

# Segmentación morfológica y clasificación de niveles para la retinopatía diabética e hipertensiva mediante imágenes oftálmicas y redes convolucionales

## Morphological segmentation and level classification for diabetic and hypertensive retinopathy through ophthalmic images and convolutional networks

María C. Niño<sup>1</sup>  Darwin E. Martínez<sup>2</sup> 

<sup>1</sup>Universidad Sergio Arboleda, Escuela de Ciencias Exactas e Ingeniería. Bogotá, Colombia

### Resumen

El objetivo principal de esta investigación es realizar la segmentación y clasificación de imágenes de fondo de retina con retinopatía diabética e hipertensiva. Se propuso una combinación de una red convolucional UNet y una ConvNet para la segmentación de máscara de vasos y la clasificación de retinopatía, respectivamente. El proceso de clasificación se basa en diez clases definidas, donde los valores que van del 0 al 4 representan la retinopatía diabética y los valores del 5 al 9 corresponden a la retinopatía hipertensiva. Los resultados aproximados en la segmentación fueron índices Jaccard de 74%, F1 de 85% y un Accuracy de 96%, y en la clasificación un Accuracy de 80%.

### Abstract

The main objective of this research is to perform segmentation and classification of diabetic and hypertensive retinopathy retinal fundus images. Using a combination of a convolutional network UNet and a ConvNet was proposed for vessel mask segmentation and retinopathy classification, respectively. The classification process relies on ten defined classes, where values ranging from 0 to 4 represent diabetic retinopathy, and values ranging from 5 to 9 correspond to hypertensive retinopathy. The approximate results in the segmentation were Jaccard index of 74%, F1 of 85%, and an Accuracy of 96%, and in the classification was Accuracy of 80%.

**Palabras clave:** Retinopatía diabética e hipertensiva, Visión por computador, Redes Neuronales Convolucionales.

**Keywords:** Diabetic and hypertensive retinopathy, Computer vision, Convolutional Neural Networks.

### ¿Cómo citar?

Niño, M.C., Martínez, D.E. Morphological Segmentation and Level Classification for Diabetic and Hypertensive Retinopathy through Ophthalmic Images and Convolutional Networks. Ingeniería y Competitividad. 2024, 26(1) e-20513181

<https://doi.org/10.25100/iyv.v26i1.13181>

Recibido: 23-08-23

Publicado: 02-02-24

### Correspondencia:

maria.nino02@usa.edu.co

Este trabajo está licenciado bajo una licencia internacional Creative Commons Reconocimiento-No Comercial-CompartirIgual4.0.



Conflicto de intereses:

Ninguno declarado



### ¿Por qué se llevó a cabo?

El estudio se realizó para abordar la difícil tarea de la detección y clasificación de la retinopatía diabética e hipertensiva. El objetivo era desarrollar un método que utilizara técnicas de aprendizaje profundo para segmentar imágenes del fondo de ojo de la retina y clasificarlas en etapas específicas de la retinopatía diabética e hipertensiva.

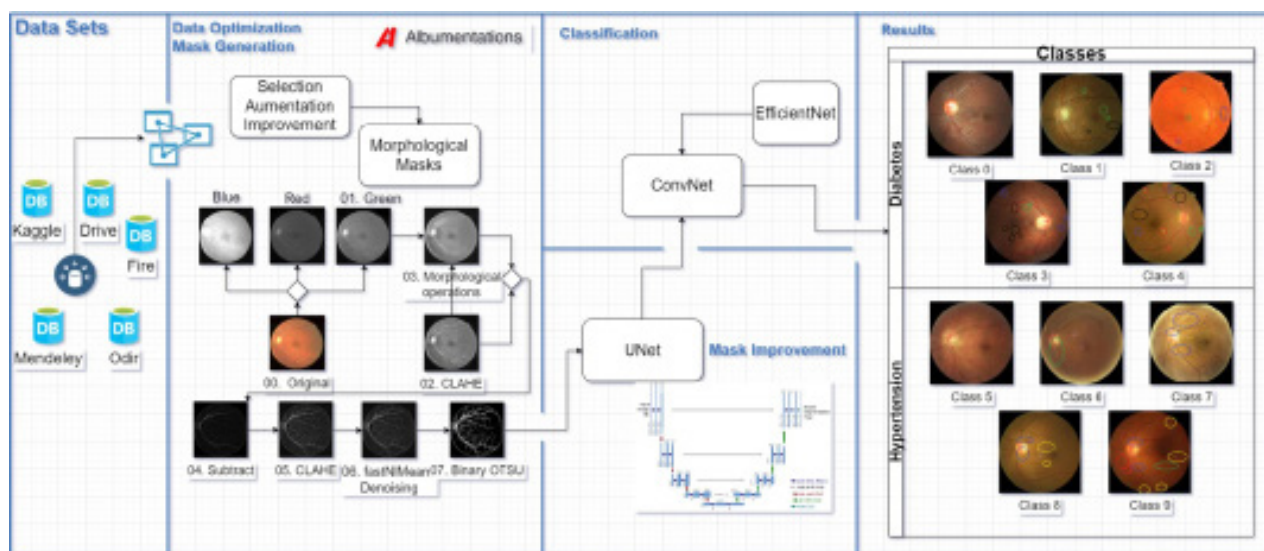
### ¿Cuáles fueron los resultados más relevantes?

El estudio propuso un método que utiliza una combinación de una red convolucional UNet para la segmentación de máscaras de vasos y una ConvNet para la clasificación de la retinopatía. El proceso de segmentación logró un índice Jaccard del 74 %, una puntuación F1 del 85 % y una precisión del 96 %. El proceso de clasificación alcanzó una precisión del 80%. La integración de las arquitecturas UNet y ConvNet mostró resultados prometedores en la segmentación y clasificación de imágenes de retinopatía diabética e hipertensiva. Los resultados demostraron tasas de precisión aproximadas del 90 % y una precisión del 80 % para la creación de máscaras morfológicas. Se utilizaron diferentes conjuntos de datos, separados por retinopatía diabética e hipertensiva, generando la máscara a través de procesos morfológicos. La utilización del modelo previamente entrenado EfficientNet mostró una mayor eficiencia y precisión en el análisis de imágenes de retinopatía.

### ¿Qué aportan estos resultados?

El estudio proporciona un método sólido para la segmentación y clasificación de imágenes de retinopatía diabética e hipertensiva, crucial para la detección y el tratamiento tempranos. La integración de las arquitecturas UNet y ConvNet ofrece un enfoque integral, combinando las fortalezas de ambas redes. La creación de máscaras morfológicas personalizadas para el conjunto de datos de retinopatía diabética aborda la variabilidad en las imágenes de retinopatía diabética, mejorando el proceso de segmentación. El estudio demuestra el potencial de las técnicas de aprendizaje profundo en el análisis de imágenes médicas, específicamente para la detección de retinopatía, lo que podría conducir a mejores resultados para los pacientes y procesos de diagnóstico más eficientes. Las investigaciones futuras pueden aprovechar estos resultados explorando mayores mejoras en el proceso de clasificación, incluido el uso de técnicas de aprendizaje por transferencia y estrategias avanzadas de aumento de datos.

### Graphical Abstract



## Introducción

Las enfermedades crónicas o de largo plazo empeoran con el tiempo, creando a menudo la necesidad de tratamiento y cuidado a largo plazo [1]. Según la Organización Panamericana de la Salud estas enfermedades causan la muerte de 41 millones de personas cada año, lo que equivale al 71% de las muertes que ocurren en el mundo. En la Región de las Américas, hay 5.5 millones de muertes por estas enfermedades cada año [1]. Así mismo, resulta preocupante el hecho de que en la región de las Américas se reporten aproximadamente 1.6 millones de muertes anuales debido a enfermedades cardiovasculares. Actualmente, la diabetes y la hipertensión son las enfermedades con mayor prevalencia entre la población. Estas condiciones dan lugar a complicaciones significativas, que afectan especialmente al sistema vascular, ya que causan graves daños a órganos como riñones, corazón, cerebro y ojos [2]. En el caso del ojo, existe un riesgo significativo de daño a la vasculatura retiniana debido a la presión aumentada, lo que eventualmente puede llevar al desarrollo de retinopatía. La Tabla 1 muestra una breve descripción de las principales partes del ojo y las consecuencias para la salud debido a la retinopatía diabética e hipertensiva respectivamente. Muchas de estas alteraciones ocurren en diferentes áreas o partes del ojo, a menudo identificadas como cuadrantes. Estas áreas se delimitan mediante dos líneas: una vertical que divide el ojo en mitades nasal y temporal, y otra horizontal que separa la parte superior de la inferior [3].

Tabla 1 Descripción de las principales partes del ojo y sus efectos debido a la retinopatía diabética e hipertensiva. Los síntomas que están relacionados con la retinopatía diabética se presentarán en cursiva, los síntomas de la retinopatía hipertensiva se presentarán subrayados y aquellos que aplican a ambas condiciones se presentarán en negrita.

