



Modelo de Machine Learning para la Identificación de Pólipos en Imágenes de Colonoscopias

Jairo Enrique Alba Talero
Código: 8980176

Proyecto Aplicado para optar al título de Magíster en Ciencia de Datos

Director: Dr. Diego Luis Linares Ospina

Codirector: Dr. Gloria Inés Álvarez

FACULTAD DE INGENIERÍA Y CIENCIAS
MAESTRÍA EN CIENCIA DE DATOS
SANTIAGO DE CALÍ, NOVIEMBRE 03 DE 2025

Agradecimientos

Quiero dar un gran agradecimiento a mi pareja por su apoyo a lo largo de este proceso de mi formación académica.

También, extendiendo un profundo agradecimiento a mis tutores de trabajo de grado, los docentes Diego Luis Linares Ospina y Gloria Inés Álvarez, por su acompañamiento, dedicación y compromiso constantes, fundamentales para llevar a buen término cada fase de este proyecto.

Resumen

Los pólipos en el colon son acumulaciones celulares que, aunque en muchos casos son inofensivas, pueden evolucionar hacia lesiones precancerosas o cancerosas, lo que hace fundamental su detección temprana mediante colonoscopia. Sin embargo, un porcentaje significativo de pólipos puede pasar desapercibido durante el procedimiento, lo que motiva el desarrollo de herramientas de apoyo basadas en inteligencia artificial. En este trabajo se desarrolló un sistema de aprendizaje profundo para asistir la interpretación de imágenes de colonoscopia.

Se procesaron imágenes obtenidas en procedimientos médicos, que fueron seleccionadas de acuerdo con las etiquetas brindadas por expertos, garantizando así la calidad del conjunto de datos. Posteriormente, se implementaron diferentes arquitecturas de redes neuronales convolucionales (CNNs) y estrategias de transferencia de aprendizaje, con el fin de evaluar su desempeño en la clasificación automática de pólipos.

Los experimentos incluyeron una CNN entrenada desde cero, su versión optimizada y dos variantes de transferencia de aprendizaje basadas en VGG16. Los resultados mostraron que la CNN inicial alcanzó una exactitud del 94 %, mejorando al 96 % con la optimización. Los modelos de transferencia superaron estos valores: el modelo con capas densas entrenables obtuvo un 95 % de exactitud, mientras que la variante con ajuste fino en el bloque cinco alcanzó el mejor desempeño global, logrando un 97 % de exactitud. Además se evaluó cada modelo con otras métricas equilibradas como: precisión, recall (sensibilidad) y F1-score.

Finalmente, se desarrolló una aplicación web con Streamlit que permite cargar imágenes, preprocesarlas automáticamente y obtener la predicción del modelo en tiempo real. Este prototipo asegura accesibilidad, reproducibilidad y potencial escalabilidad hacia entornos hospitalarios, consolidándose como una herramienta de apoyo clínico en la detección temprana de pólipos y la prevención del cáncer colorrectal.

Palabras clave: Pólipo, colonoscopia, aprendizaje profundo, imágenes, redes neuronales convolucionales (CNNs), clasificación.

Abstract

Colonic polyps are cellular accumulations that, although often harmless, can evolve into pre-cancerous or cancerous lesions, making their early detection through colonoscopy crucial. However, a significant percentage of polyps may go unnoticed during the procedure, which motivates the development of artificial intelligence-based support tools. In this work, a deep learning system was developed to assist in the interpretation of colonoscopy images.

Images obtained from medical procedures were processed and selected according to labels provided by experts, thus ensuring the quality of the dataset. Subsequently, different convolutional neural network (CNN) architectures and transfer learning strategies were implemented to evaluate their performance in the automatic classification of polyps.

The experiments included a CNN trained from scratch, its optimized version, and two transfer learning variants based on VGG16. The results showed that the initial CNN achieved an accuracy of 94%, improving to 96% with optimization. The transfer learning models surpassed these values: the model with trainable dense layers obtained 95% accuracy, while the variant with fine-tuning in block five achieved the best overall performance, reaching 97% accuracy. Additionally, each model was evaluated with other balanced metrics such as precision, recall (sensitivity), and F1-score.

Finally, a web application was developed with Streamlit that allows users to upload images, pre-process them automatically, and obtain the model's prediction in real-time. This prototype ensures accessibility, reproducibility, and potential scalability to hospital environments, consolidating itself as a clinical support tool for the early detection of polyps and the prevention of colorectal cancer.

Keywords: Polyp, colonoscopy, deep learning, images, convolutional neural networks (CNNs), classification.

Índice general

1. Introducción	1
2. Descripción del Problema	2
2.1. Planteamiento del Problema	2
2.2. Formulación del Problema	3
2.3. Objetivo General	3
2.4. Objetivos Específicos	3
3. Marco Teórico y Antecedentes de la Literatura	5
3.1. Pólipos Colorrectales: Bases Clínicas y Clasificación	5
3.1.1. Pólipos Colorrectales	5
3.1.2. Clasificación Histológica	5
3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales	8
3.2.1. Aprendizaje Profundo	8
3.2.2. Redes Neuronales Convolucionales	8
3.2.3. Aprendizaje por Transferencia	11
3.2.4. Métricas de Desempeño de Modelos en Aprendizaje Automático	12
3.3. Antecedentes	13
4. Preparación de Datos	16
4.1. Obtención de los datos	16
4.2. Descripción del conjunto de datos PolypGen 2021	16
4.3. Descripción del conjunto de datos KVASIR	17
4.4. Preprocesamiento de las imágenes	18
5. Desarrollo de Modelos	21
5.1. Modelo Red Neuronal Convolutiva	22
5.1.1. Modelo Optimizando parámetros	24
5.2. Modelo Transferencia de Aprendizaje Capas Densas Entrenables	24
5.3. Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables	26
6. Análisis de Resultados	29
7. Despliegue del Prototipo	37
7.1. Interfaz para interactuar con el modelo.	37
7.2. Descripción del desarrollo de la interfaz	38

Índice general	5
-----------------------	----------

8. Conclusiones y Trabajos Futuros	41
8.1. Conclusiones	41
8.2. Trabajos futuros	42
Bibliografía	43

Índice de figuras

3.1. Red neuronal convolucional para clasificación de imágenes.	9
4.1. Presencia de pólipo.	17
4.2. Sin presencia de pólipos.	17
4.3. Imagen original KVASIR.. . . .	19
4.4. Imagen modificada KVASIR.. . . .	19
4.5. Imagen original PolyGen 2021.	20
4.6. Imagen modificada PolyGen 2021.	20
5.1. Arquitectura de Red Neuronal Convolucional.	23
5.2. Arquitectura Modelo Transferencia de Aprendizaje Capas Densas Entrenables.	26
5.3. Arquitectura Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables	27
6.1. Perdida de entrenamiento y validación, Precisión de entrenamiento y validación modelo red neuronal convolucional	32
6.2. Perdida de entrenamiento y validación, Precisión de entrenamiento modelo transferencia de aprendizaje capas densas	33
6.3. Perdida de entrenamiento y validación, Precisión de entrenamiento modelo transferencia de aprendizaje bloque cinco y capas densas entrenables	35
7.1. Presentación interfaz.	39
7.2. Clasificación de imagen presencia de pólipo.	40
7.3. Clasificación de imagen sin presencia de pólipo.	40

Índice de cuadros

3.1. Matriz de confusión para la clasificación de pólipos.	12
4.1. Distribución de imágenes combinadas por conjunto, clase y fuente de origen	20
5.1. Hiperparámetros explorados en el modelo convolucional	21
5.2. Configuración de hiperparámetros del experimento 25	21
6.1. Comparación del desempeño de los modelos en imágenes con presencia de pólipos	29
6.2. Comparación del desempeño de los modelos en imágenes sin presencia de pólipos	30
6.3. Exactitud global y promedio macro obtenidos por los diferentes modelos	30
6.4. Matriz de confusión del modelo Red Neuronal Convolucional.	31
6.5. Matriz de confusión del modelo transferencia de aprendizaje capas densas entrenables	33
6.6. Matriz de confusión del modelo con transferencia de aprendizaje y capas densas entrenables	34

Introducción

La colonoscopia es uno de los métodos más eficaces para la detección temprana de pólipos colorrectales[1]. Sin embargo, las limitaciones inherentes al procedimiento y a la experiencia de los especialistas pueden llevar a la omisión de lesiones importantes, comprometiendo la eficacia del diagnóstico de lesiones precancerosas y cancerosas en esa porción del tracto digestivo que a veces se pasan por alto. En este contexto, el uso de tecnologías de inteligencia artificial, específicamente el aprendizaje profundo, ha mostrado un gran potencial para asistir en la identificación y clasificación de pólipos en imágenes de colonoscopias [2].

Este proyecto tuvo como objetivo desarrollar un prototipo funcional de un sistema automatizado basado en machine learning, enfocado en la detección de pólipos en imágenes de colonoscopias. Para lograrlo, se abordaron los siguientes objetivos específicos: primero, se realizó un procesamiento previo de las imágenes con el fin de optimizar su calidad para el análisis; luego, se entrenaron modelos de aprendizaje profundo que permitieron clasificar dichas imágenes con alta precisión; finalmente, se evaluó el desempeño de los modelos desarrollados y se construyó una interfaz interactiva para facilitar la interacción de los usuarios con el sistema.

Este trabajo proporciona, modelos de clasificación de alto rendimiento ajustados mediante técnicas de optimización de hiperparámetros y una interfaz web que permite a los usuarios acceder a las predicciones del sistema. Se generó este documento, en el que se describe detalladamente, el proceso de obtención y procesamiento de las imágenes de colonoscopias, así como el proceso de entrenamiento de tres modelos de aprendizaje profundo: un modelo base basado en redes neuronales convolucionales en el cual luego fueron optimizados los hiperparámetros, y los otros dos basados en transferencia de aprendizaje; modelo VGG16 con las últimas tres capas densas entrenables y otro dejando el quinto bloque y las capas densas entrenables para la clasificación de imágenes de colonoscopias en dos clases, presencia de pólipos o tejidos sin la presencia de pólipos, posteriormente se evaluó el desempeño de los modelos desarrollados y finalmente se desarrolló una interfaz para interactuar con el modelo.

El desarrollo de estos modelos no solo buscan mejorar la precisión diagnóstica, sino también servir como un apoyo significativo en el campo de la gastroenterología, contribuyendo así a la prevención efectiva del cáncer colorrectal.

Descripción del Problema

2.1. Planteamiento del Problema

Los pólipos en el colon son protuberancias que ocurren en la luz del colon que generalmente son esporádicos o parte de otros síndromes [3]. Estos pólipos se clasifican como diminutos si miden ≤ 5 mm de diámetro, pequeños si miden entre 6 y 9 mm de diámetro, y grandes si miden ≥ 1 cm de diámetro. Adicionalmente, se pueden clasificar en pólipos neoplásicos (adenomatosos) y no neoplásicos (hiperplásicos, inflamatorios, hamartomatosos)[4].

En este sentido, la colonoscopia desempeña un papel crucial en la detección y eliminación de precursores neoplásicos colorrectales, pues en el caso de una enfermedad o síndrome gastroenterológico, es de importancia detectar un tumor en estadios tempranos o, más aún, detectar el pólipo en estado adenomatoso (premaligno o con alto riesgo de cáncer), para prevenir complicaciones[5]. Sin embargo, a pesar de las mejoras significativas producidas en los endoscopios, hasta un 21 % de los pólipos no son visualizados durante la colonoscopia[6] y por ello existen limitaciones y variaciones en el desempeño de los endoscopistas, lo que lleva a lesiones pasadas por alto y resultados subóptimos [7].

La tasa de adenomas que pasan desapercibidos en una colonoscopia puede estar relacionada con factores propios del paciente dentro de los cuales se encuentra la preparación inadecuada o algunas condiciones anatómicas [8]; de otra parte están los factores propios de las lesiones, que en ocasiones son sutiles, como las serradas, planas o deprimidas, otro factor es la localización compleja de los pólipos, difíciles de identificar en la colonoscopia de vista frontal[9]; por último, están los factores relacionados con la experticia del profesional, la técnica utilizada y la fatiga asociada con el sobreagendamiento de procedimientos. Todo ello compromete la calidad de la colonoscopia[10].

Los métodos como el aprendizaje profundo particularmente en la colonoscopia, prometen mejorar la tasa de detección de adenomas y caracterizar mejor las lesiones, pues permiten mantener la atención en el procedimiento, así como mejorar la semiología de todo hallazgo en tiempo real, y podrían ser de gran valor para el trabajo del personal de salud [11].

Teniendo en cuenta lo anterior, el diagnóstico asistido por computador (DAC) constituye una herramienta con gran potencial para ayudar a los endoscopistas en las tareas de detección y clasificación histológica de los pólipos colorrectales [2]. En los últimos años se han descrito diferentes tecnologías y ha aumentado la evidencia sobre su potencial utilidad, lo que ha generado grandes

expectativas en las sociedades científicas [6].

Para poder desarrollar sistemas de DAC eficientes, el contenido de los vídeos e imágenes de colonoscopia debe ser analizado minuciosamente por expertos en procesamiento de imágenes, que definirán modelos que permitan buscar aquellas lesiones o indicadores definidos por los médicos [7]. Por lo anterior, la aplicación de aprendizaje profundo en colonoscopias se considera una alternativa de solución a la problemática anteriormente planteada. Estos métodos implican la detección y localización de pólipos y deciden si hay presencia o ausencia de un pólipo en la imagen, y habitualmente están diseñados para alertar a los endoscopistas con un sonido o marcador cuando se sospecha la presencia de un pólipo durante la realización de una colonoscopia [12].

2.2. Formulación del Problema

La pregunta central de esta investigación fue: **¿Cómo identificar la presencia de pólipos en una imagen de colonoscopia a través de un modelo de Aprendizaje Automático (machine learning)?**

Para lograrlo, se plantearon las siguientes preguntas de sistematización:

1. ¿Cómo procesar las imágenes de colonoscopias para utilizarlas de entrada a modelos de clasificación?
2. ¿Cómo realizar un entrenamiento de modelos de Machine Learning para identificar la presencia de pólipos en una imagen de colonoscopia?
3. ¿Cómo se puede medir el desempeño de los modelos de Machine Learning?
4. ¿Cómo los usuarios pueden utilizar los modelos?

Para responder a las preguntas anteriores se ejecutaron los siguientes objetivos.

2.3. Objetivo General

Desarrollar un prototipo funcional de un sistema para la identificación de pólipos en imágenes de colonoscopias.

2.4. Objetivos Específicos

- Realizar un procesamiento previo de las imágenes de colonoscopias.
- Entrenar modelos de aprendizaje profundo para la clasificación de imágenes de colonoscopias.

