros	Probabilidad (p)
HorizontalFlip	—	0.50
VerticalFlip	—	0.50
RandomBrightnessContrast	brightness_limit=0.2, contrast_limit=0.2	0.40
GaussNoise	—	0.20
MotionBlur	blur_limit=3	0.20
ElasticTransform	alpha=80, sigma=4	0.20

Cuadro 8.2: Parámetros usados en `A.Compose` para transformaciones adicionales.

Este segundo bloque de transformaciones se configuró de manera que, en cada pasada, se aplique de forma aleatoria una o varias de las operaciones descritas, respetando siempre un rango clínicamente plausible. De esta manera, aunque las imágenes resultantes presenten variaciones en brillo, contraste o deformación, se procura mantener intactos los bordes definidos, los patrones cromáticos característicos de las lesiones (por ejemplo, el color rojo intenso en casos de sangrado) y las texturas propias de cada patología. En ningún caso se aplicaron transformaciones que pudieran generar artefactos irreales—por ejemplo, distorsiones exageradas o cambios de color que modifiquen la naturaleza misma de la lesión—, ya que ello introduciría sesgos adicionales y dificultaría el aprendizaje de rasgos clínicos válidos.

A continuación, se muestran ejemplos visuales de algunas de estas técnicas aplicadas a imágenes representativas de las categorías patológicas incluidas en el entrenamiento:

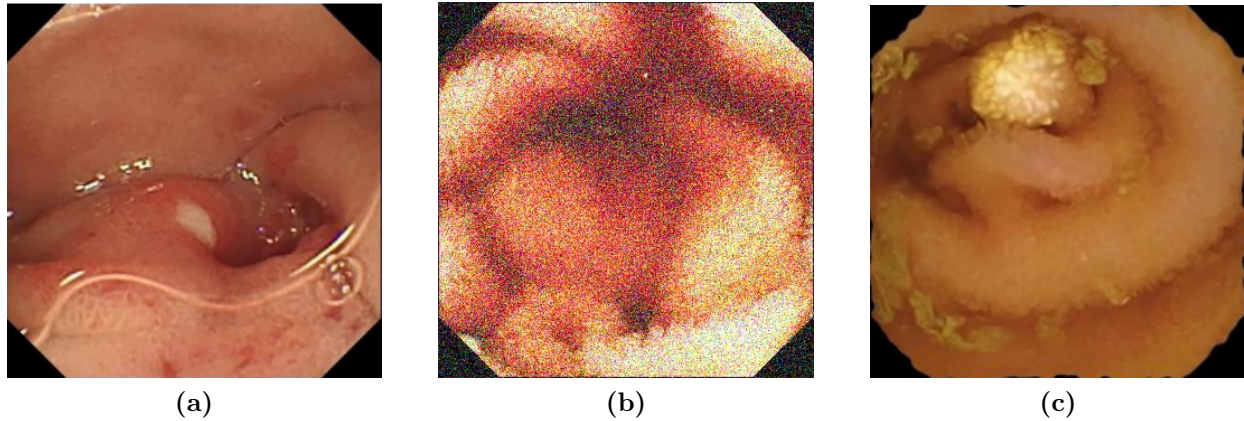


Figura 8.1: a) Lesión ulcerosa rotada 180°: esta operación no altera los bordes característicos ni el patrón de erosión, pero introduce variabilidad en la orientación. b) Muestra de hemorragia con adición de ruido gaussiano y ajustes de brillo/contraste: simula condiciones de baja calidad de imagen para evaluar la robustez del modelo. c) Linfangiectasia sometida a una deformación elástica ligera: recrea variaciones anatómicas suaves en la forma de los vasos linfáticos.

Cada clase patológica se incrementó mediante estas técnicas hasta alcanzar un total de 3 200 imágenes. Para mantener la proporcionalidad del conjunto de datos y evitar la sobresaturación de la clase *Mucosa normal*, se realizó un muestreo aleatorio de 3 200 instancias de dicha categoría (de un total de más de 34 000 disponibles). De esta forma, la siguiente tabla resume la distribución final tras la aumentación inicial:

Cuadro 8.3: Distribución de imágenes tras aumento de datos (fase inicial).

Clase	Cantidad original	Cantidad tras aumento
Mucosa normal	34 338	3 200 (muestreo aleatorio)
Sangre	446	3 200
Linfangiectasia	592	3 200
Angiectasia	866	3 200
Úlcera	854	3 200

8.2. Evaluación del conjunto aumentado

La estrategia de aumento de datos permitió reducir considerablemente el desequilibrio entre clases, proporcionando al modelo un conjunto más balanceado para el entrenamiento. Los resultados obtenidos con este conjunto fueron positivos, y en general, las transformaciones aplicadas resultaron efectivas. No obstante, un análisis más detallado evidenció que algunas imágenes sintéticas presentaban artefactos o distorsiones geométricas poco realistas, lo que podría introducir ruido en el entrenamiento y afectar la interpretación de lesiones reales. Por esta razón, aunque se presentan los resultados del modelo entrenado con datos aumentados, posteriormente se optó por evaluar el

desempeño utilizando únicamente las imágenes originales. Esta decisión se basó en que, además de obtener métricas ligeramente superiores, el trabajo con imágenes reales ofreció una experiencia más coherente con un entorno clínico auténtico.

A continuación, se incluyen las curvas de rendimiento (pérdida y exactitud en entrenamiento y validación), la matriz de confusión y la tabla de métricas (precisión, recall, F1-Score y exactitud) correspondientes al modelo entrenado sobre el conjunto con aumentación de datos. En este experimento se realizó un único ciclo de entrenamiento utilizando la arquitectura DenseNet121 con semilla 42, empleando la misma configuración de hiperparámetros y la misma cabeza de salida que en el ensayo posterior sin aumentación de datos. Posteriormente, se entrenó el mismo modelo sin aplicar ninguna técnica de aumentación y, como se detalla en el capítulo de validación de métricas, el rendimiento sin aumentación resultó ligeramente superior, aunque por un margen reducido.

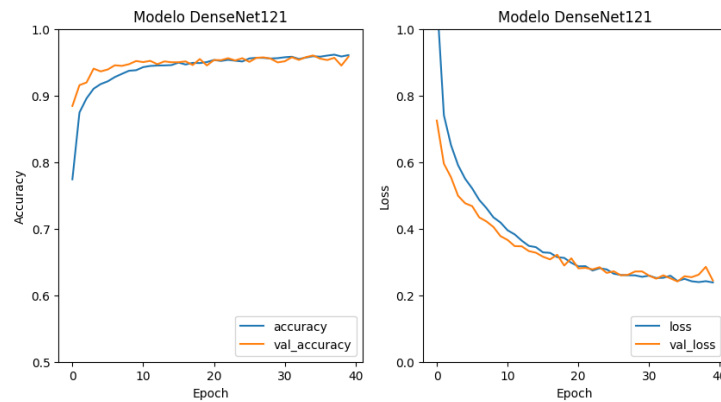


Figura 8.2: Curvas de pérdida y exactitud en entrenamiento y validación (DenseNet121).

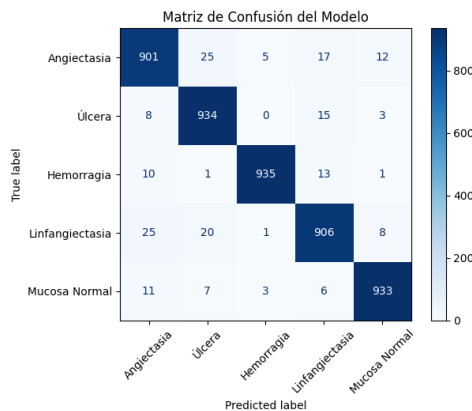


Figura 8.3: Matriz de confusión para el conjunto de prueba (DenseNet121).

Cuadro 8.4: Métricas de rendimiento para el modelo entrenado con imágenes aumentadas.

Configuración	Precisión	Recall	F1-Score	Exactitud
Con aumentación de datos	0.96	0.96	0.96	0.96

8.3. Conjunto Final Seleccionado

Dado que el proyecto tenía como prioridad preservar la calidad diagnóstica de los datos, se decidió descartar el uso de imágenes aumentadas en las fases posteriores. Esta decisión fue reforzada por el hecho de que, al entrenar nuevamente los modelos utilizando únicamente imágenes originales, se obtuvieron resultados más consistentes y clínicamente interpretables. Si bien el conjunto de datos resultante era más reducido y mantenía cierto grado de desbalance, la mejora observada en los resultados validó esta elección como un compromiso razonable entre cantidad y realismo.

En consecuencia, el preprocesamiento final se realizó exclusivamente sobre imágenes originales, manteniendo la integridad visual de las muestras. Para atenuar parcialmente el desequilibrio entre clases sin recurrir a imágenes sintéticas, se seleccionaron aleatoriamente 1 000 imágenes de la clase *Mucosa normal*. Las clases patológicas conservaron su número original de instancias, tal como se muestra en la Tabla 8.5:

Cuadro 8.5: Distribución final de imágenes (sin aumento de datos).

Clase	Cantidad de imágenes
Mucosa normal	1 000
Sangre	446
Linfangiectasia	592
Angiectasia	866
Úlcera	854

Esta configuración final equilibró mejor la fidelidad clínica de las imágenes con la necesidad de representar adecuadamente las clases minoritarias, permitiendo al modelo aprender patrones reales sin depender de distorsiones sintéticas.

8.4. Ajustes Técnicos (Resolución, Color y Normalización)

Desde el punto de vista técnico, las imágenes se conservaron en su resolución original de 336×336 píxeles. Esta decisión respondió a la necesidad de mantener intactos los detalles morfológicos y texturales relevantes para el análisis de lesiones, especialmente en un contexto clínico donde pequeñas variaciones visuales pueden ser significativas para la interpretación diagnóstica.

No se aplicaron técnicas adicionales de reducción de ruido ni algoritmos de realce de bordes, ya que las imágenes originales presentaban una calidad suficientemente alta y sin artefactos relevantes que justificaran intervenciones de este tipo. Se optó por preservar la estructura visual tal como fue

capturada, evitando transformaciones que pudieran alterar o suavizar características importantes para la red neuronal.

Respecto al espacio de color, todas las imágenes se mantuvieron en formato RGB, sin convertir a escala de grises ni a espacios alternativos como HSV o YCbCr. Esta elección se basó en la premisa de que la información cromática podría contener pistas útiles para distinguir entre distintos tipos de lesiones —como variaciones en el tono de la mucosa o presencia de sangrado—, y que eliminarla podría empobrecer la representación semántica del contenido visual.

Finalmente, como paso común en el preprocesamiento de entradas para redes neuronales, se normalizaron los valores de los píxeles al rango $[0, 1]$. Esta operación contribuye a estabilizar los gradientes durante el entrenamiento, mejora la velocidad de convergencia del modelo y reduce el riesgo de inestabilidades numéricas, especialmente en arquitecturas profundas.

Una vez establecida la calidad, balance y formato de las imágenes mediante un preprocesamiento cuidadosamente definido, se procedió a la implementación de los modelos de clasificación. Esta siguiente etapa constituye el núcleo técnico del sistema automatizado, y comprende tanto la elección de arquitecturas de redes neuronales profundas como la configuración de sus parámetros para las tareas de identificación de patologías e hitos anatómicos. En el siguiente capítulo, se describe en detalle este proceso.

Implementación modelos de deep learning

Como parte central del sistema automatizado de análisis de imágenes endoscópicas, en este capítulo se describen en detalle las dos redes neuronales convolucionales (CNN) implementadas: una destinada a la clasificación de patologías y otra enfocada en la detección de hitos anatómicos. A lo largo de esta sección, se explicarán las arquitecturas seleccionadas, el procedimiento de entrenamiento (incluyendo la dinámica de semillas y la distribución de datos), los ajustes específicos para cada tarea y, finalmente, se presentará la base para el análisis de resultados en el siguiente capítulo.

Antes de profundizar en detalles técnicos, se describen las arquitecturas CNN escogidas (DenseNet121, DenseNet169 y DenseNet201) y se justifica por qué resultan adecuadas para cada tarea específica. En segundo lugar, se explica el procedimiento de entrenamiento, haciendo énfasis en la estrategia de semillas, la estratificación de datos y los hiperparámetros seleccionados. En tercer lugar, se detallan los cambios realizados en la capa de salida de cada red para adaptar el modelo a su respectivo problema –ya sea clasificación multiclase o binaria– y, con ello, se sientan las bases para el análisis cuantitativo y cualitativo de resultados que se desarrollará en el Capítulo 10.

9.1. Visión General de las Arquitecturas DenseNet

Para ambas tareas se seleccionaron tres variantes de la familia DenseNet: **DenseNet121**, **DenseNet169** y **DenseNet201**. A continuación se profundiza en sus características estructurales, las razones para su elección y una comparativa entre ellas.

En primer lugar, la característica distintiva de DenseNet radica en que cada capa recibe, como entrada, las salidas de todas las capas anteriores, formando un grafo muy conectado mediante concatenaciones de mapas de características. Este diseño aporta varias ventajas fundamentales. Al concatenar salidas de capas previas se mitiga el problema del *vanishing gradient*, asegurando que las capas iniciales continúen aprendiendo representaciones útiles incluso en redes muy profundas. Cada capa reutiliza mapas de características de bajo y alto nivel, mejorando la capacidad para detectar patrones sutiles, como pequeñas hemorragias o cambios en la mucosa. Además, la concatenación constante ejerce una regularización implícita que reduce la redundancia de parámetros y contribuye a mitigar el sobreajuste, algo especialmente importante cuando se dispone de conjuntos de datos moderados.

Cada *DenseBlock* está compuesto por una serie de unidades básicas en la forma: *BatchNormalization* → ReLU → convolución 3×3 . Tras cada unidad, los mapas de características generados se

concatenan con los mapas existentes, y este proceso se repite dentro del bloque. Entre *DenseBlocks*, las *Transition Layers* aplican una convolución 1×1 (para reducir el número de canales) seguida de un *AveragePooling* 2×2 , controlando así el crecimiento de la red. Esta combinación mantiene bajo el número total de parámetros en comparación con arquitecturas tradicionales de profundidad similar, permite extraer simultáneamente características de diferentes escalas —esencial cuando las lesiones endoscópicas varían en tamaño, forma y textura— y facilita el reuso de información de capas tempranas en niveles más profundos, potenciando la detección de patrones finos, como bordes irregulares o texturas anómalas.

A continuación, se describen brevemente las tres variantes consideradas, resaltando sus fortalezas y limitaciones:

- **DenseNet121**

Consta de cuatro *DenseBlocks* con (6, 12, 24, 16) capas respectivamente, con aproximadamente 8 millones de parámetros y un tamaño de modelo de ~ 33 MB. Su principal ventaja radica en que el tiempo de entrenamiento e inferencia es más reducido y el requerimiento de memoria GPU es menor, lo que lo hace ideal para experimentos iniciales y despliegues en entornos con recursos limitados. Sin embargo, su capacidad de representación es menor en comparación con variantes más profundas, lo cual podría resultar un inconveniente si las diferencias entre clases son muy sutiles.

- **DenseNet169**

Integra *DenseBlocks* de (6, 12, 32, 32), con aproximadamente 14 millones de parámetros y un tamaño de modelo de ~ 57 MB. Ofrece un equilibrio intermedio entre capacidad de representación y complejidad computacional, adecuado cuando se requieren detalles moderadamente complejos sin incrementar excesivamente los tiempos de entrenamiento. Se recomienda cuando se busca un balance entre precisión y velocidad de cómputo.

- **DenseNet201**

Incluye *DenseBlocks* de (6, 12, 48, 32), totalizando cerca de 20 millones de parámetros y un tamaño de modelo de ~ 80 MB. Su mayor profundidad le permite modelar patrones muy sutiles, lo que resulta idóneo cuando las diferencias entre clases son extremas o poco evidentes. Como contrapartida, su mayor consumo de memoria y tiempos de entrenamiento/inferencia elevan los requisitos de hardware para su uso efectivo.

Característica	DenseNet121	DenseNet169	DenseNet201
Número de capas	121	169	201
Parámetros aproximados	8 millones	14 millones	20 millones
Tamaño del modelo	~ 33 MB	~ 57 MB	~ 80 MB

Cuadro 9.1: Comparativa de variantes DenseNet para clasificación de imágenes médicas.

Los tres modelos ofrecen un rango de profundidades y capacidades que permiten evaluar el compromiso entre precisión y eficiencia. En la práctica, los experimentos indicarán cuál variante resulta más estable y precisa para la clasificación de patologías y cuál para la detección de hitos anatómicos.

9.2. Selección de Semillas y Procedimiento por Ciclo

Para evitar que los resultados dependieran de una única división de datos o de una única inicialización de pesos, se definió un esquema de entrenamiento en cinco ciclos por cada arquitectura, utilizando cinco semillas distintas: 42, 123, 2025, 1 y 999. De esta manera, cada variante DenseNet se entrenó cinco veces, una con cada semilla, para ambas tareas (patologías e hitos anatómicos), sumando un total de 30 experimentos exhaustivos.

Cada semilla controla la función `train_test_split` con estratificación por clase, lo que garantiza que el 70 % de las imágenes se destine a entrenamiento y el 30 % a validación, manteniendo la misma proporción de cada etiqueta. Al variar la semilla, cambia la selección puntual de las imágenes dentro de cada subconjunto, simulando ligeros cambios en la base de datos. Asimismo, la semilla también determina la inicialización aleatoria de las nuevas capas densas añadidas a la cabeza de la red, de modo que se evalúa la sensibilidad de la arquitectura a la aleatorización interna. Una variante que muestre baja variabilidad en sus métricas a lo largo de las cinco semillas se interpretará como más estable.

El flujo de cada ciclo se realiza de la siguiente forma: primero, se aplica `train_test_split` con estratificación (70 % entrenamiento / 30 % validación), fijando la permutación de índices mediante la semilla correspondiente para crear subconjuntos balanceados. A continuación, se carga el modelo DenseNet (121, 169 o 201) preentrenado en ImageNet y se reemplaza la cabeza original por la estructura específica de la tarea. Luego, se configuran los hiperparámetros de entrenamiento, tales como *learning rate*, *batch size* (siempre 32) y número máximo de épocas (100); las funciones de pérdida (`SparseCategoricalCrossentropy` para patologías y `BinaryCrossentropy` para hitos); y el optimizador (Adam con la tasa de aprendizaje correspondiente). También se activa el mecanismo de *early stopping* (paciencia = 5, `min_delta` = 0.001, restaurando los mejores pesos). Durante el entrenamiento se registran en cada época las curvas de pérdida (`loss`) y exactitud (`accuracy`) tanto en entrenamiento como en validación, monitoreando la brecha entre ambas para detectar sobreajuste o estancamiento. Finalmente, al concluir el entrenamiento (o interrumpirse por *early stopping*), se evalúa el modelo en el conjunto de validación para obtener las métricas de precisión, *recall* y F1-Score por clase, así como la exactitud global. Estos resultados se almacenan para calcular posteriormente

la media (μ) y desviación estándar (σ) de cada métrica.

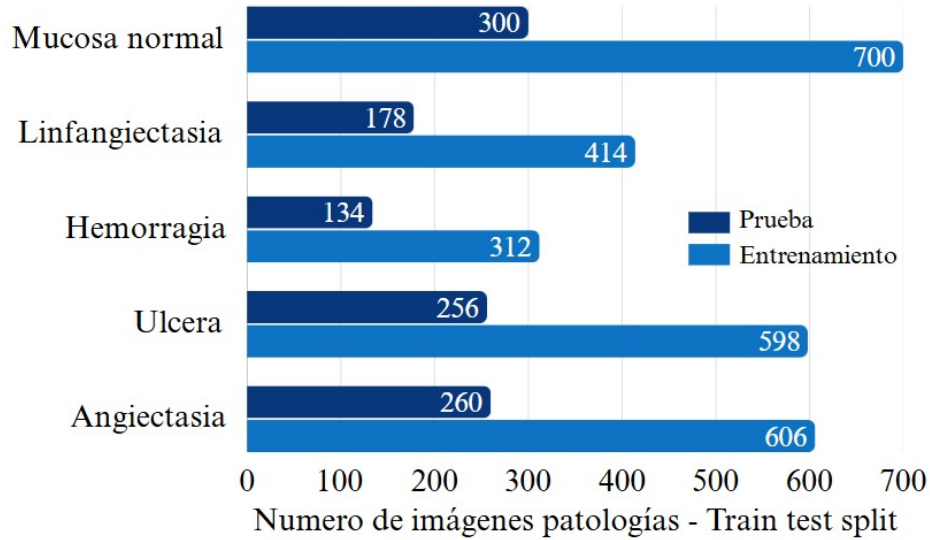


Figura 9.1: Distribución porcentual de la base de datos — Patologías (70 % entrenamiento / 30 % validación).

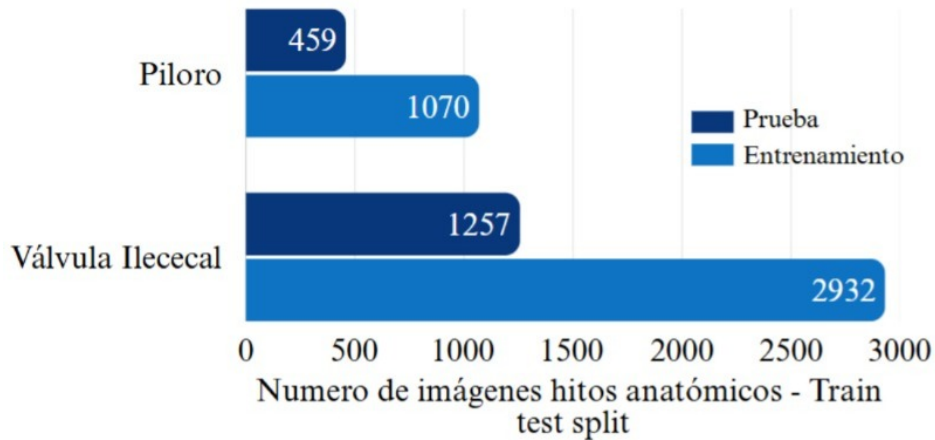


Figura 9.2: Distribución porcentual de la base de datos — Hitos Anatómicos (70 % entrenamiento / 30 % validación).

Con estos dos gráficos (Figuras 9.1 y 9.2) se verifica que la estratificación se realizó correctamente antes de cada ciclo, asegurando que cada clase mantenga su proporción en los subconjuntos de entrenamiento y validación.

9.3. Personalización de la Cabeza de Red para Cada Tarea

Para optimizar el rendimiento de los modelos en cada tarea específica —clasificación de patologías e identificación de hitos anatómicos—, se ajustó cuidadosamente la arquitectura base DenseNet mediante modificaciones personalizadas en sus capas superiores. La estrategia principal consistió en reemplazar la capa de salida original por una estructura densa especialmente diseñada para las necesidades particulares de cada tarea. Esto permitió mantener una arquitectura robusta y ligera, potenciando la generalización y mejorando la discriminación entre clases mediante una extracción eficiente y precisa de características relevantes desde los mapas de activación producidos por la red.

9.3.1. Clasificación de Patologías (Cinco Clases)

La tarea de clasificación de patologías consiste en distinguir entre tejido sano y cuatro tipos de lesiones clínicas (angiectasia, úlceras, pólipos, sangre y mucosa normal). Para ello, la cabeza de red incorporó, en primer lugar, una capa de `GlobalAveragePooling2D` tras el último `DenseBlock`, la cual convierte los mapas de características en un vector de dimensión fija. A continuación, se añadió una capa densa de 512 neuronas con activación ReLU, regularización $L_2 = 5 \times 10^{-4}$ y `dropout = 0.4`. Posteriormente, se incorporó una segunda capa densa de 256 neuronas con activación ReLU, regularización $L_2 = 5 \times 10^{-4}$, `BatchNormalization` y `dropout = 0.3`. Finalmente, la capa de salida consta de 5 neuronas con activación `softmax`, una para cada categoría (tejido sano más las cuatro clases de lesiones).

La configuración de entrenamiento incluyó la función de pérdida `SparseCategoricalCrossentropy`, el optimizador `Adam` con `learning rate = 5 \times 10^{-4}`, y el mecanismo de `early stopping` con paciencia = 5 y `min_delta = 0.001`, restaurando los mejores pesos. El tamaño de lote seleccionado fue 32 y el número máximo de épocas se fijó en 100. Esta combinación busca equilibrar la capacidad de aprendizaje de la arquitectura con la necesidad de evitar el sobreajuste en un conjunto de datos moderado, utilizando regularización y `dropout` para reforzar la generalización.

Componente	Detalles
Cabeza de red	
GlobalAveragePooling2D	Convierte los mapas de características en un vector de dimensión fija.
Capa densa (512)	Activación ReLU; regularización $L_2 = 5 \times 10^{-4}$; <i>dropout</i> = 0.4.
Capa densa (256)	Activación ReLU; regularización $L_2 = 5 \times 10^{-4}$; BatchNormalization ; <i>dropout</i> = 0.3.
Capa de salida	5 neuronas con activación softmax (tejido sano + cuatro lesiones).
Configuración de entrenamiento	
Función de pérdida	SparseCategoricalCrossentropy
Optimizador	Adam con learning rate = 5×10^{-4}
Early stopping	Paciencia = 5; min_delta = 0.001; restaurar mejores pesos
Batch size	32
Épocas máximas	100

Cuadro 9.2: Configuración de la cabeza de red y parámetros de entrenamiento para clasificación de patologías.

9.3.2. Detección de Hitos Anatómicos (Clasificación Binaria)

Para determinar la posición aproximada dentro del intestino delgado —zona proximal versus zona distal— se entrenaron modelos cuya cabeza identifica dos hitos anatómicos: el píloro y la válvula ileocecal. Primero se incluyó una capa de **GlobalAveragePooling2D** tras el último *DenseBlock*, seguida inmediatamente por una capa de **BatchNormalization** para estabilizar las activaciones. Luego, se incorporó una capa densa de 512 neuronas con activación ReLU, regularización $L_2 = 5 \times 10^{-4}$ y *dropout* = 0.4; posteriormente, una capa densa de 256 neuronas con activación ReLU, regularización $L_2 = 5 \times 10^{-4}$, **BatchNormalization** y *dropout* = 0.3. Finalmente, la capa de salida consta de 1 neurona con activación **sigmoid**, que genera la probabilidad de presencia del hito en la imagen.

La función de pérdida seleccionada fue **BinaryCrossentropy** y el optimizador **Adam** con **learning rate** = 1×10^{-3} . El *early stopping* se configuró con paciencia = 5 y **min_delta** = 0.001, restaurando los mejores pesos. El tamaño de lote fue de 32 y el número máximo de épocas se mantuvo en 100. Esta configuración atiende a la necesidad de equilibrar precisión y sensibilidad para minimizar tanto falsos positivos como falsos negativos, facilitando una convergencia más rápida en este problema binario.