Órgano	Definición	Impacto de la Retinopatía
Córnea	Estructura compleja, fibrosa, transparente y delgada, refracta la luz y junto con el cristalino producen una imagen invertida reducida [4].	Cambios morfológicos y funcionales de la córnea [5]
Canal Hialoide	Es un área central transparente del ojo que cruza el vítreo desde la papila o disco óptico hasta el cristalino [4].	En la etapa proliferativa, hay una mayor tensión en esta parte debido a la degeneración macular [6]. <u>Esclerosis reactiva debido a la acumulación de material hialino</u> [7].
Retina	Capa fotosensible que recubre la parte interna de la parte posterior del globo ocular[8]. Está compuesta por la papila o disco óptico, la mácula y los vasos sanguíneos retinales [9].	En la etapa proliferativa, se demostraron niveles elevados de homocisteína en la retina [10]. <u>Vasoconstricción difusa o focal, extravasación debido al aumento de la permeabilidad vascular</u> [7].

Mácula	Mientras que el resto de la retina proporciona visión periférica, la mácula permite al ojo humano ver en detalle [9].	Ocurren hemorragias, produciendo exudación con edema retiniano o exudados duros [7]. El principio de la No Proliferativa se caracteriza por un aumento de la permeabilidad vascular, y en el edema macular diabético de PDR [10].
Nervio Óptico	Es la primera sección de la vía que va desde el ojo hasta el cerebro. Es un área aproximadamente circular en el centro de la retina [9].	En la etapa temprana de la Diabetes con una morfología no afectada que sugiere degeneración del nervio óptico [11]. <u>Disco pálido, borrosidad de los márgenes papilares y cavernosos</u> [7].
Vasos Sanguíneos	Se encuentran en las capas superiores y supervisan la alimentación de la retina junto con los vasos sanguíneos coroideos [9].	Estas alteraciones en los vasos pueden llevar a la isquemia del tejido retiniano [12].

Actualmente, las imágenes de fondo de ojo se utilizan para detectar y evaluar enfermedades oculares, como el glaucoma y la retinopatía; algunos autores la consideran la mejor técnica para estudiar el interior del ojo incluyendo la retina, el nervio óptico, la mácula y el polo posterior [13]. En la Figura 1, una fotografía muestra un ejemplo de la retina de un paciente sano, resaltando tres estructuras oculares distintas: la papila o disco óptico, la mácula y los vasos sanguíneos retinianos [9]. Los oftalmólogos observan y toman imágenes del fondo de ojo para examinar posibles lesiones o anomalías [9]. Para diagnosticar tanto la retinopatía diabética como la hipertensiva, el oftalmólogo realiza un examen inicial que implica una evaluación integral de los ojos [13].

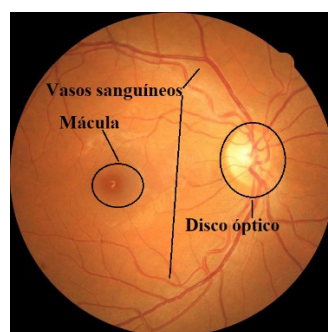


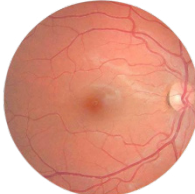
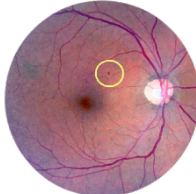
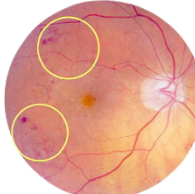
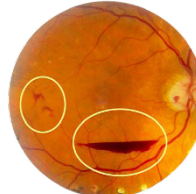
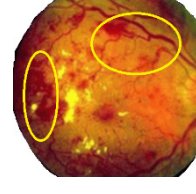
Figura 1 La imagen retiniana del ojo derecho de un paciente sano resalta el disco óptico, la mácula y los vasos sanguíneos, elementos principales para el diagnóstico de la retinopatía.

### Retinopatía Diabética (DR)

Esta enfermedad se considera grave y es la principal causa de discapacidad visual asociada a la diabetes mellitus. Se trata de la complicación microvascular más prevalente de esta enfermedad, con el potencial de ocasionar severas afectaciones en la retina, lo que eventualmente conduce a la pérdida de visión [14]. La Retinopatía Diabética (DR) afecta los vasos sanguíneos en la retina, el tejido fotosensible ubicado en el fondo del ojo. Esta condición se manifiesta en dos etapas conocidas como Retinopatía Diabética No

Proliferativa (NPDR) y Retinopatía Diabética Proliferativa (PDR). En la etapa proliferativa, se observan cambios significativos en el grosor de la retina y en el área macular, lo que puede dar lugar a la formación de nuevos vasos sanguíneos anormales. Estos vasos anormales, a su vez, pueden provocar hemorragias oculares y aumentar el riesgo de desprendimiento de la retina, y por ende puede ocasionar la pérdida de la visión. Por otro lado, en la etapa no proliferativa, es frecuente la presencia de hinchazón en la retina debido al daño en los vasos sanguíneos más pequeños. En casos más graves, puede desarrollarse edema macular, que consiste en la acumulación de líquido en la mácula, la parte central de la retina responsable de la visión detallada. Este edema macular, si no se controla adecuadamente, puede progresar hacia la etapa proliferativa [15]. La Tabla 2 describe las características clave de esta enfermedad y sus clases:

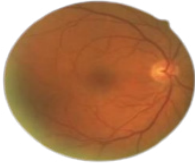
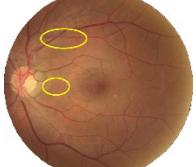
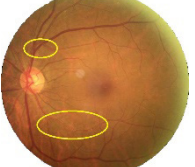
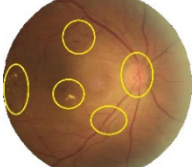
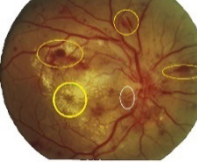
Tabla 2 Principales características de la retinopatía diabética en las diferentes clases, acompañadas por imágenes de esta clasificación; las características principales de cada nivel están encerradas en un círculo amarillo.

Clase 0	Clase 1	Clase 2	Clase 3	Clase 4
				
No se evidencia ninguna anomalía. Menos de un 1% de probabilidad de desarrollar PDR [16].	Medía NPDR, microaneurismas como la única anomalía encontrada[16].	La NPDR moderada contiene pequeñas hemorragias o microaneurismas [16].	NPDR severa. 20 o más hemorragias en los cuatro cuadrantes [16].	PDR: Etapa más avanzada. Se desarrollan nuevos vasos sanguíneos frágiles y anormales [16].

### Retinopatía Hipertensiva (HR)

La retinopatía hipertensiva (HR) es un daño vascular retiniano causado por la hipertensión, especialmente en pacientes con niveles moderados de esta condición. El examen de la imagen del fondo de ojo permite identificar estos cambios morfológicos, tales como microaneurismas, exudados duros amarillos, exudados blandos (manchas de algodón), hemorragias, mácula, disco óptico, cabeza del nervio óptico y un aumento en el número de vasos sanguíneos, como la constricción arteriolar o la esclerosis [17]. La Tabla 3 describe las características principales de esta enfermedad y sus clases:

Tabla 3 Características principales de la retinopatía hipertensiva en las diferentes clases, acompañadas de imágenes de esta clasificación. Las características principales de cada nivel están encerradas en un círculo amarillo.

Clase 0	Clase 1	Clase 2	Clase 3	Clase 4
				
Sin signos visibles de retinopatía hipertensiva [18].	Hipertensión benigna. Esclerosis a un nivel moderado de las arteriolas retinianas [19].	Esclerosis a un nivel moderado de las arteriolas retinianas. Nivel de presión arterial más elevado en comparación con la clase 1 [19].	Manchas algodonosas y hemorragias. También síntomas de lesiones espásticas de las arteriolas [19].	Hipertensión maligna. Todos los síntomas anteriores están presentes, también de edema macular [19].

### Estado del Arte

A lo largo de los años, los investigadores han llevado a cabo numerosos estudios sobre la retinopatía diabética; sin embargo, ha habido relativamente poca investigación sobre la retinopatía hipertensiva [20]. En este último caso, la mayoría de los estudios se han centrado en la segmentación de los vasos sanguíneos. Chala et al. proponen el uso de redes neuronales convolucionales profundas para la segmentación de los vasos sanguíneos retinianos, inspirados en arquitecturas populares como FCN, U-net, ya que son las arquitecturas más utilizadas en la imagen médica, con una estructura multi-encoder-decoder y tres unidades principales: un codificador RGB, un codificador verde y un decodificador [21]. Mientras que el primer método emplea una UNet, Zhang et al. emplean TransUNet porque combina las fortalezas de la convolución y los mecanismos de atención multi-cabeza para capturar características locales y dependencias globales. Incorpora la extracción de puntos de referencia anatómicos auto-supervisados a través del aprendizaje contrastivo para enfatizar el aprendizaje de la morfología de los vasos retinianos [22]. En contraste, NA et al. emplean una versión modificada de la popular arquitectura U-NET junto con algoritmos de preprocesamiento propuestos para segmentar las imágenes de entrada; el modelo trabaja en parches extraídos de imágenes obtenidas después de la etapa de preprocesamiento [23].