- Evaluar el desempeño de los modelos desarrollados de identificación de pólipos en imágenes de colonoscopias.
- Desarrollar una interfaz para interactuar con el modelo.

Marco Teórico y Antecedentes de la Literatura

El cáncer colorrectal (CCR) representa una de las neoplasias malignas de mayor incidencia y mortalidad a nivel mundial (Organización Mundial de la Salud). Su desarrollo, sin embargo, sigue un curso predecible en la mayoría de los casos, originándose a partir de lesiones precursoras benignas conocidas como pólipos. Esta progresión, denominada secuencia adenoma-carcinoma, ofrece una ventana de oportunidad única para la prevención y la detección temprana.

La comprensión de las bases clínicas de los pólipos, la fisiopatología del CCR, su sistema de estadificación y los métodos diagnósticos disponibles es fundamental para el desarrollo de nuevas estrategias terapéuticas y de cribado. Este capítulo establece el marco conceptual sobre el cual se sustenta la presente investigación.

3.1. Pólipos Colorrectales: Bases Clínicas y Clasificación

3.1.1. Pólipos Colorrectales

Un pólipo colorrectal es una proliferación anómala de células en la mucosa del colon o recto que se proyecta hacia la luz intestinal. Su importancia clínica radica en que muchos de ellos, constituyen lesiones precursoras del cáncer colorrectal (CCR), uno de los cánceres más frecuentes a nivel mundial. Por esta razón, su detección y resección son pilares en la prevención del CCR.[3].

Estos crecimientos anormales se desarrollan en el revestimiento del colon o el recto. Pueden variar en forma, tamaño y características histológicas.

3.1.2. Clasificación Histológica

Los pólipos se clasifican según sus características histológicas, lo que determina su potencial de malignidad. La principal división se establece entre pólipos neoplásicos y no neoplásicos.

1. **Pólipos Neoplásicos (Adenomatosos):** Son los que poseen un verdadero potencial maligno. Constituyen el precursor directo del adenocarcinoma colorrectal. Se subclasifican en:
 - **Adenoma Tubular:** El tipo más común (aproximadamente 80-85%). Suelen ser pequeños y pediculados. Su potencial de malignización es menor en comparación con otros tipos [13].

- **Adenoma Velloso:** Representan el 5-10 % de los adenomas. Tienden a ser más grandes, sésiles (planos) y presentan un mayor riesgo de displasia de alto grado y transformación a carcinoma [4].
- **Adenoma Tubulovelloso:** Contienen una mezcla de características tubulares y vellosas (20-80 % de componente velloso). Su riesgo de malignización es intermedio [4].
- **Pólipos Serrados:** Un grupo heterogéneo que ha ganado importancia en los últimos años. Incluyen el adenoma serrado sésil (ASS), que tiene un alto potencial de malignización a través de una vía molecular diferente a la secuencia adenoma-carcinoma clásica [14].

2. **Pólipos No Neoplásicos:** Generalmente no se consideran precursores del cáncer. Incluyen:

- **Pólipos Hiperplásicos:** Son los pólipos no neoplásicos más comunes. Suelen ser pequeños (<5 mm), se localizan frecuentemente en el recto y el sigmoides y, tradicionalmente, no se considera que tengan potencial maligno, aunque pueden ser difíciles de distinguir de las lesiones serradas [3].
- **Pólipos Inflamatorios:** Son también conocidos como pseudopólipos y no son verdaderas neoplasias. Se forman como resultado de la regeneración de la mucosa en áreas de inflamación crónica severa, típicamente en pacientes con enfermedad inflamatoria intestinal como la colitis ulcerosa o la enfermedad de Crohn[15].
- **Pólipos Hamartomatosos:** Son crecimientos desorganizados de tejido maduro y nativo de la pared colónica. Aunque los pólipos en sí mismos tienen un bajo potencial maligno, a menudo están asociados a síndromes genéticos (como el síndrome de Peutz-Jeghers o la poliposis juvenil) que conllevan un riesgo incrementado de desarrollar cáncer colorrectal y otros tipos de cáncer[16].

Por lo tanto aunque la clasificación de los pólipos es amplia, el interés clínico se centra en la capacidad de las lesiones neoplásicas (adenomas y serrados) para transformarse. Este proceso de transformación maligna, en el cual el pólipo benigno se convierte en una lesión invasora, es lo que define al cáncer colorrectal(CCR) propiamente dicho. El cáncer colorrectal es una neoplasia maligna que se origina en las células epiteliales de la mucosa del colon o del recto. El tipo histológico más frecuente, en más del 95 % de los casos, es el adenocarcinoma[17]. Este se desarrolla a partir de la acumulación de una serie de alteraciones genéticas y epigenéticas en las células del epitelio colónico, lo que conduce a una proliferación celular descontrolada, invasión de tejidos adyacentes y, eventualmente, capacidad de metastatizar a órganos distantes.

La secuencia adenoma-carcinoma describe el modelo de progresión molecular más aceptado, donde mutaciones en genes supresores de tumores (como APC) y oncogenes (como KRAS) transforman gradualmente la mucosa normal en un adenoma de bajo grado, luego de alto grado, y finalmente en un carcinoma invasor[18].

El diagnóstico del CCR y sus lesiones precursoras se basa en una combinación de métodos de cribado y pruebas diagnósticas confirmatorias.

1. **Métodos de Cribado:** destinados a la detección de la enfermedad en población asintomática. Incluyen pruebas de sangre oculta en heces (inmunoquímicas o basadas en guayaco) y pruebas de ADN fecal. Un resultado positivo en estas pruebas requiere la realización de un estudio confirmatorio.
2. **Métodos Diagnósticos**
 - **Colonoscopia:** Es el estándar de oro (gold standard) para el diagnóstico del CCR y los pólipos[3]. Permite la visualización directa de toda la mucosa del colon y del recto. Su principal ventaja es que es una herramienta tanto diagnóstica como terapéutica: permite no solo identificar lesiones sospechosas, sino también tomar biopsias para el análisis histopatológico y realizar la resección de los pólipos (polipectomía) en el mismo procedimiento.
 - **Biopsia y Estudio Histopatológico:** La confirmación definitiva de cáncer solo puede obtenerse mediante el examen microscópico de una muestra de tejido (biopsia) por parte de un patólogo. Este análisis confirma la presencia de células malignas y determina el tipo histológico y el grado de diferenciación del tumor.
 - **Estudios de Imagen para la Estadificación:** Una vez confirmado el diagnóstico de CCR, se realizan estudios de imagen para determinar la extensión de la enfermedad (estadificación clínica). La tomografía computarizada (TC) de tórax, abdomen y pelvis es el estudio estándar para detectar metástasis a distancia. La resonancia magnética (RM) de pelvis es fundamental para la estadificación locorregional del cáncer de recto[18].

El cáncer colorrectal (CCR) constituye una de las principales causas de mortalidad por cáncer a nivel mundial. Según estimaciones de la Agencia Internacional para la Investigación sobre el Cáncer (IARC) y la Organización Mundial de la Salud (GLOBOCAN), ocupa el tercer lugar en frecuencia y el segundo en mortalidad [19], reflejando la fuerte influencia de factores de riesgo modificables como los hábitos dietéticos, la obesidad, el consumo de alcohol y el sedentarismo.

En América Latina y el Caribe, el CCR se ubica como la segunda causa de muerte por cáncer, después del de pulmón. Este hallazgo es consistente con la transición epidemiológica de la región, en la que los cambios en los patrones alimenticios y de estilo de vida han incrementado su carga. Además, se observa una tendencia al alza en países de ingresos medios, lo que genera un importante desafío para los sistemas de salud que aún enfrentan brechas en programas de detección temprana y acceso oportuno a tratamiento [19].

La incidencia de pólipos colorrectales aumenta con la edad, siendo más comunes en personas mayores de 50 años. Por lo tanto la detección temprana de pólipos es crucial para prevenir el desarrollo de cáncer colorrectal

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

Dado que el diagnóstico se realiza principalmente mediante colonoscopia, la efectividad de ella depende de la habilidad del gastroenterólogo para identificar y eliminar estos crecimientos con la resección inmediata del pólipo. La tasa de detección de pólipos puede ser mejorada mediante el uso de tecnologías de asistencia, como sistemas de inteligencia artificial que ayudan en la interpretación de las imágenes obtenidas durante el procedimiento.

Para la identificación de los pólipos el aprendizaje automático o Machine Learning (ML), en particular el aprendizaje profundo o Deep Learning, es una herramienta valiosa en el análisis de imágenes médicas debido a su capacidad para aprender y hacer predicciones a partir de grandes volúmenes de datos, por lo tanto se enuncian a continuación algunos conceptos.

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

3.2.1. Aprendizaje Profundo

El aprendizaje profundo, Deep Learning en inglés, es un campo del aprendizaje automático (Machine Learning), que se engloba, a su vez, dentro de la inteligencia artificial. Las técnicas de Deep Learning se encargan de extraer, por sí solas, las características más relevantes de los datos originales para resolver el problema, además de aprender la transformación que hay que aplicar a estas características para dar la salida esperada [20].

En el contexto de la detección de pólipos en colonoscopias, el aprendizaje profundo se utiliza para: automatizar la detección de pólipos: reduciendo la carga de trabajo del médico y mejorando la consistencia del diagnóstico. Mejorar la precisión diagnóstica: detectando pólipos que pueden ser pasados por alto por el ojo humano [21].

Para dicha detección se hace uso de técnicas de aprendizaje profundo tales como redes neuronales convolucionales y redes neuronales mediante el aprendizaje por transferencia las cuales son clases de redes neuronales profundas que han demostrado ser especialmente efectivas para el procesamiento y análisis de datos visuales. La arquitectura de estas redes neuronales están diseñadas para aprovechar la estructura espacial de las imágenes, lo que las hace ideales para la detección de pólipos en colonoscopias. A continuación se enunciará los aspectos más relevantes.

3.2.2. Redes Neuronales Convolucionales

Una red neuronal convolucional (CNN) es un modelo de aprendizaje profundo diseñado para procesar datos con estructura en forma de grilla, como las imágenes. Su principal fortaleza es la capacidad de extraer automáticamente características relevantes y combinarlas para la clasificación. Para ello, la imagen de entrada pasa por capas convolucionales que aplican filtros para detectar patrones como bordes, formas y texturas, seguidas de funciones de activación (ReLU, Sigmoide,

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

Tanh, Softmax) que permiten aprender representaciones más complejas. Posteriormente, se utilizan operaciones de pooling que reducen la dimensionalidad conservando la información esencial y eliminando ruido, lo que hace más eficiente el procesamiento[22].

En las capas más profundas, la red aprende representaciones cada vez más abstractas: las primeras capturan rasgos simples como líneas, las intermedias identifican estructuras más complejas y las finales reconocen objetos completos. Finalmente, las capas totalmente conectadas combinan la información aprendida y funcionan como un clasificador que, a través de funciones de salida como softmax (multiclase) o sigmoid (binaria), devuelve la probabilidad de que la imagen pertenezca a cada categoría. En analogía, una CNN actúa como un médico que observa primero rasgos básicos, luego estructuras más detalladas y, con toda la información, emite un diagnóstico final. Su estructura básica se presenta en 3.1

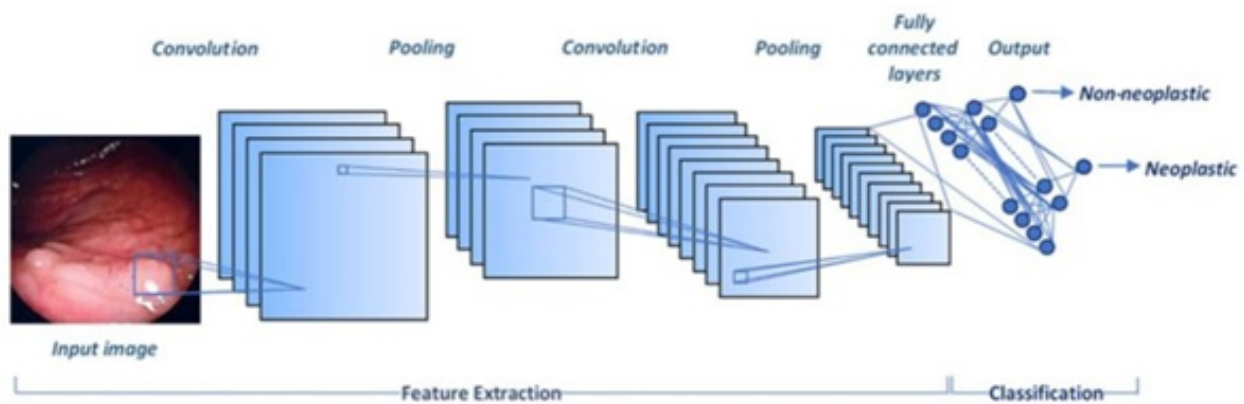


Figura 3.1: Red neuronal convolucional para clasificación de imágenes.

Fuente: tomada de [23]

A continuación se da la descripción de cada termino presente en una red neuronal convolucional.

- **Convolución:** operación matemática fundamental que consiste en aplicar un filtro (también llamado kernel) a una imagen de entrada para crear un mapa de características (feature map). Este mapa resalta patrones específicos en la imagen, como bordes, texturas, formas o colores [24]. Cada filtro tiene un conjunto de pesos que aprende durante el entrenamiento. Diferentes filtros pueden aprender a detectar diferentes características.
- **Filtros (Kernels):** son matrices pequeñas (por ejemplo, de 3x3 o 5x5) que se aplican sobre la imagen para extraer características. Por ejemplo, un filtro podría detectar bordes horizontales, mientras que otro podría detectar bordes verticales[24].
- **Mapas de Características. (Feature Maps):** cuando un filtro se desplaza sobre la imagen, (o una copia previa), produce un mapa de características, este mapa contiene la respuesta del

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

filtro en cada región de la imagen. Esencialmente, resalta las regiones donde el filtro detecta una característica particular[24].

- **Pooling. (Agrupamiento):** después de la Convolución, se puede aplicar la operación de pooling para reducir la dimensionalidad del mapa de características. El tipo más común, es el max pooling que tomará el valor máximo en una pequeña región de la imagen, reducido así la resolución pero manteniendo las características importantes[24].
- **Capa de activación:** después de la convolución, las redes aplican una función no lineal como Relu (Rectified Linear Unit). Esto introduce no linealidad en el modelo, lo que permite que la red aprenda relaciones complejas[24].

Sin funciones de activación, las redes neuronales serían esencialmente una cadena de operaciones lineales, lo cual mantendría la salida como una transformación lineal de la entrada. Esto limitaría su capacidad para aprender patrones complejos y tomar decisiones basadas en datos con estructuras no triviales. La introducción de funciones no lineales permite que la red neuronal modele relaciones mucho más complejas.

