

DISEÑO DE ALGORITMO PARA LA SEGMENTACIÓN DE DISCO ÓPTICO EN  
IMÁGENES DE FONDO DE RETINA

CRISTIAN JAVIER GÓMEZ MONTAÑA

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE  
TELECOMUNICACIONES  
BUCARAMANGA

2026

DISEÑO DE ALGORITMO PARA LA SEGMENTACIÓN DE DISCO ÓPTICO EN  
IMÁGENES DE FONDO DE RETINA

CRISTIAN JAVIER GÓMEZ MONTAÑA

Tesis presentada como requisito parcial para la obtención del título de  
Ingeniero Electrónico

Director

Said Pertuz, PhD

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE  
TELECOMUNICACIONES  
BUCARAMANGA

2026

## TABLA DE CONTENIDO

	<b>p.</b>
INTRODUCCIÓN	9
1 OBJETIVOS	12
2 TRABAJO PREVIO RELACIONADO	13
3 METODOLOGÍA	16
3.1 Base de datos	16
3.2 Arquitectura de segmentación	16
3.3 Entrenamiento y ajuste de hiperparámetros	18
3.4 Medidas de desempeño	19
4 RESULTADOS	20
5 CONCLUSIÓN	24
6 RECOMENDACIONES	25
BIBLIOGRAFÍA	26
ANEXOS	28

## LISTA DE FIGURAS

	<b>p.</b>
Figura 1 Imagen de fondo de retina con delimitación del disco óptico mediante un contorno circular en color azul.	10
Figura 2 Arquitectura U-Net implementada para la segmentación del disco óptico, compuesta por una etapa de codificación y una de decodificación con conexiones de concatenación, finalizando con una convolución 1×1 con activación sigmoide para generar el mapa de segmentación.	17
Figura 3 Histograma de los coeficientes de Dice. El entrenamiento se programó para 20 épocas; sin embargo, se detuvo en la época 16 debido al criterio de early stopping.	21
Figura 4 Resultados de la segmentación del disco óptico. (a) Imágenes originales provenientes de las bases de datos. (b) Máscaras de referencia proporcionadas por dichas bases. (c) Predicción generada por la red U-Net. (d) Segmentación final obtenida a partir del modelo.	22

## LISTA DE TABLAS

	<b>p.</b>
Tabla 1     Resultados de las métricas de desempeño del modelo U-Net durante el entrenamiento, validación y prueba.	20
Tabla 2     Comparación de métricas de desempeño con otros métodos de segmentación del disco óptico.	23
Tabla 3     Resumen de las capas y dimensiones de salida del modelo U-Net.	28

## ANEXOS

	<b>p.</b>
Anexo A    Resumen de las capas y dimensiones de salida del modelo U-Net.	28

## RESUMEN

**TÍTULO:** DISEÑO DE ALGORITMO PARA LA SEGMENTACIÓN DE DISCO ÓPTICO EN IMÁGENES DE FONDO DE RETINA \*

**AUTOR:** CRISTIAN JAVIER GÓMEZ MONTAÑA \*\*

**PALABRAS CLAVE:** GLAUCOMA; SEGMENTACIÓN DEL DISCO ÓPTICO; IMÁGENES DE FONDO DE RETINA; REDES NEURONALES CONVOLUCIONALES; U-NET.

**DESCRIPCIÓN:** El glaucoma es la principal causa de ceguera irreversible en el mundo. Su carácter crónico y silencioso dificulta la detección temprana. La segmentación del disco óptico en imágenes de fondo de retina se considera el paso preliminar en el desarrollo de la herramienta de diagnóstico para la detección temprana del glaucoma. En este trabajo se desarrolló un modelo de segmentación automática basado en la arquitectura U-Net, utilizando imágenes de retina de la base de datos pública de Origa-ligth. Para mejorar la capacidad del modelo, se aplicaron técnicas de aumento de datos y se evaluó su rendimiento mediante el coeficiente de similitud de Dice (DSC) y la intersección sobre la unión (IoU). Experimentos con conjunto de prueba de 650 imágenes muestran que el modelo consiguió una IoU de 0.568 y un DSC de 0.721.

---

\* Tesis

\*\* Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director: Said Pertuz, PhD.

## ABSTRACT

**TITLE:** ALGORITHM DESIGN FOR OPTIC DISC SEGMENTATION IN RETINAL FUNDUS IMAGES \*

**AUTHOR:** CRISTIAN JAVIER GÓMEZ MONTAÑA \*\*

**Keywords:** GLAUCOMA; OPTIC DISC SEGMENTATION; FUNDUS IMAGES; CONVOLUTIONAL NEURAL NETWORKS; U-NET.

**Description:** Glaucoma is the leading cause of irreversible blindness worldwide. Its chronic and asymptomatic nature makes early detection difficult. Optic disc segmentation in fundus images is considered a preliminary step in the development of diagnostic tools for the early detection of glaucoma. In this work, an automatic segmentation model based on the U-Net architecture was developed using retinal images from the public ORIGA-light database. To enhance the model's performance, data augmentation techniques were applied, and its effectiveness was evaluated using the Dice Similarity Coefficient (DSC) and Intersection over Union (IoU). Experiments conducted on a test set of 650 images show that the model achieved an IoU of 0.568 and a DSC of 0.721.