Algunas de las redes utilizadas en la clasificación de la retinopatía incluyen: Abbast et al. propusieron el uso de la red DenseNet para detectar cinco etapas de la retinopatía hipertensiva basadas en una arquitectura de capas densas semánticas e instanciadas y una estrategia de pre-entrenamiento [18]. Adicionalmente, Nguyen et al. emplean un conjunto de modelos de redes neuronales convolucionales (CNN) EOPSO-CNN basado en el algoritmo de optimización de enjambre de partículas ortogonales (OPSO) para realizar la detección y clasificación de la retinopatía diabética [24]. Asimismo, Takahashi et al. utilizaron una red neuronal profunda GoogLeNet completamente aleatorizada modificada, utilizando un área retiniana que no se visualiza habitualmente en la oftalmoscopia y otro IA que sugiere directamente tratamientos y determina pronósticos para la clasificación de la retinopatía diabética [25]. Por otra parte, Dutta et al. proponen un modelo de conocimiento automatizado para identificar los principales antecedentes de la retinopatía diabética. El

modelo ha sido entrenado con tres tipos: retropropagación NN, Red Neuronal Profunda (DNN) y Red Neuronal Convolutiva (CNN) [26].

### Conjunto de datos

Los conjuntos de datos consisten en diferentes fuentes disponibles en Kaggle, Drive, ChaseDB1 y Mendeley. Se realizó un preprocesamiento en el conjunto de datos para mejorarlo, ya que las imágenes tenían diferentes formatos y algunas contenían un marco negro grueso. Posteriormente, se utilizaron diversas técnicas de visión por computadora y aprendizaje profundo para realizar la segmentación y clasificación de las diferentes etapas de las enfermedades.

La Tabla 4 muestra las diferentes fuentes de información, incluyendo la cantidad total, la división en conjuntos de entrenamiento y validación, el formato y el tipo de archivo asociado con cada conjunto, al igual que sus etiquetas correspondientes. Hay dos tipos de etiquetas utilizadas en el conjunto de datos. El primer tipo es un archivo CSV empleado en el conjunto de datos de Kaggle, que proporciona información sobre la clase de enfermedad y especifica si corresponde al ojo derecho o izquierdo. En el caso del conjunto de datos de Odir, también se incluyen la edad y el género del paciente. El segundo tipo de etiqueta representa los vasos sanguíneos en formato de imagen, específicamente PNG.

Tabla 4 Descripción del conjunto de datos de las diferentes fuentes utilizadas en la investigación, resaltando la cantidad, la división en los diferentes conjuntos y el formato de las imágenes

Tipo	Fuente	Total				
Diabetes	Kaggle [27]	88.702	35.126	53.576	JPEG	CSV
	Drive [28]	40	20	20	TIF	
	Fire [29]	268	134	134	JPG	
	Mendeley [30]	100	50	50	JPG	
	Odir [31]	8000	7000	1000	JPG	CSV

Para hacer frente a los desequilibrios inherentes en el conjunto de datos, particularmente la representación disminuida de imágenes de retinopatía hipertensiva y una distribución desigual entre las clases, con un sesgo pronunciado hacia datos saludables, fue necesario implementar un proceso de selección cuidadoso. Esto implicó tanto la reducción en el número de imágenes de diabetes como la aplicación de aumentación de datos, especialmente en las clases de enfermedades avanzadas. La Tabla 5 muestra el conjunto de datos refinado.

Tabla 5 Descripción de la distribución de datos utilizada en los diferentes conjuntos después de la selección y aumento de las imágenes para equilibrar el número total de imágenes a utilizar.

Conjunto	Cantidad	Diabetes	Hipertensión
Entrenamiento	23.900	11.902	11.998
Validación	5.996	2.998	2.998
Pruebas	7.672	3.836	3.836
Total	37.568	18.736	18.832

## Preprocesamiento de imágenes

El preprocesamiento de imágenes implica la reducción del ruido y el ajuste de la forma de las imágenes antes de su análisis. Por esa razón, se implementaron diversas técnicas para optimizar los conjuntos de datos como la Ecuación de Histograma Adaptativa Limitada al Contraste (CLAHE, por sus siglas en inglés), que es una variante de la Ecuación de Histograma Adaptativa (AHE) [32]. Opera en pequeñas regiones de la imagen, llamadas "mosaicos", en lugar de en toda la imagen [32]. Además, se implementó un recorte adicional de bordes y una mejora en el redondeo de algunas imágenes debido a que presentaban un formato más alargado. Asimismo, se realizó un redimensionamiento a diferentes tamaños cuadrados con el fin de evaluar cuál de ellos permitía una extracción de características más efectiva en las redes neuronales convolucionales utilizadas.

## Metodología

En el ámbito del análisis de imágenes médicas, la fusión de arquitecturas avanzadas de aprendizaje profundo ha demostrado ser una estrategia poderosa para el diagnóstico integral de enfermedades [33]. Este estudio aprovecha la sinergia entre una U-Net para la segmentación precisa de los vasos sanguíneos del ojo y una Red Neuronal Convolucional (ConvNet) con la arquitectura de EfficientNet, conocida por su efectividad. El objetivo es abordar simultáneamente la retinopatía diabética y la retinopatía hipertensiva, distinguiendo entre sus cuatro niveles de gravedad. Este enfoque integrado ofrece varias ventajas. En primer lugar, la precisión de segmentación lograda por la U-Net mejora la extracción de características, permitiendo que la ConvNet realice clasificaciones más informadas y contextualmente conscientes. En segundo lugar, la arquitectura eficiente de EfficientNet optimiza los recursos computacionales, lo que permite un entrenamiento y una inferencia más rápidos y eficientes en términos de recursos. Finalmente, los resultados se generan en un archivo CSV donde se relaciona la imagen a evaluar con el nivel de la enfermedad, siendo del 0 al 4 los niveles correspondientes a la retinopatía diabética y del 5 al 9 la retinopatía hipertensiva. La Figura 2 muestra el proceso de fusión de las dos redes y la integración de las máscaras morfológicas.



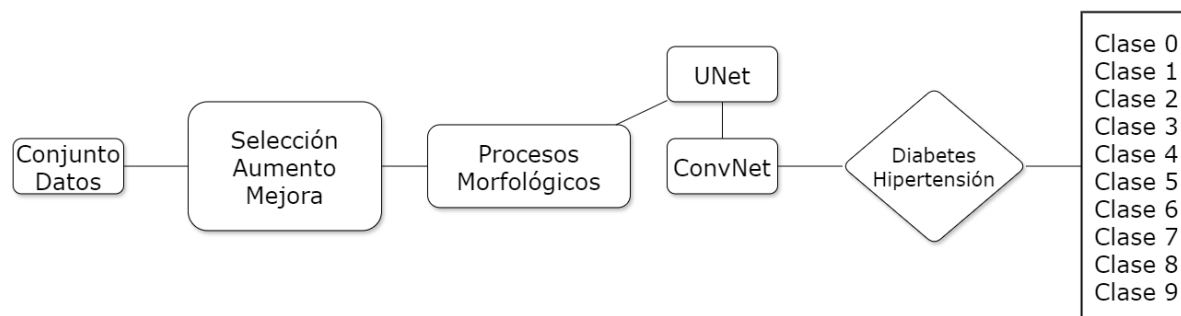


Figura 2 El diagrama ilustra un proceso secuencial para clasificar la retinopatía diabética e hipertensiva. Comienza con la mejora de la imagen, seguida por la creación de máscaras morfológicas y la segmentación de los vasos sanguíneos utilizando el modelo U-Net.

Esta red presenta una estructura codificador-decodificador compuesta por dos partes [34]. La diferencia con la estructura codificador-decodificador tradicional es que hay conexiones que copian los datos que se están procesando en el codificador al decodificador para facilitar la generación de la máscara [34]. La Figura 3 presenta la arquitectura U-Net propuesta por Ronneberger et al. [35] para vasos sanguíneos [35].