Componente	Detalles
Cabeza de red	
GlobalAveragePooling2D	Convierte los mapas de características en un vector de dimensión fija.
BatchNormalization	Estabiliza las activaciones inmediatamente después de GlobalAveragePooling2D.
Capa densa (512)	Activación ReLU; regularización $L_2 = 5 \times 10^{-4}$; <i>dropout</i> = 0.4.
Capa densa (256)	Activación ReLU; regularización $L_2 = 5 \times 10^{-4}$; BatchNormalization; <i>dropout</i> = 0.3.
Capa de salida	1 neurona con activación sigmoid (presencia del hito).
Configuración de entrenamiento	
Función de pérdida	BinaryCrossentropy
Optimizador	Adam con <i>learning rate</i> = 1×10^{-3}
Early stopping	Paciencia = 5; <i>min_delta</i> = 0.001; restaurar mejores pesos
Batch size	32
Épocas máximas	100

Cuadro 9.3: Configuración de la cabeza de red y parámetros de entrenamiento para detección de hitos anatómicos.

Una vez entrenados los modelos DenseNet con sus respectivas configuraciones y variantes arquitectónicas, era necesario evaluar su desempeño de forma objetiva, repetible y robusta. Para ello, se diseñó una estrategia de validación basada en múltiples ejecuciones controladas por semillas aleatorias, que permitió analizar no solo los resultados promedio, sino también la estabilidad del modelo ante distintos escenarios. En el próximo capítulo se presenta dicha validación, comparando el rendimiento de los modelos por arquitectura y por tarea, y justificando la elección final de cada uno.

Validación del rendimiento de los modelos

En este capítulo se presenta la validación de métricas obtenidas a partir de las arquitecturas DenseNet evaluadas en dos tareas fundamentales: la clasificación de patologías observadas en las imágenes de cápsula endoscópica y la detección de hitos anatómicos (píloro y válvula ileocecal) a partir de imágenes estáticas. Con el objetivo de garantizar la robustez y reproducibilidad de los resultados, cada variante (DenseNet121, DenseNet169 y DenseNet201) fue sometida a múltiples ciclos de entrenamiento con particiones distintas de los datos, empleando un conjunto de semillas predefinidas. Este enfoque permitió medir no solo el desempeño absoluto de cada modelo, sino también la variabilidad intrínseca derivada de la inicialización y la división aleatoria de los conjuntos de entrenamiento y prueba.

10.1. Dinámica de Entrenamiento con Semillas

Como se detalló en el capítulo anterior, una parte fundamental de la estrategia de validación consistió en entrenar cada modelo bajo cinco condiciones de inicialización distintas, controladas mediante semillas aleatorias (42, 123, 2025, 1 y 999). Esta metodología permitió no solo evaluar el rendimiento medio de cada arquitectura, sino también cuantificar su estabilidad ante cambios leves en la partición de datos y la inicialización de pesos.

Cada ciclo de entrenamiento repitió la misma configuración experimental en cuanto a arquitectura, hiperparámetros y funciones de pérdida, variando únicamente la semilla activa. Al final de cada ciclo se evaluó el modelo sobre el conjunto de prueba y se calcularon las métricas principales: precisión, sensibilidad (recall), F1-Score por clase y exactitud global.

Este enfoque aporta dos ventajas clave: por un lado, permite comparar de forma objetiva el desempeño entre arquitecturas bajo condiciones controladas; por otro, ofrece una estimación de la variabilidad inherente al modelo mediante el cálculo de la desviación estándar (σ) para cada métrica. Esta medida de dispersión es especialmente importante en el ámbito clínico, donde no solo importa la precisión promedio, sino también la consistencia del sistema ante diferentes muestras o ejecuciones.

Los resultados consolidados (media y desviación estándar) de las cinco ejecuciones por arquitectura y tarea se presentan en la siguiente sección. Estas tablas permitirán identificar qué modelos ofrecen no solo el mejor desempeño absoluto, sino también la mayor robustez, lo que justifica su eventual selección para integrar el sistema automatizado.

10.2. Resultados por Modelo y Tarea

A continuación se presentan los resultados obtenidos para cada arquitectura (DenseNet121, DenseNet169 y DenseNet201) aplicada a las dos tareas principales del sistema: la clasificación de cinco tipos de patologías y la detección de hitos anatómicos intestinales, específicamente el píloro y la válvula ileocecal. El desempeño de los modelos fue evaluado de forma sistemática en cinco ciclos de entrenamiento independientes, cada uno utilizando una semilla diferente para controlar la aleatoriedad en la partición de datos y la inicialización de pesos.

Cada tabla reporta las métricas de evaluación **precisión**, **recall**, **F1-Score** y **accuracy**, calculadas sobre el conjunto de prueba para cada una de las clases involucradas en la tarea correspondiente. Para cada métrica se muestran el promedio (μ) y la desviación estándar (σ) obtenidos tras los cinco ciclos. Estas medidas permiten cuantificar tanto el desempeño absoluto de los modelos como su estabilidad entre ejecuciones.

Este análisis detallado clase por clase es crucial para identificar patrones de error, fortalezas específicas de cada arquitectura y posibles sesgos en el aprendizaje. Asimismo, los resúmenes finales de métricas globales comparan los tres modelos en cada tarea, facilitando la selección de la arquitectura más eficiente y robusta para ser integrada en el sistema final de predicción.

10.2.1. DenseNet201 – Patologías (Cinco Ciclos)

Este cuadro muestra las métricas promedio (precisión, recall y F1-score, con sus desviaciones estándar) obtenidas por DenseNet201 tras cinco ciclos de entrenamiento en la tarea de clasificación de las cinco categorías de patología. Además, se incluye al final la precisión global del modelo en cada ciclo.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.976 \pm 0.010	0.982 \pm 0.015	0.978 \pm 0.007
Úlcera	0.968 \pm 0.012	0.986 \pm 0.005	0.976 \pm 0.005
Hemorragia	0.992 \pm 0.004	0.982 \pm 0.007	0.986 \pm 0.005
Linfangiectasia	0.972 \pm 0.010	0.976 \pm 0.012	0.974 \pm 0.008
Mucosa Normal	0.968 \pm 0.012	0.948 \pm 0.012	0.954 \pm 0.008
Precisión global (Accuracy)			0.972 \pm 0.004

Cuadro 10.1: DenseNet201 – Clasificación de patologías (Cinco ciclos).

10.2.2. DenseNet169 – Patologías (Cinco Ciclos)

En esta tabla se resumen los resultados promedio de DenseNet169 para la clasificación de las mismas cinco patologías, tras cinco ejecuciones independientes. Se detallan los valores medios y desviaciones estándar de precisión, recall y F1-score para cada clase, y al final la precisión global.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.976 \pm 0.008	0.972 \pm 0.010	0.974 \pm 0.005
Úlcera	0.978 \pm 0.007	0.978 \pm 0.004	0.976 \pm 0.005
Hemorragia	0.996 \pm 0.005	0.986 \pm 0.010	0.990 \pm 0.006
Linfangiectasia	0.962 \pm 0.007	0.970 \pm 0.009	0.964 \pm 0.005
Mucosa Normal	0.954 \pm 0.010	0.954 \pm 0.008	0.954 \pm 0.010
Precisión global (Accuracy)			0.970 \pm 0.004

Cuadro 10.2: DenseNet169 – Clasificación de patologías (Cinco ciclos).

10.2.3. DenseNet121 – Patologías (Cinco Ciclos)

A continuación se exponen las métricas medias obtenidas por DenseNet121 en la clasificación de patologías tras cinco ciclos de experimentación. Se presentan la precisión, recall y F1-score para cada categoría, y al final la precisión media global del modelo.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.962 \pm 0.010	0.974 \pm 0.005	0.968 \pm 0.007
Úlcera	0.982 \pm 0.013	0.978 \pm 0.010	0.978 \pm 0.004
Hemorragia	0.992 \pm 0.007	0.986 \pm 0.014	0.990 \pm 0.011
Linfangiectasia	0.964 \pm 0.016	0.970 \pm 0.006	0.966 \pm 0.005
Mucosa Normal	0.960 \pm 0.011	0.950 \pm 0.011	0.952 \pm 0.007
Precisión global (Accuracy)			0.970 \pm 0.004

Cuadro 10.3: DenseNet121 – Clasificación de patologías (Cinco Ciclos).

10.2.4. DenseNet201 – Hitos Anatómicos (Cinco Ciclos)

Este cuadro recoge el desempeño de DenseNet201 en la detección de los hitos “píloro” y “válvula ileocecal” tras cinco rondas de entrenamiento y validación. Se muestran los promedios y desviaciones estándar de las métricas de precisión, recall y F1-score por clase, y al final la precisión global.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.992 \pm 0.012	0.982 \pm 0.007	0.986 \pm 0.005
Válvula Ileocecal	0.994 \pm 0.005	0.998 \pm 0.004	0.996 \pm 0.005
Precisión global (Accuracy)			0.992 \pm 0.002

Cuadro 10.4: DenseNet201 – Detección de hitos anatómicos (Cinco ciclos).

10.2.5. DenseNet169 – Hitos Anatómicos (Cinco Ciclos)

En esta tabla se presentan los resultados promedio de DenseNet169 para distinguir los hitos anatómicos principales del tracto digestivo. Se incluyen las métricas de precisión, recall y F1-score (con sus desviaciones estándar) para cada hito, así como la precisión global media obtenida tras cinco ciclos.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.988 ± 0.004	0.982 ± 0.010	0.986 ± 0.005
Válvula Ileocecal	0.996 ± 0.005	0.998 ± 0.004	0.996 ± 0.005
Precisión global (Accuracy)			0.991 ± 0.002

Cuadro 10.5: DenseNet169 – Detección de hitos anatómicos (Cinco ciclos).

10.2.6. DenseNet121 – Hitos Anatómicos (Cinco Ciclos)

A continuación se muestran los resultados promedio de DenseNet121 en la tarea de identificar el píloro y la válvula ileocecal. Los valores indicados corresponden a las métricas de precisión, recall y F1-score con sus desviaciones estándar, y a la precisión global promediada tras cinco ejecuciones.

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.996 ± 0.005	0.982 ± 0.012	0.986 ± 0.005
Válvula Ileocecal	0.996 ± 0.005	1.000 ± 0.000	0.996 ± 0.005
Precisión global (Accuracy)			0.993 ± 0.003

Cuadro 10.6: DenseNet121 – Detección de hitos anatómicos (Cinco Ciclos).

10.2.7. Resumen Final de Métricas Promedio

10.2.7.1. Patologías

Este resumen agrupa los resultados globales de las tres arquitecturas DenseNet (201, 169 y 121) en la clasificación de patologías. Para cada modelo se reporta la precisión media y su desviación estándar, así como los valores promedio de recall y F1-score en las cinco ejecuciones.

Modelo	Accuracy ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
DENSENET201	0.972 ± 0.004	0.975 ± 0.004	0.974 ± 0.003
DENSENET169	0.970 ± 0.004	0.972 ± 0.003	0.972 ± 0.004
DENSENET121	0.970 ± 0.004	0.972 ± 0.005	0.971 ± 0.004

Cuadro 10.7: Resumen Final – Clasificación de patologías.

10.2.7.2. Hitos Anatómicos

La tabla final sintetiza el desempeño comparativo de las tres versiones de DenseNet en la detección de hitos anatómicos. Se incluyen las precisiones medias y sus desviaciones estándar, junto con los valores promedio de recall y F1-score obtenidos tras cinco ciclos de evaluación para cada modelo.

Modelo	Accuracy ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
DENSENET201	0.992 ± 0.002	0.990 ± 0.003	0.991 ± 0.005
DENSENET169	0.991 ± 0.002	0.990 ± 0.004	0.991 ± 0.005
DENSENET121	0.993 ± 0.003	0.991 ± 0.006	0.991 ± 0.005

Cuadro 10.8: Resumen Final – Detección de hitos anatómicos.

10.3. Análisis y Justificación de Selección de Modelos

Con base en las métricas detalladas en las tablas anteriores, se analizan a continuación las razones para seleccionar los modelos finales en cada tarea.

Por tanto, **DenseNet201** se selecciona como modelo definitivo para la clasificación de patologías, dado que maximiza el accuracy promedio y reduce la variabilidad entre ejecuciones.

10.3.1. Clasificación de Patologías con DenseNet201

Según la Tabla 10.7, DenseNet201 obtuvo:

Accuracy $0,972 \pm 0,004$, Recall $0,975 \pm 0,004$, F1-Score $0,974 \pm 0,003$.

Este rendimiento supera ligeramente al de DenseNet169 y DenseNet121, tanto en accuracy (0.970) como en recall (0.972) y F1-Score (0.972/0.971), con desviaciones estándar inferiores. Por tanto, DenseNet201 se eligió para clasificación de patologías por su mayor sensibilidad (recall) y consistencia entre ciclos.

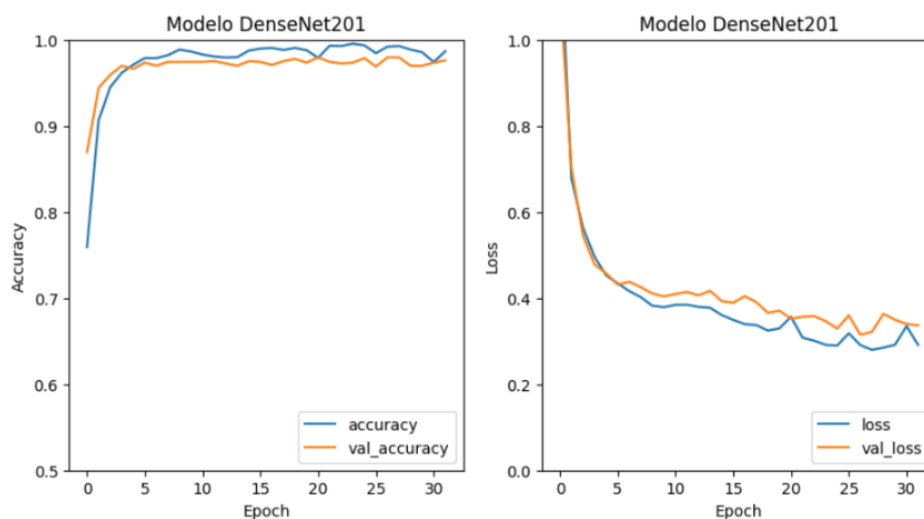


Figura 10.1: Curvas de pérdida y exactitud (train/val) del mejor ciclo de DenseNet201 en clasificación de patologías.

En la Figura 10.1 se observa que, en el mejor ciclo seleccionado, las curvas de entrenamiento y validación convergen rápidamente hacia una baja pérdida (loss) y alta exactitud. El “gap” entre ambas curvas es mínimo a partir de la época 3, indicando que no hay sobreajuste significativo. Esto corrobora la estabilidad mostrada por la pequeña desviación estándar en las métricas de la Tabla 10.7.

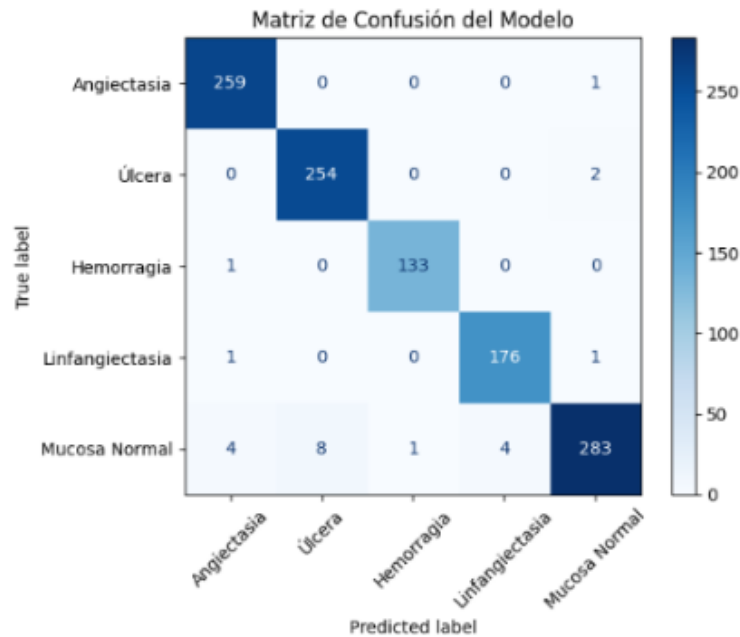


Figura 10.2: Matriz de confusión del mejor ciclo de DenseNet201 en clasificación de patologías.

La Figura 10.2 muestra la matriz de confusión correspondiente: todas las clases principales (Angiectasia, Úlcera, Hemorragia, Linfangiectasia y Mucosa Normal) presentan tasas de clasificación superiores al 97%. En particular, Hemorragia alcanza cerca de 99% de elementos correctamente identificados, lo que respalda el recall promedio de 0.975 ± 0.004 .

10.3.2. Detección de Hitos Anatómicos con DenseNet121

Según la Tabla 10.8, DenseNet121 alcanzó:

$$\text{Accuracy } 0,993 \pm 0,003, \quad \text{Recall } 0,991 \pm 0,006, \quad \text{F1-Score } 0,991 \pm 0,005.$$

Estas cifras superan las de DenseNet201 (accuracy 0.992, recall 0.990) y DenseNet169 (accuracy 0.991, recall 0.990), aunque la diferencia es pequeña. Al tener menos parámetros, DenseNet121 ofrece menor riesgo de sobreajuste y latencias de inferencia más bajas, ventaja relevante para el análisis “near real-time” de hitos.

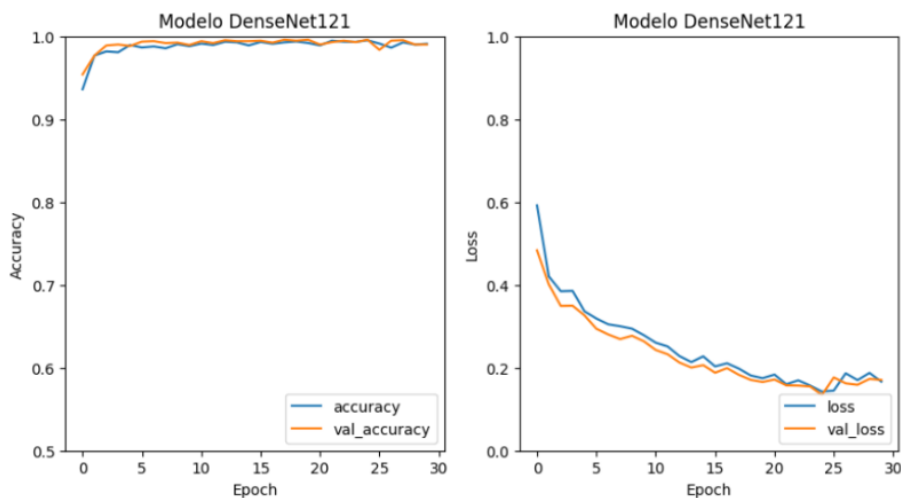


Figura 10.3: Curvas de pérdida y exactitud (train/val) del mejor ciclo de DenseNet121 en detección de hitos anatómicos.

En la Figura 10.3, las curvas muestran una rápida convergencia con pérdida casi estacionaria desde la época 2-3 y exactitud cercana al 100% en validación. El gap de validación es casi imperceptible, lo que coincide con la desviación estándar relativamente baja (± 0.003 en accuracy) y confirma que DenseNet121 generaliza bien sobre el conjunto de hitos.

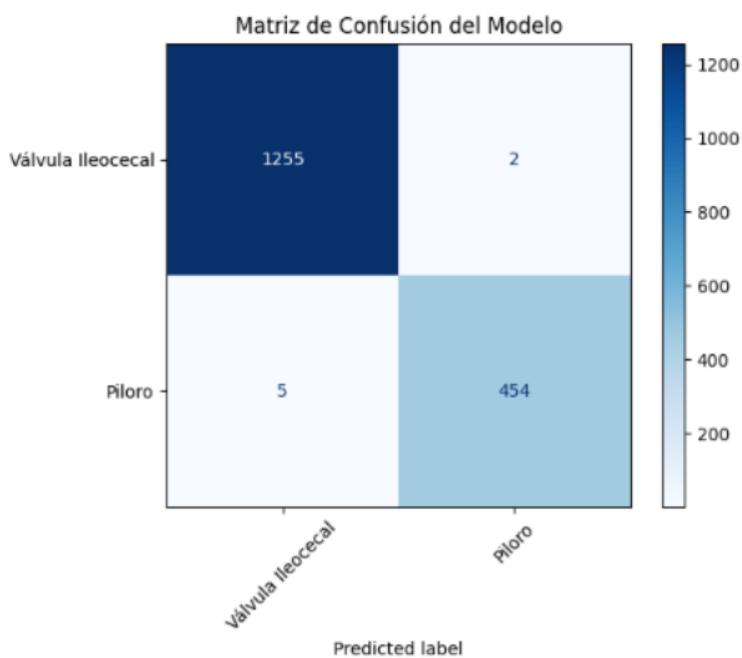


Figura 10.4: Matriz de confusión del mejor ciclo de DenseNet121 en detección de hitos anatómicos.

La Figura 10.4 evidencia que tanto Píloro como Válvula Ileocecal se clasifican con tasas superiores al 99 %. Esto respalda el recall de 0.991 ± 0.006 y la precisión casi perfecta en cada clase, justificando la elección de DenseNet121.

10.4. Análisis

- Para la **clasificación de patologías**, la arquitectura DenseNet201 ofreció la mejor combinación entre sensibilidad, precisión y estabilidad. En particular, alcanzó un *recall* promedio de 0.975, lo que refleja una alta capacidad del modelo para identificar correctamente casos positivos en todas las clases consideradas. Este desempeño es especialmente relevante en contextos clínicos, donde la minimización de falsos negativos es crítica para garantizar diagnósticos precisos y oportunos. Las curvas de pérdida (*loss*) y precisión (*accuracy*) durante el entrenamiento mostraron una convergencia rápida y estable, sin signos evidentes de sobreajuste, lo cual refuerza la robustez y generalización del modelo frente a variaciones en los datos. Además, el análisis detallado de la matriz de confusión evidenció un rendimiento sobresaliente con tasas de clasificación correctas superiores al 97 % en cada una de las cinco categorías evaluadas: Mucosa Normal, Angiectasia, Úlcera, Hemorragia y Linfangiectasia. Esto confirma que DenseNet201 no solo es efectivo en la distinción general de patologías, sino que también mantiene un equilibrio óptimo entre sensibilidad y especificidad, garantizando así su aplicabilidad en entornos clínicos reales donde es fundamental identificar correctamente las lesiones para guiar decisiones terapéuticas.
- En cuanto a la **detección de hitos anatómicos**, la arquitectura DenseNet121 destacó particularmente al alcanzar la mayor precisión global (*accuracy* promedio de 0.993) con una estructura computacionalmente más ligera y eficiente en comparación con las otras variantes evaluadas. Este aspecto es crucial en escenarios donde la rapidez y eficiencia del procesamiento son fundamentales, como en análisis clínicos en tiempo casi real. Las curvas de entrenamiento y validación mostraron un comportamiento altamente consistente, con diferencias mínimas entre ambos conjuntos, lo que evidencia una excelente capacidad de generalización y minimización del riesgo de sobreajuste. La matriz de confusión asociada mostró una clara predominancia diagonal, indicando muy pocas clasificaciones erróneas y un rendimiento particularmente notable en la identificación del píloro y la válvula ileocecal, ambas con tasas de acierto superiores al 99 %. Esta precisión en la detección anatómica es especialmente valiosa, pues permite estimar con alta fiabilidad la localización aproximada de las lesiones o patologías detectadas, incluso cuando se dispone únicamente de imágenes estáticas individuales, sin recurrir a secuencias temporales o datos contextuales adicionales. Por ello, DenseNet121 emerge como la opción óptima para esta tarea específica, combinando alta precisión diagnóstica con eficiencia operacional.

Interfaz gráfica de usuario

11.1. Fases de diseño de la interfaz gráfica

Como parte fundamental del desarrollo de este proyecto, se implementó una interfaz gráfica dirigida a los usuarios finales, específicamente a los médicos, con el propósito de facilitar la visualización e interpretación de los resultados obtenidos por el sistema. La necesidad de esta interfaz surge del reconocimiento de que muchos de los procesos de análisis de imágenes o datos médicos generan una gran cantidad de información técnica que puede ser difícil de interpretar rápidamente sin una herramienta visual adecuada. Por esta razón, se diseñó una interfaz intuitiva, clara y funcional, que permite presentar los resultados de manera estructurada y accesible, priorizando la comprensión inmediata de la información relevante.

La interfaz desarrollada para este proyecto integra de manera funcional y organizada los dos modelos desarrollados: uno enfocado en la clasificación automática de patologías visibles en las imágenes obtenidas por cápsula endoscópica, y otro orientado a la localización anatómica de dichas alteraciones dentro del intestino delgado. Esta integración permite que los médicos no sólo identifiquen la presencia de una posible patología, sino también su ubicación aproximada, lo que facilita el análisis clínico y la planificación de intervenciones. La interfaz fue pensada para brindar una experiencia de uso intuitiva, en la que los resultados de ambos modelos se presentan de forma conjunta y comprensible.

11.1.1. Bocetos y planificación de flujo de navegación

Como parte inicial del proceso de diseño, se elaboraron bocetos que permitieron definir la estructura y flujo de navegación de la interfaz. Estos esquemas fueron esenciales para anticipar la experiencia del usuario, planificar las pantallas necesarias y establecer la lógica de interacción entre las funcionalidades del sistema. La primera pantalla diseñada corresponde al ingreso de los datos básicos del paciente. Esta etapa inicial tiene como propósito asociar correctamente los resultados del análisis con un caso clínico específico, permitiendo un registro individualizado y ordenado. Tal como se muestra en la Figura 11.1, se incluyen campos para nombre, identificación, edad y médico tratante, junto con un botón de navegación que habilita el paso a la siguiente etapa del análisis.

CLASIFICACION DE IMAGENES DE CAPSULA
EDOSCOPICA DE INTESTINO DELGADO

Pontificia Universidad
JAVERIANA
Cali

DATOS DEL PACIENTE

Nombre:

Identificación:

Edad:

Medico:

Siguiente

Figura 11.1: Pantalla 1 - Boceto: ingreso de datos del paciente

A continuación, la segunda pantalla permite al usuario seleccionar el tipo de análisis a realizar, planteando de forma clara la pregunta: “¿Qué desea ingresar?”. Esta vista presenta dos botones principales: “**Una imagen**”, para análisis puntuales de una sola imagen endoscópica; y “**Estudio completo**”, que habilita la carga de múltiples imágenes de un examen completo. Esta flexibilidad fue pensada para ajustarse a distintos escenarios clínicos, desde validaciones rápidas hasta exploraciones diagnósticas más extensas.

CLASIFICACION DE IMAGENES DE CAPSULA
EDOSCOPICA DE INTESTINO DELGADO

¿QUE DESEA INGRESAR?

Una Imagen

Estudio completo

Figura 11.2: Pantalla 2 - Boceto: selección del tipo de análisis

En caso de elegir el análisis de una imagen individual, el sistema redirige al usuario a una pantalla donde se muestra la imagen cargada junto con el resultado del análisis automatizado. Se despliega la **patología detectada** y la **ubicación anatómica estimada**, generadas por los modelos entrenados. También se incorporan botones de navegación para regresar al menú anterior o a la pantalla inicial, permitiendo una interacción fluida.