---

\* Thesis

\*\* Facultad de Ingenierías Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Advisor: Said Pertuz, PhD.

## INTRODUCCIÓN

El glaucoma es una enfermedad ocular de progresión lenta que provoca un daño irreversible en la visión. Debido a que suele desarrollarse sin síntomas evidentes en sus primeras etapas, su detección temprana representa un desafío clínico significativo. Lo que genera un alto impacto en la salud pública y en la calidad de vida de los pacientes. Según la Organización Mundial de la Salud (OMS), en el año 2020 se reportaron en el mundo, 295 millones de personas con baja visión y 43 millones con ceguera, afectando al 3 % de toda la población colombiana, con un aumento del 10 % entre la población mayor de 40 años. Un oftalmólogo especializado en glaucoma (glaucomatólogo) puede llegar a detectar la enfermedad en menos de 30 minutos <sup>1</sup>.

El Glaucoma es una enfermedad asociada con el cambio de presión ocular afectando el nervio óptico. Este está conformado por fibras nerviosas encargadas de enviar los mensajes visuales al cerebro, produciendo la pérdida de visión y posible ceguera <sup>2</sup>. Una de las estrategias para detectar la progresión del glaucoma se basa en la identificación de anomalías en el disco óptico (DO), ya que esta estructura corresponde a la región donde convergen las fibras nerviosas del nervio óptico. Este análisis se realiza a partir de imágenes de fondo de retina, las cuales corresponden a fotografías de la superficie interior del ojo<sup>3</sup>, como se muestra en la Figura 1, y permiten evaluar cambios estructurales relevantes.

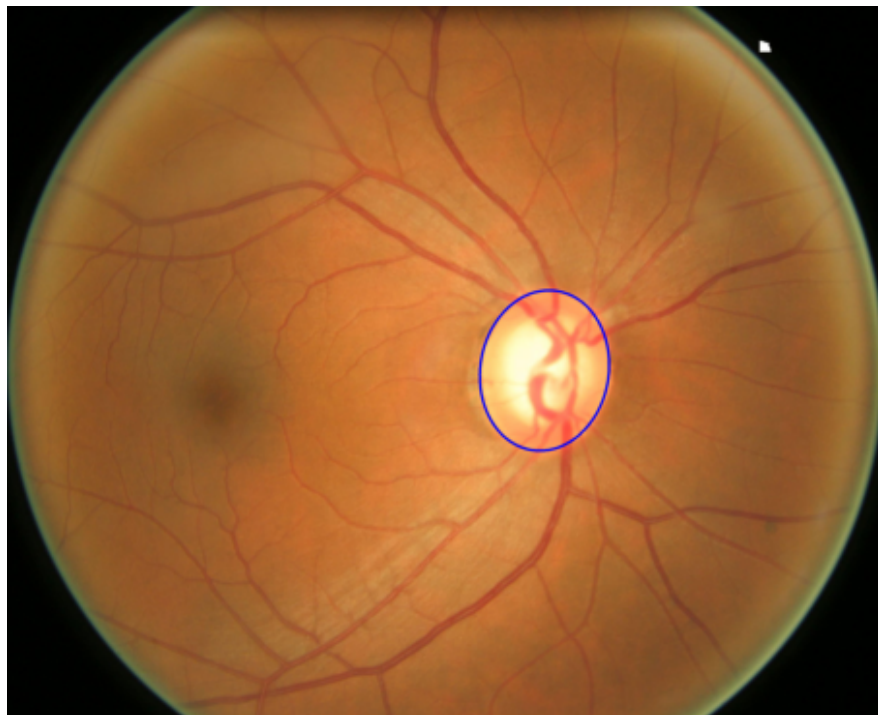
---

<sup>1</sup> Instituto Nacional para Ciegos (INCI). *12 de marzo: día del glaucoma, la causa principal de la ceguera*. 2023. URL: <https://www.inci.gov.co/blog/12-de-marzo-dia-del-glaucoma-la-causa-principal-de-la-ceguera> (visitado 21-09-2025).

<sup>2</sup> Yih-Chung Tham et al. "Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040: A Systematic Review and Meta-Analysis". En: *Ophthalmology* 121.11 (2014), págs. 2081-2090.

<sup>3</sup> Fundación Colombiana de Glaucoma. *¿Qué es el glaucoma?* 2024. URL: <http://www.glaucoma.org.co/articulos/que-es-glaucoma> (visitado 21-09-2025).

Figura 1. Imagen de fondo de retina con delimitación del disco óptico mediante un contorno circular en color azul.