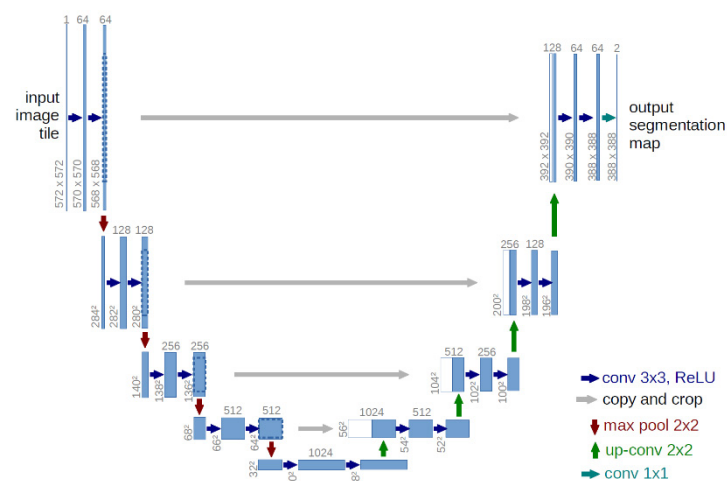


Figura 3 Arquitectura U-Net empleada por Nikhil Tomar para la segmentación de vasos sanguíneos en retinopatía hipertensiva. La flecha roja que señala hacia abajo representa el proceso de max pooling, que reduce a la mitad el tamaño de la imagen. La flecha verde es una convolución transpuesta, una técnica de aumento de muestreo que amplía el tamaño de las imágenes. Cada recuadro azul corresponde a un mapa de características multi-canal. El número de canales se indica en la parte superior del recuadro. El tamaño x-y se proporciona en el borde inferior izquierdo del recuadro. Los recuadros blancos representan mapas de características copiados. Las flechas indican las diferentes operaciones [35].

La investigación también se basó en el código desarrollado por Nikhil Tomar [36]. La primera parte de este desarrollo en la etapa de preprocesamiento, utilizando un aumento de datos con la biblioteca de código abierto "Albumentations", la segunda etapa de esta red está compuesta por cuatro capas de codificador. Una capa de cuello de botella, que contiene pocos nodos en comparación con las capas anteriores. La idea es que esta vista comprimida solo contenga la información "útil" para poder reconstruir los vasos [37].

Finalmente, las cuatro capas de decodificador reconstruyen la imagen original, duplicando el tamaño entre cada capa. La validación del modelo se realiza con el índice de Jaccard, el puntaje f1, recall, precisión y las funciones de accuracy.

De igual manera, la validación del proceso se lleva a cabo utilizando las funciones "Dice Loss"; este enfoque aborda eficazmente el desequilibrio entre el primer plano y el fondo [38]. "Dice BCE Loss" calcula probabilidades y compara cada salida de clase real basada en la pérdida de distribución de Bernoulli, la cual compara cada entrada y salida con probabilidades predichas, que pueden ser 0 o 1 [39].

### Redes neuronales convolucionales para el reconocimiento visual onvNet

Durante la última década, el campo del reconocimiento visual ha logrado una transición exitosa de las funciones de ingeniería al diseño de arquitecturas (ConvNet) [40]. ConvNets representativos como VGGNet, Inception, ResNe(X)t, DenseNet, MobileNet, EfficientNet y RegNet se centraron en diferentes aspectos de precisión y rendimiento, y popularizaron muchos principios de diseño útiles [40]. Las ConvNets tienen varios sesgos inductivos integrados que las hacen apropiadas para una amplia variedad de aplicaciones de visión por computadora [40]. La Figura 4 muestra la red convolucional propuesta por Person para ejecutar la clasificación de la retinopatía diabética [41].

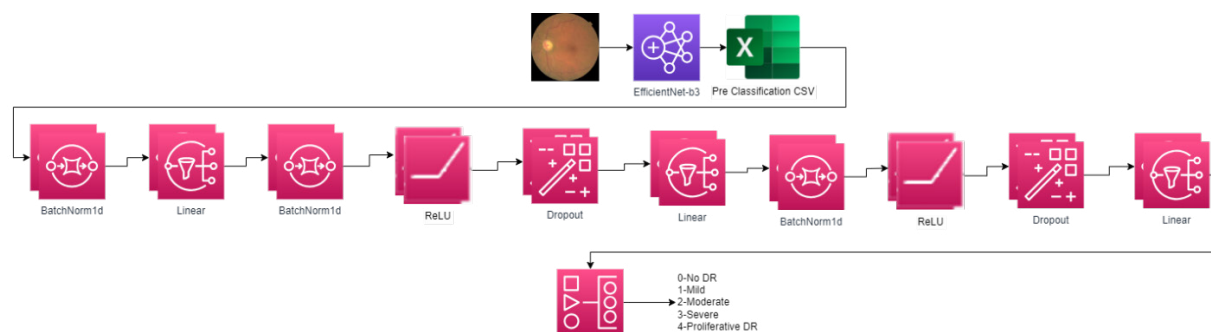


Figura 4 Arquitectura de redes convolucionales para la clasificación de la retinopatía diabética propuesta por Aladdin Person. Fuente de la imagen: elaboración propia a partir de la arquitectura original en [41].

La arquitectura de la red está compuesta por el modelo de transferencia de aprendizaje EfficientNet porque es más eficiente que la mayoría de los otros modelos de CNN [42]. Hay ocho modelos de esta arquitectura, cada uno con siete bloques. Estos bloques también tienen un número variable de sub-bloques cuyo número aumenta desde EfficientNetB0 hasta EfficientNetB7. El número total de capas en EfficientNet-B0 es de 237, en EfficientNet-B7 es de 813 [43]. El modelo propuesto por Aladdin Person utiliza el modelo EfficientNet-B3, el cual utiliza un método de escalado compuesto que crea diferentes modelos en la red neuronal convolucional [44]. En la mayoría de las exploraciones, se observan mejores niveles de precisión al usar los modelos superiores EfficientNetB7; no obstante, para la investigación, se obtuvieron los mejores resultados al usar el modelo B0 que incluye 4,049,564 parámetros [45]. La Tabla 6 presenta las capas, canales y arquitectura del modelo EfficientNet-B3 y B0, donde IRC representa la conexión residual invertida utilizada por el bloque MBConv. En las redes neuronales convolucionales (CNN) tradicionales, el número de canales generalmente aumenta a medida que la red se hace más profunda, lo que, a su vez, puede resultar en un aumento significativo en los requisitos de computación y memoria [46]. Para abordar

este problema, se introduce el bloque IRC en EfficientNet. Al utilizar el bloque IRC, se logra un equilibrio entre la profundidad, el ancho y la resolución del modelo, optimizando el compromiso entre el tamaño del modelo y la precisión. La conexión de salto dentro del bloque IRC mejora el flujo de gradiente y ayuda a entrenar redes más profundas de manera más efectiva. Un Bloque Residual Invertido, a veces llamado Bloque MBConv, es un tipo de bloque residual utilizado para modelos de imagen que utiliza una estructura invertida por razones de eficiencia.

Utiliza los mismos módulos MBConv1 y MBConv6 que EfficientNetB0 [47]. Fue propuesto originalmente para la arquitectura CNN MobileNetV2 [48]. La etapa representa el diseño típico en una red ConvNet, donde las capas a menudo se dividen en varias etapas [49].

Tabla 6 Descripción de la arquitectura EfficientNet-B3 y B0 resaltando las capas, la resolución y el número de canales.

Escenario i	EfficientNetB3				EfficientNetB0			
	Operador $\hat{F}_i$	Resolución $\hat{H}_i \times \hat{W}_i$	#Canales $\hat{C}_i$	#Capas $\hat{L}_i$	Operador $\hat{F}_i$	Resolución $\hat{H}_i \times \hat{W}_i$	#Canales $\hat{C}_i$	#Capas $\hat{L}_i$
1	Conv3x3+BN	300 x 300	40	1	Conv3x3	224 x 224	32	1
2	MBConv1, k3x3	112 x 112	24	2	MBConv1, k3x3	112 x 112	16	1
3	MBConv6, k3x3	112 x 112	32	1	MBConv6, k3x3	112 x 112	24	2
4	MBConv6, k3x3, IRC	112 x 112	32	2	MBConv6, k5x5	56 x 56	40	2
5	MBConv6, k5x5	56 x 56	48	1	MBConv6, k3x3	28 x 28	80	3
6	MBConv6, k5x5, IRC	28 x 28	48	2	MBConv6, k5x5	14 x 14	112	3
7	MBConv6, k3x3	28 x 28	96	1	MBConv6, k5x5	14 x 14	192	4
8	MBConv6, k3x3, IRC	14 x 14	96	4	MBConv6, k3x3	7 x 7	320	1
9	MBConv6, k5x5	14 x 14	136	1	Conv1x1 & Pooling & FC	7 x 7	1280	1
10	MBConv6, k5x5, IRC	14 x 14	136	3				
11	MBConv6, k5x5	7 x 7	332	1				
12	MBConv6, k5x5, IRC	7 x 7	332	4				
13	MBConv6, k3x3	7 x 7	384	2				
14	Conv1x1 & Pooling & FC	7 x 7	1536	1				

Después del uso del modelo EfficientNet, el uso de una capa "Lineal" o completamente conectada ayuda a cambiar la dimensionalidad de la salida de la capa anterior para que el modelo pueda definir fácilmente la relación entre los valores de los datos en los que trabaja el modelo [50]. La capa lineal es el bloque de construcción central de casi todos los modelos de redes neuronales modernas, incluso las redes convolucionales típicamente alimentan una o más capas lineales después del procesamiento por convoluciones [51].