- **Capa convolucional** representa el componente central y la principal innovación de las Redes Neuronales Convolucionales, siendo la responsable de la extracción jerárquica de características a partir de los datos de entrada. Su operación fundamental, la convolución, consiste en el desplazamiento de un filtro o kernel a través de las dimensiones espaciales. El resultado de esta operación genera un mapa de características, que indica la respuesta del filtro en cada posición espacial, resaltando la presencia de patrones específicos como bordes, texturas o formas. Mediante el uso de múltiples filtros en una misma capa, el modelo es capaz de aprender a detectar un conjunto diverso de características de manera simultánea, formando una representación rica y distribuida de la información visual[22].
- **Capas totalmente conectadas. (Fully connected layers o DENSE):** después de varias capas convolucionales y de pooling al final de la red. Se añaden las capas conectadas. Estas capas se usan para tomar las características extraídas por las capas convolucionales y producir una predicción final. (Por ejemplo, clasificar una imagen en una categoría)[24].

- **Entrenamiento de una Red Neuronal Convolucional**

Las Redes Neuronales Convolucionales (CNN) se entrenan utilizando el algoritmo de retropropagación (backpropagation) y el descenso de gradiente. Durante el entrenamiento la red ajusta los pesos de los filtros y las conexiones para minimizar el error de predicción en un conjunto de datos[22].

Es el proceso en el que la red aprende a hacer predicciones precisas a partir de los datos de entrenamiento. Este proceso implica ajustar los pesos (parámetros) de la red de manera que minimice el error entre las predicciones de la red y las respuestas reales (o etiquetas) en los datos.

Un flujo típico de una CNN puede ser:

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

1. **Entrada:** una imagen, por ejemplo 64x64 píxeles en tres canales rojo, verde y azul (RGB).
2. **Capa convolucional:** aplica varios filtros para extraer características, pueden haber múltiples capas convolucionales seguidas.
3. **Capa de Pooling:** reduce la dimensionalidad, del mapa de características.
4. **Capas adicionales:** más capas convolucionales y pooling pueden añadirse para extraer características más complejas.
5. **Capa totalmente conectada (DENSE):** Conecta todas las características, extraídas a neuronas que se encargan de clasificar o hacer Predicciones.
6. **Salida:** Una predicción, como una probabilidad. para cada clase a un problema clasificación de imágenes. [22]

3.2.3. Aprendizaje por Transferencia

El aprendizaje por transferencia (transfer learning) constituye una estrategia ampliamente utilizada en el ámbito del aprendizaje profundo, mediante la cual un modelo previamente entrenado en una tarea general, normalmente sobre un conjunto de datos de gran escala como ImageNet, es reutilizado y adaptado para abordar un nuevo problema específico. Este enfoque se fundamenta en la premisa de que las capas iniciales de una red neuronal convolucional son capaces de capturar características de las imágenes, tales como bordes, formas y texturas, que resultan útiles en una amplia gama de dominios, mientras que las capas finales pueden ajustarse para representar de manera más precisa los patrones propios de la tarea particular, por ejemplo, la clasificación de pólipos en imágenes médicas[25].

Desde una perspectiva práctica, el aprendizaje por transferencia ofrece ventajas significativas al permitir reducir los tiempos de entrenamiento, optimizar el uso de recursos computacionales y mejorar el rendimiento del modelo, especialmente en escenarios donde se dispone de un número limitado de datos etiquetados. La adaptación del modelo puede realizarse mediante la congelación de las capas preentrenadas y el reentrenamiento exclusivo de las últimas capas densas, o a través de un ajuste más fino (fine-tuning) que implica la actualización de los pesos de toda la red, posibilitando una mayor especialización en la tarea objetivo.

En el caso particular de la clasificación de imágenes de colonoscopias, el aprendizaje por transferencia resulta altamente pertinente debido a la escasez de conjuntos de datos de gran escala y alta calidad en este campo. Las CNN preentrenadas pueden aprovechar las representaciones generales previamente aprendidas y adaptarlas para identificar estructuras propias del colon, como mucosa normal y lesiones características, entre ellas los pólipos colorrectales. Este proceso puede implementarse mediante dos estrategias principales: (i) la extracción de características, en la cual se congelan las capas convolucionales del modelo preentrenado y únicamente se entrenan las capas finales con las imágenes de colonoscopias; y (ii) el ajuste fino (fine-tuning), que implica reentrenar

3.2. Fundamentos del Aprendizaje Profundo en la Detección de Pólipos Colorrectales

parcial o totalmente la red con el nuevo conjunto de datos, permitiendo una mayor especialización y sensibilidad en la detección de lesiones[25].

3.2.4. Métricas de Desempeño de Modelos en Aprendizaje Automático

Teniendo en cuenta los valores reales y los predichos por el modelo es decir las clases positivas y negativas obtenidas por la red y las esperadas, se obtienen cuatro métricas en un problema de clasificación.

		Clasificación Real	
		Pólipos	No Pólipos
Clasificación Predicha	Pólipos	TP	FP
	No Pólipos	FN	TN

Cuadro 3.1: Matriz de confusión para la clasificación de pólipos.

Si la clase asignada por la red como pólipos coincide con la clase pólipos esperada (real) se cuenta como un **verdaderos positivos** (True Positive TP). Los Verdaderos negativos, (true negatives TN), ocurre si la clase asignada por la red coinciden con la clase no pólipos y la clasificación esperada (real) es no pólipos. Un **falso positivo** (FP) se produce cuando la red indica una clasificación pólipo y la clasificación esperada (real) es no pólipo. Un **falso negativo** (FN) se produce cuando la red indica una clasificación no pólipo cuando la clasificación esperada (real) es pólipo.[26]

La matriz anterior 3.1 se denomina **matriz de confusión** y a partir de ella se construyen la mayoría de las métricas más utilizadas para la evaluación de los modelos algunas de ellas son:

- **Exactitud** o accuracy en inglés. Esta métrica da el ratio de datos clasificados correctamente (tanto para la clase positiva como negativa):

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

- **Recall o sensibilidad** en inglés. Con esta métrica se pone el foco en lo bien que lo hace el modelo con los datos de la clase positiva. Da la razón de datos de la clase positiva que han sido bien clasificados, es decir la proporción de verdaderos positivos entre todos los casos reales positivos[26].

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.2)$$

- **Precisión.** Esta métrica es utilizada para saber la razón de datos que son realmente de la clase positiva respecto del total de ejemplos que la red ha clasificado como positivos, es decir

la proporción de verdaderos positivos entre todas las predicciones positivas[26].

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.3)$$

- **Especificidad** o specificity en inglés. Similar a recall solo que esta métrica se centra en los datos de la clase negativa, es decir la proporción de verdaderos negativos entre todos los casos reales negativos[26].

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (3.4)$$

- **F1-score**, es la media armónica entre precision (qué tan correctas son las predicciones positivas) y recall (qué tanto de lo positivo real detecta el modelo)[20].

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.5)$$

Esta evalúa el balance entre falsos positivos y falsos negativos

A continuación se presenta algunas características los pólipos los cuales son precursores del cáncer colorectal.

3.3. Antecedentes

Para el desarrollo del proyecto fue esencial conocer los estudios previos realizados a nivel mundial y publicados en revistas indexadas. Estos estudios ofrecen una base firme para establecer los antecedentes más relevantes.

Se utilizó Google Académico para identificar los artículos de mayor importancia en el área de modelos de segmentación y clasificación de imágenes de colonoscopias. Estos artículos constituyeron un inicio de bastante importancia para encontrar los antecedentes más relevantes de la investigación en curso.

La clasificación y segmentación de imágenes es esencial en el análisis de imágenes médicas, ya que permite identificar y delimitar regiones de interés (ROI) en diversas imágenes, como órganos, lesiones y tejidos. Una segmentación precisa es crucial para múltiples aplicaciones clínicas, como el diagnóstico de enfermedades, la planificación de tratamientos y el seguimiento de la progresión de las enfermedades. Los modelos de Machine Learning han mostrado un gran potencial en la segmentación de imágenes de colonoscopias debido a su capacidad para aprender características complejas y ofrecer resultados precisos y eficientes.

Dentro de los artículos identificados se destacan los siguientes trabajos de investigación en relación con el tema de análisis propuesto:

- **Automated classification of colon polyps in endoscopic image data**

En el artículo los autores desarrollaron un sistema automatizado para clasificar los pólipos de colon en hiperplasias y adenomas según las estructuras de los vasos en imágenes endoscópicas, lo que logró una alta precisión de predicción del 96,2%. El sistema se basó en la segmentación de vasos, extrayendo 22 características destinadas a describir las complejas topologías vasculares de los pólipos. Se aplicaron tres estrategias de selección de características para identificar las más relevantes. El sistema fue evaluado en 286 imágenes NBI de pólipos pequeños (de 10 mm o menos), provenientes de una base de datos representativa. El análisis histopatológico se utilizó como referencia (verdad fundamental) para evaluar el rendimiento del sistema [27]. Este artículo fue altamente relevante y de gran utilidad para el desarrollo del proyecto ya que ambos comparten un enfoque similar en cuanto al uso de imágenes médicas para la detección y clasificación automatizada de pólipos en el colon.

- Por otro lado **Deep Learning for Classification of Colorectal Polyps on Whole-slide Images**,

en este artículo los autores desarrollaron un método basado en aprendizaje profundo para clasificar con precisión diferentes tipos de pólipos colorrectales en imágenes completas, lo que puede ayudar a los patólogos con la caracterización y el diagnóstico de pólipos colorrectales para una detección más eficiente y precisa del cáncer colorrectal. Utilizaron imágenes de pólipos colorrectales completas anotadas como estándares de referencia para entrenar y probar modelos de aprendizaje profundo. Implementaron una arquitectura ResNet modificada con atajos adicionales para abordar el problema del gradiente de desaparición. Preprocesaron las imágenes reescalando, normalizando y aplicando técnicas de aumento de datos como vibración de color y rotación [28]. Este trabajo fue particularmente útil para el proyecto propuesto, ya que ambos enfoques comparten la meta de utilizar modelos de aprendizaje profundo para la clasificación de pólipos colorrectales.

- Por su parte en el artículo **Classification of colon polyps in NBI endoscopy using vascularization features**.

El resumen de este artículo es que presenta un algoritmo de clasificación basado en computadora que puede distinguir entre dos tipos de pólipos de colon, hiperplasias y adenomas, con una alta precisión de aproximadamente el 90% basada en características extraídas de la vascularización de los pólipos. - Segmentación de vasos utilizando dos algoritmos diferentes - Esqueletización de los vasos detectados y su uso como puntos de semilla para el algoritmo Fast Marching para segmentar todo el lumen del vaso - Extracción de características de la segmentación de vasos - Uso de estas características para clasificar los pólipos como adenomas o hiperplasias, con una precisión de clasificación reportada de alrededor del 90% en pruebas de dejar uno afuera en una base de datos de 56 pólipos [29]. También fue altamente relevante para el proyecto desarrollado, ya que ambos emplearon imágenes endoscópicas para la clasificación de pólipos y comparten el enfoque en la extracción de características específicas de los pólipos para mejorar la precisión del diagnóstico.

En el artículo **BASIC (BLI Adenoma Serrated International Classification) classification for colorectal polyp characterization with blue light imaging**, se describe el

desarrollo y la validación de un nuevo sistema de clasificación llamado BASIC (BLI Adenoma Serrated International Classification) para caracterizar pólipos colorrectales utilizando imágenes de luz azul (BLI) con y sin aumento óptico. Crearon una videoteca de 48 videos/imágenes fijas de 24 pólipos/cáncer verificados histológicamente.

Siete endoscopistas experimentados revisaron los videos/imágenes e identificaron posibles descriptores. Los descriptores se categorizaron y resumieron utilizando una metodología Delphi modificada. Los endoscopistas revisaron de forma independiente 36 videos/imágenes adicionales y se evaluó el acuerdo interobservador utilizando la estadística AC1 de Gwet [30]. Se reviso este artículo pues se penso en trabajos futuros implementar modelos en tiempo real que sirvan para detección de pólipos a través de videos.

- **Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks** el artículo describe el desarrollo y la evaluación de un sistema de red neuronal convolucional (CNN) que puede detectar y clasificar automáticamente los pólipos colorrectales utilizando imágenes endoscópicas, con el objetivo de apoyar la estrategia de resear y descartar para el manejo de pólipos pequeños durante la colonoscopia. Utilizaron una arquitectura CNN profunda llamada SSD para detectar y clasificar pólipos colorrectales, entrenaron la CNN en un gran conjunto de datos de 16,418 pólipos colorrectales e imágenes normales luego validaron la CNN entrenada en un conjunto independiente de 7,077 imágenes de colonoscopia y evaluaron el desempeño de la CNN en la detección y clasificación de los pólipos colorrectales, incluidas las medidas de sensibilidad, valor predictivo positivo y velocidad de procesamiento [31]. Este artículo resulto muy pertinente para el proyecto propuesto, dado que ambos se propusieron como meta el uso de redes neuronales convolucionales (CNN) para la detección y clasificación automática de pólipos colorrectales en imágenes endoscópicas.

Preparación de Datos

4.1. Obtención de los datos

Las imágenes utilizadas para el desarrollo de esta investigación se obtuvieron de la plataforma Kaggle, se utilizaron dos conjuntos de datos especializados en imágenes de colonoscopias: KVA-SIR y PolypGen 2021. Ambos datasets contienen imágenes de colonoscopia con clasificaciones de pólipos y tejidos sin presencia de pólipos, lo cual los convierte en recursos fundamentales para el entrenamiento y evaluación de modelos de clasificación en aplicaciones clínicas. La elección de estos conjuntos se fundamenta en su calidad, disponibilidad pública para estudios de clasificación dentro del campo médico.

A continuación se da una descripción más detallada de estos dataset de imágenes.

El dataset PolypGen 2021 fue introducido como parte del desafío internacional PolypGen Challenge, en el marco del congreso MICCAI 2021. Aporta imágenes de colonoscopias, capturadas con diferentes dispositivos y en condiciones clínicas diversas, lo que favorece la generalización de los modelos entrenados.

4.2. Descripción del conjunto de datos PolypGen 2021

El conjunto de datos PolypGen 2021 constituye uno de los recursos públicos más completos y de acceso abierto para el estudio de la detección y segmentación de pólipos colorrectales. Está conformado por imágenes fijas, secuencias de vídeo, así como fotogramas positivos y negativos, recopilados en un consorcio de seis hospitales distribuidos en Europa y África (Noruega, Francia, Reino Unido, Egipto e Italia). En total, el dataset incluye imágenes, y secuencias de vídeo, en cuanto a imágenes 3.762 tienen presencia de pólipos y 4.275 imágenes de tejidos sin presencia de pólipos [32], con dimensiones correspondientes a 1350 píxeles de alto por 1080 píxeles de ancho. Esta diversidad se refleja no solo en el origen geográfico de los pacientes, sino también en el empleo de diferentes sistemas endoscópicos y protocolos clínicos, lo que contribuye a la variabilidad técnica y poblacional y favorece la generalización de los modelos de aprendizaje profundo entrenados con este conjunto.