En este contexto, el desarrollo de métodos computacionales para la segmentación del DO resulta fundamental, ya que permite apoyar el diagnóstico oportuno y contribuir a la prevención de la progresión de la enfermedad. Para realizar la segmentación del DO se empleaban métodos manuales y técnicas clásicas de procesamiento de imágenes, los cuales presentaban ciertas limitaciones ante variaciones de iluminación, ruido y diferencias anatómicas. Con la aparición de técnicas de aprendizaje automático y aprendizaje profundo, se han desarrollado métodos más avanzados para el análisis y segmentación de imágenes médicas, mejorando los resultados en la detección del glaucoma <sup>4</sup>.

La U-Net, es una de las arquitecturas de redes neuronales convolucionales (CNN) más sólidas, versátiles y efectivas para la segmentación de imágenes biomédicas, ya que mantiene un equilibrio entre precisión y eficiencia, siendo más accesible en términos de entrenamiento y recursos de hardware<sup>5</sup>. Su diseño basado en un encoder – decoder con conexiones de salto, permite recuperar información espacial de alta resolución logrando delinear con mayor precisión estructuras pequeñas y complejas como el DO.

En base a esto, se propone un método de segmentación del DO basado en la arquitectura U-Net, empleando el conjunto de datos de *Origa (Ligth)* <sup>6</sup>, al cual se le realiza un preprocesamiento aplicando operaciones morfológicas y estandarización en las imágenes. En este documento se describe el proceso de diseño e implementación de la red neuronal, y se reportan los resultados experimentales de desempeño.

---

<sup>4</sup> Yuming Jiang et al. "JointRCNN: A Region-Based Convolutional Neural Network for Optic Disc and Cup Segmentation". En: *IEEE Transactions on Biomedical Engineering* 67.2 (2020), págs. 335-343.

<sup>5</sup> Huazhu Fu et al. "Disc-Aware Ensemble Network for Glaucoma Screening From Fundus Image". En: *IEEE Transactions on Medical Imaging* 37.11 (2018), págs. 2493-2501.

<sup>6</sup> sshikamaru. *Glaucoma Detection*. 2020. URL: <https://www.kaggle.com/datasets/sshikamaru/glaucoma-detection> (visitado 21-09-2025).

## 1. OBJETIVOS

### Objetivo general

- Diseñar un algoritmo para la segmentación del disco óptico en imágenes de fondo de retina.

### Objetivos específicos

1. Comparar algoritmos del de estado del arte para la segmentación del disco óptico en imágenes de fondo de retina.
2. Diseñar un algoritmo que permita la segmentación automática del disco óptico mediante la integración de algoritmos del estado del arte.
3. Evaluar el desempeño del algoritmo propuesto mediante el uso de métricas de evaluación en bases de datos públicas.

## 2. TRABAJO PREVIO RELACIONADO

En la literatura se han propuesto diversos enfoques para la segmentación automática del disco óptico y la copa óptica en imágenes de fondo de retina, con el objetivo de apoyar el diagnóstico temprano del glaucoma. Cheng et al. (2013)<sup>7</sup> proponen un método para la segmentación del disco y la copa óptica basado en la clasificación de superpíxeles, enfocado en emplear características de color, textura y la localización anatómica para clasificar cada pixel logrando diferenciar la región del disco de la región no discal. De manera similar Zahoor y Fraz (2017)<sup>8</sup> presentan un método rápido para la segmentación del disco óptico basado en la transformada polar, aprovechando la forma aproximadamente circular de esta estructura anatómica del disco. El procedimiento consiste en tomar la región de interés y transformar las coordenadas cartesianas a coordenadas polares, donde el contorno del disco se representa con una estructura lineal, facilitando su detección mediante el análisis de bordes.

No obstante, los enfoques basados en procesamiento tradicional de imágenes presentan dificultades para adaptarse a la variabilidad propia de las imágenes de fondo de retina, ya que dependen de características definidas manualmente y de supuestos previos sobre la forma y ubicación del disco óptico. Esta dependencia los hace sensibles a cambios en iluminación, contraste y presencia de patologías, además de limitar su capacidad para abordar de manera conjunta la segmentación de la copa óptica y el disco óptico en bases de datos diversas.

---

<sup>7</sup> Jun Cheng et al. "Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening". En: *IEEE Transactions on Medical Imaging* 32.6 (2013), págs. 1019-1032.

<sup>8</sup> Muhammad Nauman Zahoor y Muhammad Moazam Fraz. "Fast Optic Disc Segmentation in Retina Using Polar Transform". En: *IEEE Access* 5 (2017), págs. 12293-12300.