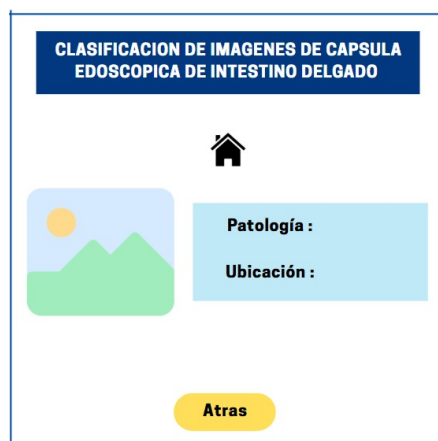


Figura 11.3: Pantalla 3 - Boceto: análisis de imagen individual

Por otro lado, si el usuario selecciona la opción de “Estudio completo”, el sistema despliega una vista general con los resultados agregados. Esta pantalla resume las **patologías detectadas**, su **ubicación estimada** (proximal o distal) y el **número de imágenes asociadas** a cada una. Esta funcionalidad proporciona una visión clínica estructurada y facilita la revisión global de los hallazgos. Al igual que en las demás vistas, se incluyen opciones de navegación para mantener la coherencia y accesibilidad durante el uso del sistema.

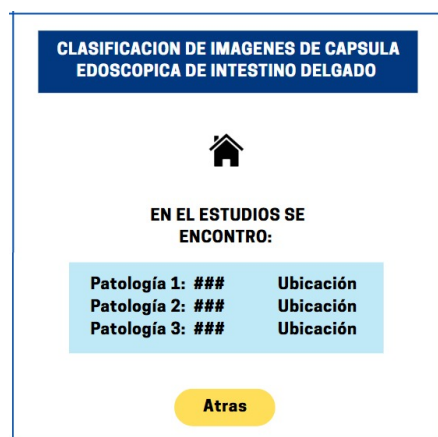


Figura 11.4: Pantalla 4 - Boceto: resumen de estudio completo

Finalmente, el diseño contempla la posibilidad de hacer clic sobre cada patología detectada para desplegar las imágenes asociadas a dicha clase. Esta funcionalidad permite al profesional revisar de forma visual los casos identificados, facilitando la validación del diagnóstico y el análisis detallado por categoría clínica.

11.2. Implementación y estructura de la interfaz gráfica

11.2.1. Principios de diseño y herramientas empleadas

Con el objetivo de facilitar la interacción entre el usuario y el sistema de diagnóstico, se desarrolló una interfaz gráfica intuitiva y funcional. Esta interfaz permite la carga, análisis y visualización de imágenes de cápsula endoscópica de manera amigable, sin requerir conocimientos técnicos en programación o aprendizaje profundo. Su propósito principal es servir como puente entre los modelos de inteligencia artificial desarrollados y los profesionales médicos que interpretan los resultados, garantizando así un flujo de trabajo eficiente. La solución fue diseñada con énfasis en la claridad visual, la accesibilidad funcional y la rapidez de procesamiento, elementos esenciales en contextos clínicos donde la agilidad diagnóstica es fundamental.

El desarrollo de la interfaz se realizó en Visual Studio, empleando el módulo Tkinter de Python para la construcción de ventanas, formularios y navegación entre secciones. Se adoptó una arquitectura orientada a objetos, basada en clases, con una estructura modular que facilita tanto la escalabilidad del sistema como su mantenimiento. Cada clase encapsula una funcionalidad específica dentro del flujo de trabajo clínico, permitiendo una separación clara entre la lógica operativa y la presentación gráfica, lo cual favorece un diseño ordenado, robusto y adaptable.

11.2.2. Estructura de clases de la interfaz

Las clases que componen la interfaz fueron organizadas de forma modular, asignando a cada una una función específica dentro del flujo clínico de la aplicación:

- **Clase principal:** responsable de inicializar la aplicación, cargar los modelos entrenados (clasificación de patologías y detección de hitos anatómicos), y gestionar la transición entre las distintas pantallas. Esta clase centraliza el control del sistema y garantiza el funcionamiento de los elementos globales.
- **Clase de ingreso de datos:** permite registrar la información clínica del paciente y del especialista tratante (nombre, número de identificación, edad y médico responsable). Esta clase asegura que cada análisis esté correctamente vinculado a un caso clínico individual, facilitando la generación de historiales personalizados.
- **Clase del menú principal:** actúa como panel de navegación general del sistema. Desde esta sección, el usuario puede acceder a las distintas funcionalidades: análisis de imágenes individuales, revisión del historial, análisis de estudios completos, entre otros. Su organización favorece la fluidez y la lógica del flujo de trabajo clínico.
- **Clase de análisis individual de imágenes:** habilita la selección de una imagen específica para su análisis mediante los modelos de *deep learning*. Los resultados (tipo de patología y localización anatómica estimada) se presentan en una nueva ventana con formato visual optimizado. Se incorporan indicadores de color y mensajes dinámicos para resaltar el nivel de confianza de las predicciones.

- **Clase para análisis de estudio completo:** permite el procesamiento automático de múltiples imágenes correspondientes a un único estudio. Esta clase analiza de forma integral el trayecto intestinal capturado por la cápsula endoscópica, incorporando filtros por tipo de lesión y un árbol interactivo que facilita la revisión de hallazgos organizados por categoría.
- **Clase de historial de imágenes:** administra los casos previamente analizados y organizados por paciente. Desde esta sección es posible consultar estudios anteriores, comparar resultados, y realizar un seguimiento clínico continuo. Además, permite reanalizar imágenes específicas, funcionalidad útil para validar diagnósticos frente a nuevas versiones de los modelos o ante dudas clínicas puntuales.

Cada clase definida en la interfaz contiene un conjunto de funciones internas que permiten cumplir con los objetivos asignados a su módulo. Estas funciones gestionan tareas como la carga y el procesamiento de imágenes, la interacción con los modelos de predicción, el manejo de entradas por parte del usuario, la navegación entre pantallas y la presentación estructurada de los resultados obtenidos. Además, se implementó un sistema de manejo de errores que previene interrupciones durante el análisis, complementado con mensajes dinámicos que proporcionan retroalimentación visual clara e inmediata.

11.3. Desarrollo visual y flujo de la interfaz

11.3.1. Pantallas principales y navegación

La interfaz gráfica implementada para la aplicación de clasificación de imágenes de cápsula endoscópica del intestino delgado fue diseñada con un enfoque centrado en la experiencia del usuario, priorizando la claridad de los campos de entrada y el cumplimiento de principios de usabilidad clínica. Cada elemento visual fue cuidadosamente estructurado para facilitar la interacción tanto del personal médico como técnico, manteniendo una estética profesional y funcional. Se emplearon paletas de colores sobrias y botones con retroalimentación visual inmediata, con el propósito de transmitir confianza y simplicidad en su uso.

Al iniciar la aplicación, la primera pantalla que se despliega corresponde al registro de los datos del paciente y del especialista tratante. Esta ventana incluye campos para ingresar el nombre, número de identificación, edad y médico responsable, junto con un botón de validación que permite continuar al análisis. Su diseño se alinea con el boceto original planteado en la etapa de prototipado, respetando la estructura planeada y su función principal como puerta de entrada al sistema.



CLASIFICACIÓN DE IMÁGENES DE CÁPSULA ENDOSCÓPICA DE INTESTINO DELGADO



DATOS DEL PACIENTE

Nombre:

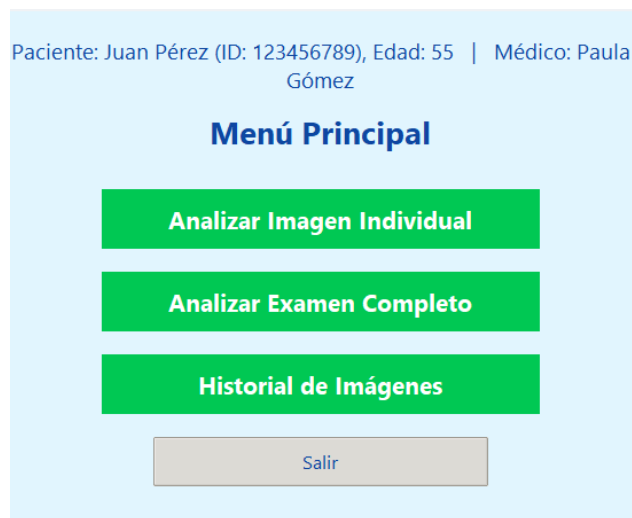
Identificación:

Edad:

Médico:

Figura 11.5: Pantalla 1 - Registro del Paciente

A continuación, se presenta el **menú principal** de la aplicación, que aparece inmediatamente después del registro del paciente. En esta pantalla se muestra de manera clara la información básica del paciente en la parte superior, seguida de tres botones principales que estructuran el flujo de trabajo clínico. El botón **“Analizar Imagen Individual”** permite cargar una imagen aislada de cápsula endoscópica para su evaluación. El botón **“Analizar Examen Completo”** está destinado al procesamiento masivo de múltiples imágenes de un estudio completo. Finalmente, el botón **“Historial de Imágenes”** da acceso a los registros anteriores del paciente, facilitando la revisión de exámenes previos. Esta organización promueve una navegación intuitiva y secuencial a lo largo del uso de la aplicación.



Paciente: Juan Pérez (ID: 123456789), Edad: 55 | Médico: Paula Gómez

Menú Principal

Figura 11.6: Pantalla 2 - Menú Principal

Al seleccionar el botón **“Analizar Imagen Individual”**, el sistema redirige a una pantalla donde se visualiza la imagen cargada para su análisis. En la parte superior se presentan los datos clínicos del paciente (nombre, identificación y edad), junto con el nombre del médico

tratante. En el centro se ubica la imagen endoscópica cargada, mostrada en un tamaño adecuado para su interpretación. En la parte inferior, se despliegan los resultados generados por el sistema: la **patología detectada** y la **ubicación estimada** dentro del intestino delgado (por ejemplo, zona proximal o distal). En los casos donde no se identifica una patología o el modelo muestra baja confianza en la localización anatómica, se emite un mensaje de alerta.



Figura 11.7: Pantalla 3 - Análisis de Imagen Individual

Cuando el usuario selecciona el botón “**Analizar Examen Completo**”, se accede a una interfaz dedicada al análisis masivo de imágenes. Esta sección organiza automáticamente las imágenes clasificadas según su categoría patológica, permitiendo al especialista explorar visualmente cada grupo (por ejemplo, *úlceras*, *angiectasias*, *hemorragia*, entre otras). Cada clase puede desplegar sus imágenes correspondientes, acompañadas por la **ubicación estimada** (proximal o distal), lo que agiliza la identificación de patrones clínicos y zonas afectadas. Visualmente, esta pantalla incluye un título descriptivo, una barra de progreso que indica el

estado del procesamiento por lotes, y un panel inferior donde se agrupan miniaturas de las imágenes clasificadas. Cada miniatura posee una etiqueta informativa sobre la patología detectada y su localización estimada. Al hacer doble clic sobre una categoría, se abre una nueva ventana con las imágenes ampliadas organizadas en formato de galería, lo cual permite una revisión visual detallada de cada caso.

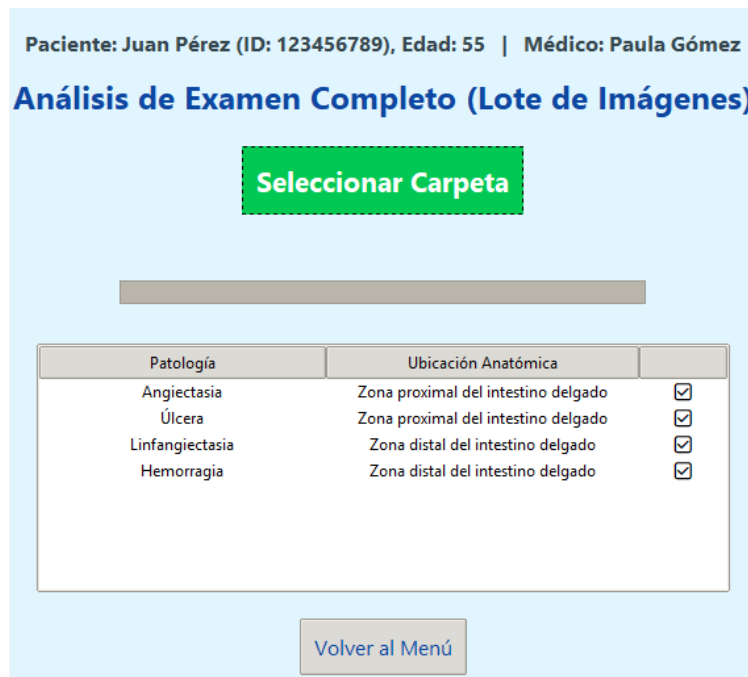


Figura 11.8: Pantalla 4 - Análisis de Estudio Completo

A modo de ejemplo, en la Figura 11.8, se muestra una imagen endoscópica clasificada como úlcera en la zona proximal. Se observan lesiones erosivas con bordes irregulares, mucosa con pliegues prominentes y presencia de burbujas, características propias del duodeno proximal.

En la Figura 11.8, se presenta el resumen de un análisis por lotes, visualizado como una tabla con las clases patológicas detectadas y su ubicación estimada. En este caso, se identificaron *angiectasia* y *úlcera* en la zona proximal, mientras que *linfangiectasia* y *hemorragia* fueron ubicadas en la zona distal. Estos resultados se generan a partir de la combinación del modelo de clasificación de patologías y el modelo de localización anatómica, que se apoya en imágenes de referencia del píloro y la válvula ileocecal. Al hacer clic sobre una fila de la tabla, se abre una ventana con las imágenes asociadas, presentadas en galería visual. Cada imagen está acompañada de su diagnóstico individual y la localización estimada, permitiendo al especialista validar los hallazgos y comparar casos similares con facilidad.



Figura 11.9: Galería desplegada al seleccionar la categoría “Linfangiectasia” en el análisis por lote.

La Figura 11.9 muestra un conjunto de imágenes clasificadas como linfangiectasia. En ellas se identifican formaciones nodulares blanquecinas sobre la mucosa, textura homogénea y disminución en el número de pliegues, características compatibles con localización distal. Esta vista facilita el análisis agrupado de casos similares, lo cual fortalece la validación clínica y la consistencia del modelo.

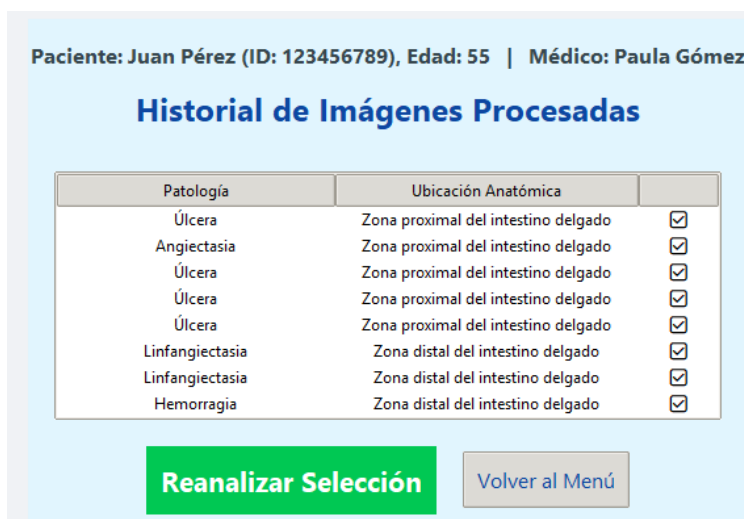


Figura 11.10: Pantalla de historial de imágenes procesadas

Por último, el botón “**Historial de Imágenes**” da acceso a una sección donde se visualizan todas las imágenes analizadas durante la sesión. Esta funcionalidad permite rastrear los hallazgos clínicos, revisar resultados anteriores y realizar comparaciones temporales. La pantalla correspondiente muestra una tabla con las imágenes previamente procesadas, en la que se incluye el diagnóstico por patología y la ubicación anatómica estimada. Esta vista admite la selección individual de imágenes para ser reanalizadas mediante el botón “**Reanalizar Selección**”, lo cual es útil cuando se desea confirmar un resultado, aplicar nuevas versiones del modelo o realizar una validación manual adicional.

11.4. Resumen y relevancia clínica

En conjunto, la interfaz gráfica desarrollada representa un puente entre la complejidad técnica de los modelos de inteligencia artificial y la práctica clínica diaria, facilitando la adopción de herramientas de análisis automatizado en entornos reales. Su diseño modular, claridad visual y lógica de navegación aseguran que los profesionales médicos puedan utilizar el sistema de manera eficiente, confiable y adaptable a futuras mejoras.

Conclusiones

El desarrollo del sistema automatizado basado en deep learning para el análisis de imágenes de cápsula endoscópica del intestino delgado permitió cumplir de manera satisfactoria el objetivo general del proyecto. Se construyó una solución funcional capaz de identificar diferentes patologías y estimar la ubicación anatómica de las lesiones a partir de imágenes médicas, respondiendo a una necesidad clínica concreta: reducir el tiempo de análisis, minimizar los errores humanos derivados de la fatiga visual y mejorar la precisión diagnóstica en la revisión de estudios endoscópicos no invasivos. La propuesta representa una contribución significativa desde la ingeniería biomédica al campo de la gastroenterología, al articular técnicas de inteligencia artificial con aplicaciones clínicas reales.

En relación con los objetivos específicos, se logró cumplir cada uno de los planteamientos iniciales. Se seleccionó una base de datos de dominio público con imágenes representativas del intestino delgado, abarcando distintas patologías (angiectasia, mucosa normal, úlcera, hemorragia y linfangiectasia) y referencias anatómicas clave (píloro y válvula ileocecal). Estas imágenes fueron sometidas a un proceso completo de preprocesamiento, que incluyó normalización, transformaciones geométricas y aumento de datos.

Se implementaron y entrenaron tres variantes de la arquitectura DenseNet (DenseNet121, DenseNet169 y DenseNet201), las cuales fueron evaluadas tanto en la tarea de clasificación de patologías como en la identificación de hitos anatómicos. Los resultados obtenidos mostraron altos niveles de precisión, sensibilidad y especificidad, superando el 90

Asimismo, se diseñó e implementó una interfaz gráfica de usuario desarrollada en Python, utilizando la biblioteca Tkinter. Esta interfaz permite cargar imágenes individuales o estudios completos, visualizar los resultados generados por los modelos y analizar las predicciones de forma clara, comprensible y orientada al uso clínico. También se integró un formulario para el registro de datos básicos del paciente y del profesional tratante, lo que refuerza el enfoque de la herramienta como apoyo al diagnóstico médico y no únicamente como una prueba de concepto computacional.

Pese a los logros alcanzados, se identificaron áreas susceptibles de mejora para futuras etapas del proyecto. En primer lugar, se reconoce la necesidad de ampliar la base de datos con nuevas clases patológicas e imágenes provenientes de diferentes fuentes clínicas, lo cual permitiría aumentar la capacidad de generalización del sistema y disminuir posibles sesgos asociados al conjunto de entrenamiento actual. En particular, la inclusión de imágenes reales obtenidas en instituciones médicas nacionales facilitaría su validación en contextos clínicos locales.

Por otra parte, aunque el sistema utiliza los hitos del píloro y la válvula ileocecal como referencias anatómicas, la estimación de la ubicación proximal o distal de las lesiones sigue siendo aproximada. Esta limitación podría abordarse incorporando modelos que analicen secuencias temporales, como redes neuronales recurrentes tipo LSTM, o mediante el uso de datos complementarios, como el tiempo de tránsito registrado por la cápsula, si está disponible. Estas mejoras permitirían una localización más precisa, aspecto crucial para la planificación de procedimientos terapéuticos posteriores.

También se destaca la importancia de realizar pruebas piloto en entornos clínicos reales, involucrando a gastroenterólogos en la evaluación del sistema. Estas pruebas permitirían valorar la utilidad, usabilidad e interpretabilidad de la herramienta desde la práctica médica, generando retroalimentación valiosa para optimizar tanto el diseño funcional como la presentación gráfica. Adicionalmente, se plantea la posibilidad de incorporar funcionalidades complementarias, como la generación automática de reportes, exportación de resultados o integración con sistemas de historia clínica electrónica.

Finalmente, se identifica como relevante la exploración de alternativas para optimizar el sistema en contextos con recursos computacionales limitados. Si bien DenseNet201 y DenseNet169 mostraron un rendimiento sobresaliente, su aplicación en dispositivos portátiles o instituciones con menor capacidad tecnológica podría verse restringida. En este sentido, sería recomendable considerar técnicas de compresión de modelos, poda de redes o el uso de arquitecturas más ligeras que conserven un desempeño adecuado con menor demanda computacional.

En conclusión, este proyecto logró desarrollar una herramienta funcional, robusta y clínicamente orientada que explota el potencial del deep learning en beneficio del diagnóstico médico. Cumplió los objetivos propuestos y estableció una base sólida para futuras investigaciones e implementaciones en entornos reales. Si bien existen desafíos por superar, el sistema representa un avance importante hacia la automatización del análisis de imágenes de cápsula endoscópica y la optimización del trabajo del especialista en gastroenterología.

Trabajos futuros

La evolución tecnológica en el ámbito de la salud plantea la necesidad de seguir desarrollando herramientas que mejoren la precisión, eficiencia y accesibilidad del diagnóstico médico. En este contexto, los sistemas automatizados basados en inteligencia artificial representan una solución prometedora para afrontar los desafíos actuales en la revisión de imágenes clínicas. El presente proyecto, centrado en el análisis de imágenes obtenidas mediante cápsula endoscópica, abre la puerta a múltiples líneas de trabajo que permitirán optimizar su funcionamiento, ampliar su alcance y su aplicabilidad en entornos reales.

Una de las principales proyecciones del proyecto es el desarrollo de una interfaz web dinámica y segura, que facilite la recolección continua de estudios clínicos por parte de diferentes usuarios y entidades de salud. Esta plataforma permitirá la carga directa de imágenes, promoviendo la creación de una base de datos estructurada, diversa y en constante expansión. El objetivo de esta herramienta será fortalecer el entrenamiento de los modelos de deep learning, permitiendo que estos evolucionen a partir de datos reales e incrementen su capacidad de aprendizaje, adaptabilidad y rendimiento frente a nuevas condiciones o escenarios clínicos.

También se propone que el sistema sea capaz, no solo de identificar anomalías presentes en el intestino delgado tal como se hace en el presente proyecto, sino que se amplíen las posibilidades de clasificación de imágenes en todo el tracto digestivo. Para ello se plantea, tomar como base el modelo de esta investigación para desarrollar futuros modelos que integren mecanismos de estimación de ubicación, aprovechando patrones espaciales, temporales y texturales propios de cada segmento del tracto digestivo: esófago, estómago, intestino delgado y colon. Esta funcionalidad ofrecerá un valor agregado al proceso diagnóstico, al brindar información adicional que pueda ser de gran utilidad clínica para determinar la localización de las lesiones con mayor precisión, facilitando una toma de decisiones más informada.

Paralelamente, se contempla que se puede continuar alimentando el sistema con el objetivo de que permita ampliar progresivamente el número de patologías que el sistema puede reconocer de manera automática. En este trabajo, Inicialmente, se priorizaron las anomalías más frecuentes y clínicamente relevantes, como úlceras, pólipos, hemorragias y lesiones inflamatorias. Sin embargo, se espera que con la recolección de nuevos datos y su integración a la plataforma, se pueda entrenar modelos que abarquen una mayor variedad de condiciones gastrointestinales, incluyendo las patologías menos comunes como tumores pequeños o lesiones precoces difíciles de distinguir a simple vista. Esto permitirá que el sistema evolucione de una herramienta de apoyo puntual a una plataforma integral de apoyo diagnóstico para el tracto gastrointestinal.

Asimismo, se proyecta la adaptación de esta herramienta para que pueda ser utilizada en otros procedimientos endoscópicos, como la endoscopia tradicional y la colonoscopia. Esta expansión tecnológica implicará nuevos retos en cuanto al procesamiento de imágenes de diferente naturaleza, pero permitirá extender los beneficios del sistema automatizado a un mayor número de estudios clínicos. Por ello, se espera que el sistema no se limite a un tipo específico de examen, sino que se convierta en una solución transversal aplicable a diversas modalidades diagnósticas del aparato digestivo.

Este proyecto no solo busca resolver un problema puntual relacionado con la revisión manual de imágenes, sino que también se proyecta como una plataforma integral en constante evolución, capaz de adaptarse a nuevas necesidades clínicas y tecnológicas. El desarrollo de herramientas digitales, apoyadas en inteligencia artificial y basadas en el trabajo colaborativo multidisciplinario, representa un paso fundamental hacia el futuro del diagnóstico médico moderno, más preciso, automatizado y accesible.

Anexo 1 – Manual de Usuario

Este manual describe el uso del sistema de interfaz gráfica desarrollado para el análisis de imágenes obtenidas mediante cápsula endoscópica del intestino delgado. El sistema ha sido diseñado para ofrecer a los especialistas una herramienta intuitiva que facilite la visualización rápida y precisa de imágenes con posibles patologías. Además de identificar alteraciones clínicas relevantes, la interfaz también permite localizar hitos anatómicos clave, como el píloro y la válvula ileocecal, facilitando así la estimación de la ubicación proximal o distal de las anomalías detectadas. A través de esta integración, se busca optimizar el proceso diagnóstico y mejorar la eficiencia clínica, permitiendo al usuario explorar de forma clara y estructurada los resultados generados por modelos de deep learning entrenados tanto para la clasificación de patologías como para la determinación de la localización anatómica dentro del intestino delgado.

14.1. Instrucciones de uso

A continuación, se describen los pasos necesarios para utilizar correctamente la interfaz gráfica desarrollada. Esta sección guía al usuario desde el inicio de la aplicación hasta la interpretación de los resultados, abarcando el ingreso de datos, la carga de imágenes individuales o estudios completos, y la navegación por las diferentes funciones disponibles.

14.1.1. Registro de paciente y acceso al menú principal

Al iniciar la aplicación, el usuario encontrará una primera pantalla destinada al registro de los datos del paciente y del especialista tratante. Se deben ingresar información básica como el nombre, número de identificación, edad y médico responsable. Tras completar los campos, debe hacer clic en el botón *Siguiente* para continuar con el análisis. Esta etapa inicial es esencial, pues permite asociar adecuadamente los resultados generados con cada estudio individual.

CLASIFICACIÓN DE IMÁGENES DE CÁPSULA ENDOSCÓPICA DE INTESTINO DELGADO

IBS Pontificia Universidad JAVERIANA Cali

DATOS DEL PACIENTE

Nombre:

Identificación:

Edad:

Médico:

Siguiente

El formulario muestra cuatro campos de entrada para los datos del paciente: Nombre, Identificación, Edad y Médico. Cada campo tiene un cursor de texto y una flecha negra que apunta a la derecha. El campo de 'Identificación' está rodeado por un círculo rojo.

Es obligatorio completar todos los campos señalados. Estos datos garantizan un correcto registro del paciente y habilitan el acceso a las siguientes funcionalidades de la interfaz.

Nota: Ingrese únicamente letras en los campos de *Nombre* y *Médico responsable*, y solo caracteres numéricos en los campos de *Edad* e *Identificación*. El sistema no permitirá avanzar si la información no cumple con este formato.