La construcción del dataset siguió un protocolo estandarizado que garantizó la heterogeneidad en la distribución de los datos y la representatividad clínica de los casos. Cada centro participante aportó imágenes y videos de al menos 50 pacientes, recopilados y validados por gastroenterólogos

sénior. El protocolo incluyó: (i) muestreo de un solo cuadro de video por paciente desde diferentes ángulos de visión; (ii) muestreo de secuencias con marcos en los que el pólipo era visible e invisible, con espacios mínimos entre ellos; (iii) en los fotogramas únicos se incluyeron todas las instancias de pólipos detectados en un paciente, mientras que en las secuencias solo se consideró un pólipo objetivo localizado; y (iv) las secuencias positivas incorporaron tanto cuadros con pólipos como sin ellos, siempre provenientes de videos con pólipos confirmados, mientras que las secuencias negativas se obtuvieron exclusivamente de pacientes con mucosa normal[32]. Además, el dataset presenta variedad morfológica al incluir pólipos de distintos tamaños (pequeños, medianos y grandes), así como lesiones planas, ampliando así la complejidad del desafío diagnóstico y computacional. Por las características anteriores PolypGen se considera uno de los datasets más completos para la tarea de clasificación de pólipos en imágenes de colonoscopias[32].

4.3. Descripción del conjunto de datos KVASIR

El KVASIR Dataset es un conjunto de datos abierto, desarrollado por el Simula Research Laboratory en Noruega, con el objetivo de apoyar investigaciones en el área de visión por computador aplicada a la endoscopia gastrointestinal, contiene imágenes de pólipos y tejidos sin presencia de pólipos ver figura 4.1 y 4.2, las cuales fueron clasificadas por expertos médicos. Las imágenes provienen de procedimientos reales de colonoscopia y presentan una resolución de 619 x 529 píxeles. Este dataset ha sido utilizado como referencia en distintas publicaciones científicas[33].

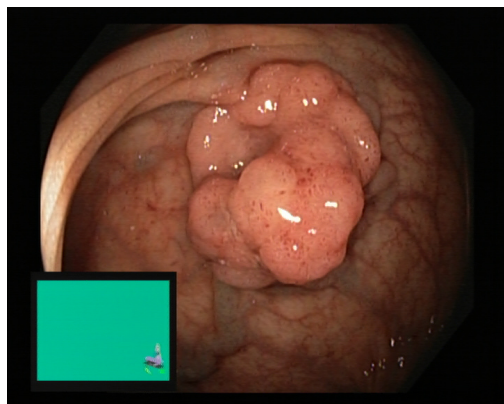


Figura 4.1: Presencia de pólipo.

Fuente: conjunto de datos KVASIR.

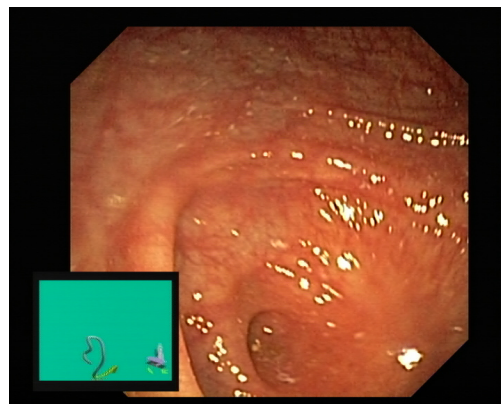


Figura 4.2: Sin presencia de pólipos.

Fuente: conjunto de datos KVASIR.

El conjunto de datos KVASIR fue desarrollado por el Simula Research Laboratory en colaboración con el Hospital Vestre Viken de Noruega y publicado en 2017 como parte de un esfuerzo por proporcionar un punto de referencia o estándar predefinido (benchmark), estandarizado y accesible para la comunidad científica. Su principal objetivo es facilitar el desarrollo, validación y compara-

ción de algoritmos de visión por computador y aprendizaje automático aplicados a la clasificación y detección de hallazgos gastrointestinales.

El dataset está compuesto por 4.000 imágenes endoscópicas que fueron seleccionadas, anotadas y validadas por especialistas en gastroenterología. Estas imágenes se organizan en ocho clases clínicas balanceadas, con aproximadamente 500 instancias por categoría, que abarcan tanto estructuras anatómicas normales como condiciones patológicas y procedimientos. Las clases incluyen: línea Z (unión gastroesofágica), píloro, ciego, esofagitis, pólipos, colitis ulcerosa, sangrado gastrointestinal y extracción con pinza. Esta organización balanceada permite que los algoritmos entrenados con KVASIR no sufran un sesgo marcado hacia clases mayoritarias, un problema común en bases de datos médicas[33].

Desde su publicación, KVASIR ha sido relevante en la investigación biomédica, convirtiéndose en un conjunto de referencia para la validación de sistemas de apoyo al diagnóstico por computadora. Se ha utilizado extensamente en estudios de clasificación de pólipos. Su disponibilidad en acceso abierto ha promovido la reproducibilidad y la comparación objetiva entre diferentes enfoques metodológicos, un aspecto crucial en la investigación en inteligencia artificial aplicada a la medicina.

El dataset KVASIR constituye un recurso de entrenamiento para algoritmos de aprendizaje automático y también una plataforma de validación para la estandarización de benchmarks internacionales en endoscopia gastrointestinal. Su diseño balanceado, validación clínica y diversidad de clases lo convierten en un punto de partida sólido para estudios orientados tanto al diagnóstico asistido como a la investigación en segmentación, clasificación y recuperación de patrones en imágenes endoscópicas.

Ambos dataset se descargaron y organizaron en directorios correspondientes para su uso en el preprocesamiento y entrenamiento del modelo. A continuación se hace una descripción del procesamiento de imágenes.

4.4. Preprocesamiento de las imágenes

Dado que muchas de las imágenes de colonoscopias incluían una marca de agua en la esquina inferior izquierda, generada por el equipo utilizado en la captura de las imágenes, se realizó un procesamiento previo. Este procesamiento consistió en añadir un recuadro negro en la esquina mencionada de dimensiones 248 x 193 píxeles, cubriendo dicha marca para evitar interferencias con el entrenamiento del modelo, como se evidencia en las figuras 4.3 y 4.4.

Para el caso del dataset PolypGen 2021 se redimensionaron las imágenes al mismo tamaño de las imágenes del KVASIR Dataset, correspondiente a 619 ancho X 529 alto y luego se añadió el recuadro negro como se observa en las figura 4.5 y 4.6.

Para llevar a cabo este proceso: se desarrolló un script en Python ejecutado en Google Colab que automatizó el preprocesamiento de las imágenes. Este proceso inició cargando las imágenes originales desde una carpeta en Google Drive, la cual contenía imágenes de colonoscopias clasificadas en formatos .jpg. Para garantizar un manejo organizado y preservar las imágenes originales, se creó automáticamente una carpeta de salida en Google Drive, destinada a almacenar las imágenes

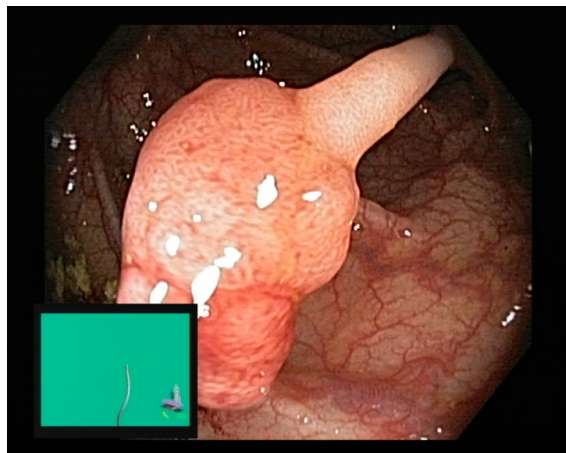


Figura 4.3: Imagen original KVASIR..

Fuente: conjunto de datos KVASIR.

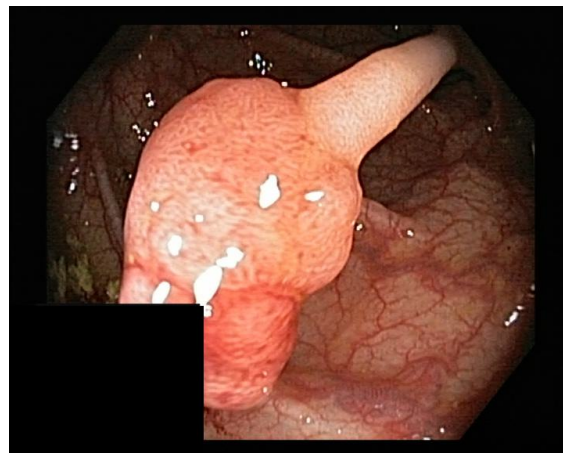


Figura 4.4: Imagen modificada KVASIR..

Fuente: adaptado de conjunto de datos KVASIR..

modificadas.

El procesamiento consistió en identificar y recorrer cada imagen en la carpeta de entrada. Posteriormente, se definió el rectángulo negro mencionado, que se posicionó en la parte inferior de cada imagen, con el objetivo de cubrir áreas que pudieran contener información que perturbara el análisis, como datos sensibles o etiquetas no relacionadas con el contenido visual. Este rectángulo se dibujó utilizando la biblioteca Pillow (PIL), y las imágenes resultantes se guardaron en la carpeta de salida manteniendo sus nombres originales.

Adicionalmente, para asegurar la calidad y verificación del procesamiento, se utilizó la biblioteca matplotlib para visualizar las imágenes modificadas, permitiendo confirmar visualmente la correcta aplicación de las transformaciones. Este procedimiento aseguró un preprocesamiento consistente y estandarizado de las imágenes de colonoscopias, facilitando el uso posterior de estas en el desarrollo de los demás objetivos, como la implementación de algoritmos de aprendizaje profundo. El proceso finalizó con todas las imágenes procesadas y organizadas en la carpeta designada, cumpliendo de manera eficiente el objetivo planteado.

Para construir un conjunto de datos balanceado que integro información de las distintas fuentes, se realizó una mezcla de imágenes provenientes de los datasets KVASIR y PolypGen2021. Inicialmente, se redimensionaron las imágenes de PolypGen2021, compuestas por 4.275 imágenes de tejido normal y 3.762 con presencia de pólipos. Posteriormente, se recopilamos imágenes para las clases normal y pólipo desde diversas carpetas. Estas imágenes fueron combinadas, mezcladas aleatoriamente y se seleccionaron 1.500 imágenes por clase, creando un nuevo dataset balanceado que fue organizado en la carpeta imagenes mezcladas. Esta carpeta se estructuró en tres subconjuntos: Train2 (con 1.050 imágenes por clase, equivalente al 70%), Val2 (150 imágenes por clase, 10%) y Test2 (150



Figura 4.5: Imagen original PolyGen 2021.

Fuente: conjunto de datos PolyGen 2021.



Figura 4.6: Imagen modificada PolyGen 2021.

Fuente: adaptado de conjunto de datos PolyGen 2021.

imágenes por clase, 20%), cada uno con sus respectivas subcarpetas normal y pólipos 4.1.

	Clase	Total	KVASIR	PolypGen2021	Proporción (%)
Train2	Pólipos	525	262	263	50 % / 50 %
	Normal	525	263	262	50 % / 50 %
Val2	Pólipos	75	37	38	50 % / 50 %
	Normal	75	38	37	50 % / 50 %
Test2	Pólipos	150	75	75	50 % / 50 %
	Normal	150	75	75	50 % / 50 %
TOTAL		1500	750	750	50 % / 50 %

Cuadro 4.1: Distribución de imágenes combinadas por conjunto, clase y fuente de origen

Desarrollo de Modelos

Se desarrollaron tres modelos: el primero basado en una red neuronal convolucional con dos capas convolucionales con 32 filtros de tamaño 3x3, dos capas densas de 128 y 1 unidad respectivamente, donde se obtuvo un accuracy de 0.94 en el conjunto test, con curvas de desempeño 6.1 y el cual fue base para luego llevar a acabo una experimentación variando los parámetros como se muestra en el cuadro 5.1

Hiperparámetro	Valores explorados	Descripción
Filtros por capa	32, 64	Número de filtros aplicados en cada capa convolucional.
Neuronas en capa densa	128, 256	Tamaño de la(s) capa(s) densa(s) posterior(es) a las convolucionales.
Tasa de Dropout	0.3, 0.5	Proporción de neuronas desactivadas durante el entrenamiento para evitar el sobreajuste.
Tamaño del kernel	3, 5	Dimensión de la ventana del filtro convolucional (por ejemplo, 3×3 o 5×5).
Capas convolucionales	2, 3, 4	Cantidad de capas convolucionales apiladas en la arquitectura.

Cuadro 5.1: Hiperparámetros explorados en el modelo convolucional

buscando optimizar hiperparámetros. En la experimentación se obtuvo el mejor modelo en el experimento 25, correspondiente a los parámetros del cuadro 5.2

Filtros	Neuronas densas	Dropout	Tamaño kernel	Capas convolucionales
64	128	0.3	3	2

Cuadro 5.2: Configuración de hiperparámetros del experimento 25

obteniéndose un accuracy de 0.96.

El segundo modelo basado en transferencia de aprendizaje teniendo en cuenta la arquitectura VGG-16, por sus siglas en ingles (Visual Geometry Group) con 16 capas con pesos entrenables, 13 capas convolucionales y 3 capas densas para este modelo se dejaron las últimas 3 capas densas

entrenables, en el cual se obtuvo un accuracy de 0.95 cuyo desempeño se puede visualizar en la figura 6.2.

Finalmente el tercero también basado en transferencia de aprendizaje teniendo en cuenta la arquitectura VGG-16 dejando el quinto bloque y las capas densas entrenables, presento un accuracy de 0.97. A continuación se hace una descripción de los diferentes modelos desarrollados:

5.1. Modelo Red Neuronal Convolutacional

Se desarrolló e implementó un modelo de red neuronal convolutacional (CNN) utilizando TensorFlow y Keras. Este modelo tuvo como propósito clasificar imágenes de colonoscopias en dos categorías: imágenes normales y aquellas que contienen pólipos. Para este fin, se utilizó Google Colab como entorno de desarrollo, con los datos organizados en carpetas las carpetas mencionadas en el capítulo anterior dentro de Google Drive.

Las imágenes se procesaron mediante la clase ImageDataGenerator de Keras, que aplicó transformaciones como reescalado al rango $[0,1]$, zoom aleatorio, volteo horizontal y transformaciones de corte. Estas técnicas de aumento de datos enriquecieron el conjunto de entrenamiento, aumentando la diversidad de ejemplos y mejorando la capacidad generalizadora del modelo.