Con el avance del aprendizaje profundo, comenzaron a explorarse enfoques basados en redes neuronales convolucionales, Sevastopolsky (2017)<sup>9</sup> propone una modificación de la arquitectura U-Net para la segmentación conjunta del disco y la copa óptica. El modelo adopta una estructura encoder-decoder que permite aprender representaciones jerárquicas directamente a partir de los píxeles de entrada. Posteriormente, Fu et al. (2018)<sup>10</sup> introducen un enfoque más avanzado mediante una red profunda multietiqueta para la segmentación conjunta del disco y la copa óptica conocida como M-Net, incorpora entradas multiescala, salidas laterales y una función de pérdida multietiqueta permitiendo capturar información contextual a diferentes niveles de resolución. Adicionalmente, emplean la transformada polar como una estrategia para normalizar la variabilidad geométrica del disco óptico mejorando la precisión en la delimitación de bordes.

En una línea similar, Gu et al. (2019)<sup>11</sup> proponen CE-Net, una arquitectura diseñada para preservar la información espacial y capturar contexto semántico de alto nivel. El modelo emplea una red residual preentrenada como codificador y un módulo extractor de contexto que combina convoluciones dilatadas densas y pooling multikernel, permitiendo una representación multiescala sin pérdida significativa de resolución.

En este contexto, los trabajos más recientes que emplean redes neuronales convolucionales muestran que la segmentación del disco óptico puede mejorar de manera notable cuando se adaptan y optimizan arquitecturas como U-Net. Las modificaciones propuestas en estos estudios permiten capturar mejor el contexto y los detalles relevantes de la

---

<sup>9</sup> A. Sevastopolsky. "Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network". En: *Pattern Recognition and Image Analysis* 27.3 (2017), págs. 618-624.

<sup>10</sup> Huazhu Fu et al. "Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation". En: *IEEE Transactions on Medical Imaging* 37.7 (2018), págs. 1597-1605.

<sup>11</sup> Zaiwang Gu et al. "CE-Net: Context Encoder Network for 2D Medical Image Segmentation". En: *IEEE Transactions on Medical Imaging* 38.10 (2019), págs. 2281-2292.

imagen, lo que se traduce en una segmentación más precisa y estable.

## 3. METODOLOGÍA

### 3.1. Base de datos

Para el desarrollo del proyecto se utilizó la base de datos pública de *Origa-light* que contiene 650 imágenes de fondo de ojo de retina con anotaciones de profesionales formados del Instituto de Investigación Ocular de Singapur <sup>6</sup>. El tamaño de las imágenes es aproximadamente de 2548x2048 píxeles en formato RGB. Se realiza un preprocesamiento a las imágenes para extraer una región de interés adaptando el tamaño de dichas imágenes a 128x128 píxeles.

### 3.2. Arquitectura de segmentación

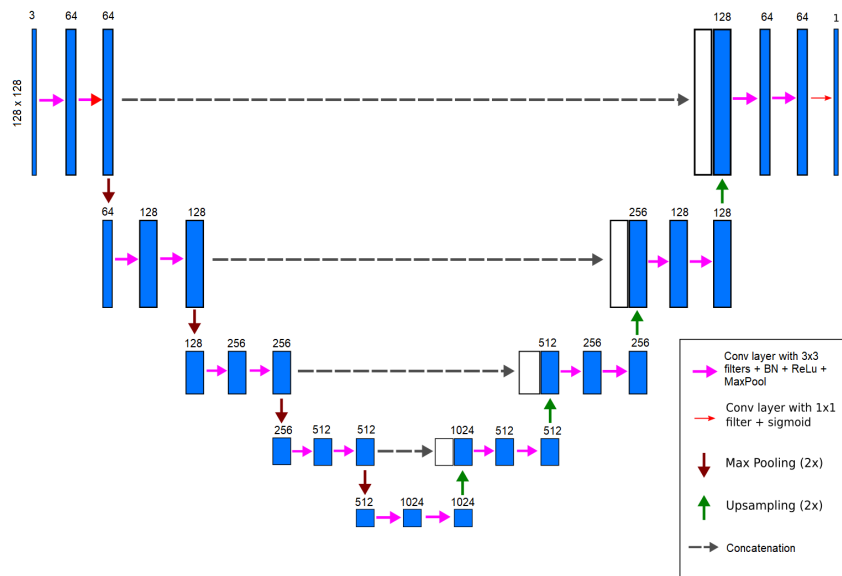
La arquitectura empleada corresponde a una red U-Net basada en la implementación del repositorio Optic-Nerve<sup>12</sup>. En este trabajo se entrenó el modelo desde cero utilizando la base de datos Origa-light, con el propósito de adaptar la red a las características específicas de las imágenes de fondo de retina. En la figura 2 se muestra la arquitectura de la red neuronal. Cabe resaltar que se realizaron modificaciones en la estructura original, particularmente en algunas capas de la red, con el fin de ajustar la profundidad y el número de filtros a los requerimientos del conjunto de datos empleado. En el Anexo 1 se presenta la tabla resumen de la arquitectura implementada, donde se detallan las capas que conforman el modelo y sus respectivas dimensiones de salida.