La función de la segunda red convolucional es actualizar los pesos de las capas finales del modelo pre-entrenado de la primera red, esta arquitectura incluye 10 capas adicionales, con una específicamente diseñada para normalización llamada "BatchNorm1d". La función principal de BatchNorm1d es estimar y normalizar la media y la varianza para cada pequeño lote de entradas procesadas por la capa. Al realizar esta normalización, ayuda a garantizar la estabilidad y consistencia de la entrada en toda la red. La ventaja de usarlo es que estabiliza el proceso de aprendizaje y reduce la cantidad de veces que son necesarias para el entrenamiento [52]. Después de cada capa convolucional, es habitual aplicar una capa de activación no lineal. La capa más popular en las redes convolucionales es ReLU, en comparación con la más tradicional sigmoide o Tanh, ya que es computacionalmente más eficiente [42]. La capa de abandono agrega ceros aleatorios a algunos elementos del tensor de entrada utilizando muestras de una distribución de Bernoulli. Este tipo de red ha demostrado ser una forma efectiva de regularizar y prevenir la adaptación de neuronas [53].

### Entrenamiento de la Red

Para ejecutar la red de segmentación Unet, se utilizaron las siguientes transformaciones en las imágenes:

Cambios de posición con las funciones "HorizontalFlip" y "VerticalFlip".

Distorsión de imagen: mediante la función "ElasticTransform", para el análisis de imágenes biomédicas, junto con la "GridDistortion"[54].

Mejora de imagen: a través de la función "Optical Distortion" y "Rotate" [42].

Después de este proceso, se ejecuta Unet, entrenando la información y creando la máscara que ConvNet va a utilizar en el proceso de clasificación.

### Cambios a las redes originales

Actualmente se utilizan muchos métodos de segmentación para extraer la región de interés, tales como el umbralizado, basado en regiones, computación suave, histograma y redes neuronales [59]. El proceso utilizado para la generación de la extracción inicial del árbol de vasos sanguíneos se presenta en la Figura 5; de igual manera, se presenta un ejemplo de estos pasos en la Figura 9. El Paso 00 representa la imagen inicial después de haber pasado por los pasos de preprocesamiento necesarios. El Paso 01 muestra la imagen dividida en tres canales RGB. El canal verde es específicamente seleccionado para un procesamiento adicional debido a su contraste superior entre los vasos y el fondo en comparación con los otros dos canales de color en imágenes de retina [60]. En el Paso 02, se aplica el contraste "CLAHE" en la imagen del canal verde, una técnica utilizada para mejorar el contraste de una imagen limitando la amplificación del ruido [32]. Se utiliza un ajuste de contraste con un límite de dos y un kernel de 11x11 píxeles. El objetivo en este paso es mejorar el contraste y la visibilidad de los vasos sanguíneos. El Paso 03, involucra la aplicación de técnicas morfológicas, específicamente erosión y dilatación, como operaciones fundamentales [61]. Comenzando

con una operación de dilatación para preservar los detalles más finos, utilizando un elemento estructurante de forma elíptica con un tamaño de 5x5 píxeles. Una primera operación de apertura ayuda a eliminar regiones de ruido pequeñas y suaviza la imagen, luego una operación de cierre utilizando un elemento de tamaño 7x7 píxeles. El cierre morfológico combina una operación de dilatación y otra de erosión [61]. Después de esto, otra operación de apertura, pero utilizando un elemento estructurante más grande de tamaño 9x9 píxeles. La siguiente operación es otro cierre, esta vez utilizando un elemento estructurante de 11x11 píxeles. Esto ayuda a cerrar brechas más grandes y suavizar aún más la imagen. En el Paso 04 se resta la imagen generada al final del proceso de morfología con la imagen mejorada en el canal verde. En el Paso 05 se aplica nuevamente CLAHE al resultado anterior para mejorar los detalles en los vasos sanguíneos. En el Paso 06 se utiliza "fastNIMeansDenoising" para limpiar el ruido. El algoritmo Fast Non-Local Means (FNLN) bajo consideración abarca múltiples parámetros. Dentro de estos parámetros, el factor de suavizado juega un papel fundamental en la definición del alcance del filtrado. Es importante tener en cuenta que este alcance de filtrado afecta directamente la nitidez y claridad de las imágenes [62]. Finalmente, en el Paso 07 se realiza una umbralización en la imagen sin ruido utilizando la técnica "THRESH BINARY" y "THRESH\_OTSU". Esto transforma la imagen en una representación binaria, asignando valores de píxeles por debajo de cierto umbral a 0 y valores por encima del umbral a 255.

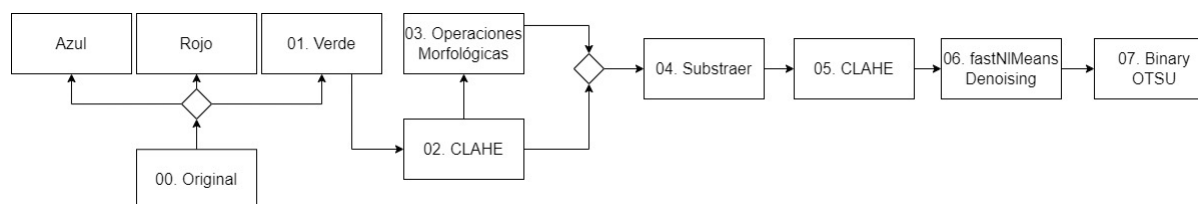


Figura 5 Descripción de los pasos para extraer las características del árbol vascular utilizando diversas técnicas de visión por computadora, las cuales luego se emplean para crear máscaras en la red UNet.

Adicionalmente, se amplió la clasificación para incluir 10 categorías distintas. Esto implicó categorizar los datos etiquetados en grupos específicos: los niveles 0-4 representan retinopatía diabética, y los niveles 5-9 representan retinopatía hipertensiva.

## Resultados

Los primeros resultados corresponden a la fase de preprocesamiento, que abarca el mejoramiento de la imagen y la creación de máscaras morfológicas. Estos pasos fundamentales sientan las bases para la posterior segmentación y clasificación de enfermedades.

La Figura 6 muestra la transformación de las imágenes después de aplicar el proceso de recorte y mejora de la redondez. La imagen a la izquierda corresponde a la imagen original obtenida del conjunto de datos de Kaggle, mientras que la imagen a la derecha representa el resultado después de recortar los bordes y redimensionarla. Es importante destacar que la imagen modificada conserva la proporción relativa de la redondez en comparación con las otras imágenes

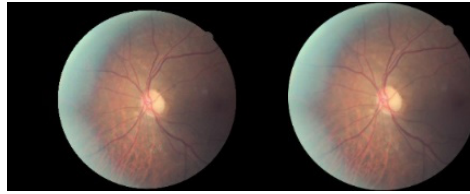


Figura 6. La imagen de la izquierda es original del conjunto de datos de Kaggle, y la imagen de la derecha es después de recortar los bordes adicionales y redimensionar sin alterar el radio del fondo de ojo.

La Figura 7 muestra algunos ejemplos de los resultados obtenidos después de aplicar las transformaciones de las imágenes que permiten aumentar el conjunto de datos, incluyendo las transformaciones de "HorizontalFlip", "ElasticTransform", "Distorsión de Cuadrícula", "Distorsión Óptica" y "Rotación".

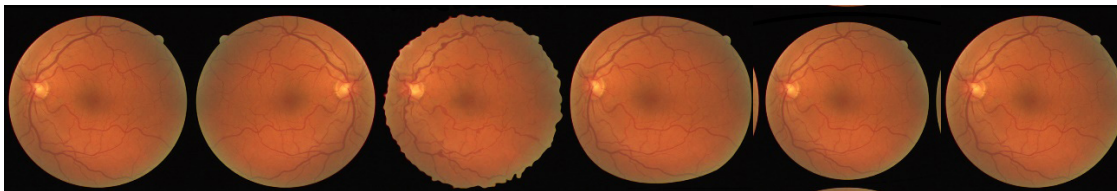


Figura 7 Resultado del aumento de imágenes utilizadas en la red UNet, la primera imagen corresponde a la original, y las demás corresponden a los aumentos mediante la biblioteca Albumentation.

La Figura 8 muestra el resultado obtenido al aumentar los datos en la red ConvNet. Se observa un aumento en el número de imágenes generadas, donde algunas muestran rotaciones horizontales y verticales, al igual que los resultados obtenidos de la red anterior. La función "ColorJitter" introduce variaciones en el color a lo largo de las imágenes, mientras que las últimas muestran la utilización de CLAHE para mejorar el contraste. Por otra parte, la aplicación de la función "Sharpen" contribuye a bordes más pronunciados, lo que resulta en una mejora en los detalles de la imagen.