Posterior al registro, el usuario será dirigido al **Menú Principal**, desde donde podrá iniciar el análisis de imágenes, revisar estudios anteriores o salir de la aplicación.

Paciente: Paula Cruz (ID: 123456789), Edad: 44 | Médico: Gabriel Gomez

Menú Principal

Analizar Imagen Individual

Analizar Examen Completo

Historial de Imágenes

Salir

El menú principal muestra tres botones verdes con texto blanco: 'Analizar Imagen Individual', 'Analizar Examen Completo' y 'Historial de Imágenes'. Cada botón tiene una flecha negra que apunta a la izquierda. El botón 'Salir' es gris y está rodeado por un círculo rojo.

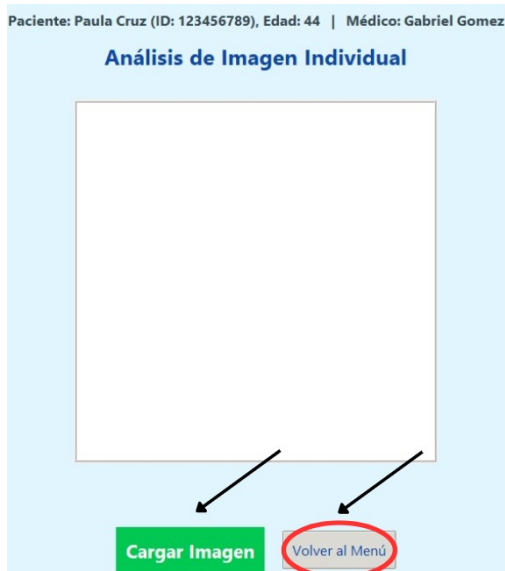
Nota: El botón *Salir* ubicado en la parte inferior permite cerrar la aplicación de forma segura.

14.1.2. Opciones principales del menú

En el Menú Principal, el usuario encontrará las siguientes opciones para continuar con el análisis:

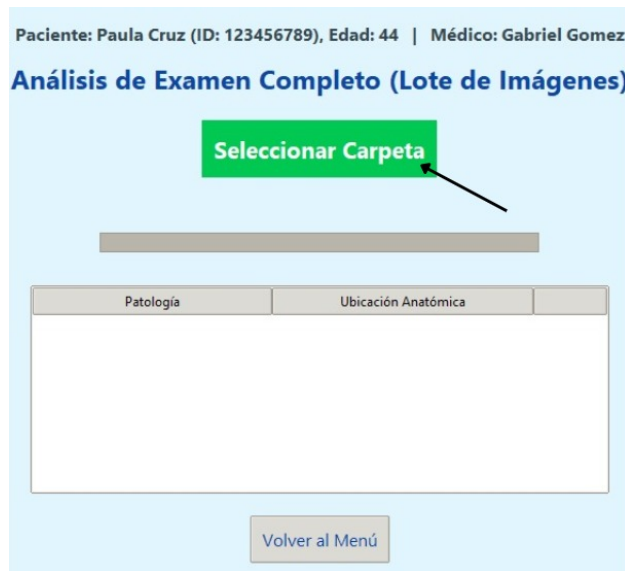
14.1.2.1. Analizar Imagen Individual

Permite cargar una imagen aislada para su análisis inmediato. Al seleccionar el botón *cargar imagen*, se abrirá una ventana para elegir el archivo (formato compatible: .jpg, .png). Una vez cargada, el sistema procesará la imagen y mostrará en pantalla el diagnóstico estimado junto con la categoría anatómica correspondiente.

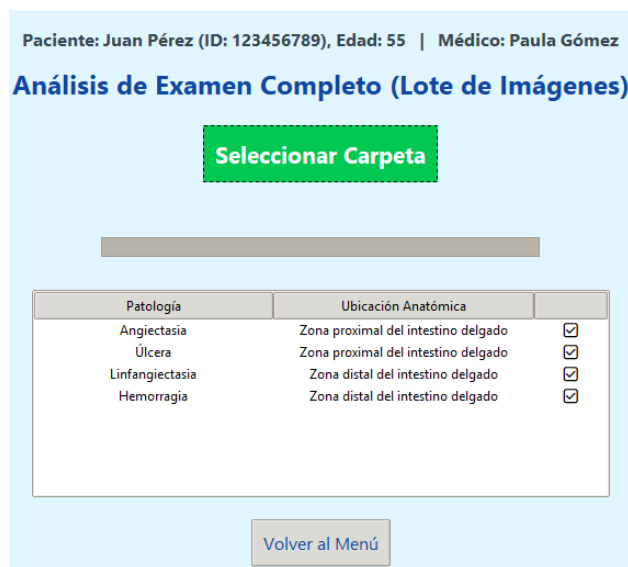


14.1.2.2. Analizar Examen Completo

Esta opción está diseñada para analizar múltiples imágenes de un mismo paciente. Al seleccionarla, se mostrará el botón *seleccionar carpeta*, donde podrá cargar la carpeta con el conjunto de imágenes endoscópicas.



Después de seleccionar la carpeta correspondiente al estudio, el sistema procesará automáticamente todas las imágenes. Posteriormente, se mostrarán los resultados de aquellas en las que se detectó alguna patología, organizados en una tabla que indica la clase patológica y la ubicación anatómica estimada.



Al hacer clic sobre cualquiera de las patologías listadas, se abrirá automáticamente una ventana emergente mostrando las imágenes específicas en las que se ha detectado la patología seleccionada. Esto permite revisar de forma detallada las evidencias visuales y facilita la interpretación clínica.



14.1.2.3. Historial de Imágenes

Permite acceder a los resultados de las imágenes analizadas durante la sesión actual. El historial incluye tanto imágenes individuales como estudios completos, mostrando los resultados tal como fueron generados inicialmente.

Paciente: Juan Pérez (ID: 123456789), Edad: 55 | Médico: Paula Gómez

Historial de Imágenes Procesadas

Patología	Ubicación Anatómica	
Úlcera	Zona proximal del intestino delgado	<input checked="" type="checkbox"/>
Angiectasia	Zona proximal del intestino delgado	<input checked="" type="checkbox"/>
Úlcera	Zona proximal del intestino delgado	<input checked="" type="checkbox"/>
Úlcera	Zona proximal del intestino delgado	<input checked="" type="checkbox"/>
Úlcera	Zona proximal del intestino delgado	<input checked="" type="checkbox"/>
Linfangiectasia	Zona distal del intestino delgado	<input checked="" type="checkbox"/>
Linfangiectasia	Zona distal del intestino delgado	<input checked="" type="checkbox"/>
Hemorragia	Zona distal del intestino delgado	<input checked="" type="checkbox"/>

[Reanalizar Selección](#) [Volver al Menú](#)

Nota: En cualquiera de las secciones, si selecciona el botón *Volver al Menú*, será redirigido nuevamente al Menú Principal. Esto le permite regresar para seleccionar otra opción o consultar el historial sin reiniciar la aplicación.

14.2. Guía de solución de problemas

En esta sección encontrará respuestas a las preguntas más frecuentes sobre el uso del sistema, orientadas a resolver dudas y facilitar la interacción autónoma con la interfaz.

- **¿Puedo volver a analizar la misma imagen?**
Sí. Puede cargar nuevamente la imagen y repetir el análisis sin restricciones.
- **¿Qué sucede si se carga una imagen incorrecta?**
Puede seleccionar una nueva imagen. El sistema no continuará con el análisis hasta que se cargue una imagen válida.
- **¿Dónde se almacenan los resultados generados?**
Actualmente, el sistema no cuenta con almacenamiento permanente. Esta función está contemplada para futuras versiones.
- **¿Qué tipo de imágenes acepta el sistema?**
Imágenes en formatos estándar (.jpg, .jpeg, .png) correspondientes a estudios de cápsula endoscópica del intestino delgado.
- **¿Puedo cargar más de una imagen al mismo tiempo?**
Sí, mediante la opción *Analizar Examen Completo* puede seleccionar una carpeta con múltiples imágenes.
- **¿Qué pasa si no ingreso todos los datos del paciente?**
Todos los campos obligatorios deben estar diligenciados para habilitar el análisis.
- **¿La aplicación funciona sin conexión a internet?**
Sí. Tanto la interfaz como los modelos están integrados localmente.
- **¿Cómo salgo del sistema de forma segura?**
Haga clic en el botón *Salir* del menú principal para cerrar la aplicación correctamente.

Anexo 2 – Informe de Rendimiento de Modelos

Este anexo presenta el informe detallado del rendimiento de los modelos de deep learning desarrollados para la clasificación automática de patologías y la detección de hitos anatómicos en imágenes de cápsula endoscópica del intestino delgado. Se incluyen tanto los resultados cuantitativos obtenidos tras cinco ciclos de entrenamiento con diferentes semillas como las representaciones visuales asociadas al desempeño de cada modelo en las distintas tareas propuestas. El objetivo principal es proporcionar una visión completa del comportamiento de los modelos, permitiendo comparar su capacidad de generalización y robustez.

Para evitar que los resultados dependieran de una única división de datos o de una sola inicialización de pesos, se implementó un esquema de entrenamiento basado en cinco ciclos por cada arquitectura, empleando cinco semillas distintas: 42, 123, 2025, 1 y 999. De este modo, cada variante DenseNet fue entrenada cinco veces, una por cada semilla, tanto para la tarea de clasificación de patologías como para la detección de hitos anatómicos. Este enfoque permitió realizar un total de 30 experimentos exhaustivos, proporcionando un marco de evaluación más robusto y permitiendo analizar la variabilidad y estabilidad de los modelos bajo diferentes condiciones de aleatoriedad.

14.2.1. DenseNet201 – Clasificación de Patologías

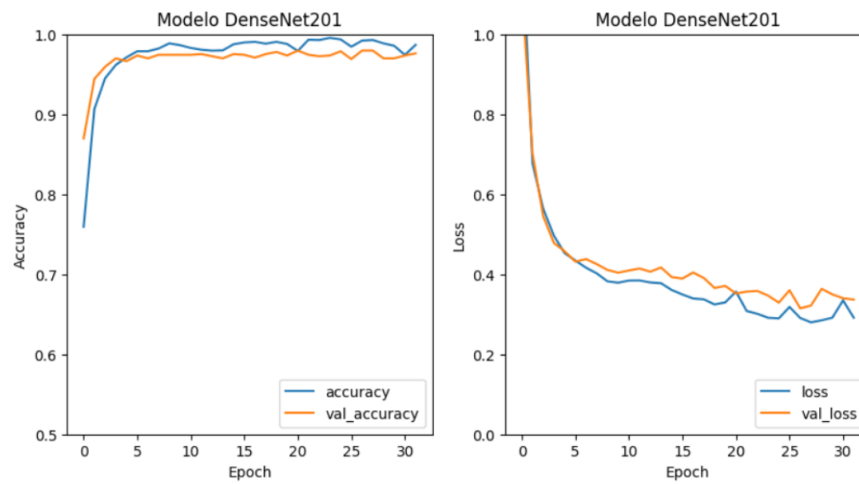
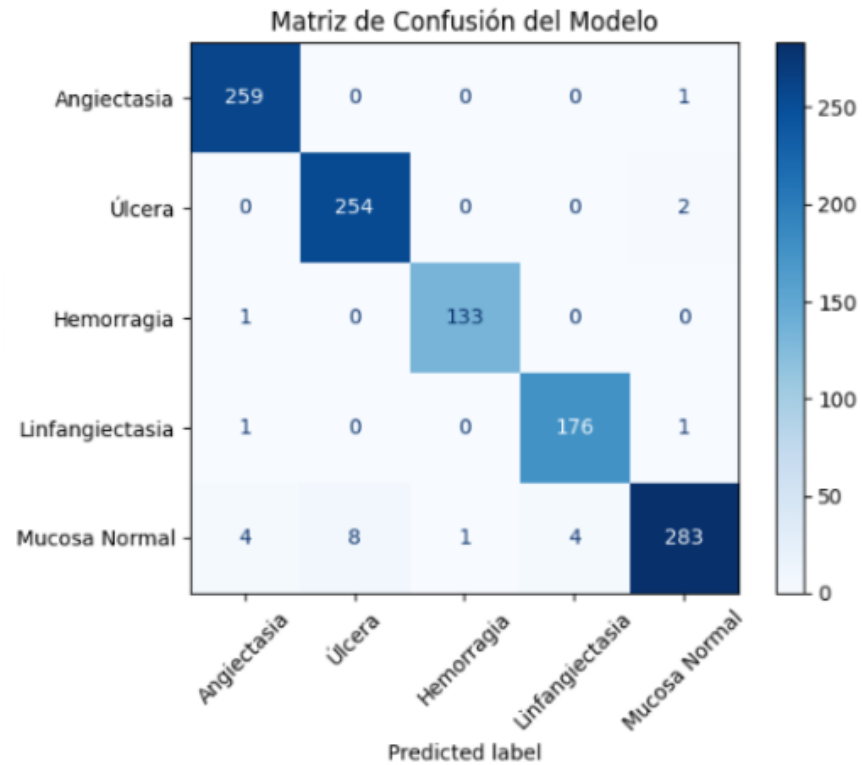
14.2.1.1. Resumen de métricas promedio (5 ciclos)

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.976 ± 0.010	0.982 ± 0.015	0.978 ± 0.007
Úlcera	0.968 ± 0.012	0.986 ± 0.005	0.976 ± 0.005
Hemorragia	0.992 ± 0.004	0.982 ± 0.007	0.986 ± 0.005
Linfangiectasia	0.972 ± 0.010	0.976 ± 0.012	0.974 ± 0.008
Mucosa Normal	0.968 ± 0.012	0.948 ± 0.012	0.954 ± 0.008
Precisión global (Accuracy)			0.972 ± 0.004

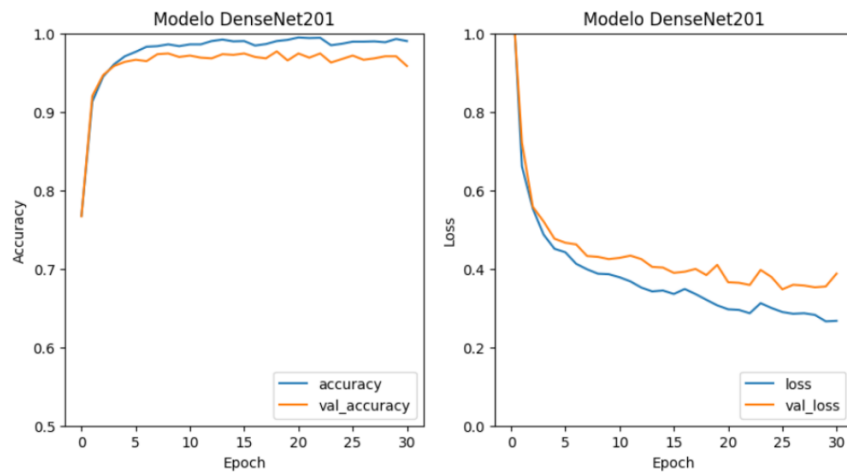
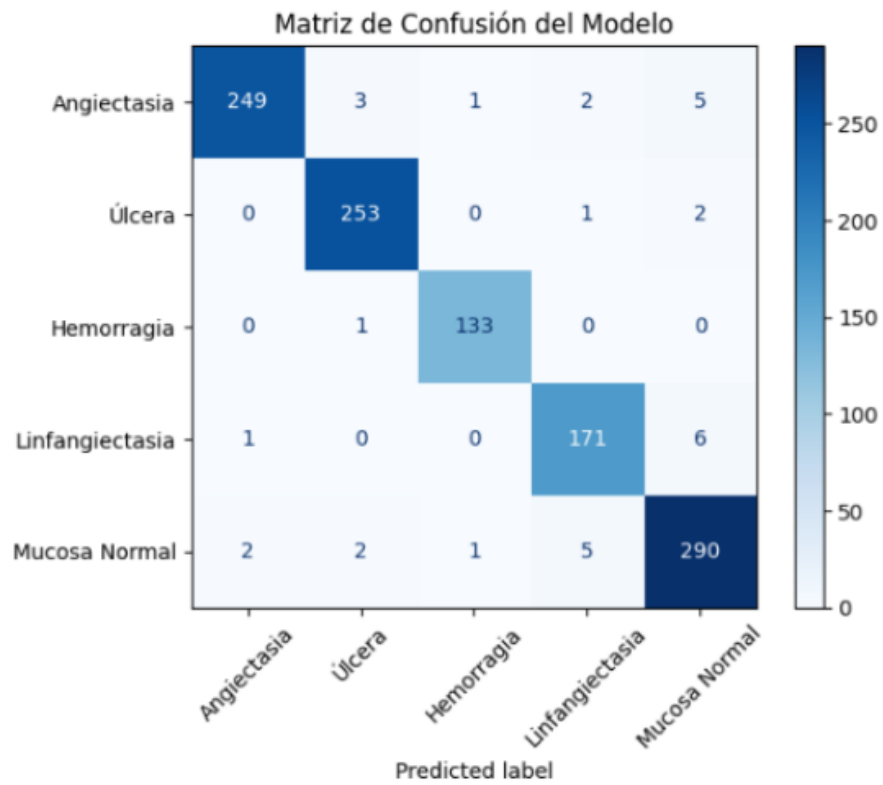
Luego de este resumen de métricas promedio para la tarea de clasificación de patologías con DenseNet201, se presentan los resultados individuales correspondientes a cada uno de los cinco ciclos de entrenamiento y validación. Para cada ciclo, se incluyen la matriz de confusión y las curvas de accuracy y loss, lo que permite observar el comportamiento del modelo bajo diferentes inicializaciones y analizar la estabilidad de su desempeño en la detección de patologías.

14.2.2. Resultados por ciclo

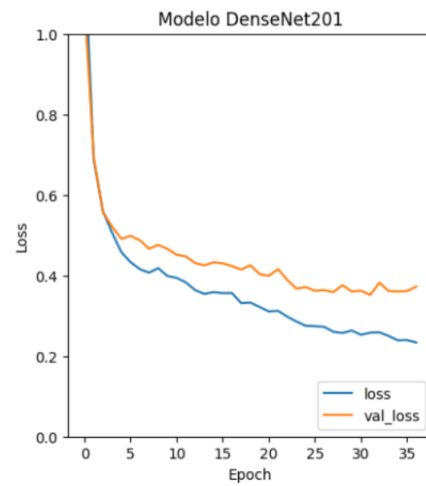
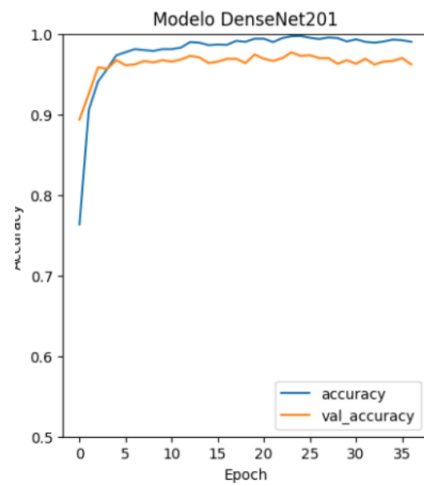
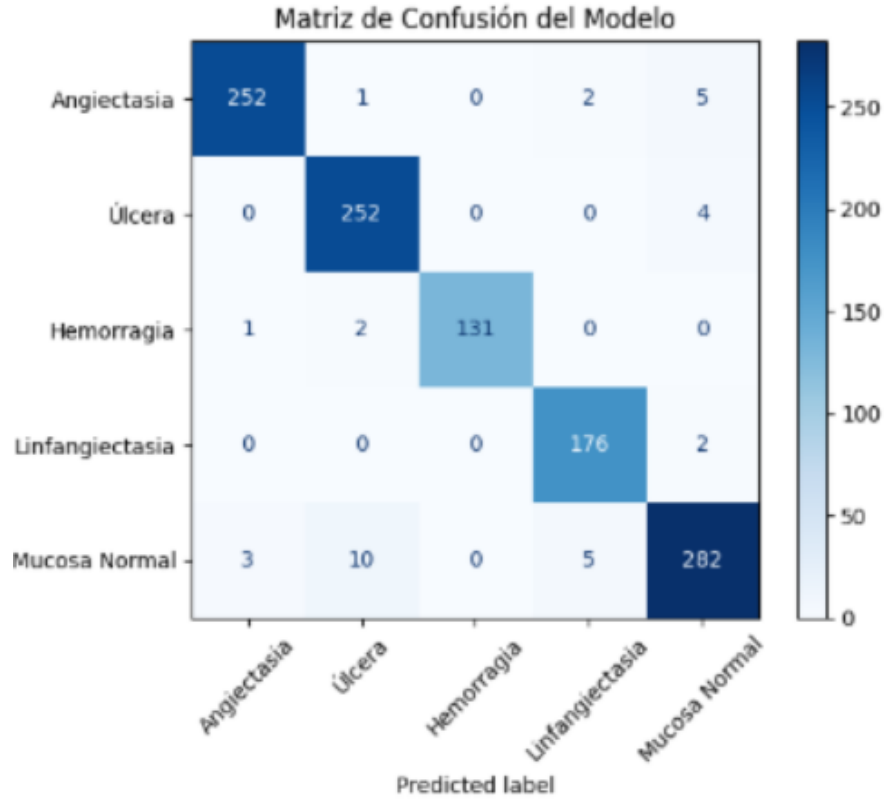
Ciclo 1 (semilla = 42)



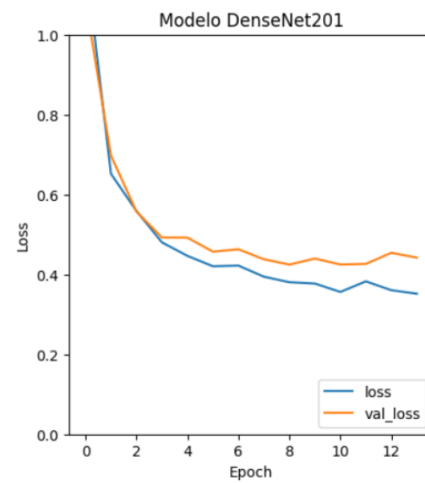
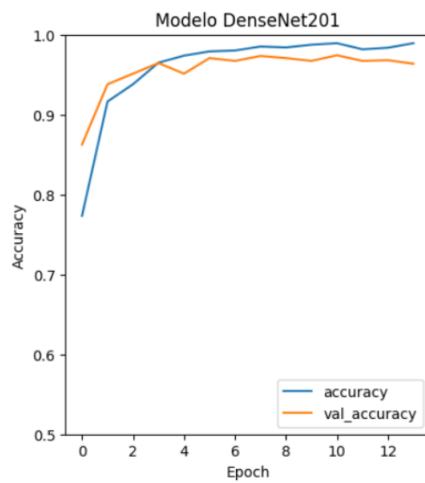
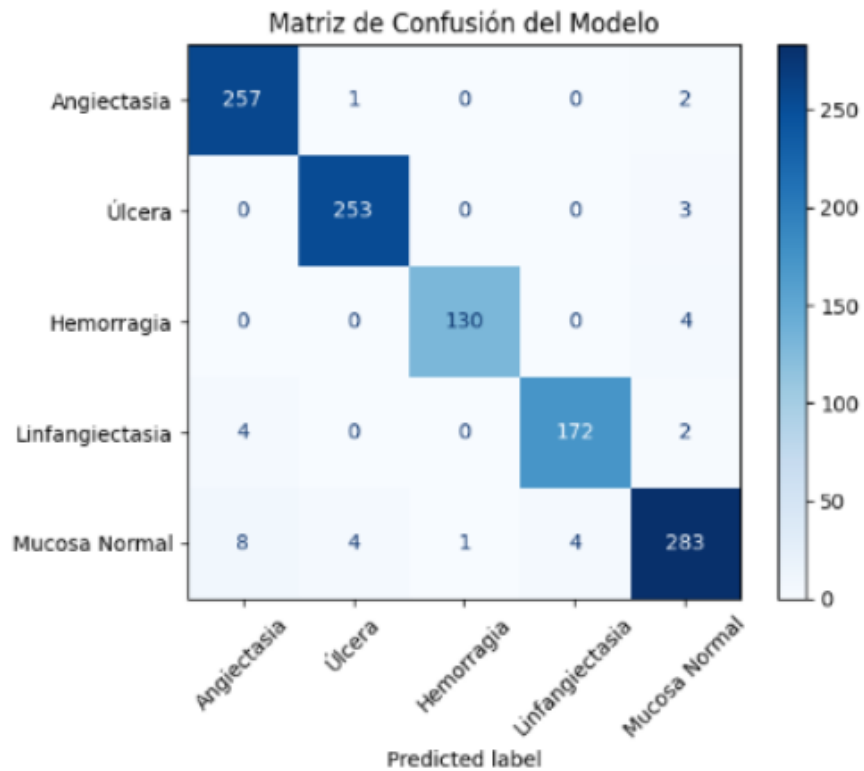
- Ciclo 2 (semilla = 123)



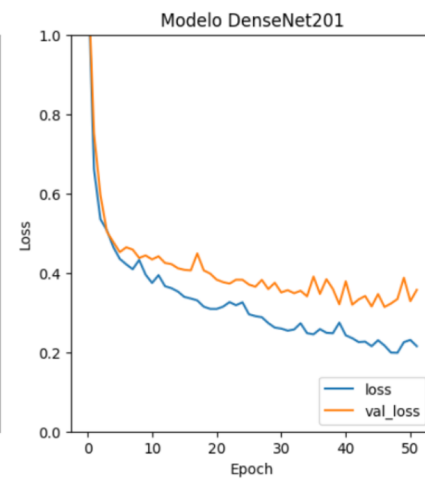
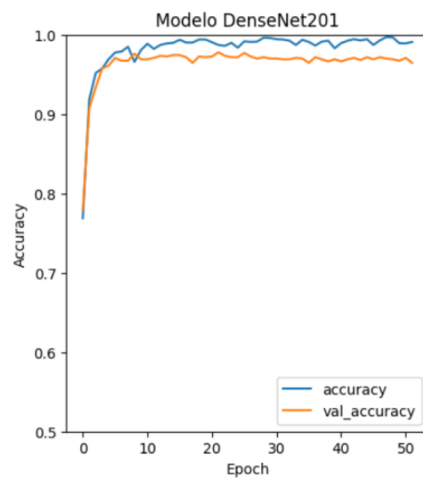
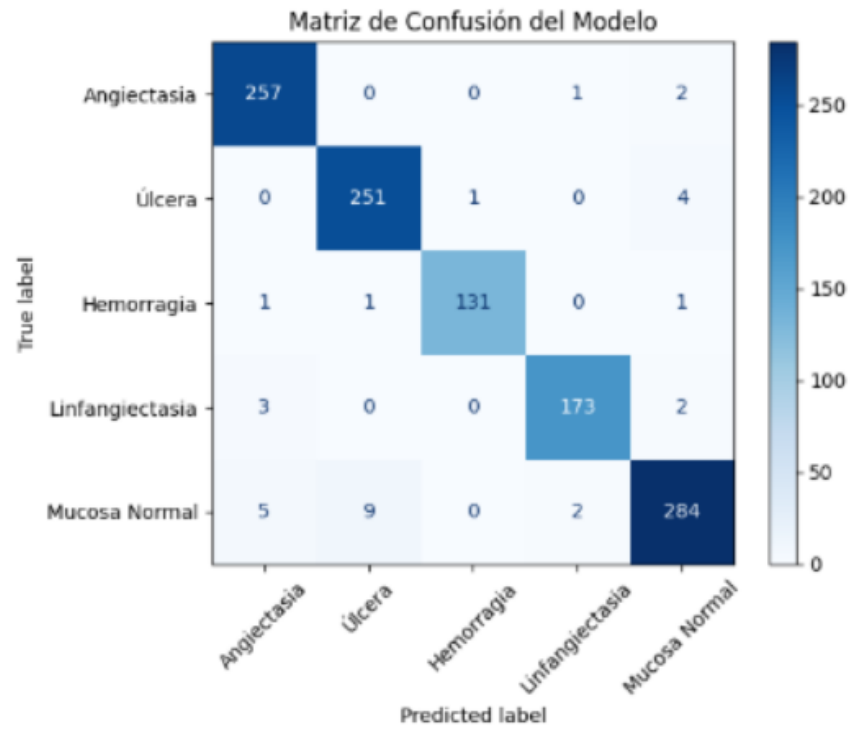
Ciclo 3 (semilla = 2025)