Esta red presenta una estructura simétrica compuesta por dos partes principales: un codificador (encoder) y un decodificador (decoder), conectados mediante un bloque in-

---

<sup>12</sup> ssykira. *optic-nerve*. Repositorio en GitHub. s.f. URL: <https://github.com/ssykira/optic-nerve> (visitado 21-09-2025).

Figura 2. Arquitectura U-Net implementada para la segmentación del disco óptico, compuesta por una etapa de codificación y una de decodificación con conexiones de concatenación, finalizando con una convolución 1×1 con activación sigmoide para generar el mapa de segmentación.



termedio (bridge). En el encoder, la red extrae características más abstractas mediante capas de convolución seguidas de una normalización (batchnormalization) y funciones de activación ReLU, reduciendo la dimensionalidad a través de capas de max pooling. El bloque bridge actúa como una capa de conexión que consolida la característica de mayor nivel obtenidas por el encoder. El decoder realiza un proceso inverso: utiliza capas de upsampling y deconvolución para recuperar la resolución espacial, combinando las características obtenidas por el encoder, lo que permite conservar la información estructural importante. Adicional a esto, se incorporó un bloque de aumento de datos (*data augmentation*), que aplica transformaciones aleatorias como giros, rotaciones y cambios de contraste, con el fin de mejorar la generalización del modelo y reducir el riesgo de sobreajuste.

### 3.3. Entrenamiento y ajuste de hiperparámetros

El entrenamiento y la evaluación del modelo se realizaron en Google Colab, aprovechando sus recursos de cómputo en la nube (GPU) y su compatibilidad con bibliotecas como TensorFlow y Keras. El modelo fue entrenado utilizando el optimizador Adam, el cual ajusta los pesos de manera adaptativa para acelerar la convergencia y mejorar la estabilidad del proceso de aprendizaje. Como función de pérdida se empleó binary cross-entropy, adecuada para tareas de segmentación binaria, permitiendo distinguir entre las regiones correspondientes al disco óptico y el fondo de la imagen.

Adicionalmente, se incorporó un bloque de aumento de datos (data augmentation), que aplica transformaciones aleatorias como volteos horizontales y verticales, rotaciones, variaciones de zoom y ajustes de contraste. Estas transformaciones se realizaron de manera dinámica durante el entrenamiento. A partir de las 325 imágenes originales destinadas al conjunto de entrenamiento y validación, el modelo procesó aproximadamente 6.500 muestras aumentadas a lo largo de 20 épocas, fortaleciendo su capacidad de generalización y reduciendo el riesgo de sobreajuste.

Para prevenir el sobreajuste, se implementó la técnica de Early Stopping, la cual detiene el entrenamiento cuando la pérdida de validación no presenta mejoras durante cinco épocas consecutivas. Aunque se estableció un máximo de 20 épocas, el entrenamiento se detuvo en la época 16 debido a la activación de este criterio.

La base de datos total fue dividida inicialmente en una proporción 50–50, asignando 325 imágenes al conjunto de prueba (test) y 325 imágenes al conjunto de entrenamiento y validación (train/validation). Posteriormente, este último conjunto se subdividió en una proporción del 80 % para entrenamiento y 20 % para validación.

### 3.4. Medidas de desempeño

Para evaluar el desempeño del modelo de segmentación, se utilizaron diferentes métricas que permiten medir su precisión y capacidad de generación las cuales fueron la exactitud o *accuracy* ( $A$ ), sensibilidad o *sensitivity* ( $Sen$ ), especificidad o *specificity* ( $Spe$ ), intersección sobre unión o *Intersection over Union* ( $IoU$ ) y el coeficiente de Dice o *Dice Coefficient* ( $Dice$ ), que se definen como:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Sen = \frac{TP}{TP + FN} \quad (2)$$

$$Spe = \frac{TN}{TN + FP} \quad (3)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (4)$$

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (5)$$

donde  $TP$  representa los verdaderos positivos (*True Positive*),  $TN$  los verdaderos negativos (*True Negative*),  $FP$  los falsos positivos (*False Positive*) y  $FN$  los falsos negativos (*False Negative*).

## 4. RESULTADOS

En la tabla 1, se ilustran los valores finales de las métricas de desempeño del modelo de U-Net. El modelo alcanzó una exactitud de 0.9912 en el entrenamiento, un 0.9893 en validación y 0.9889 en prueba, se alcanzó un coeficiente de Dice promedio de 0.767, 0.749 y 0.721 respectivamente lo que indica una buena superposición entre la predicción y la máscara del dataset . En la métrica de IoU se obtiene un 0.624, 0.602 y 0.568 respectivamente reflejando una segmentación consistente. Para la sensibilidad se alcanzó un 0.82 para entrenamiento, 0.89 para validación y un 0.74 para prueba y en cuanto a la especificidad un 0.9966, 0.9933 y 0.9976 respectivamente.