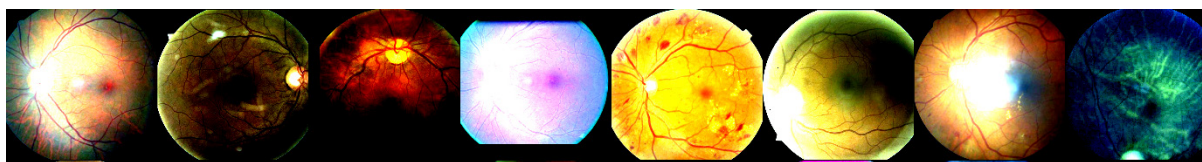


Figura 8 Resultado mostrado en un conjunto de imágenes de la ampliación generada en la primera parte de la red ConvNet, de varias imágenes seleccionadas por el algoritmo durante el proceso de entrenamiento.

Los resultados que se muestran a continuación corresponden al proceso de creación de una máscara morfológica y segmentación de los vasos sanguíneos. Después de realizar la operación morfológica, las imágenes enmascaradas se utilizaron para entrenar la red UNet. Estas máscaras se generaron porque los conjuntos de datos de diabetes no tenían esta

información. La Figura 9 muestra un ejemplo de la creación de las máscaras del árbol vascular y el proceso de Unet, que se utilizará en la clasificación de las imágenes en las diferentes clases y niveles de la enfermedad.

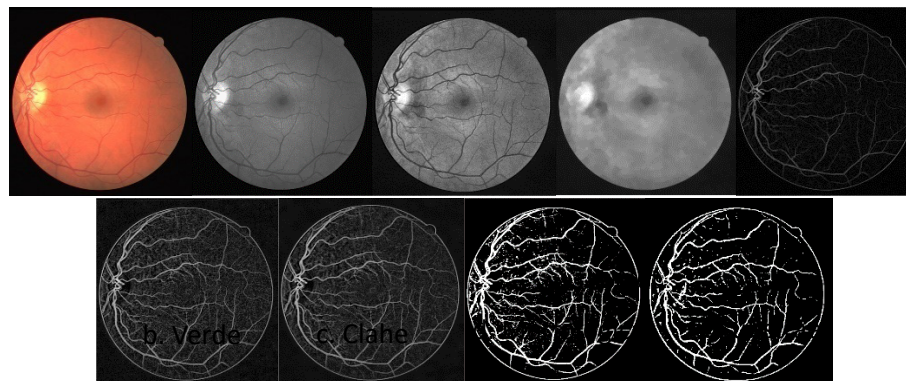


Figura 9. Proceso de resultado de segmentación, la imagen a. corresponde al fondo original de la retina. Las imágenes b a h son el resultado de aplicar el proceso de generación de máscaras. La imagen i muestra la máscara generada con la red UNet.

Para evaluar el grado de similitud entre las máscaras generadas y las originales, se volvió esencial un proceso de validación. Este procedimiento implicó una comparación exhaustiva entre las máscaras originales y aquellas producidas utilizando el método morfológico descrito anteriormente. Los hallazgos demostraron tasas de accuracy aproximado del 90% y una precisión del 80% respectivamente. Para obtener una comprensión más completa de esta similitud, se empleó una matriz de confusión para cuantificar con precisión el número de píxeles que muestran una correspondencia coincidente dentro de los dos conjuntos de máscaras. La Figura 10 muestra la comparación entre la máscara creada y la original, así como el número de píxeles que fueron seleccionados correctamente.

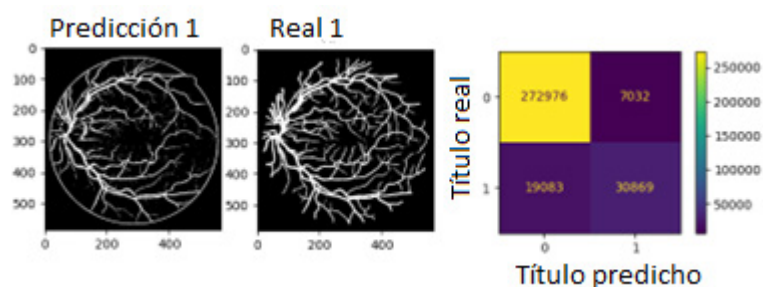


Figura 10 Ejemplo de la comparación de las máscaras creadas mediante el proceso morfológico y las máscaras originales de los conjuntos de datos de hipertensión. La imagen a la izquierda es la máscara creada, la imagen en el centro es la máscara original y la imagen a la derecha muestra en detalle la matriz de confusión.

Para evaluar el rendimiento de los algoritmos de aprendizaje automático, los analistas suelen utilizar la métrica de precisión o "accuracy". Esta métrica es preferida debido a su capacidad para evaluar de manera efectiva las capacidades de generalización de los clasificadores [63]. El Índice de Jaccard, también conocido como Intersección sobre Unión (IoU), es una estadística utilizada para medir la similitud o incluso la diversidad entre conjuntos [64]. Este enfoque de índice de Jaccard también se ha utilizado en modelos de visión por computadora y reconocimiento de objetos para medir la similitud entre dos imágenes, generalmente, la imagen predicha y su imagen real [64].

Los resultados obtenidos en la red UNet con las máscaras generadas de imágenes morfológicas muestran un Índice de Jaccard del 74%, una F1 del 85%, y una Accuracy del 96%. Por otro lado, en la red ConvNet, los resultados para la clasificación de las imágenes en las dos enfermedades y en los diferentes niveles de la enfermedad muestran un 80% de accuracy. Además de esto, en lugar de optar por conjuntos de datos comúnmente utilizados como DRIVE o CHASEDB1, se seleccionó un conjunto de datos diabético que se aparta del enfoque convencional al carecer de máscaras preexistentes. Finalmente, la selección sobre el uso de UNet en lugar de otros se debe a que este está especialmente adaptado para la segmentación de objetos biomédicos, especialmente justificado por la eficiencia en el entrenamiento con conjuntos de datos limitados, lo cual es esencial en este contexto [65].

Finalmente, se presentan los resultados obtenidos en el proceso de clasificación. La Figura 11 muestra la matriz de confusión y la información de la Tabla 7 con la distribución de la imagen para el conjunto de validación. Muestra cómo se identificó cada una de las clases en relación con la etiqueta real y predicha. En la matriz, se puede ver cómo la mayor parte del tiempo se logró una buena clasificación, y donde hubo el mayor error fue en el límite de una clase u otra. Esta prueba se realiza utilizando el modelo de transferencia de aprendizaje EfficientNet-B0 y la red ConvNet, con los incrementos y preprocesamiento previamente mencionados en las imágenes. La diferencia de colores no se toma en comparación con todas las clases, sino con la misma clase porque hay clases con menos imágenes en total, como es el caso con las clases más avanzadas de la enfermedad.

Tabla 7 contiene información de la matriz de confusión sobre el número total de imágenes según la clase, junto con las variables para la validación de los resultados en las 10 clases creadas.

Clase	Precisión	Recall	F1 Score	# Imágenes
0	0.89	0.95	0.92	980
1	0.74	0.85	0.79	772
2	0.65	0.61	0.63	502
3	0.77	0.69	0.73	526
4	0.70	0.50	0.58	218
5	0.90	0.95	0.92	980
6	0.78	0.86	0.82	772
7	0.67	0.64	0.66	502
8	0.86	0.71	0.78	526
9	1.00	0.73	0.84	218



	0	1	2	3	4	5	6	7	8	9
0	934	46	0	0	0	0	0	0	0	0
1	108	659	5	0	0	0	0	0	0	0
2	3	190	308	1	0	0	0	0	0	0
3	0	1	163	362	0	0	0	0	0	0
4	0	0	1	109	108	0	0	0	0	0
5	0	0	0	0	46	933	1	0	0	0
6	0	0	0	0	0	107	664	1	0	0
7	0	0	0	0	0	0	181	321	0	0
8	0	0	0	0	0	0	0	155	371	0
9	0	0	0	0	0	0	0	0	59	159

Figura 11. Matriz de confusión de los resultados de ConvNet para las 9 clases de retinopatía, con niveles del 0 al 4 para la diabetes y niveles del 5 al 9 para la hipertensión, utilizando el modelo EfficientNet-B0. El eje Y muestra las etiquetas reales y el eje X muestra las etiquetas predichas por la red.

Los resultados de los modelos de aprendizaje fueron comparados, obteniendo mejores resultados al utilizar el modelo B0 como primer paso. La Tabla 8 muestra las diferencias en precisión y pérdida al finalizar la ejecución del modelo preentrenado y la segunda red convolucional. La columna de número de éxitos representa el número de clases correctamente clasificadas sobre el número total de imágenes tomadas.

Tabla 8. Tabla comparativa de la clasificación de las 10 clases de retinopatía utilizando los modelos de transferencia de aprendizaje B3 y B0 y la segunda red.