La red CNN diseñada comprendió las siguientes capas principales:

Capas convolucionales: dos capas convolucionales con 32 filtros y un tamaño de kernel de 3×3 , cada una seguida de una función de activación ReLU. Estas capas permitieron extraer características espaciales relevantes de las imágenes.

Operaciones de max-pooling: dos capas de max-pooling con un tamaño de ventana 2×2 , con saltos de tamaño 2, utilizadas para reducir la dimensionalidad y conservar las características más importantes.

Capa de aplanamiento (Flatten): para transformar las matrices de características en un vector unidimensional, facilitando su entrada en las capas densas.

Capas densas: Una capa intermedia completamente conectada con 128 unidades y función de activación ReLU, seguida de una capa de salida con una unidad y activación sigmoide para clasificar las imágenes en dos categorías (tejidos con presencia de pólipos y tejidos sin presencia de pólipos).

La arquitectura mencionada se puede visualizar en la figura 5.1

El entrenamiento se llevó a cabo durante 50 épocas, utilizando un tamaño de lote de 32 imágenes. Se emplearon callbacks como ModelCheckpoint para guardar automáticamente el mejor modelo basado en el desempeño en el conjunto de validación, y EarlyStopping para detener el entrenamiento si no se observaban mejoras significativas durante 10 épocas consecutivas.

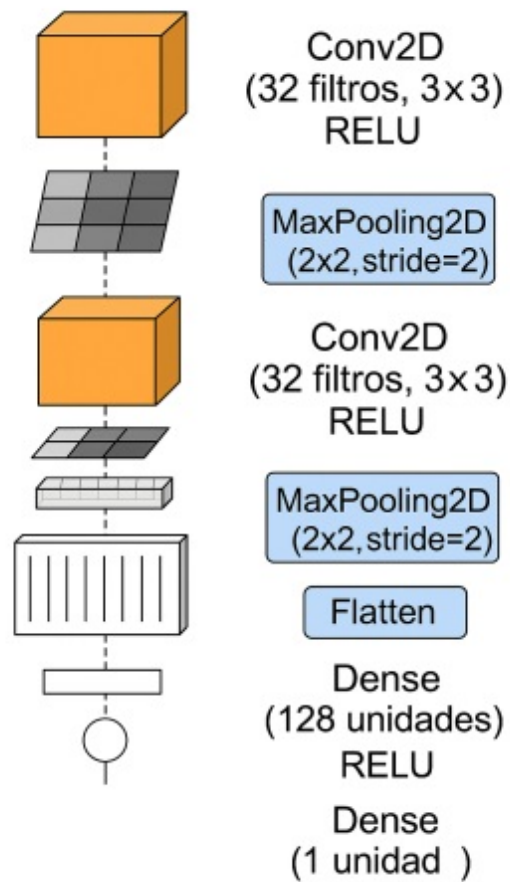


Figura 5.1: Arquitectura de Red Neuronal Convolutacional.

Fuente propia.

La arquitectura del modelo y su desempeño se evaluaron gráficamente mediante las curvas de pérdida y precisión, tanto para los conjuntos de entrenamiento como de validación. Estas curvas mostraron una mejora consistente a lo largo de las épocas, validando la capacidad del modelo para aprender y generalizar.

El modelo final, junto con las gráficas y métricas de desempeño, se almacenó en Google Drive, garantizando su accesibilidad para futuras evaluaciones y refinamientos. Este desarrollo representa un paso clave en la tesis, cumpliendo con el objetivo de entrenar modelos de aprendizaje profundo para abordar un problema crítico en el análisis de imágenes médicas.

5.1.1. Modelo Optimizando parámetros

Luego se evaluaron el desempeño de diferentes configuraciones de redes neuronales convolucionales (CNN), se diseñó un experimento de búsqueda de hiperparámetros. Con este propósito, se establecieron combinaciones de valores para cinco parámetros fundamentales en la arquitectura de la red: número de filtros convolucionales: 32 y 64, número de neuronas en la capa densa: 128 y 256, tasa de dropout: 0.3 y 0.5, tamaño del kernel: 3 y 5 y número de capas convolucionales: 2, 3 y 4 cuadro 5.1. La combinación de estos hiperparámetros generó un total de $2 \times 2 \times 2 \times 2 \times 3 = 48$ configuraciones experimentales.

Cada modelo fue implementado en TensorFlow/Keras, utilizando como entrada imágenes de 64×64 píxeles con tres canales de color (RGB). La arquitectura general consistió en múltiples capas convolucionales y de max pooling, seguidas de una capa flatten, una capa densa con activación ReLU, una capa de dropout y finalmente una capa de salida con activación sigmoide para la clasificación binaria. El modelo se compiló con el optimizador Adam, que combina las ventajas de RMSprop y momentum para un ajuste eficiente de los parámetros, empleando como función de pérdida la entropía cruzada binaria, adecuada para problemas de clasificación binaria.

Para mejorar la generalización y evitar sobreajuste, se incluyeron dos callbacks: ModelCheckpoint, encargado de guardar el modelo en la época con mejor desempeño en el conjunto de validación y EarlyStopping, configurado con una paciencia de 5 épocas, que detuvo el entrenamiento cuando no se observó mejora adicional en la métrica de validación y restauró los pesos correspondientes al mejor resultado.

Cada configuración se entrenó durante un máximo de 20 épocas, utilizando conjuntos de entrenamiento y validación previamente definidos. Finalizado el entrenamiento, cada modelo se evaluó en el conjunto de validación, registrando las métricas de precisión (accuracy) y pérdida (loss).

Los resultados de las 48 configuraciones se almacenaron en un DataFrame de Pandas, lo que permitió organizar y comparar los modelos en función de la precisión de validación. Tras la ejecución de los 48 experimentos, el modelo con mejor desempeño en el conjunto de validación correspondió a la configuración con 64 filtros convolucionales, una capa densa de 128 neuronas, una tasa de dropout de 0.3, un tamaño de kernel de 3×3 y dos capas convolucionales ver cuadro 5.2. Este modelo utilizó como entrada imágenes de dimensión 64×64 píxeles en formato RGB.

La arquitectura seleccionada logró una precisión superior en validación respecto a las demás configuraciones evaluadas, mostrando un adecuado balance entre complejidad y capacidad de generalización.

5.2. Modelo Transferencia de Aprendizaje Capas Densas Entrenables

Otro modelo desarrollado fue un clasificador de imágenes de colonoscopias basado en la arquitectura VGG16, utilizando la técnica de aprendizaje por transferencia. Este enfoque consistió en

aprovechar una red previamente entrenada en un gran conjunto de datos en este caso, ImageNet un repositorio que contiene más de 14 millones de imágenes etiquetadas en 1,000 categorías de objetos del mundo real, como animales, vehículos, herramientas y escenas cotidianas, para extraer características visuales generales, y luego se adaptó dicha red a la clasificación médica.

Para este modelo se aplicó una estrategia de transferencia sin ajuste fino, es decir que todas las capas convolucionales del modelo VGG16 fueron congeladas, desde la entrada (input layer) hasta el último bloque, es decir el hasta la capa de pooling del quinto bloque. Esta decisión garantiza que los pesos previamente aprendidos por la red durante su entrenamiento con millones de imágenes naturales no se vean modificados. En consecuencia, la red actúa exclusivamente como un extractor de características estático.

Las imágenes de colonoscopias fueron procesadas a través de esta red convolucional preentrenada, la cual generó mapas de características que describieron patrones complejos. Estos mapas fueron posteriormente enviados a una nueva sección del modelo compuesta por capas densas entrenables, específicamente diseñadas para interpretar esas representaciones y realizar la clasificación correspondiente.

La arquitectura del modelo fue la siguiente: tras la capa de salida del modelo base, se incorporaron una capa de aplanamiento (Flatten), que convierte los mapas de características en un vector unidimensional y tres capas densas:

- Una capa Dense(256) con activación ReLU, que permitió al modelo capturar combinaciones no lineales de las características extraídas.
- Una capa adicional Dense(128) con activación ReLU, que actuó como refinamiento intermedio de la representación aprendida.
- Una capa de salida Dense con activación softmax, que permitió predecir la probabilidad de pertenencia a una de las dos clases: pólipos o normales. Ver figura 5.2

Todas estas capas densas fueron inicializadas desde cero y entrenadas completamente, ya que forman la parte específica del modelo encargada de adaptar el conocimiento general de VGG16 a la tarea concreta de clasificación médica.

Además el modelo fue compilado utilizando: el optimizador Adam, por su adaptabilidad y eficiencia.

La función de pérdida categorical crossentropy, que se emplea en problemas de clasificación multiclase (binaria con codificación one-hot).

5.3. Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables

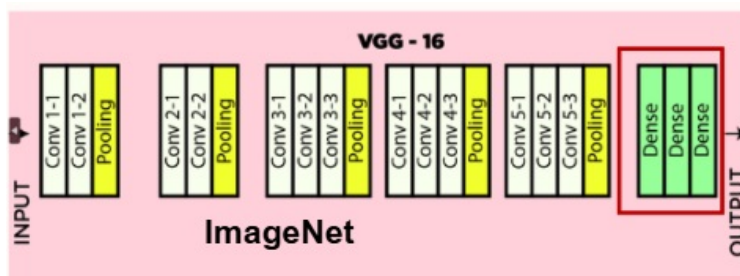


Figura 5.2: Arquitectura Modelo Transferencia de Aprendizaje Capas Densas Entrenables.

Imagen tomada de: <https://tinyurl.com/4969p9hc>.

Se utilizó la métrica de evaluación accuracy (exactitud), que mide la proporción de predicciones correctas durante el entrenamiento y la validación. Con respecto al aumento de datos y procesamiento, para maximizar la generalización del modelo y reducir el riesgo de sobreajuste, se aplicaron técnicas de aumentación de datos sobre el conjunto de entrenamiento mediante la herramienta ImageDataGenerator. Las transformaciones aplicadas incluyeron: normalización de píxeles entre 0 y 1, rotación aleatoria de hasta 20 grados, volteo horizontal aleatorio, que simuló la variabilidad espacial en la adquisición de las imágenes.

El conjunto de datos estaba organizado en directorios independientes para entrenamiento, validación y prueba. Todas las imágenes fueron redimensionadas a 224x224 píxeles, acorde con el tamaño de entrada requerido por VGG16, y procesadas en lotes de 32 imágenes.

El hecho de congelar completamente las capas convolucionales permitió reducir significativamente el tiempo de entrenamiento y la necesidad de recursos computacionales, ya que únicamente se ajustan los pesos de las capas densas finales.

5.3. Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables

Para el tercer modelo de clasificación de imágenes de colonoscopias, se utilizó una estrategia de aprendizaje por transferencia a partir de la arquitectura VGG16, también preentrenada sobre el conjunto de datos ImageNet.

En este caso se utilizó la siguiente estrategia:

- Se congelaron todas las capas convolucionales desde bloque 1 hasta el bloque 4, lo que implica que sus pesos permanecieron fijos durante todo el entrenamiento.
- Se dejaron entrenables las capas convolucionales del bloque cinco y las capas densas. Esta elección estratégica permitió ajustar las capas más profundas del extractor de características

5.3. Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables

a las particularidades presentes en las imágenes médicas del dominio específico, sin perder la robustez de las representaciones más generales extraídas por las capas más superficiales. Como se ve en la figura 5.3

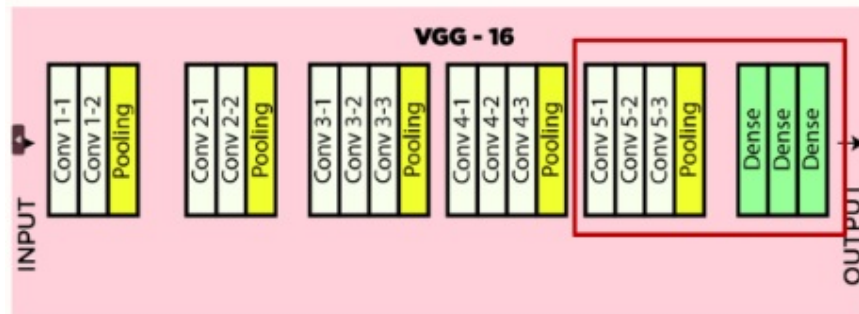


Figura 5.3: Arquitectura Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables

Imagen tomada de: <https://tinyurl.com/4969p9hc>.

Esta configuración de arquitectura combinó estabilidad y especialización: el conocimiento general aprendido de millones de imágenes naturales en ImageNet se mantuvo intacto en las primeras capas, mientras que el bloque cinco y las capas densas añadidas fueron ajustadas específicamente al contexto clínico.

Capas densas, sobre la salida del bloque convolucional preentrenado y parcialmente ajustado, se añadieron dos capas densas entrenables:

- Una capa Dense(256 neuronas) con activación ReLU, que actúo como una capa intermedia para capturar patrones complejos a partir de las características extraídas.
- Una capa de salida Dense(1 neurona) con activación sigmoid, que genero la probabilidad de pertenencia a una de las dos clases objetivo: tejidos con presencia de pólipos o tejidos sin presencia de pólipos.

Este diseño permitió que la red se enfocara en refinar y traducir las representaciones visuales generadas por VGG16 hacia una decisión de clasificación clínicamente relevante. En cuanto a compilación y entrenamiento, en el modelo se utilizo:

- Optimizador Adam, elegido por su eficiencia computacional y adaptabilidad durante el proceso de entrenamiento.

5.3. Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables

- Función de pérdida categorical crossentropy, adecuada para problemas de clasificación multi-clase (en este caso, binaria con codificación one-hot).
- La métrica de evaluación accuracy, que permitió monitorear el rendimiento del modelo en función de su capacidad de clasificación correcta.

Para el preprocesamiento y aumento de datos en pro de mejorar la generalización del modelo y reducir el riesgo de sobreajuste, se aplicaron técnicas de aumentación de datos (data augmentation) durante el entrenamiento. A través de ImageDataGenerator, las imágenes fueron sometidas a transformaciones como: escalado de píxeles a valores entre 0 y 1, rotaciones aleatorias de hasta 20 grados, volteo horizontal aleatorio. Estas transformaciones simulaban variaciones comunes en las imágenes del mundo real y permitió al modelo aprender de un conjunto más amplio de posibles escenarios visuales sin necesidad de aumentar la cantidad real de datos disponibles. Además las imágenes fueron redimensionada a 224x224 píxeles, acorde a la entrada esperada por VGG16, y procesada en lotes de 32 imágenes durante el entrenamiento.

El ajuste de las últimas capas convolucionales permitió que el modelo refinara su capacidad de detección de patrones visuales relevantes en el contexto médico. Esto implicó una mejora en la especialización del modelo sin necesidad de reentrenar completamente la red desde cero, lo cual es beneficioso en contextos donde los recursos computacionales o los datos disponibles son limitados.