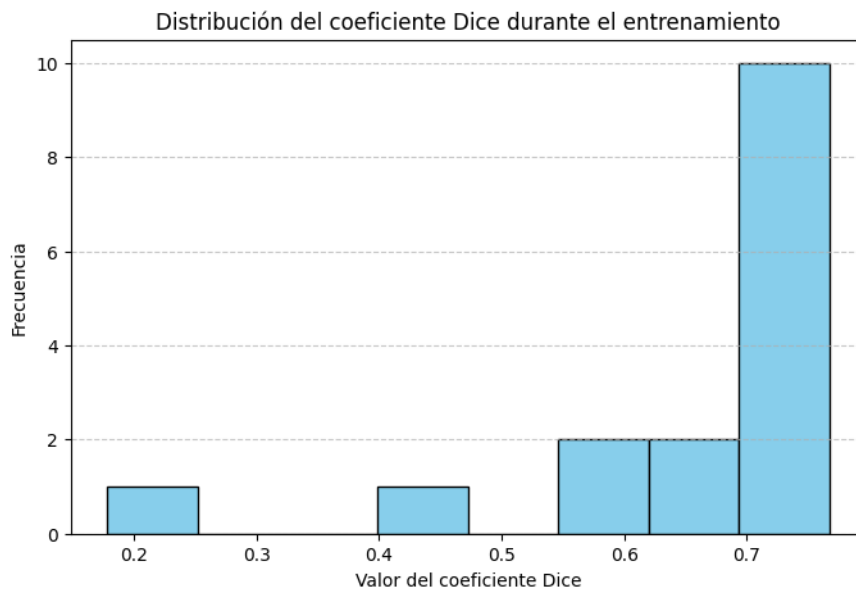
Tabla 1. Resultados de las métricas de desempeño del modelo U-Net durante el entrenamiento, validación y prueba.

<b>Dataset</b>	<b>Acc</b>	<b>Dice</b>	<b>IoU</b>	<b>Sen</b>	<b>Spe</b>	<b>Loss</b>
Entrenamiento	0.9912	0.7673	0.6241	0.8231	0.9966	0.0218
Validación	0.9893	0.7491	0.6029	0.8926	0.9933	0.0269
Prueba (Test)	0.9889	0.7217	0.5684	0.7401	0.9976	0.0277

En la figura 3 se muestra la distribución del coeficiente de Dice alcanzado durante el entrenamiento del modelo. La mayoría de los valores se concentran entre 0.7 y 0.8, lo que indica que el modelo logró un nivel de solapamiento alto entre las segmentaciones predichas y las máscaras reales. Solo unos pocos valores se encuentran por debajo de 0.5, lo que sugiere que el entrenamiento fue consistente y que el modelo aprendió progresivamente a segmentar de manera más precisa.

Los resultados de la segmentación se muestran en la figura 4, el límite del disco óptico detectado por la metodología propuesta se muestra en la tercera imagen donde la región blanca corresponde al disco óptico detectado y en la cuarta imagen se muestra la segmentación final.

Figura 3. Histograma de los coeficientes de Dice. El entrenamiento se programó para 20 épocas; sin embargo, se detuvo en la época 16 debido al criterio de early stopping.



En la Tabla 2 se presentan los resultados de desempeño del modelo propuesto en comparación con otros métodos de segmentación del disco óptico. Aunque la sensibilidad obtenida fue menor respecto a algunos trabajos, las métricas de especificidad y exactitud reflejan un modelo robusto con una buena capacidad para evitar falsos positivos. Cabe aclarar que esta comparación se incluye únicamente con fines de discusión, ya que los experimentos reportados en los diferentes estudios no fueron realizados bajo las mismas condiciones ni con las mismas bases de datos, por lo que los resultados no son estrictamente equivalentes.

Figura 4. Resultados de la segmentación del disco óptico. (a) Imágenes originales provenientes de las bases de datos. (b) Máscaras de referencia proporcionadas por dichas bases. (c) Predicción generada por la red U-Net. (d) Segmentación final obtenida a partir del modelo.