- Ciclo 4 (semilla = 1)



Ciclo 5 (semilla = 999)



14.3. DenseNet201 – Detección de Hitos Anatómicos

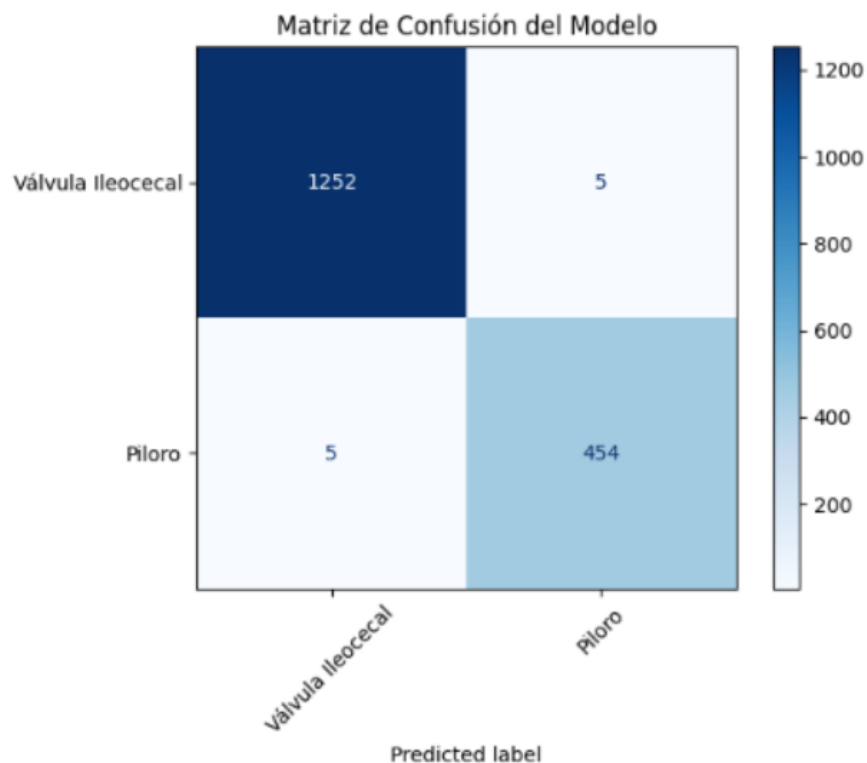
14.3.1. Resumen de métricas promedio (5 ciclos)

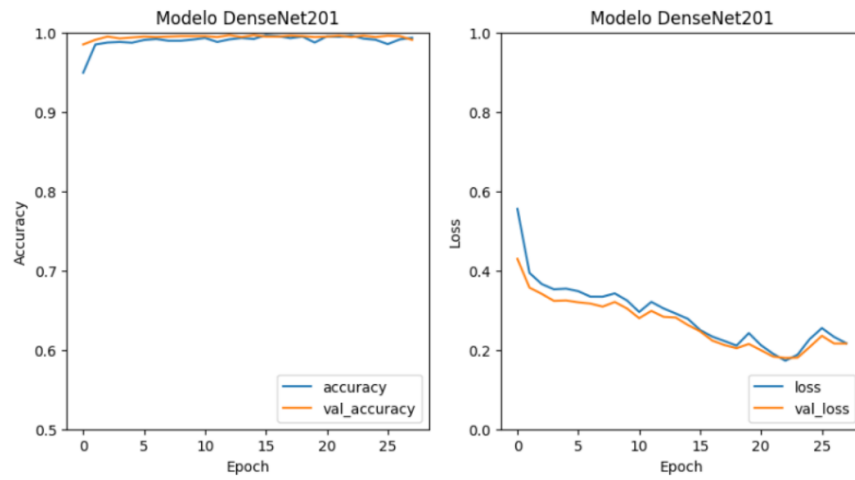
Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.992 ± 0.012	0.982 ± 0.007	0.986 ± 0.005
Válvula Ileocecal	0.994 ± 0.005	0.998 ± 0.004	0.996 ± 0.005
Precisión global (Accuracy)			0.992 ± 0.002

Finalizado el resumen de métricas promedio para la detección de hitos anatómicos con DenseNet201, a continuación se muestran los resultados individuales de cada uno de los cinco ciclos realizados. En cada ciclo, se presentan la matriz de confusión y las curvas de accuracy y loss, facilitando así el análisis detallado de la precisión y consistencia del modelo en la identificación de los hitos anatómicos.

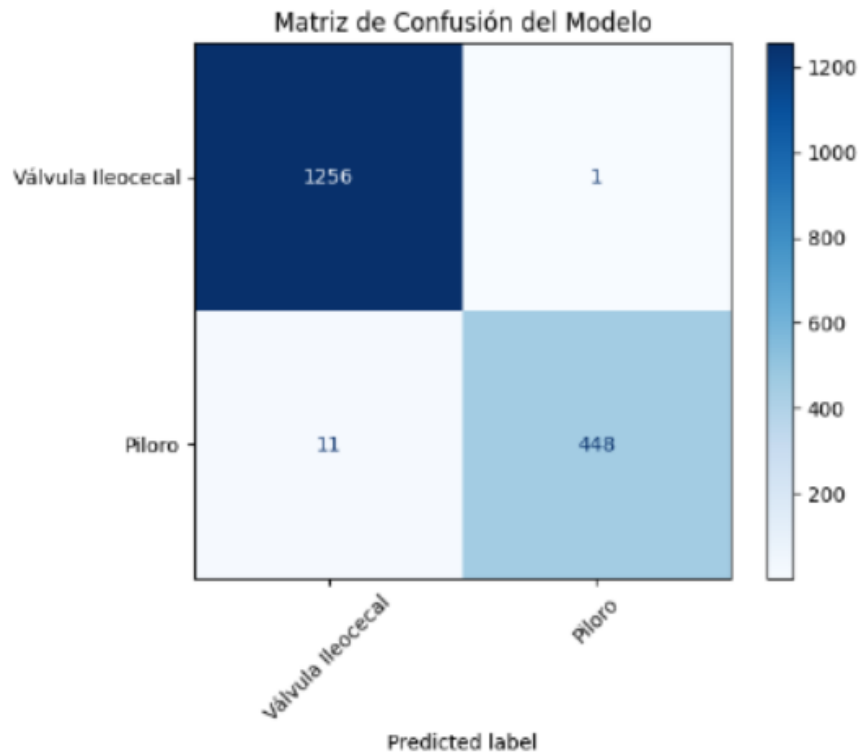
14.3.2. Resultados por ciclo

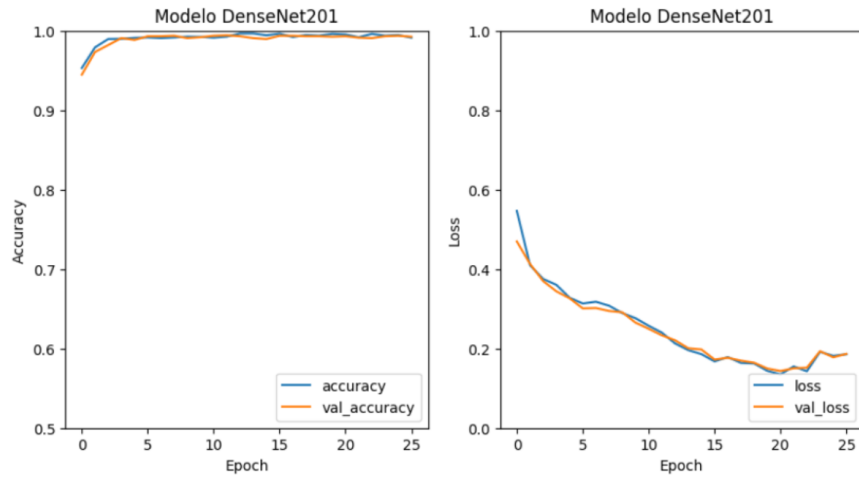
Ciclo 1 (semilla = 42)



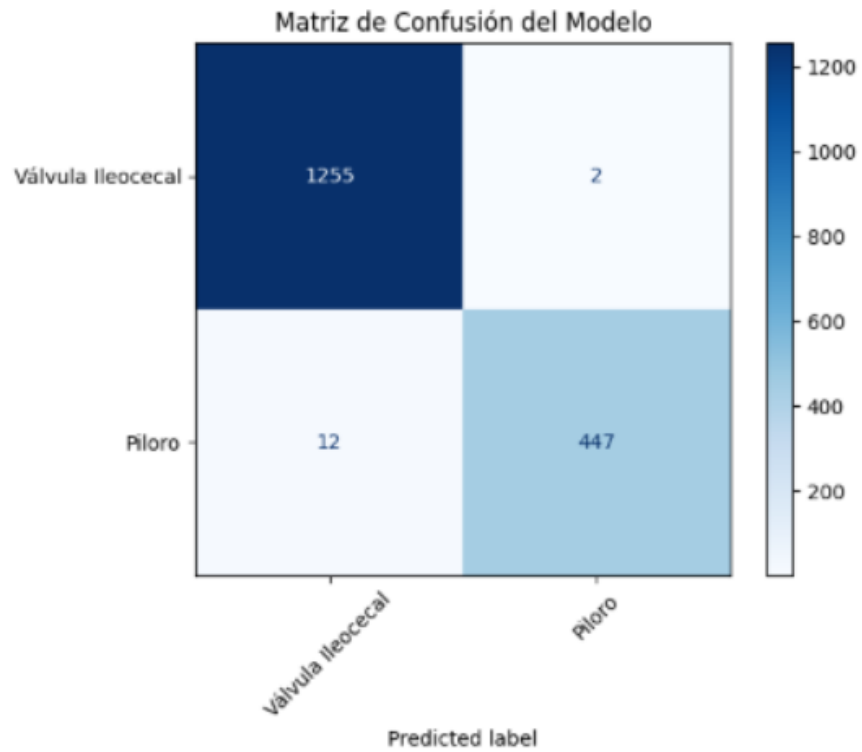


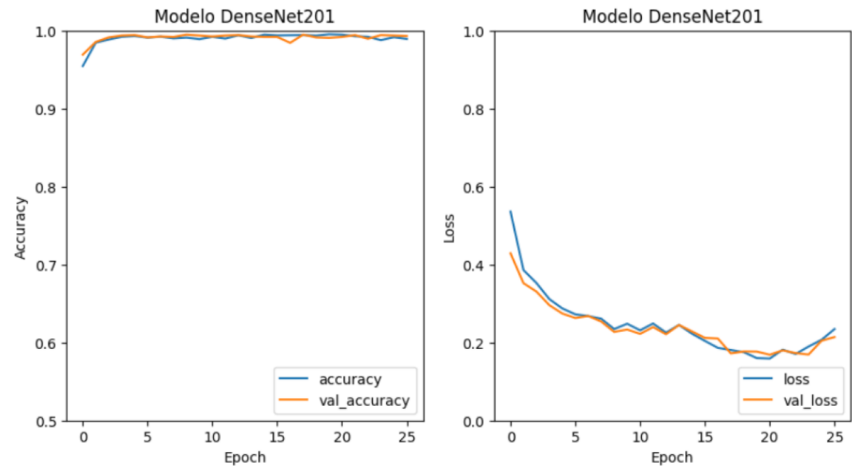
Ciclo 2 (semilla = 123)



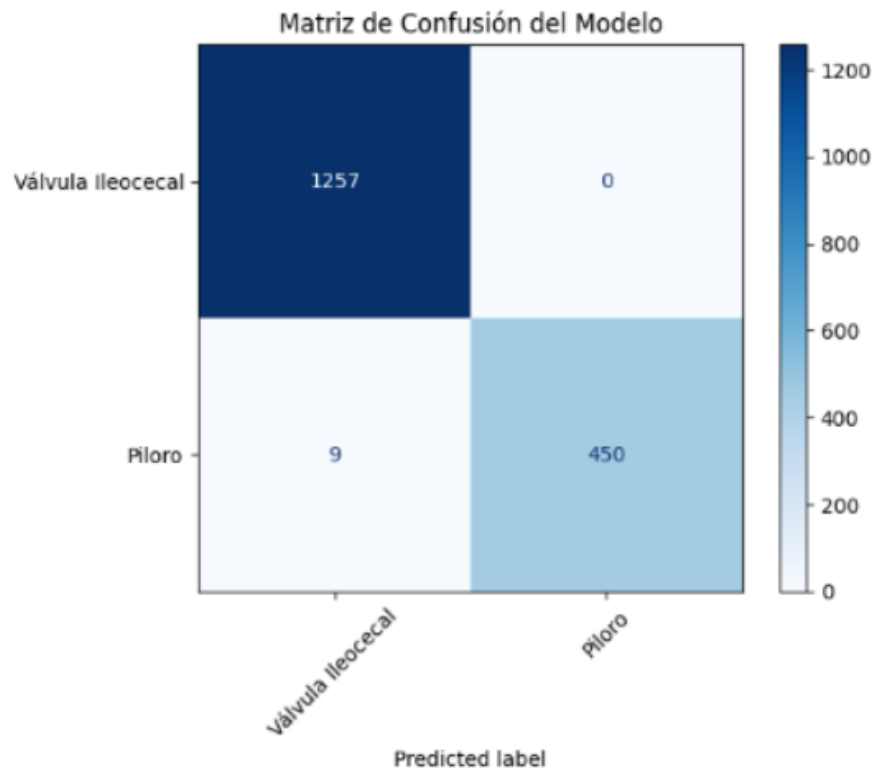


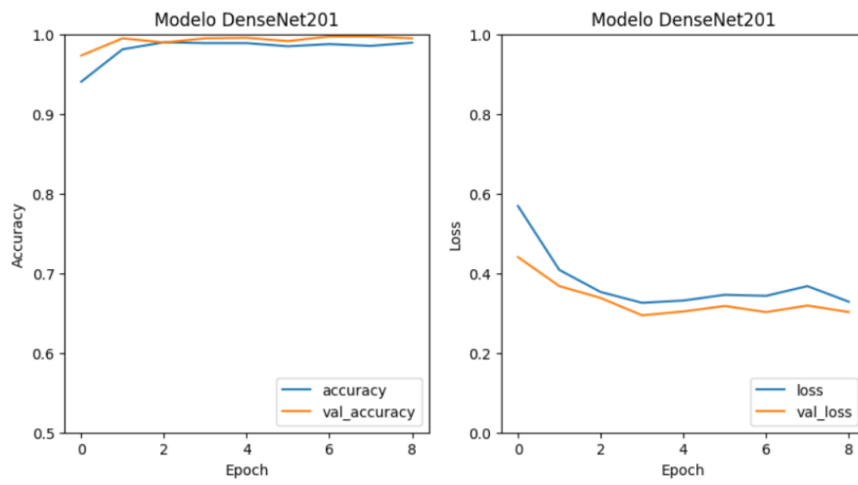
Ciclo 3 (semilla = 2025)



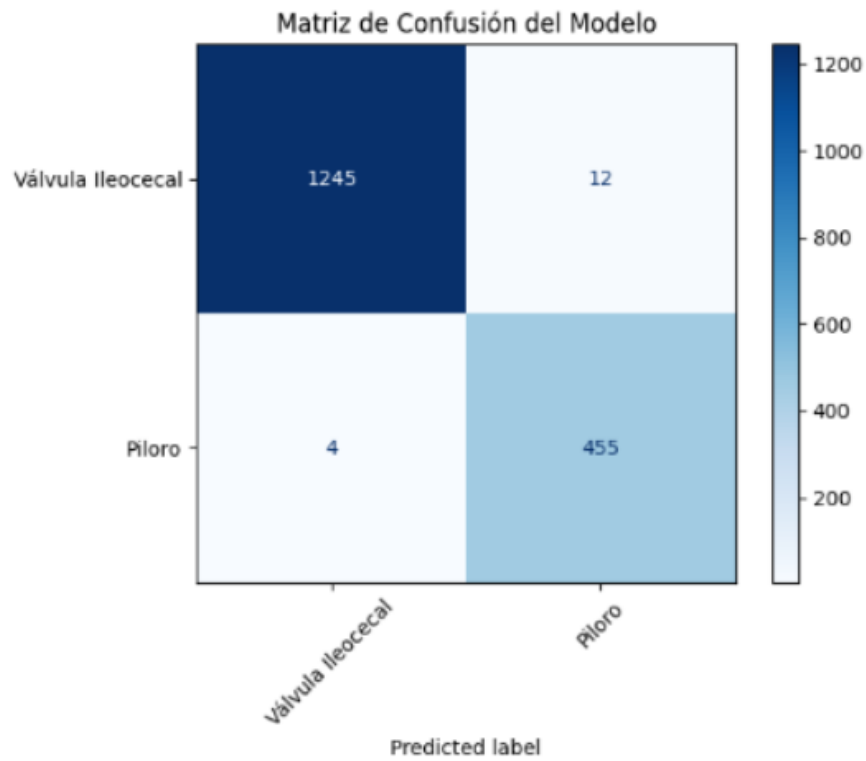


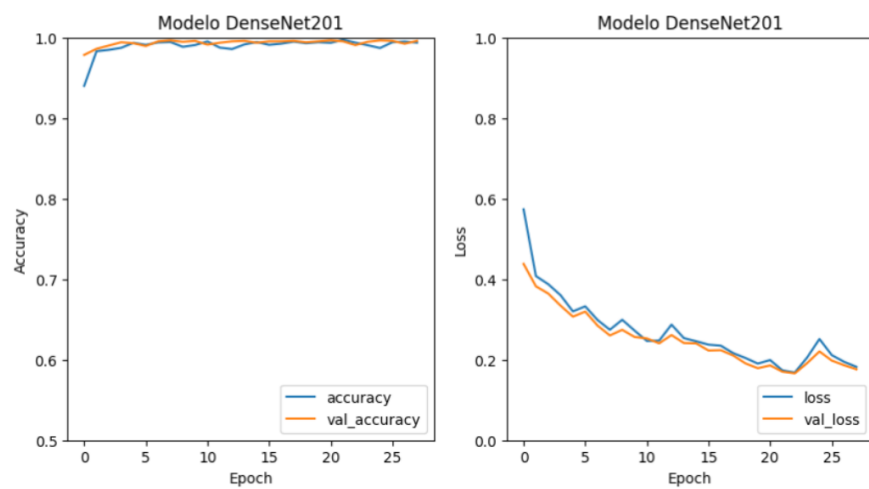
Ciclo 4 (semilla = 1)





Ciclo 5 (semilla = 999)





14.4. DenseNet169 – Clasificación de Patologías

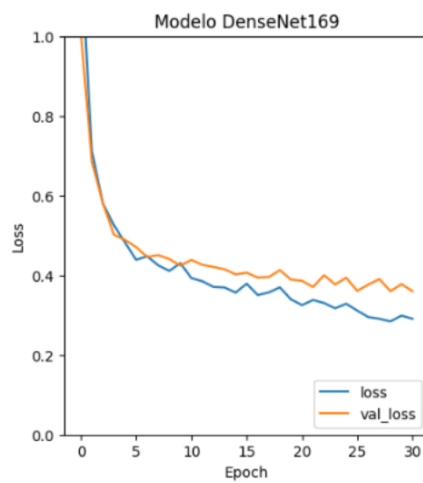
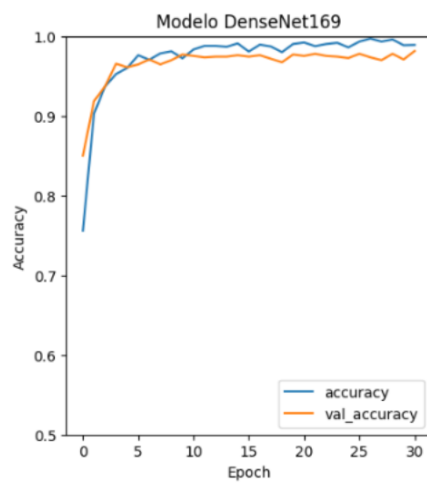
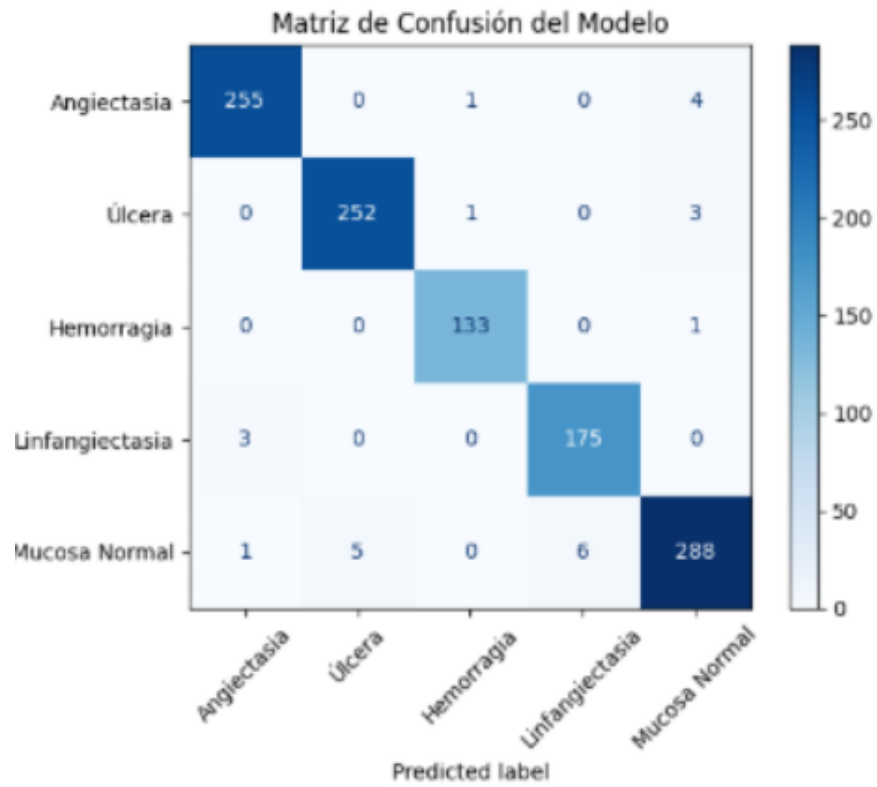
14.4.1. Resumen de métricas promedio (5 ciclos)

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.976 ± 0.008	0.972 ± 0.010	0.974 ± 0.005
Úlcera	0.978 ± 0.007	0.978 ± 0.004	0.976 ± 0.005
Hemorragia	0.996 ± 0.005	0.986 ± 0.010	0.990 ± 0.006
Linfangiectasia	0.962 ± 0.007	0.970 ± 0.009	0.964 ± 0.005
Mucosa Normal	0.954 ± 0.010	0.954 ± 0.008	0.954 ± 0.010
Precisión global (Accuracy)			0.970 ± 0.004

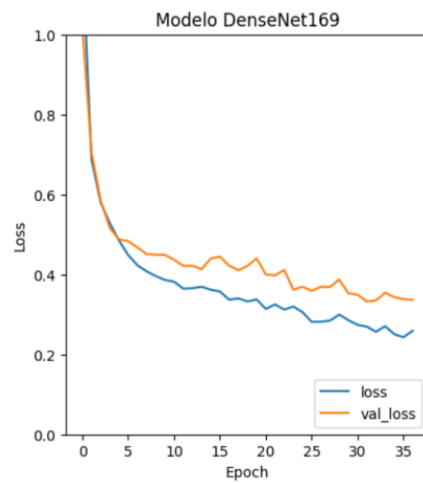
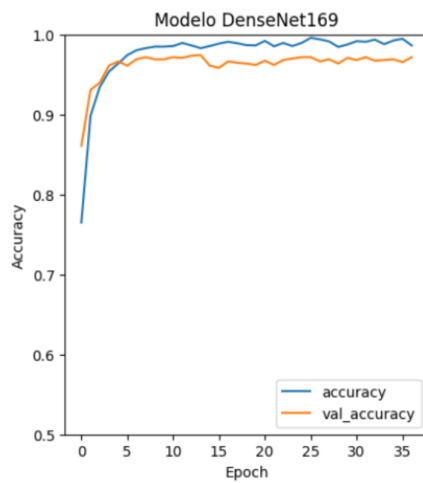
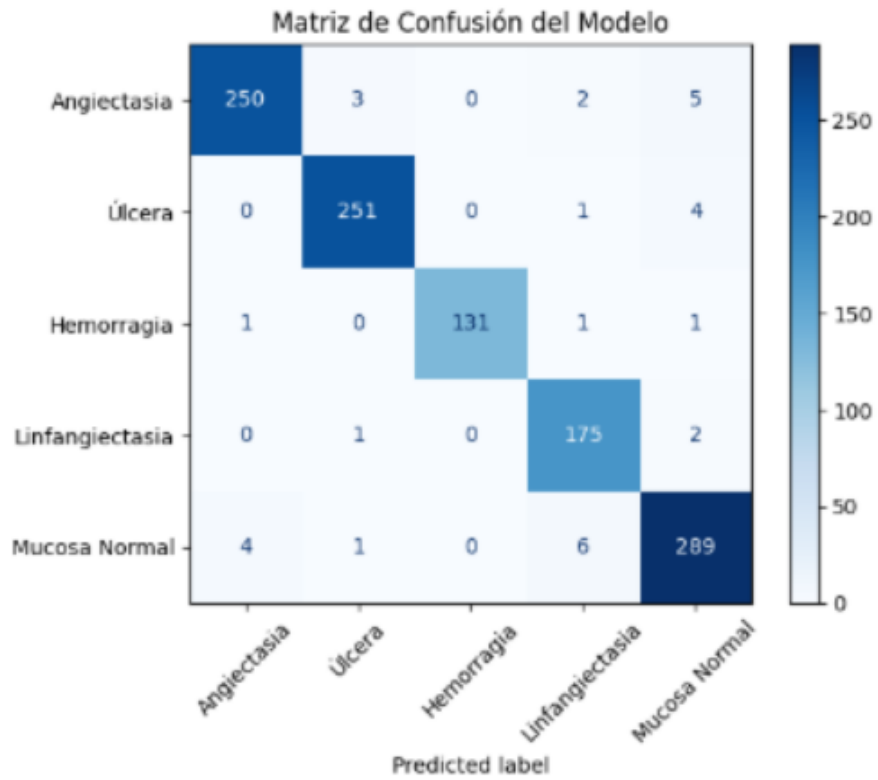
Tras la presentación del resumen de métricas promedio para la tarea de clasificación de patologías con DenseNet169, se exponen los resultados individuales obtenidos en los cinco ciclos de entrenamiento y validación. En cada caso, se reporta la matriz de confusión junto con las curvas de accuracy y loss, lo que permite evaluar el rendimiento específico del modelo en cada ciclo y la robustez de su comportamiento ante diferentes condiciones de inicialización.