Análisis de Resultados

Ya que la evaluación del desempeño de los modelos desarrollados constituye una etapa fundamental en cualquier sistema de diagnóstico asistido por inteligencia artificial. A continuación se presentan los resultados obtenidos de los modelos mencionados, se analizaron en función de métricas para evaluar el desempeño las cuales fueron la precisión, que mide la proporción de predicciones positivas correctas; el recall o sensibilidad, que cuantifica la capacidad del modelo para identificar correctamente los casos positivos; y el F1-score, que representa el equilibrio entre precisión y recall al calcular su media armónica. Adicionalmente, se reportó el soporte, correspondiente al número de instancias reales de cada clase en el conjunto de prueba, y la exactitud global (accuracy), que expresa el porcentaje total de clasificaciones correctas.

Obteniendo los siguientes resultados para los cuatro modelos desarrollados:

	Imágenes con presencia de pólipos		
	Precisión	Recall	F1-score
Modelo Red Neuronal Convolutacional	0,96	0,92	0,94
Mejor Modelo Obtenido Luego de la Estimación de Parámetros	0,96	0,96	0,96
Modelo Transferencia de Aprendizaje Capas Densas Entrenables	0,94	0,97	0,96
Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables	0,96	0,98	0,97

Cuadro 6.1: Comparación del desempeño de los modelos en imágenes con presencia de pólipos

	Imágenes sin presencia de pólipos		
	Precisión	Recall	F1-score
Modelo Red Neuronal Convolutacional	0,92	0,96	0,94
Mejor Modelo Obtenido Luego de la Estimación de Parámetros	0,96	0,96	0,96
Modelo Transferencia de Aprendizaje Capas Densas Entrenables	0,97	0,94	0,95
Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables	0,98	0,96	0,97

Cuadro 6.2: Comparación del desempeño de los modelos en imágenes sin presencia de pólipos

	Desempeño Global	
	Accuracy	Macro avg
Red Neuronal Convolutacional	0,94	0,94
Mejor Modelo Obtenido Luego de la Estimación de Parámetros	0,96	0,96
Transferencia de Aprendizaje Capas Densas Entrenables	0,95	0,95
Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables	0,97	0,97

Cuadro 6.3: Exactitud global y promedio macro obtenidos por los diferentes modelos

El modelo de **Red Neuronal Convolutacional** alcanzó un desempeño sólido, con una exactitud global del 94 % en la clasificación de imágenes endoscópicas. El reporte de clasificación muestra métricas balanceadas entre ambas clases: la categoría tejido sin presencia de pólipos obtuvo una precisión de 0.92 y un recall(sensibilidad) de 0.96(Cuadro 6.2), mientras que la categoría tejido con presencia pólipos alcanzó una precisión de 0.96 y un recall(sensibilidad) de 0.92, lo que evidencia una ligera asimetría en el rendimiento según la clase evaluada. El f1-score se mantuvo en 0.94 para ambas categorías (Cuadro 6.1), indicando un equilibrio adecuado entre precisión y sensibilidad.

Con respecto a la matriz de confusión obtenida a partir del conjunto de imagenes de prueba del modelo de clasificación de imágenes de colonoscopia presentada en cuadro 6.4 . Los resultados evidencian un desempeño robusto del sistema en la discriminación entre imágenes normales y aquellas que presentan pólipos colorrectales.

El modelo clasificó correctamente 288 imágenes sin presencia de pólipos(verdaderos negativos) y 276 imágenes con presencia de pólipos (verdaderos positivos). En cuanto a los errores de clasificación, se identificaron 12 falsos positivos, correspondientes a imágenes sin presencia de pólipos erróneamente clasificadas como presencia de pólipos, y 24 falsos negativos, asociados a imágenes

con presencia de pólipos que no fueron detectadas por el modelo.

Desde el punto de vista clínico, los falsos negativos representan el error de mayor relevancia, ya que implican la no identificación de lesiones potencialmente premalignas, lo que podría retrasar el diagnóstico oportuno y el tratamiento adecuado.

Real	Predicción	
	Sin presencia de pólipos	Con presencia de pólipos
Sin presencia de pólipos	288	12
Con presencia de pólipos	24	276

Cuadro 6.4: Matriz de confusión del modelo Red Neuronal Convolutacional.

Los resultados anteriores reflejan la capacidad del modelo para identificar patrones relevantes en las imágenes, aunque con un desempeño ligeramente inferior frente al modelo basado en transferencia de aprendizaje, lo cual sugiere que la CNN desde cero requiere un mayor ajuste o mayor cantidad de datos para igualar el nivel de generalización alcanzado por enfoques más avanzados.

Las curvas de entrenamiento y validación (Figura 6.1) muestran una disminución progresiva de la pérdida a lo largo de las 30 épocas, alcanzando valores inferiores a 0.1 en ambos conjuntos. Paralelamente, la precisión de entrenamiento y validación se incrementó de manera sostenida, superando el 95 % desde la época 15 y manteniéndose estable en las últimas iteraciones. El comportamiento similar de ambas curvas sugiere que el modelo logró un buen equilibrio entre ajuste y generalización, sin evidencias marcadas de sobreajuste.

En conjunto, estos resultados indican que el modelo es efectivo para la tarea de clasificación binaria de imágenes de colonoscopia, con un rendimiento sólido en ambas clases y un potencial clínico relevante, aunque con margen de mejora en la reducción de falsos negativos.

Como se menciona en el capítulo desarrollo de modelos, luego de la experimentación el mejor desempeño en el conjunto de validación correspondió a la configuración con 64 filtros convolucionales, dos capas convolucionales, un tamaño de kernel de 3x3, una capa densa de 128 neuronas y una tasa de dropout de 0.3. Con esta arquitectura, el modelo alcanzó una exactitud global del 96 %, mostrando un balance óptimo entre clases. El reporte de clasificación evidencia que tanto la clase tejidos sin presencia de pólipos como la clase tejidos con presencia de pólipos alcanzaron valores de precisión, recall y F1-score de 0.96 (Cuadro 6.1) y (Cuadro 6.2), lo que refleja un desempeño consistente en la identificación de ambas categorías.

Estos resultados indican que la configuración seleccionada logra un adecuado compromiso entre capacidad de generalización y complejidad, consolidándose como la más robusta dentro de las variantes evaluadas.

El modelo basado en transferencia de aprendizaje con VGG16 y capas densas entrenables (**Modelo Transferencia de Aprendizaje Capas Densas Entrenables**) alcanzó una exactitud global del 95 % en el conjunto de prueba, mostrando un rendimiento balanceado entre ambas clases. Para la clase tejidos sin presencia de pólipos obtuvo una precisión de 0.97, un re-

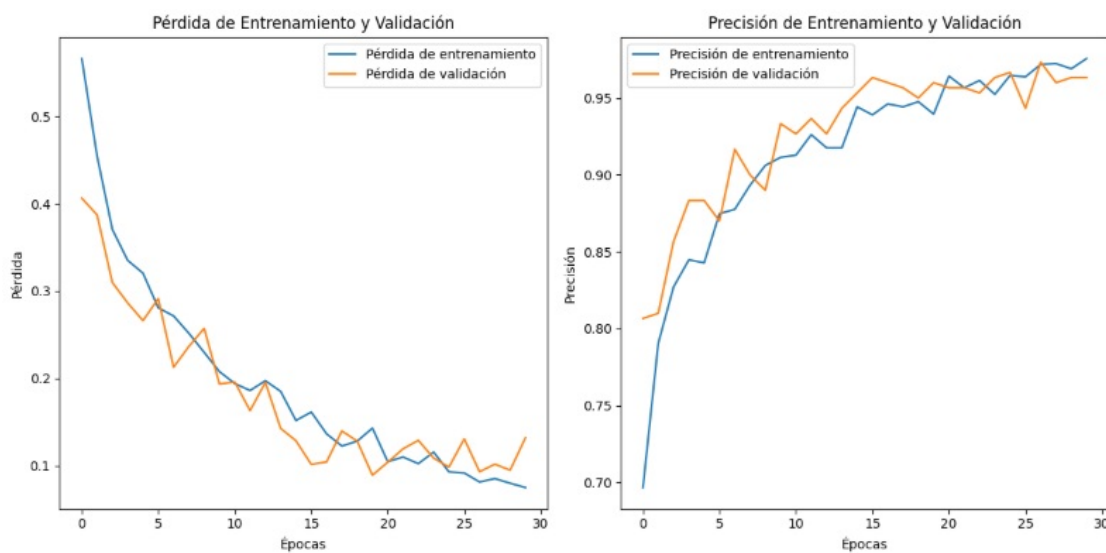


Figura 6.1: Pérdida de entrenamiento y validación, Precisión de entrenamiento y validación modelo red neuronal convolucional

Fuente: propia

call(sensibilidad) de 0.94 y un F1-score de 0.95(Cuadro 6.2), lo que indica que la mayoría de los casos tejidos con presencia de pólipos fueron correctamente identificados, aunque con algunos falsos positivos(10). En la clase con presencia de pólipos, el desempeño fue complementario, con una precisión de 0.94, un recall(sensibilidad) de 0.97 y un F1-score de 0.96 (Cuadro 6.1), reflejando la capacidad del modelo para detectar la mayoría de los casos positivos, aunque con un número reducido de falsos negativos(17), obtenido de la matriz de confusión obtenida para este modelo 6.5 que evidencia un desempeño sobresaliente en la identificación de pólipos colorrectales y tejido sin presencia de pólipos. Del total de imágenes evaluadas, el modelo clasificó correctamente 283 imágenes sin presencia de pólipos (verdaderos negativos) y 290 imágenes con presencia de pólipos (verdaderos positivos).

En cuanto a los errores de clasificación, se identificaron 17 falsos positivos, correspondientes a imágenes sin presencia de pólipos clasificadas erróneamente como con presencia de pólipos, y 10 falsos negativos, asociados a imágenes con presencia de pólipos que no fueron detectadas por el sistema.

Desde una perspectiva clínica, el bajo número de falsos negativos es particularmente relevante, ya que refleja una elevada capacidad del modelo para detectar lesiones potencialmente premalignas, reduciendo el riesgo de omitir pólipos durante el proceso de evaluación endoscópica.

Real	Predicción	
	Sin presencia de pólipos	Con presencia de pólipos
Sin presencia de pólipos	283	17
Con presencia de pólipos	10	290

Cuadro 6.5: Matriz de confusión del modelo transferencia de aprendizaje capas densas entrenables

En términos agregados, tanto el promedio macro como el ponderado de las métricas alcanzaron valores cercanos a 0.95–0.96, confirmando la robustez y consistencia del modelo en ambas clases.

Las curvas de pérdida y precisión (Figura 6.2) muestran un descenso rápido de la pérdida durante las primeras épocas de entrenamiento, seguido de una estabilización, con valores inferiores a 0.1 tanto en entrenamiento como en validación. De forma paralela, la precisión se incrementó de manera sostenida, superando el 90% desde las primeras iteraciones y estabilizándose en torno al 95–97%. El comportamiento similar de las curvas de entrenamiento y validación evidencia que el modelo logró un adecuado nivel de generalización, sin presentar sobreajuste significativo.

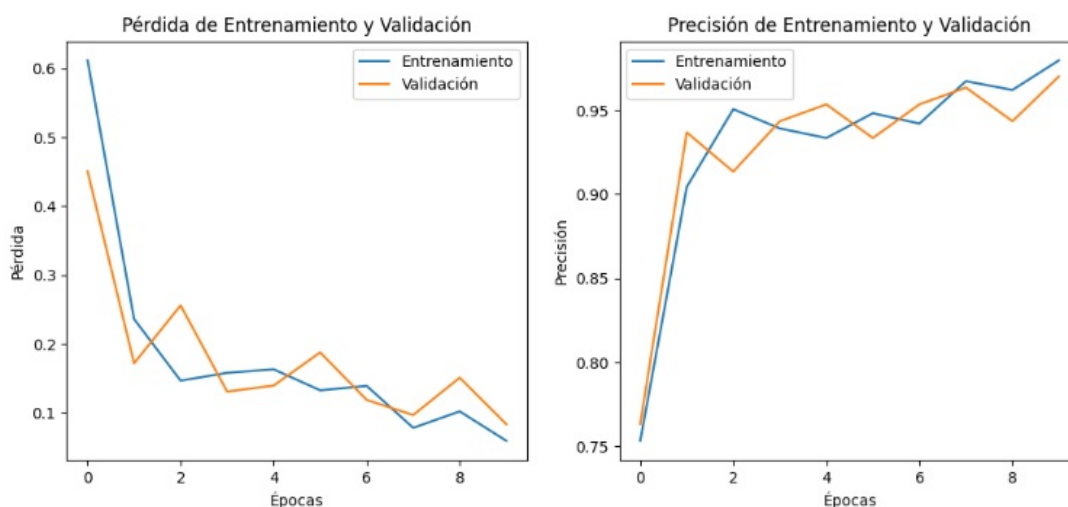


Figura 6.2: Pérdida de entrenamiento y validación, Precisión de entrenamiento modelo transferencia de aprendizaje capas densas

Fuente: propia

En conjunto, los resultados sugieren que el uso de transferencia de aprendizaje con capas densas entrenables constituye una estrategia eficaz para la clasificación de imágenes de colonoscopia, alcanzando un alto rendimiento con un número limitado de épocas y con buena capacidad para discriminar entre imágenes de tejidos con presencia de pólipos y con clase tejidos sin presencia de pólipos.

En el modelo de **Transferencia de Aprendizaje con Bloque Cinco y Capas Densas**

Entrenables, se evidencia un desempeño altamente satisfactorio. Los resultados de las curvas de precisión y pérdida por época muestran una rápida convergencia, alcanzando valores de precisión superiores al 95 % desde las primeras iteraciones y estabilizándose en torno al 97 %, mientras que la pérdida disminuye progresivamente hasta valores cercanos a 0.08, lo que refleja un proceso de entrenamiento estable y eficiente. El reporte de clasificación respalda estos hallazgos, con métricas de precisión de 0.98, recall de 0.96 y f1-score de 0.97 (Cuadro 6.2) en la clase sin presencia de pólipos y en la clase presencia de pólipos: precisión de 0.96, recall de 0.98 y f1-score de 0.97 (Cuadro 6.1) , y una exactitud global del 97 %.

Con respecto a la matriz de confusión correspondiente Modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables 6.6 evidencia un desempeño altamente eficiente en la identificación de pólipos colorrectales y tejido sin presencia de pólipos. Del total de imágenes evaluadas, el modelo clasificó correctamente 289 imágenes sin presencia de pólipos (verdaderos negativos) y 294 imágenes con presencia de pólipos (verdaderos positivos).

En relación con los errores de clasificación, se registraron 11 falsos positivos, correspondientes a imágenes sin presencia de pólipos clasificadas erróneamente como pólipos, y 6 falsos negativos, asociados a imágenes con presencia de pólipos que no fueron detectadas por el sistema.