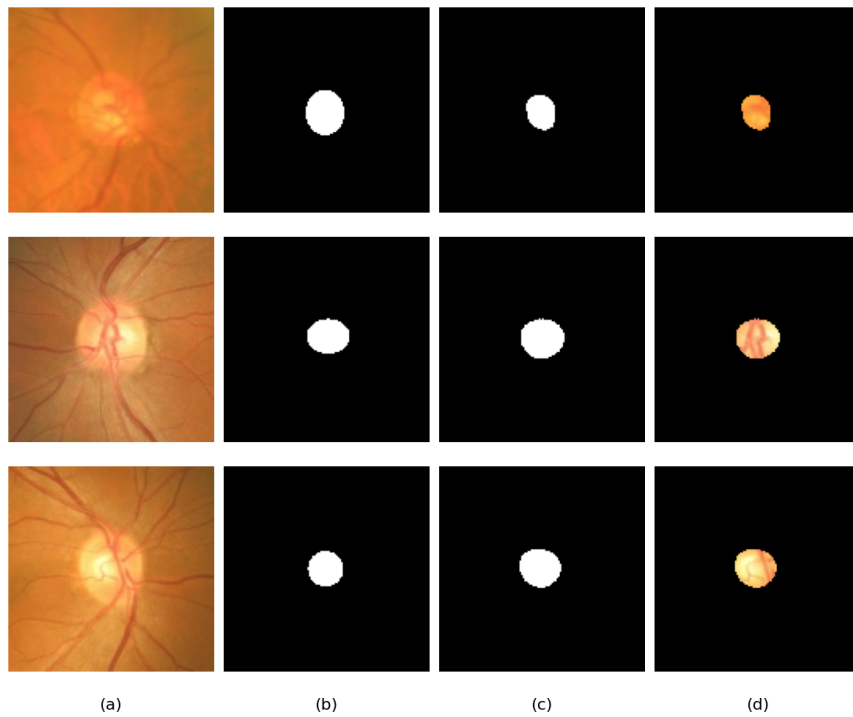


Tabla 2. Comparación de métricas de desempeño con otros métodos de segmentación del disco óptico.

<b>Método</b>	<b>Acc</b>	<b>Sen</b>	<b>Spe</b>	<b>IoU</b>	<b>Dice</b>
Fast Optic Disc <sup>13</sup>	<b>0.9986</b>	0.9384	<b>0.9994</b>	-	-
Disc-Aware <sup>14</sup>	-	0.8478	0.8380	-	-
Optic Disc <sup>15</sup>	-	-	-	<b>0.8900</b>	<b>0.9400</b>
CE-Net <sup>16</sup>	0.9904	<b>0.9800</b>	-	-	-
<b>Nuestro método</b>	0.9893	0.7401	0.9976	0.5684	0.7217

## 5. CONCLUSIÓN

El modelo basado en la arquitectura U-Net demostró un desempeño satisfactorio en la segmentación del disco óptico, obteniendo resultados consistentes en las etapas de entrenamiento y validación. Las métricas de exactitud, sensibilidad, especificidad, intersección sobre la unión (IoU) y coeficiente de Dice evidencian que el modelo logra identificar de manera precisa gran parte de la región correspondiente al disco óptico en imágenes RGB.

Al comparar el desempeño con arquitecturas más complejas, como Fast Optic Disc y CE-Net, se observó que el modelo propuesto alcanza valores de exactitud y especificidad cercanos, aunque presenta una sensibilidad ligeramente inferior. Este comportamiento puede estar asociado a la variabilidad de las imágenes empleadas, la cantidad de datos disponibles y las condiciones de preprocesamiento.

En términos generales, la U-Net se consolida como una alternativa eficiente y equilibrada para este tipo de tareas, al ofrecer un buen rendimiento con un costo computacional moderado. Los resultados obtenidos constituyen una base sólida para el desarrollo de futuras investigaciones orientadas al mejoramiento de modelos de segmentación automática y al apoyo en la detección temprana de enfermedades oculares como el glaucoma.

## 6. RECOMENDACIONES

Se recomienda ampliar el conjunto de datos mediante técnicas de aumento más variadas, con el fin de mejorar la capacidad de generalización del modelo. Asimismo, podría evaluarse el uso de diferentes funciones de pérdida para optimizar la segmentación y aumentar la sensibilidad.

Para trabajos futuros, se sugiere explorar arquitecturas que incorporen mecanismos de atención y validar el modelo en bases de datos adicionales, con el propósito de fortalecer su aplicabilidad en escenarios clínicos reales.

## BIBLIOGRAFÍA

- Cheng, Jun et al. "Superpixel Classification Based Optic Disc and Optic Cup Segmentation for Glaucoma Screening". En: *IEEE Transactions on Medical Imaging* 32.6 (2013), págs. 1019-1032 (vid. pág. 13).
- Fu, Huazhu et al. "Disc-Aware Ensemble Network for Glaucoma Screening From Fundus Image". En: *IEEE Transactions on Medical Imaging* 37.11 (2018), págs. 2493-2501 (vid. pág. 11).
- Fu, Huazhu et al. "Joint Optic Disc and Cup Segmentation Based on Multi-Label Deep Network and Polar Transformation". En: *IEEE Transactions on Medical Imaging* 37.7 (2018), págs. 1597-1605 (vid. pág. 14).
- Fundación Colombiana de Glaucoma. *¿Qué es el glaucoma?* 2024. URL: <http://www.glaucoma.org.co/articulos/que-es-glaucoma> (visitado 21-09-2025) (vid. pág. 9).
- Gu, Zaiwang et al. "CE-Net: Context Encoder Network for 2D Medical Image Segmentation". En: *IEEE Transactions on Medical Imaging* 38.10 (2019), págs. 2281-2292 (vid. pág. 14).
- Instituto Nacional para Ciegos (INCI). *12 de marzo: día del glaucoma, la causa principal de la ceguera*. 2023. URL: <https://www.inci.gov.co/blog/12-de-marzo-dia-del-glaucoma-la-causa-principal-de-la-ceguera> (visitado 21-09-2025) (vid. pág. 9).
- Jiang, Yuming et al. "JointRCNN: A Region-Based Convolutional Neural Network for Optic Disc and Cup Segmentation". En: *IEEE Transactions on Biomedical Engineering* 67.2 (2020), págs. 335-343 (vid. pág. 11).
- Sevastopolsky, A. "Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network". En: *Pattern Recognition and Image Analysis* 27.3 (2017), págs. 618-624 (vid. pág. 14).
- sshikamaru. *Glaucoma Detection*. 2020. URL: <https://www.kaggle.com/datasets/sshikamaru/glaucoma-detection> (visitado 21-09-2025) (vid. págs. 11, 16).