14.4.2. Resultados por ciclo

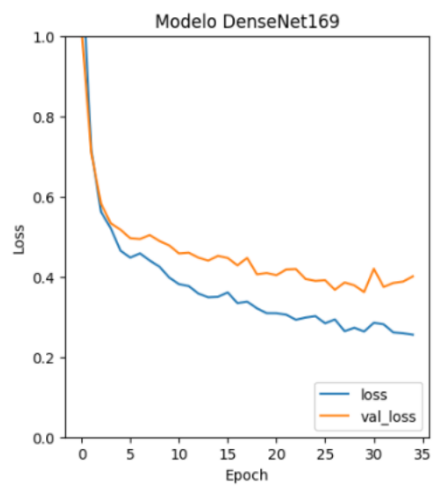
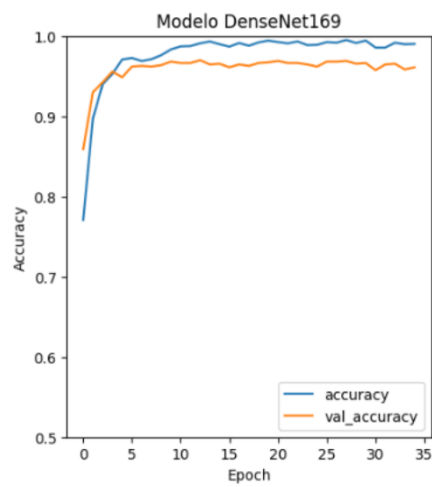
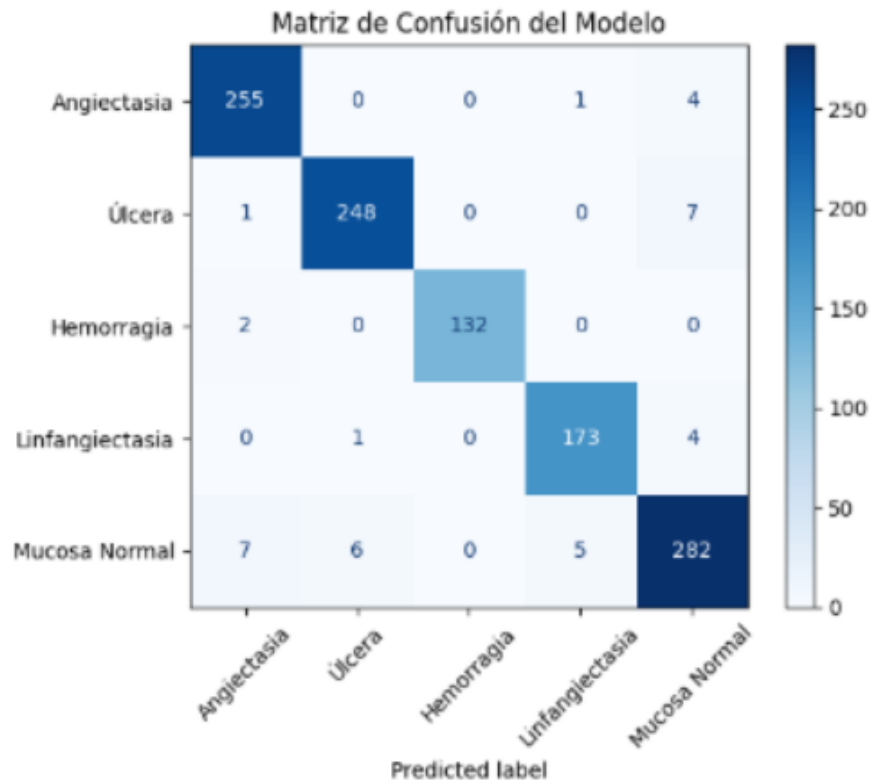
Ciclo 1 (semilla = 42)



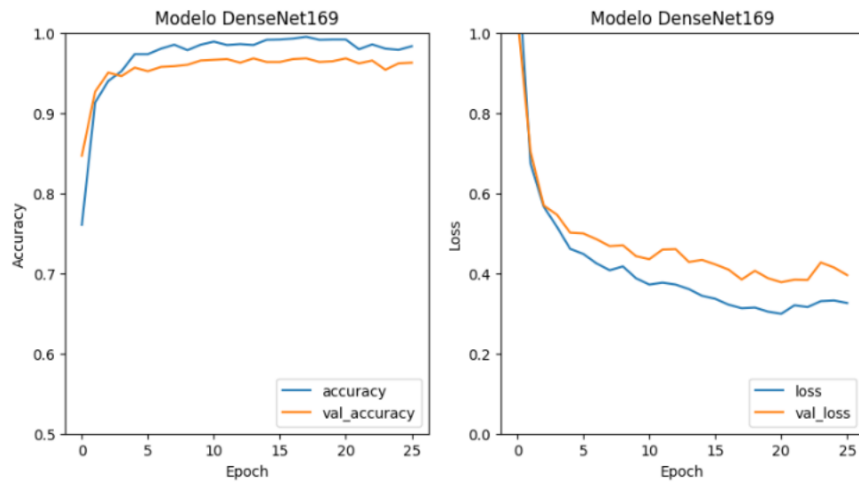
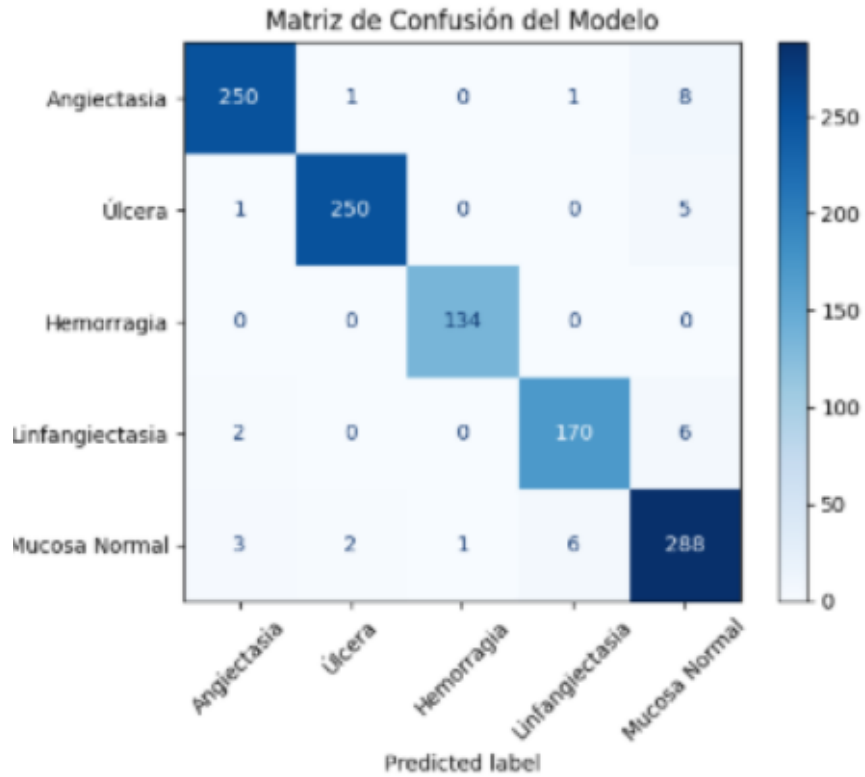
Ciclo 2 (semilla = 123)



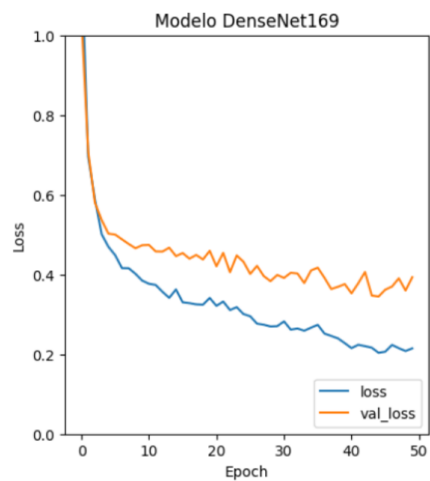
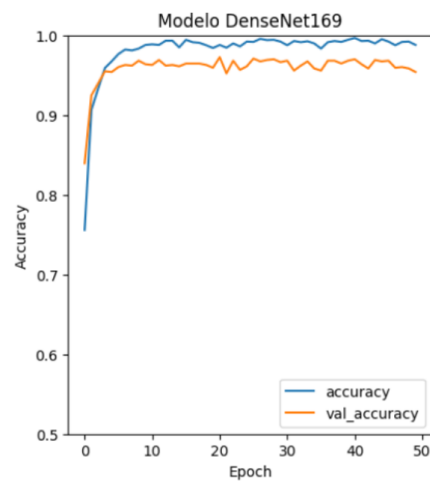
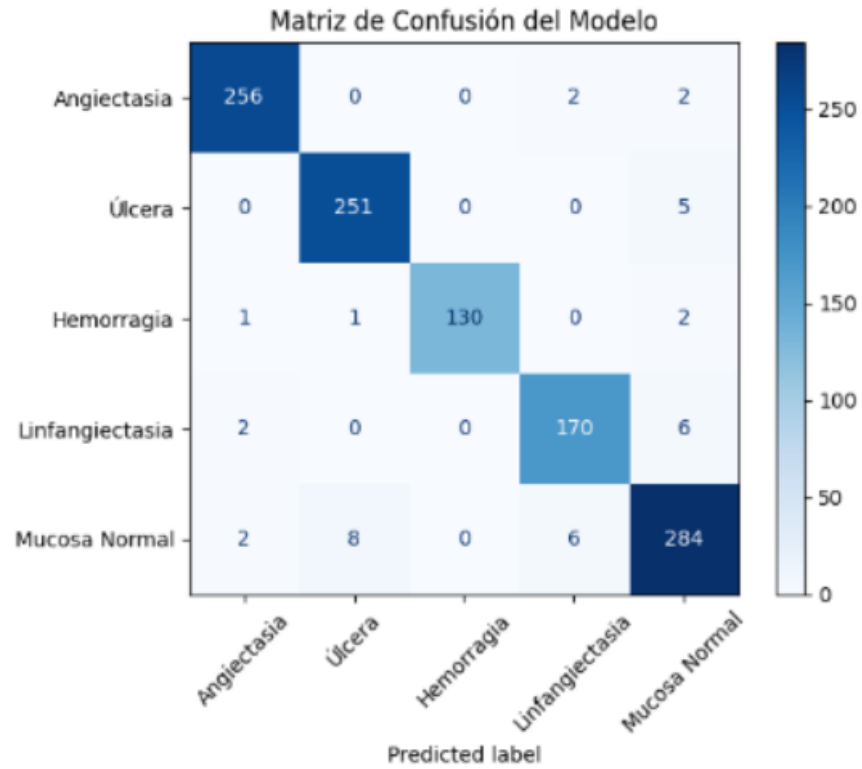
Ciclo 3 (semilla = 2025)



Ciclo 4 (semilla = 1)



Ciclo 5 (semilla = 999)



14.5. DenseNet169 – Detección de Hitos Anatómicos

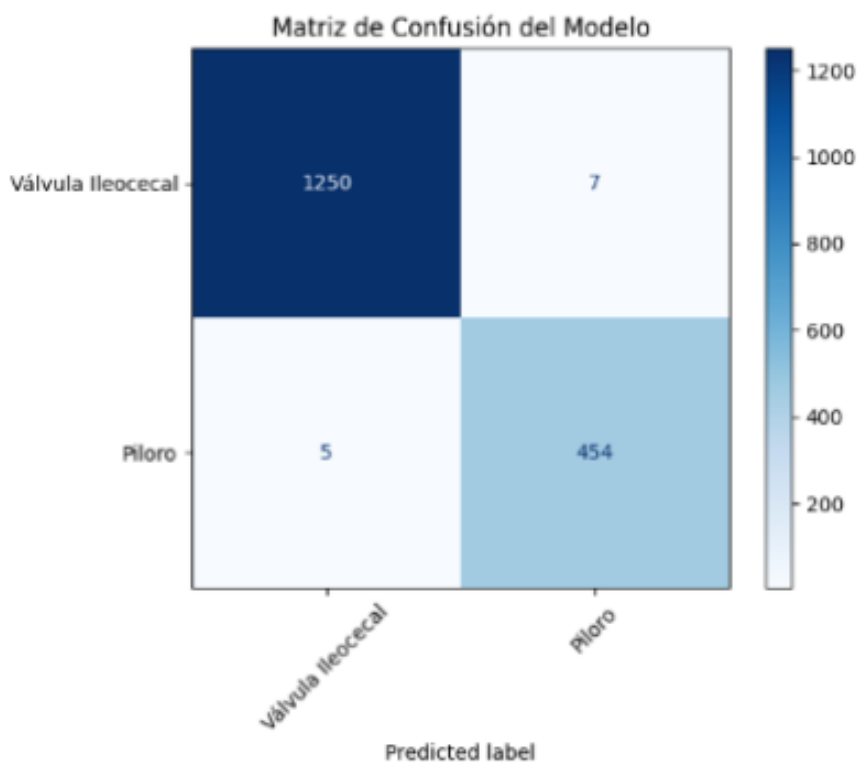
14.5.1. Resumen de métricas promedio (5 ciclos)

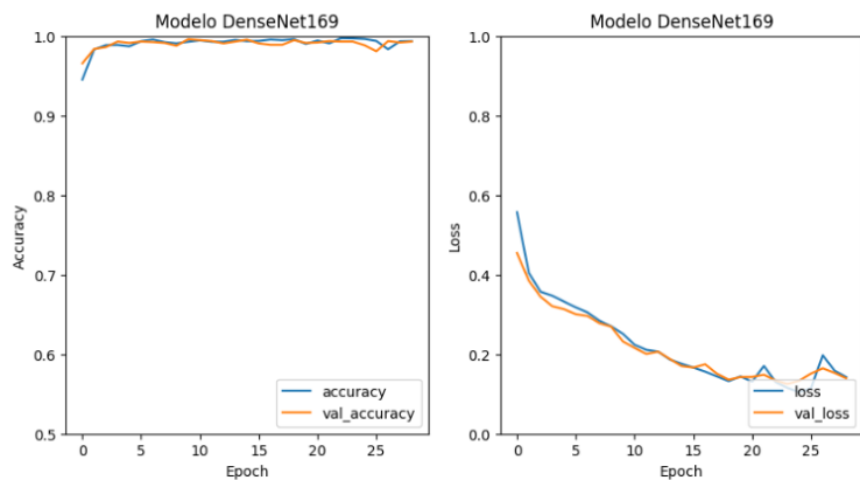
Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.988 ± 0.004	0.982 ± 0.010	0.986 ± 0.005
Válvula Ileocecal	0.996 ± 0.005	0.998 ± 0.004	0.996 ± 0.005
Precisión global (Accuracy)			0.991 ± 0.002

Posterior al resumen de métricas promedio correspondiente a la detección de hitos anatómicos con DenseNet169, se incluyen los resultados individuales para los cinco ciclos experimentales. Cada ciclo cuenta con su respectiva matriz de confusión y las curvas de accuracy y loss, lo que brinda una visión detallada sobre la estabilidad y desempeño del modelo en la tarea de localización anatómica.

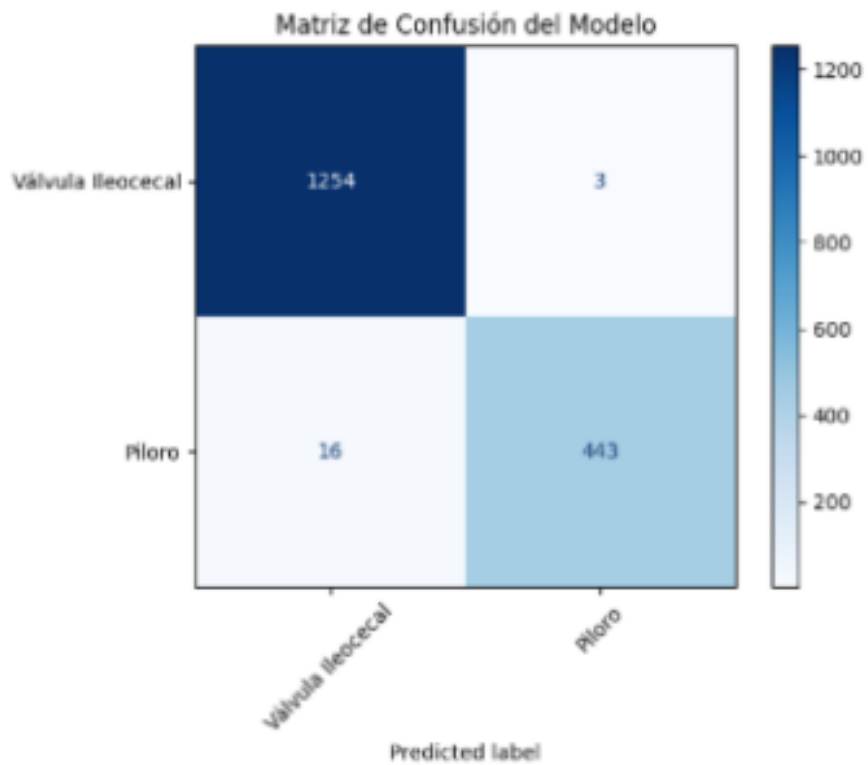
14.5.2. Resultados por ciclo

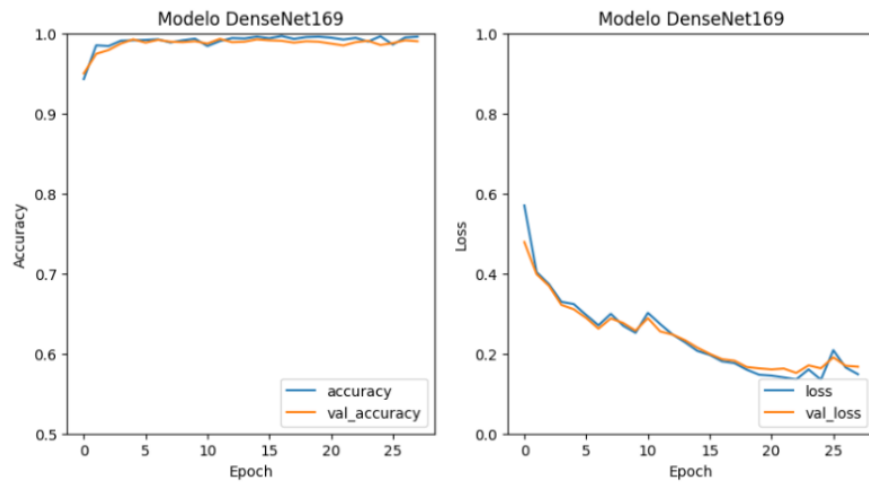
Ciclo 1 (semilla = 42)



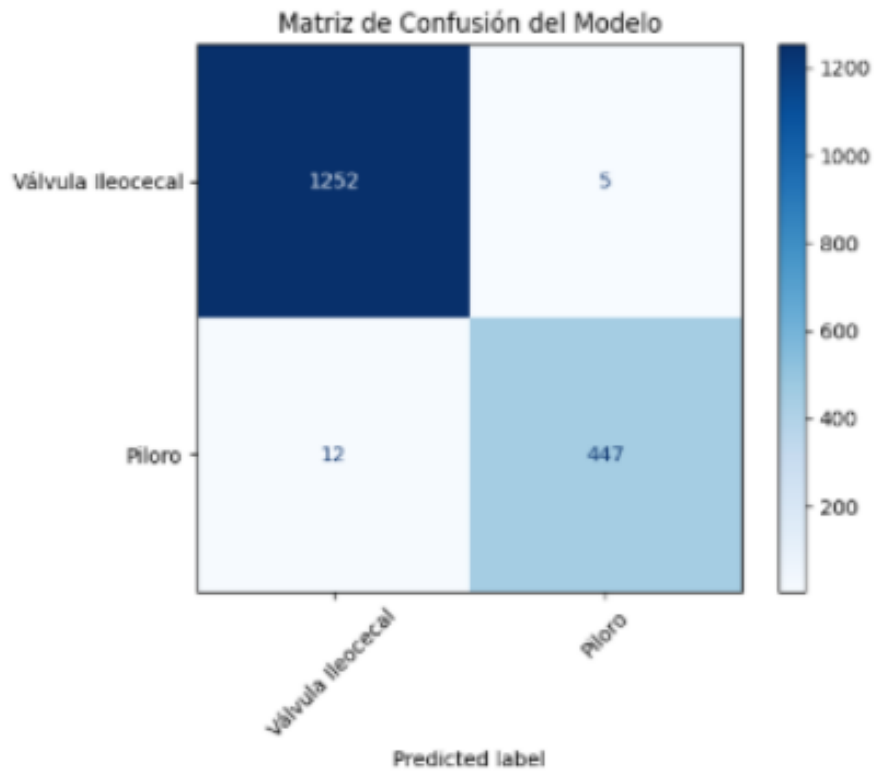


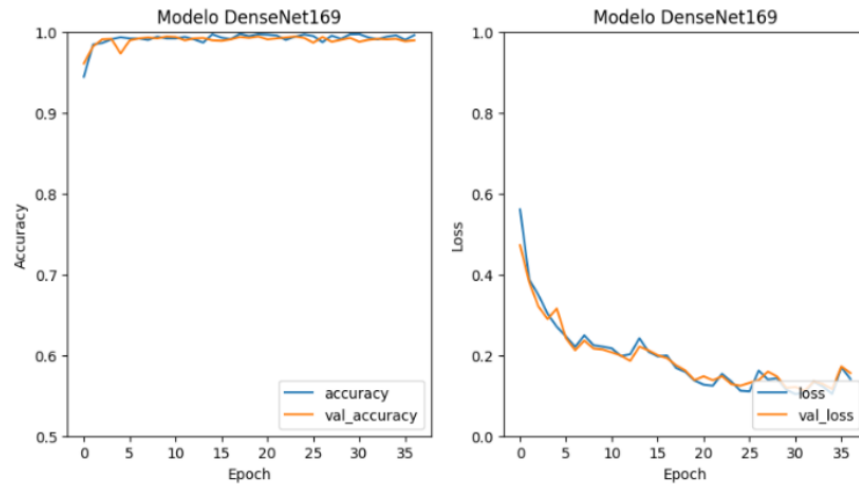
Ciclo 2 (semilla = 123)



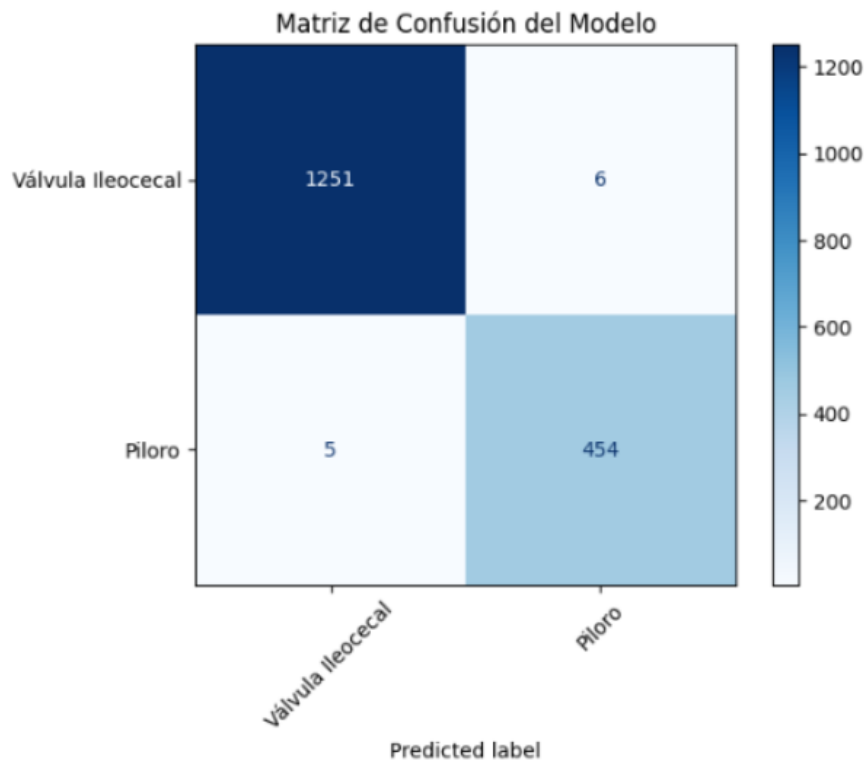


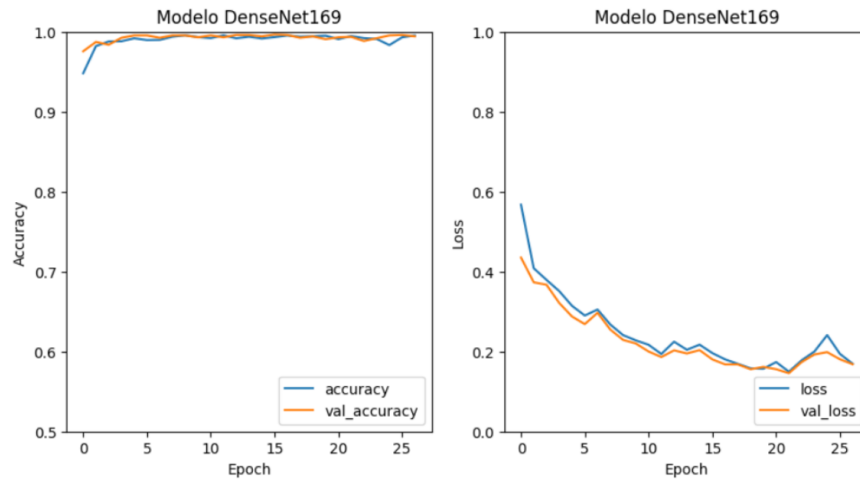
Ciclo 3 (semilla = 2025)