Desde el punto de vista clínico, el número reducido de falsos negativos resulta especialmente relevante, dado que indica una elevada capacidad del modelo para detectar lesiones potencialmente premalignas, disminuyendo el riesgo de omisión diagnóstica durante el procedimiento endoscópico.

Real	Predicción	
	Sin presencia de pólipos	Con presencia de pólipos
Sin presencia de pólipos	289	11
Con presencia de pólipos	6	294

Cuadro 6.6: Matriz de confusión del modelo con transferencia de aprendizaje y capas densas entrenables

Los resultados mencionados sugieren que el modelo presenta un adecuado balance entre sensibilidad y especificidad, logrando una generalización robusta y minimizando el riesgo de sobreajuste en la tarea de clasificación de imágenes endoscópicas.

El modelo muestra un comportamiento favorable durante el entrenamiento. En la gráfica de precisión se observa un incremento sostenido tanto en los datos de entrenamiento como en los de validación, alcanzando valores del 97 % en las últimas épocas. Esto indica que el modelo fue capaz de generalizar adecuadamente sin evidenciar caídas bruscas o sobreajuste significativo. En cuanto a la pérdida, se aprecia una disminución progresiva en ambas curvas, con una tendencia estable y convergente hacia valores cercanos a cero (Figura 6.3). La cercanía entre las curvas de entrenamiento y validación en ambas métricas refleja un buen ajuste del modelo, lo que sugiere que la arquitectura implementada logró un balance adecuado entre complejidad y capacidad de generalización, garantizando un rendimiento robusto en la tarea de clasificación.

Realizando la comparación de los modelos desarrollados (una red neuronal convolucional de-

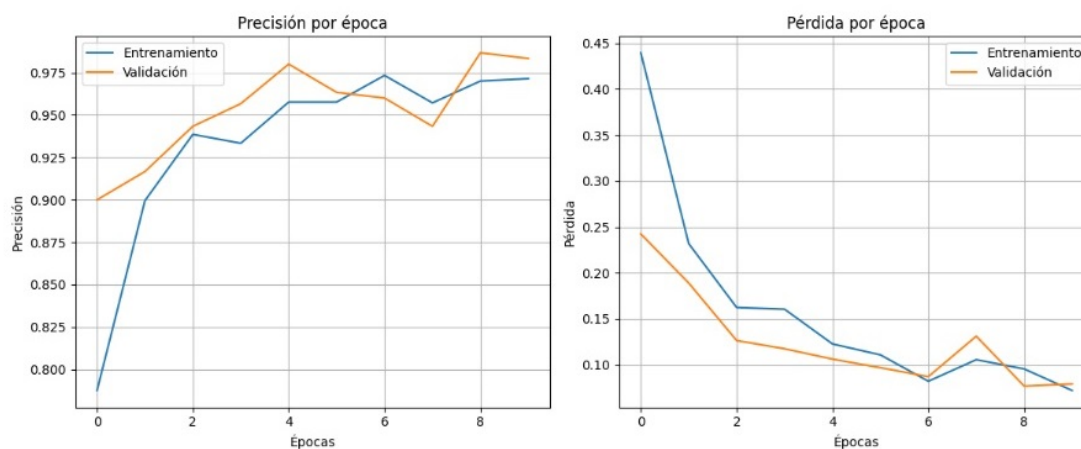


Figura 6.3: Pérdida de entrenamiento y validación, Precisión de entrenamiento modelo transferencia de aprendizaje bloque cinco y capas densas entrenables

Fuente: propia

sarrollada desde cero, un modelo optimizado tras la estimación de parámetros, y dos variantes de modelos de transferencia de aprendizaje) se evidencia diferencias en cuanto a su capacidad para discriminar entre clase tejidos sin presencia de pólipos y aquellas que presentan pólipos. El análisis se sustenta en las métricas de precisión, recall, F1-score y accuracy, indicadores ampliamente aceptados para la evaluación de clasificadores en problemas biomédicos.

El modelo base de red neuronal convolucional (CNN) mostró un desempeño sólido, aunque con una ligera tendencia a subestimar la clase positiva (presencia de pólipo). La mejora mediante estimación de parámetros elevó todos los indicadores, mostrando un avance significativo en la estabilidad de las predicciones.

Sin embargo, los modelos de transferencia de aprendizaje demostraron un rendimiento superior: el modelo con capas densas entrenables y con la estrategia de ajuste fino en el bloque cinco, se logró el mejor desempeño global.

Teniendo en cuenta la exactitud global (Accuracy), el análisis de la métrica global de exactitud confirma la superioridad de los modelos de transferencia: CNN: 0,94, Modelo optimizado: 0,96, Transfer learning con capas densas: 0,95, Transfer learning con bloque cinco ajustado: 0,97. En cuanto al macro average (promedio macro) teniendo en cuenta la métrica de desempeño de F1-score para cada clase por separado y luego promediando estos valores de forma no ponderada, otorgando el mismo peso a cada clase fue de 0.94, 0.96, 0.95 y 0.97 respectivamente (cuadro 6.3).

Aunque la diferencia entre 0,95 y 0,97 puede parecer marginal, en aplicaciones clínicas estas me-

mejoras porcentuales se traducen en un número considerable de diagnósticos adicionales correctamente identificados.

Los resultados anteriores permiten extraer varios hallazgos importantes:

- La CNN base, aunque adecuada, presenta limitaciones inherentes en la extracción de características, requiriendo una optimización cuidadosa de hiperparámetros para mejorar su rendimiento.
- La estimación de parámetros mejora notablemente la capacidad de la CNN, confirmando la importancia de técnicas de regularización, ajuste de tasas de aprendizaje y número de filtros en las capas convolucionales.
- La transferencia de aprendizaje supera a los modelos entrenados desde cero, especialmente al aplicar ajuste fino en capas profundas (bloque cinco). Este resultado es coherente con la literatura, que destaca cómo los modelos preentrenados en grandes bases de datos (ImageNet) capturan representaciones jerárquicas transferibles al dominio biomédico.
- El balance entre precisión y recall es crucial en este contexto. Mientras que el modelo con capas densas entrenables privilegia el recall (minimiza falsos negativos), el ajuste en el bloque cinco logra un equilibrio más estable, constituyéndose en la mejor alternativa para aplicaciones clínicas.

Luego del análisis comparativo se evidencia que el modelo de transferencia de aprendizaje **con ajuste fino en el bloque cinco y capas densas entrenables** constituye la mejor opción, al alcanzar la mayor precisión, recall, F1-score y accuracy. Este modelo no solo asegura una adecuada clasificación de imágenes normales y con pólipos, sino que también ofrece un balance crítico entre sensibilidad y especificidad, aspectos esenciales en un entorno médico donde los errores de clasificación pueden tener implicaciones graves para la salud del paciente.

En síntesis, la transferencia de aprendizaje con ajuste fino profundo se consolida como la estrategia más efectiva para la identificación automática de pólipos en imágenes de colonoscopia, representando un avance significativo frente a arquitecturas entrenadas desde cero.

Despliegue del Prototipo

Con el objetivo de permitir la interacción directa con el modelo de **Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables**, se diseñó y desarrolló una aplicación web mediante la biblioteca **Streamlit**. Esta herramienta, utilizada para el despliegue de aplicaciones de ciencia de datos y aprendizaje automático, permitió la creación de una interfaz accesible, interactiva y funcional que facilita la validación práctica del modelo en tiempo real.

7.1. Interfaz para interactuar con el modelo.

La aplicación fue diseñada como un entorno de validación, en el cual los usuarios pueden cargar imágenes de colonoscopia desde su dispositivo y obtener, de manera inmediata, la predicción generada por el modelo. Una vez cargada la imagen médica en formato `.png`, `.jpg` o `.jpeg`, el sistema realiza el preprocesamiento requerido (incluyendo el redimensionamiento, normalización y ajuste del formato de entrada) para posteriormente pasarla al modelo de clasificación. Como salida, la plataforma muestra la categoría asignada (sin presencia de pólipo o presencia de pólipo).

En la interfaz, se incluyeron funcionalidades complementarias, tales como: visualización de la imagen cargada, lo que permite al usuario verificar el archivo antes del procesamiento. La interfaz fue pensada para facilitar su uso tanto a investigadores como a profesionales del área de la salud.

La elección de **Streamlit** fue debido a su facilidad de integración con modelos de aprendizaje profundo desarrollados en **TensorFlow** y **PyTorch**, su bajo costo computacional y su capacidad de ser ejecutado localmente o en servidores remotos, habilitando el acceso desde cualquier navegador. La flexibilidad obtenida posiciona a la aplicación como una herramienta idónea para demostrar la aplicabilidad del modelo en contextos reales.

Además el desarrollo de esta interfaz constituye un prototipo funcional de apoyo al diagnóstico asistido por inteligencia artificial y representa un puente entre la investigación y la práctica clínica. Al trasladar los resultados del entrenamiento del modelo a un entorno de interacción accesible, se facilita la validación por parte de especialistas médicos, se fomenta la retroalimentación interdisciplinaria y se allana el camino hacia el diseño de sistemas integrados que puedan implementarse en entornos hospitalarios. También la implementación de la aplicación con **Streamlit** materializa la propuesta de un sistema interactivo que permite evaluar en tiempo real la clasificación de imágenes de colonoscopia. Esto no solo refuerza la utilidad práctica del modelo de transferencia de aprendizaje ajustado, sino que además abre la posibilidad de su futura escalabilidad a entornos clínicos,

contribuyendo al fortalecimiento de herramientas de detección temprana de pólipos y prevención del cáncer colorrectal.

Esta aplicación implementada, contenida en el archivo `app2.py`, permite a los usuarios cargar imágenes médicas en formato `.png`, `.jpg` o `.jpeg`, visualizarlas y clasificarlas utilizando el modelo previamente entrenado y guardado. La interfaz contempla los siguientes componentes:

- **Carga del modelo entrenado:** mediante `load_model` se importa el archivo `my_keras_model.keras`.
- **Carga y visualización de imágenes:** se permite al usuario subir una imagen, que luego se muestra en la interfaz.
- **Procesamiento previo y predicción:** la imagen se redimensiona (64x64), se normaliza y se transforma en una matriz NumPy compatible con el modelo. En este caso se manejan tanto salidas binarias.
- **Interpretación y visualización del diagnóstico:** se muestra el diagnóstico en función de la clasificación de la imagen.

Esta aplicación proporciona una herramienta funcional, accesible y de fácil uso para profesionales médicos o usuarios no técnicos. El flujo básico de interacción en la aplicación es el siguiente:

1. El usuario carga una imagen de entrada desde su dispositivo.
2. La imagen es preprocesada (rescalada, normalizada y convertida a RGB).
3. El modelo realiza la predicción con base en la imagen procesada.
4. Se presentan en pantalla la clasificación correspondiente.

El desarrollo de la interfaz se llevo a cabo de la siguiente forma:

7.2. Descripción del desarrollo de la interfaz

Para garantizar la correcta ejecución y portabilidad de la aplicación, se utilizó la distribución Anaconda, en donde se generó un entorno denominado `streamlitenv`, dentro del cual se instalaron las bibliotecas necesarias: `TensorFlow/Keras`, `Streamlit`, `scikit-image`, `Pillow`, y dependencias auxiliares como `numpy` y `io`. Este procedimiento aseguró que el modelo y la interfaz pudieran ejecutarse sin conflictos de versiones con otras librerías del sistema.

El archivo principal de la aplicación, denominado `app2.py`, constituye el núcleo del sistema. En él se implementaron los siguientes componentes clave:

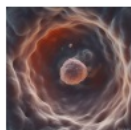
Carga del modelo entrenado: se especificó la ruta del modelo Transferencia de Aprendizaje Bloque Cinco y Capas Densas Entrenables, previamente entrenado y almacenado en formato `.keras`.

Preprocesamiento de imágenes: las imágenes cargadas por el usuario son transformadas mediante la función `preprocessimage`, este módulo garantiza compatibilidad con el modelo al: convertir imágenes a formato RGB, redimensionar las imágenes al tamaño esperado por VGG16 (224×224 píxeles) y normalizar los valores de los píxeles a un rango entre 0 y 1.

Módulo de predicción: la función `predictimage` recibe como entrada la imagen procesada y devuelve la clasificación de cada clase (sin presencia de pólipo o presencia de pólipo).

Interfaz gráfica con Streamlit: se diseñó un encabezado centralizado que resalta el nombre del clasificador y su arquitectura base (VGG16). Se incluyó un bloque introductorio con un ícono representativo cargado desde el sistema local, acompañado de un texto explicativo. Se habilitó un cargador de archivos `preprocess image.(st.file uploader)` que permite al usuario subir imágenes en formatos .png, .jpg o .jpeg ver figura 7.1. Antes de la clasificación, se aplicó un procesamiento adicional para cubrir con un rectángulo negro una sección específica de la imagen, con el fin de eliminar posibles marcas visuales o rótulos que pudieran sesgar el resultado. Al presionar el botón Clasificar, se ejecuta el diagnóstico en tiempo real y se muestra de forma clara y diferenciada: en caso de detectar pólipos, se genera una alerta visual en color rojo, en caso de clasificar sin presencia de pólipo, se despliega un mensaje en verde, figuras 7.2 y 7.3.

BIENVENIDO AL SISTEMA DE APOYO DIAGNÓSTICO PARA PÓLIPOS



Visualización artística del interior del colon

Guía de uso del sistema de clasificación

Esta aplicación utiliza el mejor modelo de inteligencia artificial desarrollado tras varios entrenamientos y pruebas de rendimiento. Su objetivo es clasificar imágenes de colonoscopias, identificando si contienen **presencia de pólipos** o **sin presencia de pólipos**.

Para usarla, cargue una imagen y el sistema realizará el análisis automáticamente, mostrando el resultado de la clasificación.

Sube una imagen de colonoscopia

Drag and drop file here
Limit 200MB per file • PNG, JPG, JPEG

Browse files

Figura 7.1: Presentación interfaz.

Fuente propia.

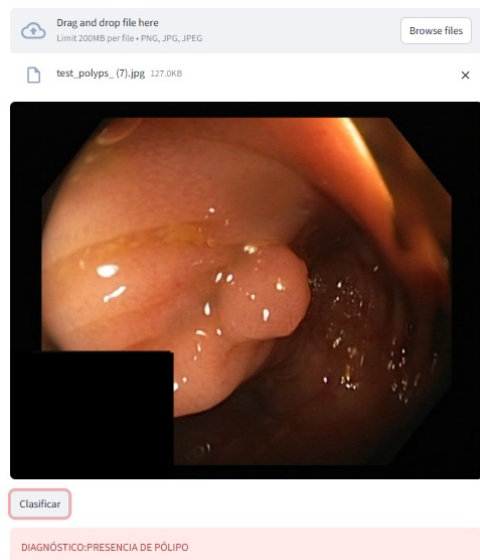


Figura 7.2: Clasificación de imagen presencia de pólipo.