ssykira. *optic-nerve*. Repositorio en GitHub. s.f. URL: <https://github.com/ssykira/optic-nerve> (visitado 21-09-2025) (vid. pág. 16).

Tham, Yih-Chung et al. "Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040: A Systematic Review and Meta-Analysis". En: *Ophthalmology* 121.11 (2014), págs. 2081-2090 (vid. pág. 9).

Zahoor, Muhammad Nauman y Muhammad Moazam Fraz. "Fast Optic Disc Segmentation in Retina Using Polar Transform". En: *IEEE Access* 5 (2017), págs. 12293-12300 (vid. pág. 13).

## ANEXOS

### Anexo A. Resumen de las capas y dimensiones de salida del modelo U-Net.

Se presenta la tabla 3 con el detalle de las capas que conforman la arquitectura del modelo U-Net utilizado en este trabajo. En ella se especifica el tipo de cada capa y sus dimensiones de salida, lo que permite observar cómo la red procesa las imágenes de entrada hasta generar la máscara final de segmentación.

Tabla 3. Resumen de las capas y dimensiones de salida del modelo U-Net.

<b>Layer (type)</b>	<b>Output Shape</b>
InputLayer	(None, 128, 128, 3)
Conv2D	(None, 128, 128, 64)
BatchNormalization	(None, 128, 128, 64)
Activation	(None, 128, 128, 64)
Conv2D_1	(None, 128, 128, 64)
BatchNormalization_1	(None, 128, 128, 64)
Activation_1	(None, 128, 128, 64)
MaxPooling2D	(None, 64, 64, 64)
Conv2D_2	(None, 64, 64, 128)
BatchNormalization_2	(None, 64, 64, 128)
Activation_2	(None, 64, 64, 128)
Conv2D_3	(None, 64, 64, 128)
BatchNormalization_3	(None, 64, 64, 128)
Activation_3	(None, 64, 64, 128)
MaxPooling2D_1	(None, 32, 32, 128)
Conv2D_4	(None, 32, 32, 256)

<b>Layer (type)</b>	<b>Output Shape</b>
BatchNormalization_4	(None, 32, 32, 256)
Activation_4	(None, 32, 32, 256)
Conv2D_5	(None, 32, 32, 256)
BatchNormalization_5	(None, 32, 32, 256)
Activation_5	(None, 32, 32, 256)
MaxPooling2D_2	(None, 16, 16, 256)
Conv2D_6	(None, 16, 16, 512)
BatchNormalization_6	(None, 16, 16, 512)
Activation_6	(None, 16, 16, 512)
Conv2D_7	(None, 16, 16, 512)
BatchNormalization_7	(None, 16, 16, 512)
Activation_7	(None, 16, 16, 512)
MaxPooling2D_3	(None, 8, 8, 512)
Conv2D_8	(None, 8, 8, 1024)
BatchNormalization_8	(None, 8, 8, 1024)
Activation_8	(None, 8, 8, 1024)
Conv2D_9	(None, 8, 8, 1024)
BatchNormalization_9	(None, 8, 8, 1024)
Activation_9	(None, 8, 8, 1024)
Conv2DTranspose	(None, 16, 16, 512)
Concatenate	(None, 16, 16, 1024)
Conv2D_10	(None, 16, 16, 512)
Conv2DTranspose_1	(None, 32, 32, 256)
Concatenate_1	(None, 32, 32, 512)
Conv2D_12	(None, 32, 32, 256)
Conv2DTranspose_2	(None, 64, 64, 128)

<b>Layer (type)</b>	<b>Output Shape</b>
Concatenate_2	(None, 64, 64, 256)
Conv2D_14	(None, 64, 64, 128)
Conv2DTranspose_3	(None, 128, 128, 64)
Concatenate_3	(None, 128, 128, 128)
Conv2D_16	(None, 128, 128, 64)
Conv2D_18	(None, 128, 128, 1)