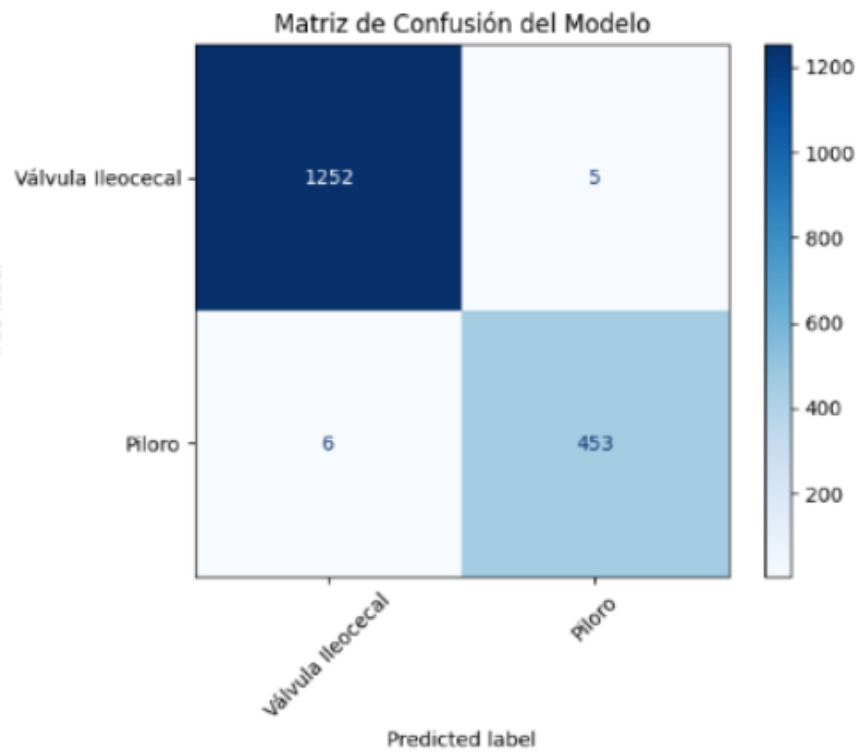


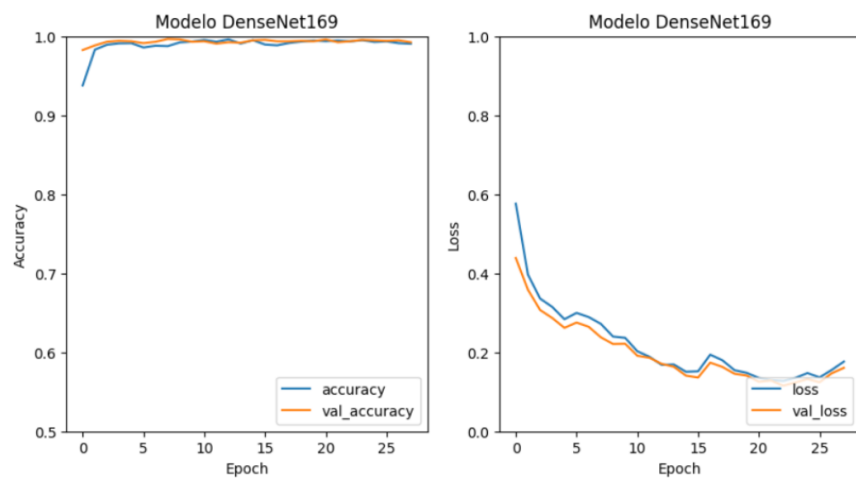
Ciclo 4 (semilla = 1)





Ciclo 5 (semilla = 999)





14.6. DenseNet121 – Clasificación de Patologías

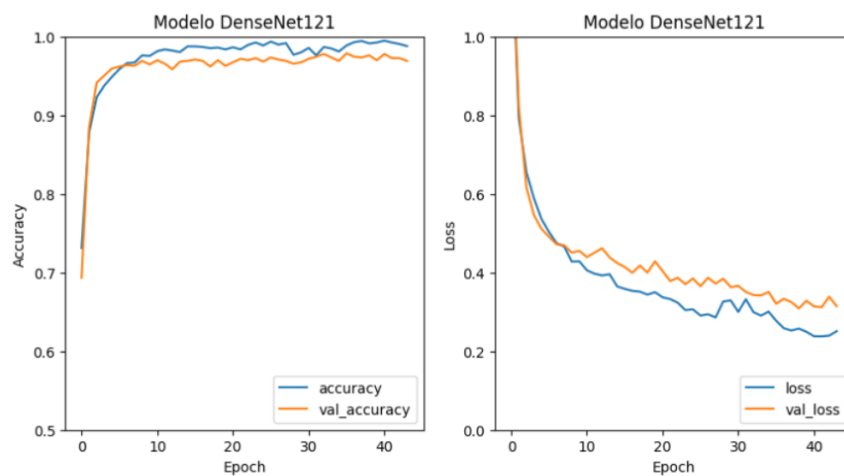
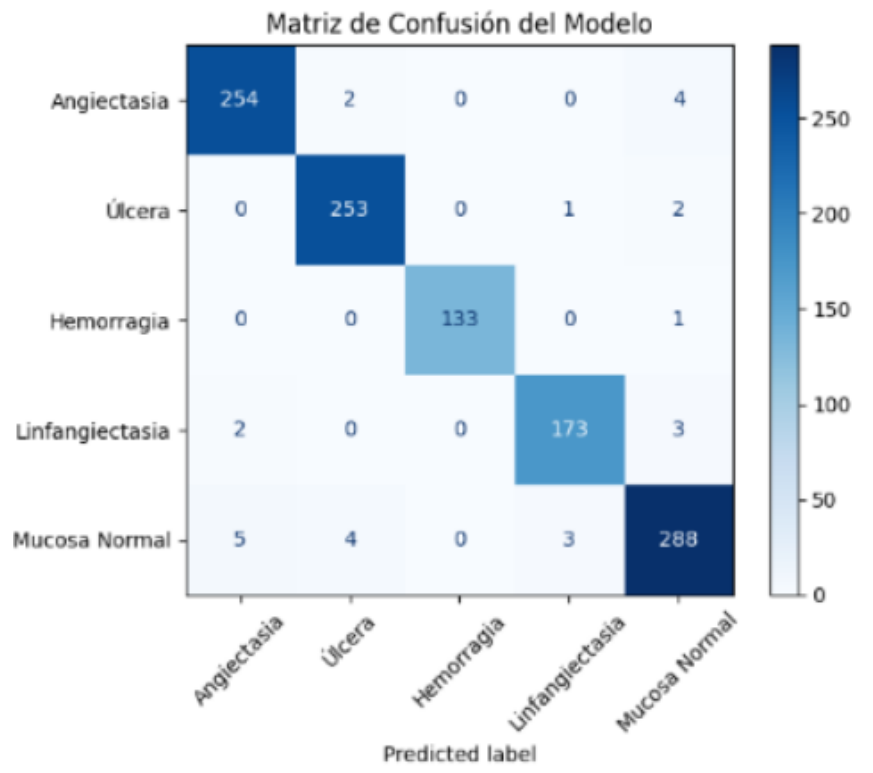
14.6.1. Resumen de métricas promedio (5 ciclos)

Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Angiectasia	0.962 ± 0.010	0.974 ± 0.005	0.968 ± 0.007
Úlcera	0.982 ± 0.013	0.978 ± 0.010	0.978 ± 0.004
Hemorragia	0.992 ± 0.007	0.986 ± 0.014	0.990 ± 0.011
Linfangiectasia	0.964 ± 0.016	0.970 ± 0.006	0.966 ± 0.005
Mucosa Normal	0.960 ± 0.011	0.950 ± 0.011	0.952 ± 0.007
Precisión global (Accuracy)			0.970 ± 0.004

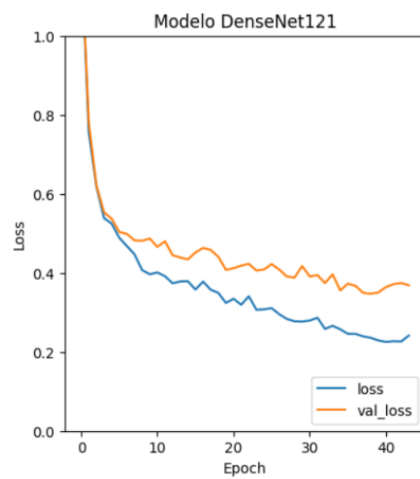
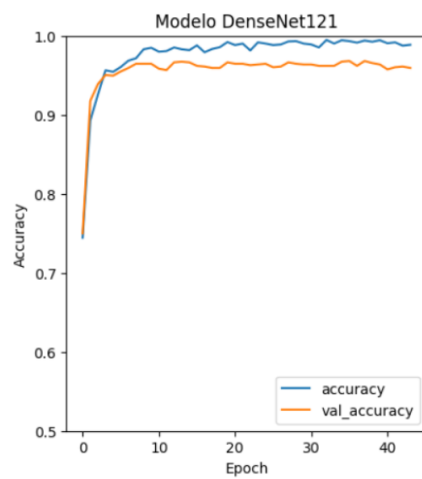
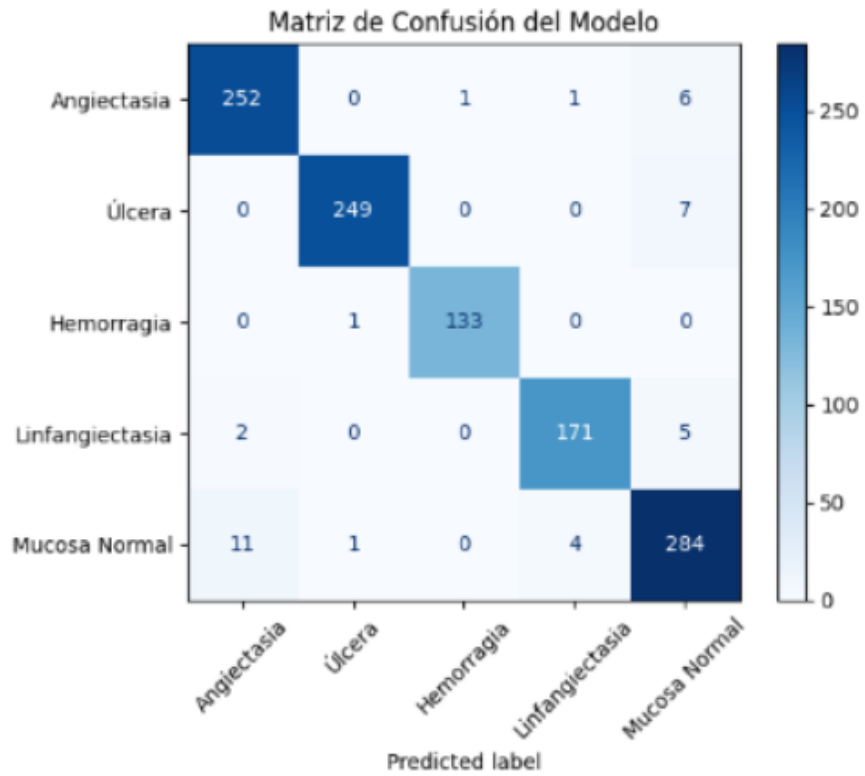
Luego de resumir las métricas promedio alcanzadas por DenseNet121 en la tarea de clasificación de patologías, se presentan los resultados detallados de cada uno de los cinco ciclos de entrenamiento y validación. Para cada ciclo, se muestran la matriz de confusión y las curvas de accuracy y loss, permitiendo examinar el comportamiento individual y la consistencia del modelo frente a diversas inicializaciones.

14.6.2. Resultados por ciclo

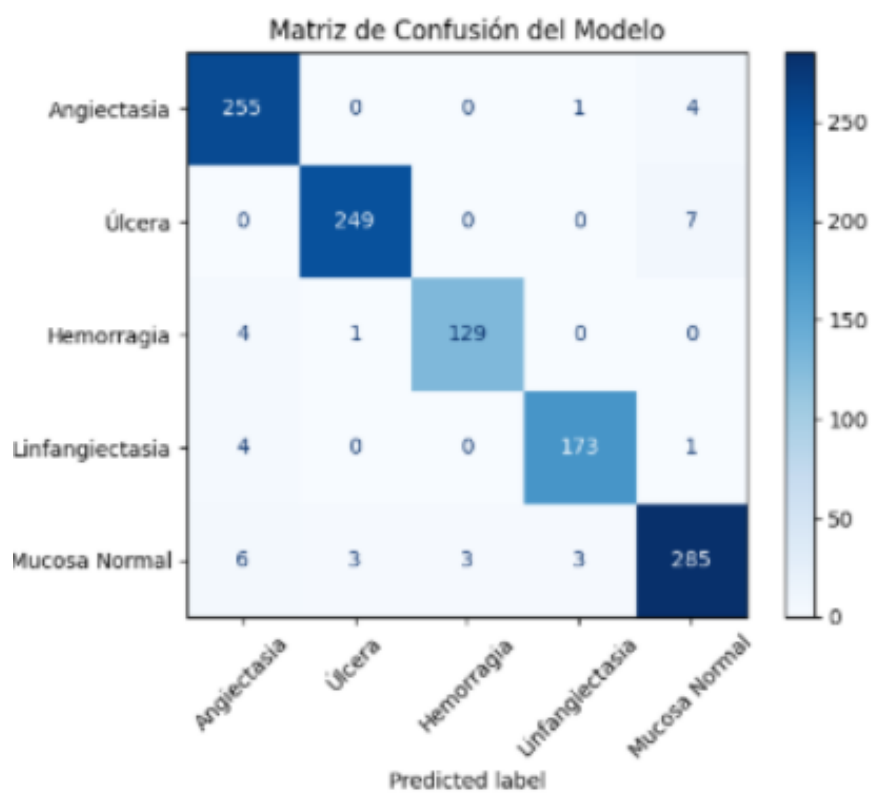
Ciclo 1 (semilla = 42)

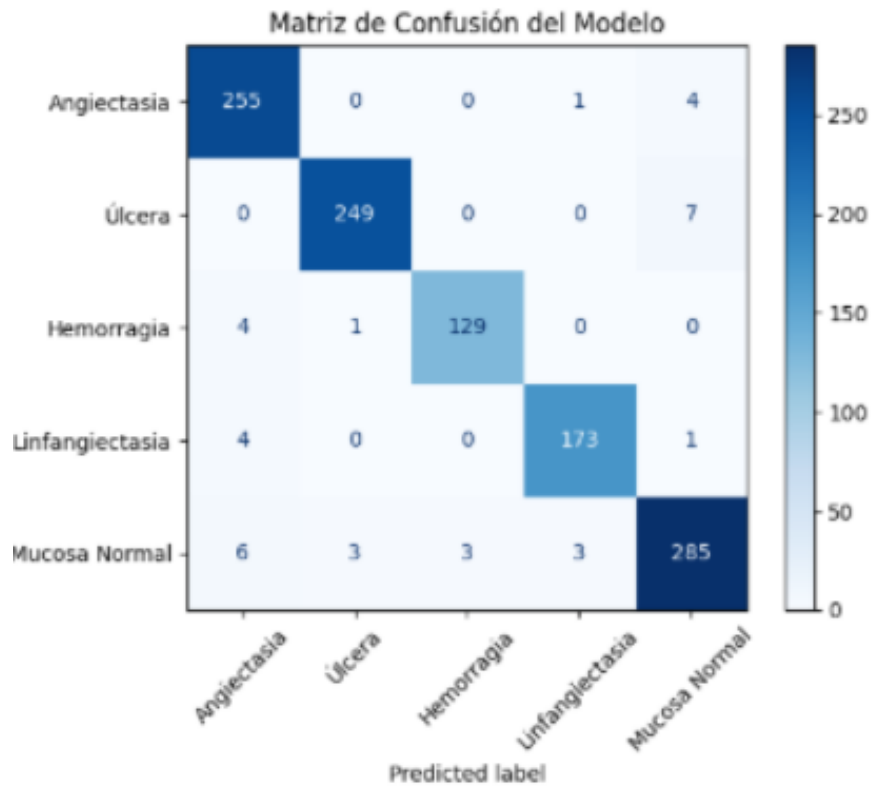


Ciclo 2 (semilla = 123)

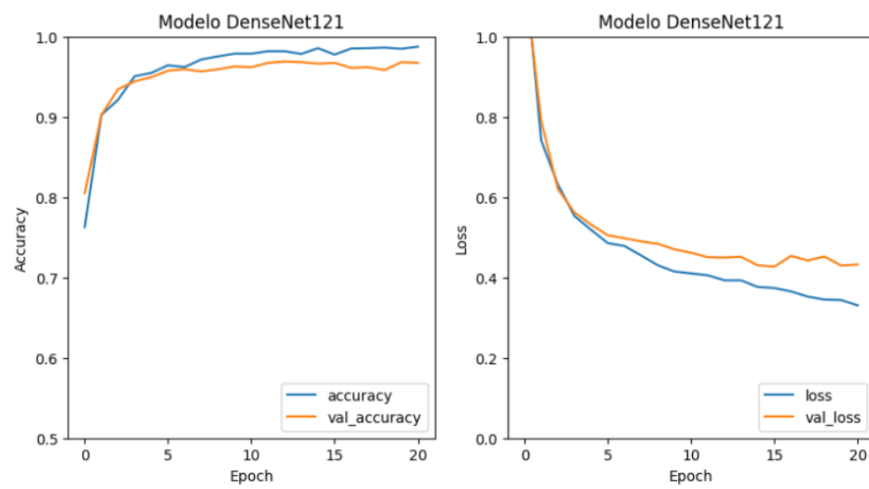


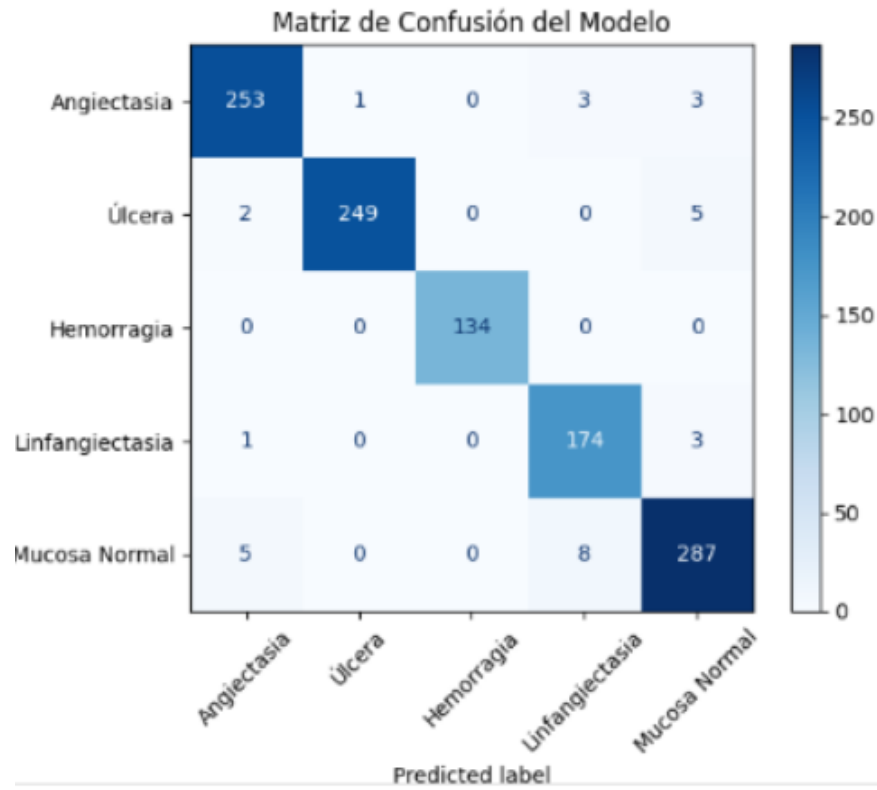
Ciclo 3 (semilla = 2025)



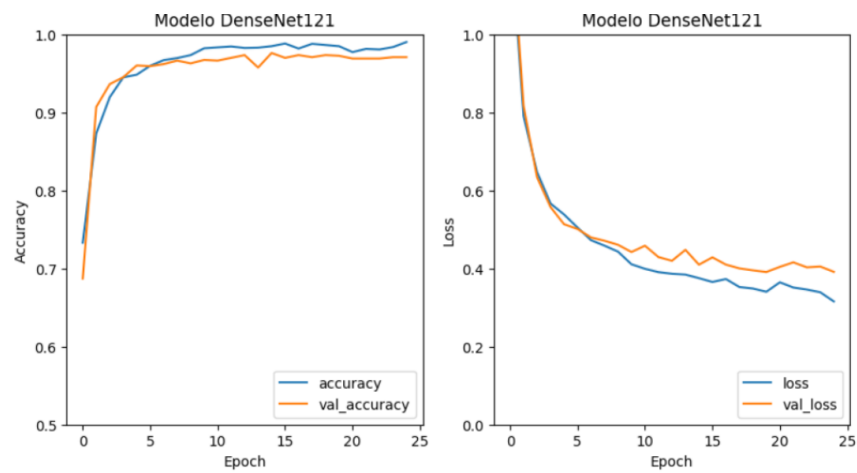


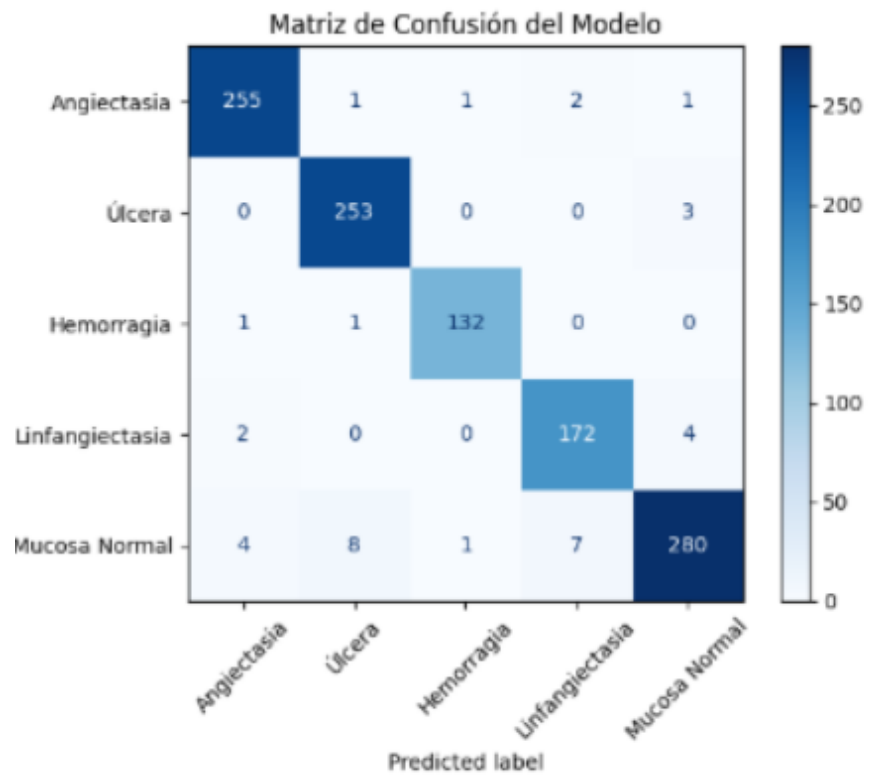
Ciclo 4 (semilla = 1)





Ciclo 5 (semilla = 999)





14.7. DenseNet121 – Detección de Hitos Anatómicos

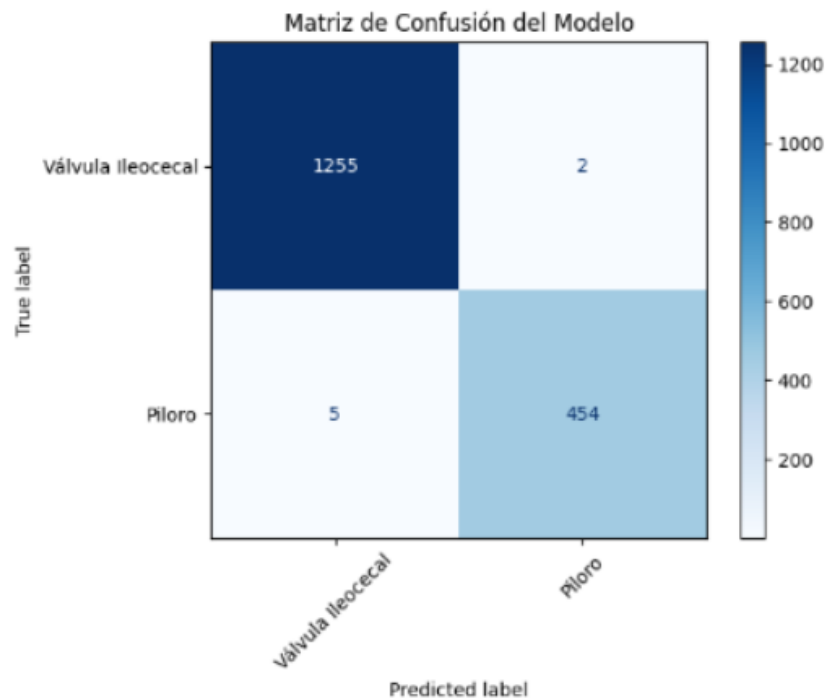
14.7.1. Resumen de métricas promedio (5 ciclos)

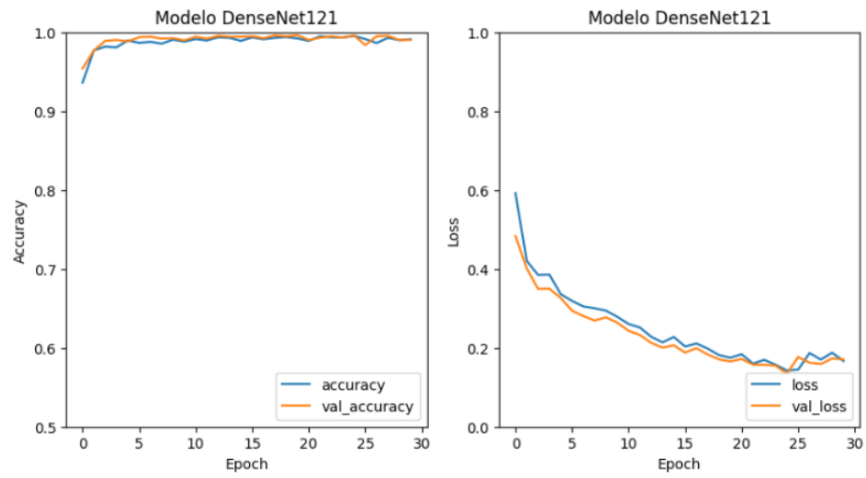
Clase	Precisión ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
Píloro	0.996 ± 0.005	0.982 ± 0.012	0.986 ± 0.005
Válvula Ileocecal	0.996 ± 0.005	1.000 ± 0.000	0.996 ± 0.005
Precisión global (Accuracy)			0.993 ± 0.003

Concluido el resumen de métricas promedio para la detección de hitos anatómicos con DenseNet121, a continuación se detallan los resultados obtenidos en los cinco ciclos de entrenamiento realizados. Cada ciclo incluye la matriz de confusión correspondiente y las curvas de accuracy y loss, lo que permite analizar en profundidad la precisión y estabilidad del modelo en la identificación de los hitos anatómicos.