Modelo	Perdida	Accuracy	# Éxitos
B0 + Segunda Red	0.35	80.37	4.819/5.996
B3 + Segunda Red	0.74	57.49	3.447/5.996

## Discusión

La implementación de este estudio se basó en dos enfoques complementarios para abordar tanto la segmentación como la clasificación de imágenes de retinopatía diabética e hipertensiva. Estos enfoques se integraron para facilitar una evaluación completa de las imágenes. Para el proceso de segmentación, se empleó una arquitectura Unet, que involucró una serie de pasos de procesamiento de imagen, incluidas operaciones morfológicas vitales. Este enfoque permitió la segmentación de conjuntos de datos que carecían de máscaras como etiquetas, mejorando la versatilidad del modelo.

En el ámbito de la investigación sobre retinopatía hipertensiva, la segmentación suele tener prioridad, utilizando conjuntos de datos provistos de máscaras como etiquetas. A diferencia del enfoque de Xu [66], centrado en la segmentación de vasos sanguíneos retinianos en conjuntos de datos establecidos como DRIVE y STARE, esta investigación implementó la

segmentación de imágenes de fondo de ojo utilizando un rango más amplio de conjuntos de datos, incluyendo aquellos destinados a la retinopatía diabética. Es crucial destacar que tanto el conjunto de datos de retinopatía diabética como el conjunto de datos de hipertensión Odir presentaron un desafío, ya que no incluían máscaras iniciales, lo que requirió la aplicación de un paso de preprocesamiento que implicaba operaciones morfológicas críticas.

En cuanto a los resultados, se logró un índice de Jaccard de aproximadamente el 74% en el proceso de segmentación. Es importante destacar que existen posibles mejoras adicionales en la segmentación al refinar el proceso de generación de máscaras, sobre todo considerando que muchas imágenes presentan variaciones en las características tonales. Por ende, se sugiere evaluar el uso del canal de escala de grises como alternativa al canal verde y mejorar las técnicas de reducción de ruido. En términos de clasificación, se logró una precisión de aproximadamente el 80%.

Las disparidades de rendimiento observadas en la Tabla 7, particularmente para las clases 1, 2, 6, 7 y 8, son reconocidas y merecen un examen más detenido. Estos desafíos destacan la complejidad de los escenarios del mundo real y la necesidad de un mayor refinamiento en el enfoque propuesto. Mirando hacia adelante a futuras investigaciones, se propone un refinamiento del proceso de clasificación trabajando con un conjunto reducido de categorías, específicamente 9 categorías a partir de 0, excluyendo la presencia de la enfermedad y los cuatro niveles de las dos enfermedades. Aun así, uno de los desafíos que se encontraron, especialmente en el caso de la retinopatía hipertensiva, es la escasez de datos. Para abordar esta limitación, se pretende explorar colaboraciones potenciales con centros médicos para ampliar el conjunto de datos. Para una mayor robustez y fiabilidad, se recomienda incorporar técnicas de validación cruzada. La utilización de la validación cruzada garantiza una evaluación más exhaustiva del rendimiento del modelo al evaluar su capacidad de generalización en diferentes subconjuntos del conjunto de datos. Este enfoque proporciona un proceso de validación integral, contribuyendo a la fiabilidad y efectividad general de la metodología propuesta.

## Conclusiones

En este estudio, se presentó un enfoque que utilizó el poder del aprendizaje profundo y las redes neuronales convolucionales tradicionales para abordar la desafiante tarea de la segmentación y clasificación de imágenes de retinopatía diabética e hipertensiva. Las principales contribuciones y hallazgos se pueden resumir de la siguiente manera: Integración de Unet y ConvNet: Introducción de una fusión de arquitecturas Unet y ConvNet, combinando las fortalezas de ambas para la segmentación y clasificación de imágenes de retinopatía diabética e hipertensiva.

Creación de máscaras morfológicas específicamente para el conjunto de datos diabéticos, porque estos no tenían las etiquetas iniciales, diferente de los conjuntos de hipertensión que se han utilizado ampliamente en otra investigación. Las características de las imágenes de retinopatía diabética pueden variar, y el uso de máscaras personalizadas garantiza que el proceso de segmentación se adapte a las características únicas presentes en estas imágenes.

Utilización del Modelo Preentrenado EfficientNet: Para mejorar la eficiencia y precisión de nuestro modelo, empleamos el modelo de aprendizaje preentrenado EfficientNet. Nuestros experimentos demostraron resultados mejorados con este modelo, mostrando su idoneidad para el análisis de imágenes de retinopatía.

Aprendizaje por Transferencia: Exploración de técnicas y arquitecturas de aprendizaje por transferencia más sofisticadas para el análisis de imágenes de retinopatía lo cual conducía a resultados aún más precisos.

Aumento de Datos: Investigación de estrategias avanzadas de aumento de datos para diversificar el conjunto de datos de entrenamiento y aumentar la robustez del modelo. En conclusión, el trabajo presenta un enfoque para el análisis de imágenes de retinopatía diabética e hipertensiva. Al combinar las fortalezas de diferentes arquitecturas de redes neuronales y la importancia en la creación de máscaras morfológicas personalizadas para el conjunto de datos de retinopatía diabética. Los esfuerzos de investigación futuros pueden avanzar aún más estas contribuciones y expandir el impacto de este trabajo en la comunidad médica.

## Referencias

1. OPS. Organización Panamericana de la Salud. [Online]. [cited 2023 09 23. Available from: <https://www.paho.org/es/temas/enfermedades-no-transmisibles>.
2. Manresa JM, Forés R, Vásquez X, Alzamora MT, Heras A, Delgado P, et al. Fiabilidad de la retinografía para la detección de retinopatía hipertensiva en Atención Primaria. *Atención Primaria*. 2020 Jun; 52(6).
3. Centro de Oftalmología Bonafonte. Assessment of Corneal Angiography Filling Patterns in Corneal Neovascularization. *Journal of Clinical Medicine*. 2023; 12(2).
4. Poříz V. Diabetic Retinopathy Detection Using Neural Networks. Tesis. Czech Technical University in Prague; 2020.
5. Jha A, VA&AAR. Association of severity of diabetic retinopathy with corneal endothelial and thickness changes in patients with diabetes mellitus. *Eye*. 2022 June; 36(6).
6. Zhou J CB. Retinal Cell Damage in Diabetic Retinopathy. *Cells*. 2023 May; 12(9).
7. Jiménez Jiménez L. Manifestaciones oftalmológicas vasculares de la hipertensión arterial. Tesis. Universidad de Sevilla, Medicina; 2020.
8. Hagos MT. Point-of-Care Diabetic Retinopathy Diagnosis: A Standalone Mobile Application Approach. Tesis. Johns Hopkins University; 2020.
9. Jiménez García J. Detección de vasos sanguíneos en retinografías mediante técnicas de procesamiento digital de imágenes. Tesis. Valladolid: Universidad de Valladolid, Escuela Técnica Superior de Ingenieros de Telecomunicación; 2017.
10. Kropp M, GO, MAea. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. *EPMA Journal*. 2023 March 01; 14(1).
11. Wu S MX. Optic Nerve Regeneration in Diabetic Retinopathy: Potentials and Challenges Ahead. *Int J Mol Sci*. 2023 January 11; 24(2).
12. Jiménez Jiménez L. Manifestaciones Oftalmológicas Vasculares de la Hipertensión Arterial. Trabajo de Grado. Sevilla: Universidad de Sevilla, Departamento de Medicina; 2020.
13. Cardona Suárez JC, Fernández Agudelo FN. Modelo de Machine Learning para clasificación de pacientes con glaucoma en la población del Valle del Cauca. Tesis. Santiago de Cali: Universidad Icesi, Ingeniería; 2022.