Fuente propia

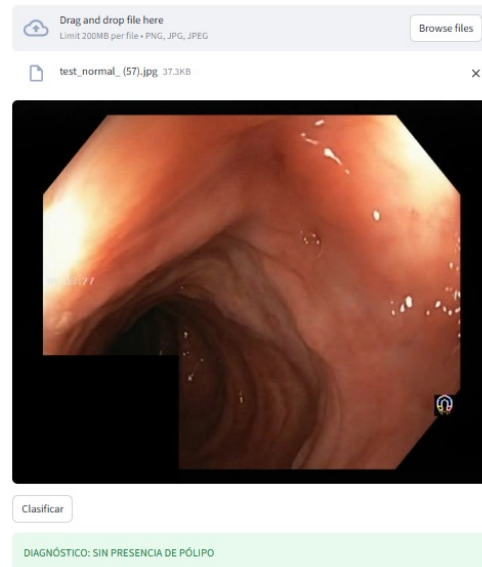


Figura 7.3: Clasificación de imagen sin presencia de pólipo.

Fuente propia

Ejecución del sistema: el código incluye el bloque estándar, lo que permite ejecutar la aplicación de forma autónoma desde el entorno streamlitenv con el comando: `streamlit run app.py`

La implementación de este prototipo representa un ejemplo práctico de transferencia de aprendizaje aplicada en un contexto biomédico. Se evidencia el uso de una arquitectura de referencia (VGG16) con ajuste fino en capas específicas y también la integración de un entorno de despliegue accesible que permite validar el modelo de manera interactiva.

El uso de Streamlit en conjunto con Anaconda asegura la reproducibilidad del experimento, la modularidad del proyecto y su futura escalabilidad, ya que la aplicación puede ser adaptada a un servidor hospitalario o a un entorno de investigación colaborativa. Con esto, se logra cerrar la brecha entre el entrenamiento experimental de modelos en laboratorio y su eventual adopción como herramientas de apoyo clínico en la detección temprana de pólipos.

Conclusiones y Trabajos Futuros

8.1. Conclusiones

1. El desarrollo del presente trabajo permitió alcanzar la totalidad de los objetivos planteados. Se estableció un protocolo de procesamiento previo para las imágenes de colonoscopia, se entrenaron diversos modelos de aprendizaje profundo y se desarrolló una interfaz funcional para interactuar con el modelo de mejor desempeño.
2. El análisis comparativo de los modelos evaluados demuestra que las arquitecturas basadas en transferencia de aprendizaje superan en desempeño a las redes convolucionales entrenadas desde cero, especialmente al aplicar ajuste fino en las capas profundas. En particular, la configuración con bloque cinco y capas densas entrenables alcanzó los valores más altos de precisión, recall, F1-score y exactitud global, consolidándose como la opción más robusta y confiable para la identificación automática de pólipos en imágenes de colonoscopia.
3. Los resultados evidencian que la optimización de hiperparámetros y la correcta selección de estrategias de entrenamiento resultan determinantes para mejorar la capacidad de generalización de los modelos. Asimismo, se resalta la importancia de mantener un equilibrio entre precisión y sensibilidad, dado que la reducción de falsos negativos es fundamental en aplicaciones médicas. En este sentido, los enfoques de transferencia de aprendizaje no solo ofrecen un rendimiento superior, sino que también representan una solución viable y clínicamente relevante en el contexto del diagnóstico asistido por inteligencia artificial.
4. La implementación de la aplicación web mediante Streamlit demuestra la viabilidad de trasladar un modelo de transferencia de aprendizaje con ajuste fino en el bloque cinco y capas densas entrenables a un entorno de validación interactivo. Esto no solo facilita la evaluación práctica en tiempo real de imágenes de colonoscopia, sino que también habilita una herramienta accesible para investigadores y profesionales de la salud, garantizando reproducibilidad, facilidad de uso y bajo costo computacional.
5. El prototipo desarrollado constituye un puente entre la investigación y la práctica clínica, al ofrecer una plataforma funcional que permite validar predicciones de manera inmediata y confiable. Además, su diseño modular y escalable abre la posibilidad de adaptación en entornos hospitalarios, consolidando así el potencial de la inteligencia artificial como apoyo en la detección temprana de pólipos y la prevención del cáncer colorrectal.

6. El desarrollo de este trabajo permitió la adquisición de un conocimiento sólido en deep learning, particularmente en el diseño, entrenamiento y validación de redes neuronales convolucionales aplicadas al análisis de imágenes médicas. La experiencia práctica en el uso de arquitecturas preentrenadas, técnicas de transferencia de aprendizaje y herramientas de despliegue como Streamlit favoreció la comprensión integral del proceso de construcción de modelos de inteligencia artificial, desde la preparación de datos hasta la implementación en un entorno interactivo.

8.2. Trabajos futuros

Como proyección de este trabajo, se plantean varias líneas de desarrollo que permitirían ampliar el alcance y la aplicabilidad del sistema:

1. Integración en entornos clínicos reales. Resulta fundamental someter la aplicación a pruebas piloto en instituciones de salud, con la participación de profesionales médicos, de manera que se pueda valorar tanto su impacto en la práctica clínica como su utilidad en la detección temprana de pólipos.
2. Despliegue en infraestructuras en la nube. La implementación en servidores remotos posibilitaría el acceso remoto, multiusuario y multiplataforma, facilitando así su adopción en escenarios de investigación colaborativa, entornos académicos y espacios hospitalarios.
3. Extensión hacia el análisis de secuencias de video. Una línea de avance especialmente relevante consiste en entrenar modelos que consideren no solo imágenes estáticas, sino también videos completos de colonoscopias, incorporando información temporal. Esto permitiría robustecer la detección de pólipos al capturar patrones dinámicos que no son perceptibles en una única imagen.

Bibliografía

- [1] A. G. Zauber, S. J. Winawer, M. J. O'Brien, I. Lansdorp-Vogelaar, M. van Ballegooijen, B. F. Hankey, W. Shi, J. H. Bond, M. Schapiro, J. F. Panish *et al.*, "Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths," *New England Journal of Medicine*, vol. 366, no. 8, pp. 687–696, 2012.
- [2] I. Barua, D. G. Vinsard, H. C. Jodal, M. Løberg, M. Kalager, Ø. Holme, M. Misawa, M. Bretthauer, and Y. Mori, "Artificial intelligence for polyp detection during colonoscopy: a systematic review and meta-analysis," *Endoscopy*, vol. 53, no. 03, pp. 277–284, 2021.
- [3] D. A. Lieberman, D. K. Rex, S. J. Winawer, F. M. Giardiello, D. A. Johnson, and T. R. Levin, "Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the us multi-society task force on colorectal cancer," *Gastroenterology*, vol. 143, no. 3, pp. 844–857, 2012.
- [4] N. Shussman and S. D. Wexner, "Colorectal polyps and polyposis syndromes," *Gastroenterology report*, vol. 2, no. 1, pp. 1–15, 2014.
- [5] M. A. Gómez-Zuleta, D. F. Cano-Rosales, D. F. Bravo-Higuera, J. A. Ruano-Balseca, and E. Romero-Castro, "Artificial intelligence techniques for the automatic detection of colorectal polyps," *Revista colombiana de Gastroenterología*, vol. 36, no. 1, pp. 7–16, 2021.
- [6] S. Kumar, N. Thosani, U. Ladabaum, S. Friedland, A. M. Chen, R. Kochar, and S. Banerjee, "Adenoma miss rates associated with a 3-minute versus 6-minute colonoscopy withdrawal time: a prospective, randomized trial," *Gastrointestinal endoscopy*, vol. 85, no. 6, pp. 1273–1280, 2017.
- [7] M. Spadaccini, J. Troya, K. Khalaf, A. Facciorusso, R. Maselli, A. Hann, and A. Repici, "Artificial intelligence-assisted colonoscopy and colorectal cancer screening: Where are we going?" *Digestive and Liver Disease*, 2024.
- [8] E. M. D. Castro, A. R. T. López, D. Aponte, J. N. R. Rodríguez, and L. C. Sabbagh, "Doble revisión de colon derecho vs revisión simple durante la colonoscopia para la detección de pólipos y adenomas de colon: revisión sistemática de la literatura," *Revista de Gastroenterología del Perú*, vol. 43, no. 4, pp. 309–318, 2023.
- [9] X. Ai, W. Qiao, Z. Han, W. Tan, Y. Bai, S. Liu, and F. Zhi, "Results of a second examination of the right side of the colon in screening and surveillance colonoscopy: a systematic review and meta-analysis," *European Journal of Gastroenterology & Hepatology*, vol. 30, no. 2, pp. 181–186, 2018.

- [10] A. O. Laiyemo, C. Doubeni, A. K. Sanderson II, P. F. Pinsky, D. S. Badurdeen, V. P. Doria-Rose, P. M. Marcus, R. E. Schoen, E. Lanza, A. Schatzkin *et al.*, “Likelihood of missed and recurrent adenomas in the proximal versus the distal colon,” *Gastrointestinal endoscopy*, vol. 74, no. 2, pp. 253–261, 2011.
- [11] O. Laudanno, B. Pucci, and S. Brayer, “Detección y vigilancia pospolipectomía de pólipos colónicos,” *Acta Gastroenterológica Latinoamericana*, vol. 52, no. 1, pp. 21–35, 2022.
- [12] J. Bernal, N. Tajkbaksh, F. J. Sanchez, B. J. Matuszewski, H. Chen, L. Yu, Q. Angermann, O. Romain, B. Rustad, I. Balasingham *et al.*, “Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge,” *IEEE transactions on medical imaging*, vol. 36, no. 6, pp. 1231–1249, 2017.
- [13] F. T. Bosman, F. Carneiro, R. H. Hruban, and N. D. Theise, *WHO classification of tumours of the digestive system.*, 2010, no. Ed. 4.
- [14] D. K. Rex, D. J. Ahnen, J. A. Baron, K. P. Batts, C. A. Burke, R. W. Burt, J. R. Goldblum, J. G. Guillem, C. J. Kahi, M. F. Kalady *et al.*, “Serrated lesions of the colorectum: review and recommendations from an expert panel,” *Official journal of the American College of Gastroenterology/ ACG*, vol. 107, no. 9, pp. 1315–1329, 2012.
- [15] R. D. Odze and J. R. Goldblum, *Odze and goldblum surgical pathology of the GI tract, liver, biliary tract and pancreas E-Book*. Elsevier Health Sciences, 2022.
- [16] J. Jass, “Classification of colorectal cancer based on correlation of clinical, morphological and molecular features,” *Histopathology*, vol. 50, no. 1, pp. 113–130, 2007.
- [17] C. Cunningham and I. Lindsey, “Colorectal cancer: management,” *Medicine*, vol. 35, no. 6, pp. 306–310, 2007.
- [18] E. R. Fearon and B. Vogelstein, “A genetic model for colorectal tumorigenesis,” *cell*, vol. 61, no. 5, pp. 759–767, 1990.
- [19] Global Cancer Observatory (GLOBOCAN). (2025) Globocan 2025: Estimated cancer incidence, mortality and prevalence worldwide. International Agency for Research on Cancer (IARC), World Health Organization. Consultado el 6 de octubre de 2025. [Online]. Available: <https://gco.iarc.fr/en>
- [20] I. P. Borrero and M. E. G. Arias, *Deep learning*. Servicio de Publicaciones de la Universidad de Huelva, 2021, vol. 19.
- [21] Z. Wu, F. Lv, C. Chen, A. Hao, and S. Li, “Colorectal polyp segmentation in the deep learning era: A comprehensive survey,” *arXiv preprint arXiv:2401.11734*, 2024.
- [22] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1, no. 2.

- [23] A. Mitsala, C. Tsalikidis, M. Pitiakoudis, C. Simopoulos, and A. K. Tsaroucha, “Artificial intelligence in colorectal cancer screening, diagnosis and treatment. a new era,” *Current Oncology*, vol. 28, no. 3, pp. 1581–1607, 2021.
- [24] E. S. Olivas, P. R. Belenguer, Q. G. Vidal, F. V. Estarich, J. V. Camisón, and J. V. Tomás, *Inteligencia artificial. Casos prácticos con aprendizaje profundo*. Ediciones de la U, 2022.
- [25] J. Wang and Y. Chen, *Introduction to transfer learning: algorithms and practice*. Springer Nature, 2023.
- [26] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.
- [27] S. Gross, S. Palm, J. J. Tischendorf, A. Behrens, C. Trautwein, and T. Aach, “Automated classification of colon polyps in endoscopic image data,” in *Medical Imaging 2012: Computer-Aided Diagnosis*, vol. 8315. SPIE, 2012, pp. 264–271.
- [28] B. Korbar, A. M. Olofson, A. P. Mirafior, C. M. Nicka, M. A. Suriawinata, L. Torresani, A. A. Suriawinata, and S. Hassanpour, “Deep learning for classification of colorectal polyps on whole-slide images,” *Journal of pathology informatics*, vol. 8, no. 1, p. 30, 2017.
- [29] T. Stehle, R. Auer, S. Gross, A. Behrens, J. Wulff, T. Aach, R. Winograd, C. Trautwein, and J. Tischendorf, “Classification of colon polyps in nbi endoscopy using vascularization features,” in *Medical Imaging 2009: Computer-Aided Diagnosis*, vol. 7260. SPIE, 2009, pp. 774–785.
- [30] R. Bisschops, C. Hassan, P. Bhandari, E. Coron, H. Neumann, O. Pech, L. Correale, and A. Repici, “Basic (bli adenoma serrated international classification) classification for colorectal polyp characterization with blue light imaging,” *Endoscopy*, vol. 50, no. 03, pp. 211–220, 2018.
- [31] T. Ozawa, S. Ishihara, M. Fujishiro, Y. Kumagai, S. Shichijo, and T. Tada, “Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks,” *Therapeutic advances in gastroenterology*, vol. 13, p. 1756284820910659, 2020.
- [32] S. Ali, D. Jha, N. Ghatwary, S. Realdon, R. Cannizzaro, O. E. Salem, D. Lamarque, C. Daul, M. A. Riegler, K. V. Anonsen *et al.*, “A multi-centre polyp detection and segmentation dataset for generalisability assessment,” *Scientific Data*, vol. 10, no. 1, p. 75, 2023.
- [33] K. Pogorelov, K. R. Randel, C. Griwodz, S. L. Eskeland, T. de Lange, D. Johansen, C. Spanpinato, D.-T. Dang-Nguyen, M. Lux, P. T. Schmidt *et al.*, “Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection,” in *Proceedings of the 8th ACM on Multimedia Systems Conference*, 2017, pp. 164–169.