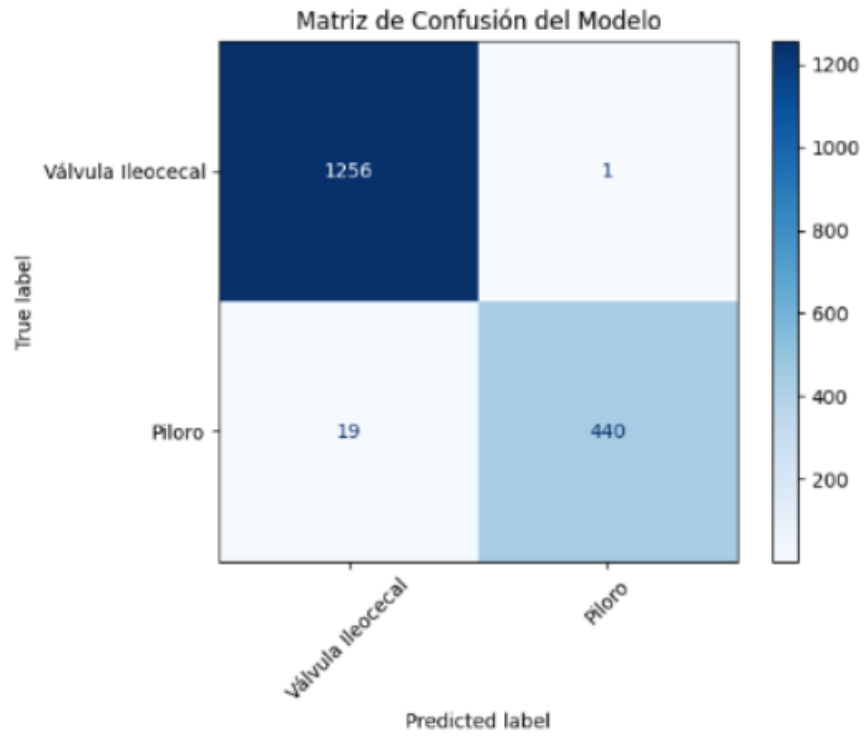
14.7.2. Resultados por ciclo

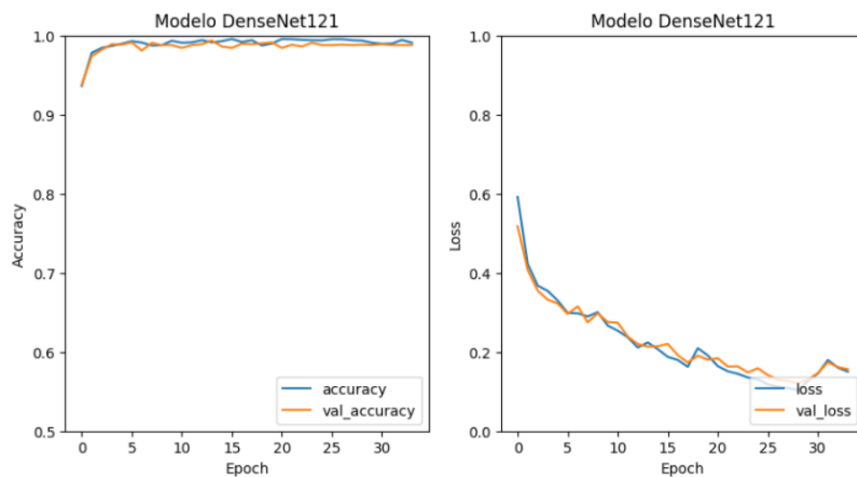
Ciclo 1 (semilla = 42)



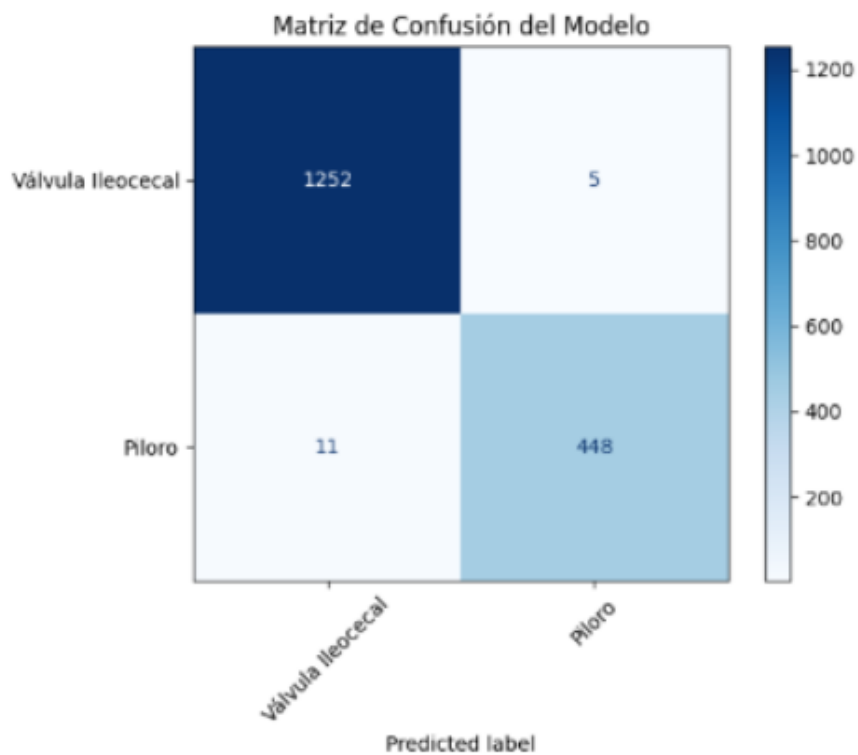


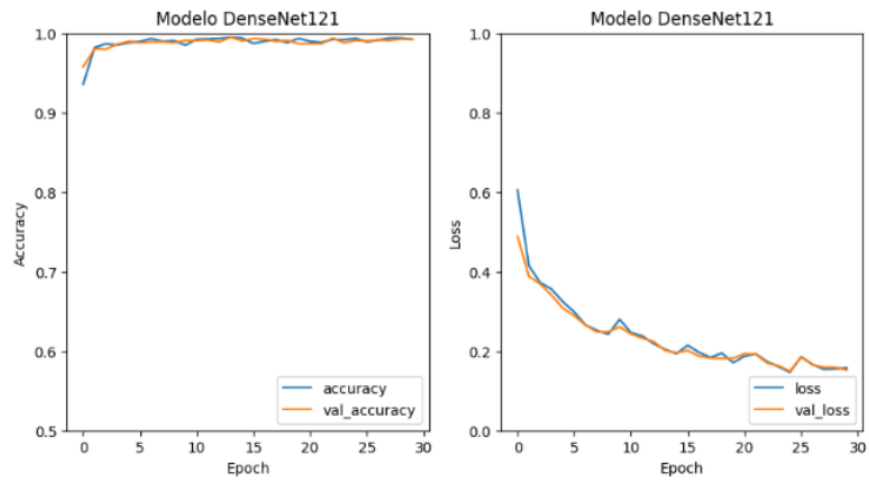
Ciclo 2 (semilla = 123)



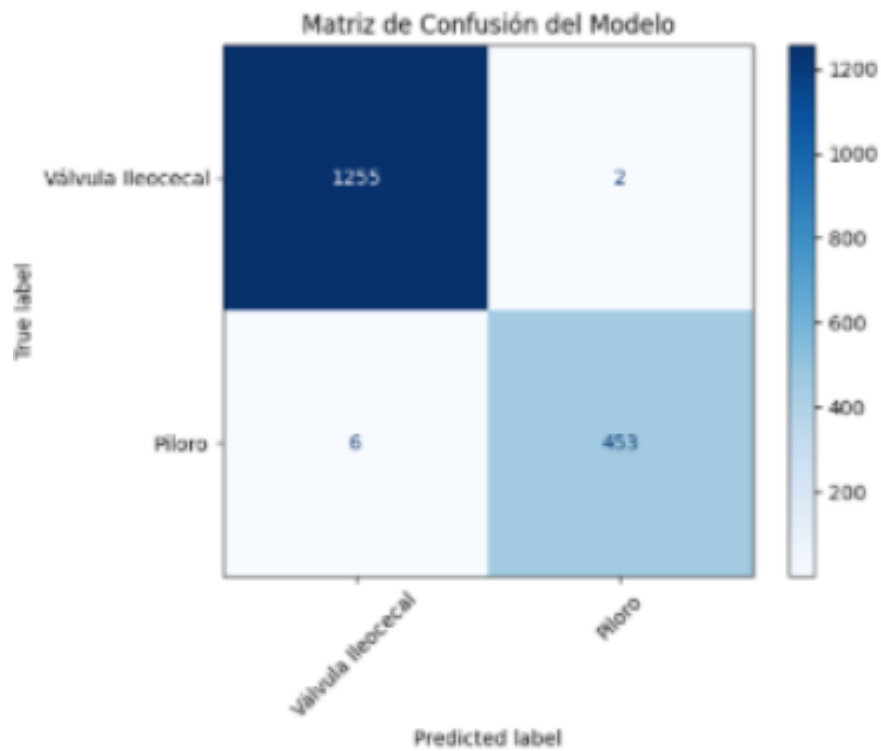


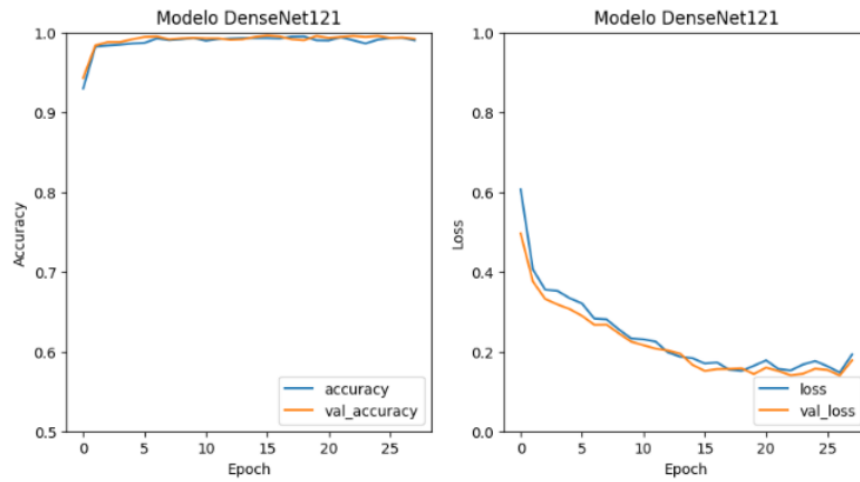
Ciclo 3 (semilla = 2025)



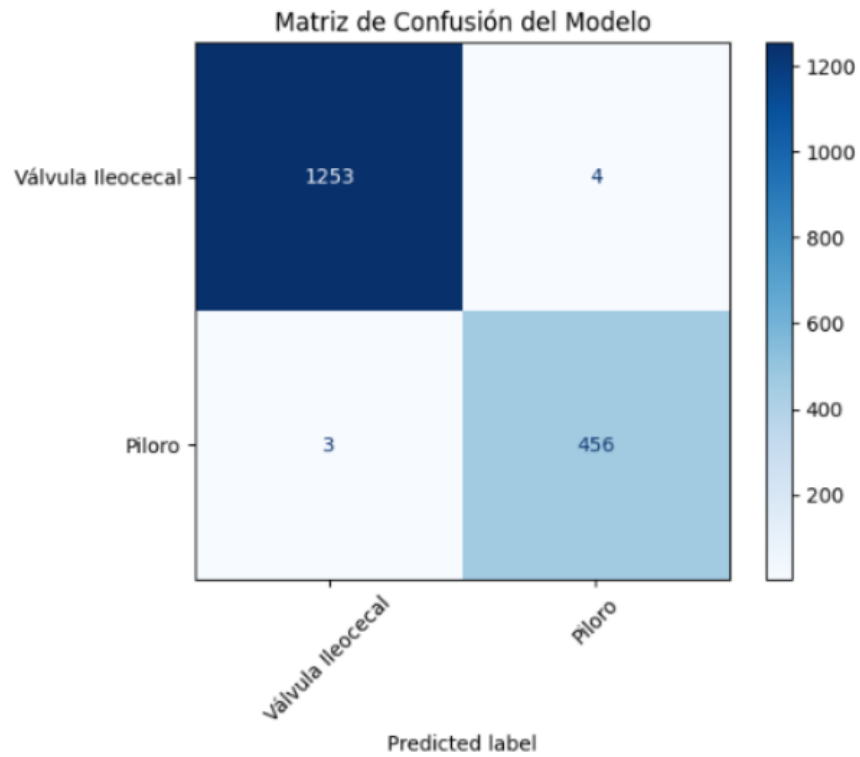


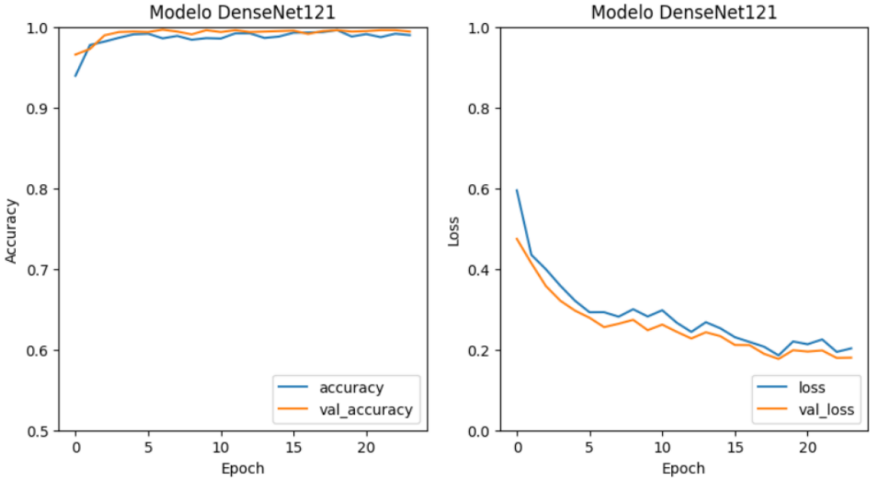
Ciclo 4 (semilla = 1)





Ciclo 5 (semilla = 999)





14.8. Resumen comparativo y discusión final

Concluidos los análisis individuales para cada modelo y tarea, en la siguiente sección se sintetizan y comparan los principales resultados obtenidos, permitiendo una evaluación global del rendimiento de las arquitecturas propuestas.

14.8.1. Resumen Final de Métricas Promedio

14.8.1.1. Patologías

Modelo	Accuracy ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
DENSENET201	0.972 ± 0.004	0.975 ± 0.004	0.974 ± 0.003
DENSENET169	0.970 ± 0.004	0.972 ± 0.003	0.972 ± 0.004
DENSENET121	0.970 ± 0.004	0.972 ± 0.005	0.971 ± 0.004

14.8.1.2. Hitos Anatómicos

Modelo	Accuracy ($\mu \pm \sigma$)	Recall ($\mu \pm \sigma$)	F1-Score ($\mu \pm \sigma$)
DENSENET201	0.992 ± 0.002	0.990 ± 0.003	0.991 ± 0.005
DENSENET169	0.991 ± 0.002	0.990 ± 0.004	0.991 ± 0.005
DENSENET121	0.993 ± 0.003	0.991 ± 0.006	0.991 ± 0.005

Los resultados muestran que las arquitecturas DenseNet evaluadas ofrecen un rendimiento notable en ambas tareas del sistema: la clasificación de patologías y la identificación de hitos anatómicos en imágenes de cápsula endoscópica del intestino delgado. Para la clasificación de patologías, DenseNet201 obtuvo los mejores resultados, alcanzando una precisión promedio (*accuracy*) de 0.972, un *recall* de 0.975 y un F1-score de 0.974, con una baja dispersión entre ciclos, lo que respalda su estabilidad y confiabilidad en la detección de alteraciones clínicas. Por su parte, en la identificación de hitos anatómicos, el modelo con mayor desempeño fue DenseNet121, que logró una precisión promedio de 0.993, un *recall* de 0.991 y un F1-score de 0.991, indicando una capacidad superior para distinguir correctamente entre los diferentes segmentos anatómicos.

En ambos casos, los valores elevados y las bajas desviaciones estándar reflejan un comportamiento robusto y consistente a lo largo de los experimentos. Además, el mejor desempeño observado en la identificación de hitos anatómicos sugiere que la distinción entre estas estructuras es más sencilla para el modelo, posiblemente debido a características visuales más marcadas respecto a las distintas patologías. Finalmente, la selección de DenseNet201 para patologías y DenseNet121 para hitos anatómicos se basó en la comparación directa de las métricas promedio obtenidas, garantizando así la mejor configuración para cada uno de los objetivos del sistema.

. Bibliografía

- [1] V. Thambawita, M. Riegler, S. Hicks, P. Halvorsen, T. de Lange, P. H. Smedsrud, and H. Borgli. (2020, Mar. 13) Kvasir-capsule: A video capsule endoscopy dataset. [Online] Available. [Online]. Available: <https://osf.io/dv2ag/>
- [2] D. Bajhaiya and S. N. Unni, “Deep learning-enabled detection and localization of gastrointestinal diseases using wireless-capsule endoscopic images,” *Biomedical Signal Processing and Control*, vol. 93, p. 106125, 2024.
- [3] S. Colleoni *et al.*, “Wireless capsule endoscopy multiclass classification using three-dimensional deep convolutional neural network model,” *BioMedical Engineering Online*, vol. 22, no. 1, 2023.
- [4] S.-J. Nam, G. Moon, J.-H. Park, Y. Kim, Y. J. Lim, and H.-S. Choi, “Deep learning-based real-time organ localization and transit time estimation in wireless capsule endoscopy,” *Biomedicines*, vol. 12, p. 1704, 2024.
- [5] X. L. Ge *et al.*, “Current status and future perspectives of artificial intelligence in capsule endoscopy,” *Journal of Digestive Diseases*, vol. 24, no. 1, pp. 1–10, Jan. 2023.
- [6] M. A. J. Ali *et al.*, “Improved image classification using deep convolutional neural networks in wireless capsule endoscopy,” *IEEE Access*, vol. 11, pp. 57 694–57 705, Jun. 2023.
- [7] J. X. Zhou, Z. Yang, D. H. Xi *et al.*, “Enhanced segmentation of gastrointestinal polyps from capsule endoscopy images with artifacts using ensemble learning,” *World Journal of Gastroenterology*, vol. 28, no. 41, pp. 5931–5943, 2022.
- [8] N. Obukhova, A. Motyko, A. Pozdeev, and B. Timofeev, “Review of noise reduction methods and estimation of their effectiveness for medical endoscopic images processing,” in *Proceedings of the XXth Conference of Open Innovations Association FRUCT*, 2018, pp. 204–210.
- [9] P. R. Slawinski, K. L. Obstein, and P. Valdastri, “Emerging issues and future developments in capsule endoscopy,” *Techniques in Gastrointestinal Endoscopy*, vol. 17, pp. 40–46, 2015.
- [10] Anon. (2024) Prevalencia poblacional de anomalías gastrointestinales detectadas mediante cápsula endoscópica. [Online]. Available: <https://www.slideshare.net>
- [11] D. Iakovidis and A. Koulaouzidis, “Software for enhanced video capsule endoscopy: challenges for essential progress,” *Nature Reviews Gastroenterology Hepatology*, vol. 12, pp. 172–186, 2015.
- [12] R. M. Gerson and L. B. Fidler, “Capsule endoscopy: Technology update and review of pertinent guidelines,” *Journal of Clinical Gastroenterology*, vol. 52, no. 4, pp. 290–297, 2018.
- [13] E. Litjens, T. Kooi, B. E. Bejnordi *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

- [14] Y. Yuan and M. Q.-H. Meng, "Deep learning for polyp recognition in wireless capsule endoscopy images," *Medical Physics*, vol. 44, pp. 1379–1389, 2017.
- [15] G. Burisch *et al.*, "The global incidence and prevalence of inflammatory bowel diseases in the 21st century: A systematic review," *The Lancet Gastroenterology Hepatology*, vol. 5, no. 1, pp. 17–30, 2020.
- [16] D. S. Weinberg *et al.*, "The gastrointestinal tract: Disorders and diagnostic innovations," *Clinical Medicine Insights: Gastroenterology*, vol. 14, pp. 1–15, 2020.
- [17] S. E. Wilson, "Small intestine: Anatomy and physiology," in *Surgical Anatomy of the Gastrointestinal Tract*, 2019, vol. 3, pp. 57–78.
- [18] P. A. Rubin, "Gastrointestinal mucosa: Structure and function in health and disease," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 325, no. 4, pp. 645–658, 2023.
- [19] S. S. Kim *et al.*, "Angiodysplasia and arteriovenous malformations of the gastrointestinal tract," *Gastrointestinal Endoscopy Clinics of North America*, vol. 34, no. 2, pp. 255–274, 2024.
- [20] A. C. of Gastroenterology. (2025) Sangrado del intestino delgado: Causas, diagnóstico y tratamiento. Accedido: 8 de febrero de 2025. [Online]. Available: <https://gi.org/patients/recursos-en-espanol/sangrado-del-intestino-delgado-causas-diagnostico-y-tratamiento/>
- [21] H. M. Freeman and M. A. Lee, "Primary and secondary intestinal lymphangiectasia," *Journal of Clinical Gastroenterology*, vol. 58, no. 1, pp. 15–25, 2023.
- [22] J. Y. Kim and T. L. Peterson, "Peptic ulcer disease: Pathophysiology and management," *The Lancet Gastroenterology & Hepatology*, vol. 9, no. 1, pp. 23–35, 2024.
- [23] S. A. de Gastroenterología. (2024) Patologías gastrointestinales comunes.
- [24] M. Clinic. (2024) Pólipos en el colon: diagnóstico y tratamiento.
- [25] W. G. Organization. (2024) Epidemiología de las enfermedades digestivas.
- [26] M. L. Fernandez-Urien *et al.*, "Double-balloon enteroscopy: Indications, diagnostic and therapeutic yield, complications, and safety," *World Journal of Gastroenterology*, vol. 16, no. 16, pp. 2313–2321, 2010.
- [27] A. T. Rex *et al.*, "Colonoscopy and endoscopic imaging of the small intestine," *Gastrointestinal Endoscopy*, vol. 84, no. 1, pp. 25–30, 2016.
- [28] G. Imaging. (2024) Pillcam capsule endoscopy. [Online]. Available: <https://www.medtronic.com/covidien/en-us/solutions/capsule-endoscopy/pillcam.html>
- [29] A. Shalmani *et al.*, "Capsule endoscopy - state of the technology and computer vision tools after the first decade," in *IntechOpen*, 2012. [Online]. Available: <https://ueaeprints.uea.ac.uk/52076/>
- [30] "Utility of capsule endoscopy in the diagnosis of crohn's disease," *Inflammatory Bowel Diseases Journal*, 2023. [Online]. Available: <https://academic.oup.com/ibdjournal/article/23/9/E43/4560696>

- [31] J. P. Ramirez *et al.*, “Clinical utility of the small bowel capsule endoscopy in the study of small bowel diseases in colombia: Results of a national registry,” *Gastroenterología y Hepatología*, vol. 47, no. 10, pp. 561–568, 2023.
- [32] J. G. Webster, *Medical Instrumentation: Application and Design*, 4th ed. New York, NY: John Wiley Sons, 2009.
- [33] H. Wang, J. Qian, and W. Yang, “Applications of deep learning in medical imaging: A review,” *Diagnostics*, vol. 11, no. 1, pp. 1–19, 2021.
- [34] A. K. Jain, P. Gupta, and A. K. Singh, “Deep learning in medical image analysis: A survey,” *Journal of Medical Systems*, vol. 44, no. 4, pp. 1–13, 2020.
- [35] (2024) Cápsula endoscópica: situación tras 15 años de su lanzamiento. [Online]. Available: <https://www.acad.es/capsula-endoscopica-situacion-tras-15-anos-de-su-lanzamiento/>
- [36] T. Aoki *et al.*, “Automatic detection of various abnormalities in capsule endoscopy videos by a deep learning-based system: a multicenter study,” *Gastrointestinal Endoscopy*, vol. 93, no. 1, 2020.
- [37] D. K. G. Dhungel, “Deep learning for medical image analysis: A review,” *Journal of Healthcare Engineering*, vol. 2018, pp. 1–9, 2018.
- [38] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [39] A. Krizhevsky, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, vol. 25, 2012.
- [40] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *International Conference on Learning Representations (ICLR)*, 2015.
- [41] M. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *European Conference on Computer Vision (ECCV)*, 2014.
- [42] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 3431–3440.
- [43] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, “Transfusion: Understanding transfer learning for medical imaging,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 3342–3352.
- [44] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.
- [45] S. Han, H. Mao, and W. J. Dally, “Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding,” in *International Conference on Learning Representations (ICLR)*, 2016. [Online]. Available: <https://arxiv.org/abs/1510.00149>

- [46] Y. Sun, C. K. Wen, S. Jin, and K. Wong, "Deep learning-based channel estimation for massive mimo with mixed-resolution adcs," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 2122–2135, 2019.
- [47] L. Guo, L. Wen, C. Li, and S. Mao, "A novel deep learning framework for industrial fault diagnosis based on densenet and transfer learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021.
- [48] J. Irvin *et al.*, "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, 2019, pp. 590–597.
- [49] R. D. Treder, T. Laueremann, and G. B. E. Schmidt, "Deep learning-based detection and classification of diabetic retinopathy using retinal fundus images," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 256, pp. 1161–1167, 2018.
- [50] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.
- [51] Z. Zhang, Q. Liu, Y. Wang, and H. Jin, "A comprehensive survey on densenet-based models in computer vision," *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–36, 2022.
- [52] Y. Chen, X. Wang, and L. Zhang, "Densenet for medical image analysis: A review," *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1235–1245, 2020.
- [53] S. R. Lewis, "The role of capsule endoscopy in evaluating small bowel disease," *Journal of Clinical Gastroenterology*, vol. 48, no. 1, pp. 4–6, 2014.
- [54] B. Jiang *et al.*, "Development and validation of a deep learning system for detection of small bowel pathologies in capsule endoscopy: a pilot study in a singapore institution," *Singapore Medical Journal*, vol. 65, no. 3, pp. 133–140, 2024.
- [55] T. Aoki *et al.*, "Comparison of clinical utility of deep learning-based systems for small bowel capsule endoscopy reading," *Journal of Gastroenterology and Hepatology*, 2023.
- [56] A. Sahafi *et al.*, "Edge artificial intelligence wireless video capsule endoscopy," *Dental Science Reports*, vol. 12, no. 1, 2022.
- [57] P. H. Smedsrud *et al.*, "Kvasir-capsule, a video capsule endoscopy dataset," *Scientific Data*, vol. 8, May 2021.
- [58] R. Leenhardt, P. Vasseur, C. Li *et al.*, "A neural network algorithm for detection of gi angiectasia during small-bowel capsule endoscopy," *Gastrointestinal Endoscopy*, vol. 89, no. 1, pp. 189–194, 2019.
- [59] A. A. Akyüz, T. Riegler, P. Halvorsen, and H. G. de Waal, "Automated detection of small bowel ulcers in wireless capsule endoscopy images," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 1, pp. 122–130, 2019.
- [60] A. H. Khan, M. H. A. Sohag, S. S. Vedaiei, M. R. Mohebbian, and K. A. Wahid, "Automatic detection of intestinal bleeding using an optical sensor for wireless capsule endoscopy,"

- in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2020, pp. 4345–4348.
- [61] K. K. Ota, T. D. Syeda-Mahmood, and R. S. Sufian, “Automated localization of pyloric transition in video capsule endoscopy sequences,” in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, Sep. 2018, pp. 2345–2349.
- [62] J. A. Salamanca-Cardona, J. G. Leiva, and R. E. Torres-Delgado, “Deep learning-based detection of ileocecal valve in capsule endoscopy images,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4567–4575, Dec. 2020.