14. Unnati V S, Tripathy K. Diabetic Retinopathy: StatPerls; 2023.
15. Sarki R. Automatic Detection Of Diabetic Eye Disease Through Deep Learning Using Fundus Images. Tesis. Victoria University; 2021.
16. Russell G. Diabetic retinal imaging: methods in automatic processing. Tesis. The University of Manchester; 2014.
17. Kropp M, Golubnitschaja O, Mazurakova A, Koklesova L, Sargheini N, Vo TTKS, et al. Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. EPMA Journal. 2023 January; 14(1).
18. Abbas Q, Ibrahim M. DenseHyper: an automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning. Multimedia Tools and Applications. 2020 November; 79.
19. Cortés Rodríguez DC. Cámara retinal: herramienta de telediagnóstico para detección de retinopatía diabética, retinopatía hipertensiva y glaucoma. Tesis. Universidad de la Salle; 2016.
20. Merchán Barrezueta MJ, Lucas Baño ES, Sánchez Escobar DA, Arellano Blacio MA. Retinopatía diabética e hipertensiva. RECIAMUC. 2023 January; 7(1): p. 290-298.
21. Chala M, Nsiri B, Hachem M, Soulaymani A, Mokhtari A, Benaji B. An automatic retinal vessel segmentation approach based on Convolutional Neural Networks. Expert Systems with Applications. 2021; 184: p. 115459.
22. Zhang H, Qiu Y, Song C, Li J. Landmark-Assisted Anatomy-Sensitive Retinal Vessel Segmentation Network. Diagnostics. 2023; 13(13).
23. NA S, Yadav AK, Akbar M, Kumar M, Yadav D. Retinal blood vessel segmentation using a deep learning method based on modified U-NET model. 2021 September.
24. Phong Thanh Nguyen VDBHKDVPTPEYGPJ. An Optimal Deep Learning Based Computer-Aided Diagnosis System for Diabetic Retinopathy. Computers, Materials & Continua. 2021; 61(3): p. 2815-2830.
25. Takahashi HaTHaAYaIYaKH. Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy. PLoS ONE. 2017; 12(6).
26. Dutta S&MB&BM&CR&INCSN. Classification of Diabetic Retinopathy Images by Using Deep Learning Models. International Journal of Grid and Distributed Computing. 2018; 11: p. 89-106.
27. Kaggle. Diabetic Retinopathy Detection. [Online].; 2015 [cited 2023 April 10. Available from: <https://www.kaggle.com/c/diabetic-retinopathy-detection/data>.
28. Drive. DRIVE (Digital Retinal Images for Vessel Extraction). [Online]. [cited 2022 06 22. Available from: <https://paperswithcode.com/dataset/drive>.
29. Fire. Kaggle. [Online].; 2020 [cited 2022 03 13. Available from: <https://www.kaggle.com/datasets/andrewmvd/fundus-image-registration>.
30. Mendeley Data. Data on Fundus Images for Vessels Segmentation, Detection of Hypertensive Retinopathy, Diabetic Retinopathy and Papilledema. [Online].; 2020 [cited 2023 04 10. Available from: <https://data.mendeley.com/datasets/3csr652p9y/2>.
31. Kaggle. Ocular Disease Recognition. [Online].; 2020 [cited 2022 February 10. Available from: <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k>.

32. Singh A. CLAHE augmentation -Ranzcr comp. [Online].; 2020 [cited 2022 11 15. Available from: <https://www.kaggle.com/code/amritpal333/clahe-augmentation-ranzcr-comp>.
33. Manizheh Safarkhani G, Hojjat M, Mehdi A. Segmentation of Retinal Blood Vessels Using U-Net++ Architecture and Disease Prediction. *Electronics*. 2022; 11(21).
34. Verdeguer Gómez J. Redes neuronales para la clasificación y segmentación de imágenes médicas. Tesis. Valencia: Universitat Politècnica de València, Departamento de Sistemas Informáticos y Comput; 2020.
35. Ronneberger O,FP,BT. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*. 2015 November 18; 9351.
36. Tomar N. Github. [Online].; 2021 [cited 2022 04 20. Available from: <https://github.com/nikhilroxtomar/Retina-Blood-Vessel-Segmentation-in-PyTorch/tree/main/UNET>.
37. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2015.
38. Zhang Y, Liu S, Li C, Wang J. Rethinking the Dice Loss for Deep Learning Lesion Segmentation in Medical Images. *Journal of Shanghai Jiaotong University (Science)*. 2021 February; 26(1): p. 93.
39. Wazir S, Moazam Fraz M. HistoSeg: Quick attention with multi-loss function for multi-structure segmentation in digital histology images. 2022 12th International Conference on Pattern Recognition Systems ICPRS. 2022 June.
40. Zhuang L, Hanzi M, Chao-Yuan W, Christoph F, Trevor D, Saining X. A ConvNet for the 2020s. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022;: p. 11966-11976.
41. Persson A. aladdinpersson/Machine-Learning-Collection. [Online].; 2020 [cited 2022 06 20. Available from: <https://github.com/aladdinpersson/Machine-Learning-Collection>.
42. Ferreras Extremo A. Estudio de algoritmos de redes neuronales convolucionales en dataset de imágenes médicas. Valladolid: Universidad de Valladolid; 2021.
43. Agarwal V. Complete Architectural Details of all EfficientNet Models. [Online].; 2020 [cited 2023 April 28. Available from: <https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142>.
44. Yasoda BY, Jagannadham DBV. Classification of Brain Tumor Using Finetuned Efficientnet. *International Journal of Creative Research Thoughts (IJCRT)*. 2022 December; 10(12).
45. Marques G, Ferreras A, de la Torre-Diez I. An ensemble-based approach for automated medical diagnosis of malaria using EfficientNet. *Multimed Tools Appl*. 2022 March 29; 81(19).
46. Alhichri H&AA&BY&AN&AN. Classification of Remote Sensing Images Using EfficientNet-B3 CNN Model with Attention. *IEEE Access*. 2021 January.
47. Amreen B, Yung-Cheol B. Lightweight EfficientNetB3 Model Based on Depthwise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images. *IEEE Access*. 2023; 11: p. 37203-37215.
48. Lasloum T, Alhichri H, Bazi Y, Alajlan N. SSDAN: Multi-Source Semi-Supervised Domain Adaptation Network for Remote Sensing Scene Classification. *Remote Sensing*. 2021 September; 13.
49. Tan M, V. Le Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks Chaudhuri KaSR, editor.; 2019.

50. PyTorch. PyTorch Linear Layer (Fully Connected Layer) Explained. [Online].; 2022 [cited 2023 April 12. Available from: <https://androidkt.com/pytorch-linear-layer-fully-connected-layer-explained/>.
51. Moczulski M, Denil M, Appleyard J, Freitas N. ACDC: A Structured Efficient Linear Layer. 2016 May.
52. Kumar B. PyTorch Batch Normalization. [Online].; 2022 [cited 2023 April 12. Available from: <https://pythonguides.com/pytorch-batch-normalization/>.
53. Educba. PyTorch Dropout. [Online].; 2022 [cited 2023 April 12. Available from: <https://www.educba.com/pytorch-dropout>.
54. Buslaev A, Parinov A, Khvedchenya E, Iglovikov V, Kalinin A. Albumentations: fast and flexible image augmentations. Computer Science > Computer Vision and Pattern Recognition. 2018 Sep; 11(2).
55. Albumentations. A list of transforms and their supported targets. [Online]. [cited 2023 04 10. Available from: [https://albumentations.ai/docs/getting\\_started/transforms\\_and\\_targets/](https://albumentations.ai/docs/getting_started/transforms_and_targets/).
56. Albumentations. Transforms (augmentations.transforms). [Online].; 2022 [cited 2022 11 15. Available from: [https://albumentations.ai/docs/api\\_reference/augmentations/transforms/](https://albumentations.ai/docs/api_reference/augmentations/transforms/).
57. Albumentations. CoarseDropout augmentation (augmentations.dropout.coarse\_dropout). [Online].; 2022 [cited 2022 11 15. Available from: [https://albumentations.ai/docs/api\\_reference/augmentations/dropout/coarse\\_dropout/](https://albumentations.ai/docs/api_reference/augmentations/dropout/coarse_dropout/).
58. Delisle P, Desrosiers C, Lombaert H. Realistic Image Normalization for Multi-Domain Segmentation. Medical Image Analysis. 2021 Dec; 74.
59. Gothwal R, Gupta S, Gupta D, Kumar A. Color image segmentation algorithm based on RGB channels. Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization. 2014;; p. 1-5.
60. Shouting Feng ZZDPQT. CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. Neurocomputing. 2020; 392: p. 268-276.
61. Baiyuan D, Gongjian W, Conghui M, Xiaoliang Y. Target recognition in synthetic aperture radar images using binary morphological operations. Journal of Applied Remote Sensing. 2016; 10(4).
62. SH K, JY K. Application of Fast Non-Local Means Algorithm for Noise Reduction Using Separable Color Channels in Light Microscopy Images. Int J Environ Res Public Health. 2021; 18(6).
63. Mohammad H, Sulaiman MN. A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process. 2015 Mar; 5.
64. Ogwok D, Ehlers EM. Jaccard Index in Ensemble Image Segmentation: An Approach. CIIS '22: Proceedings of the 2022 5th International Conference on Computational Intelligence and Intelligent. 2022 November;; p. 9-14.
65. Qin J, Liu T, Wang Z, Liu L, Fang H. GCT-UNET: U-Net Image Segmentation Model for a Small Sample of Adherent Bone Marrow Cells Based on a Gated Channel Transform Modul. Electronics. 2022; 11(22).
66. Xu S, Chen Z, Cao W, Zhang F, Tao B. Retinal Vessel Segmentation Algorithm Based on Residual Convolution Neural Network. Frontiers in Bioengineering and Biotechnology. 2021; 9.
67. Arsalan M, Owais M, Mahmood T, Woon Cho S, Ryoung Park K. Aiding the Diagnosis of Diabetic and Hypertensive Retinopathy Using Artificial Intelligence-Based Semantic Segmentation. Journal of Clinical Medicine. 2019; 8.