

Jorge Antonio Nava González

Modelo de aprendizaje profundo para detección de anomalías  
en vasos sanguíneos asociadas al glaucoma en imágenes de  
fondo de ojo

2025



Universidad Autónoma  
de Querétaro  
Facultad de Ingeniería

Modelo de aprendizaje profundo  
para detección de anomalías en  
vasos sanguíneos asociadas al  
glaucoma en imágenes de fondo de  
ojo

Tesis

Que como parte de los requisitos para obtener el  
Grado de

Maestro en

Ciencias en Inteligencia Artificial

Presenta

Jorge Antonio Nava González

Dirigido por:  
Dr. Gendry Alfonso Francia

Querétaro, Qro. a Septiembre 2025

La presente obra está bajo la licencia:  
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>



CC BY-NC-ND 4.0 DEED

Atribución-NoComercial-SinDerivadas 4.0 Internacional

### Usted es libre de:

**Compartir** — copiar y redistribuir el material en cualquier medio o formato

La licenciante no puede revocar estas libertades en tanto usted siga los términos de la licencia

### Bajo los siguientes términos:



**Atribución** — Usted debe dar [crédito de manera adecuada](#), brindar un enlace a la licencia, e [indicar si se han realizado cambios](#). Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante.



**NoComercial** — Usted no puede hacer uso del material con [propósitos comerciales](#).



**SinDerivadas** — Si [remezcla, transforma o crea a partir](#) del material, no podrá distribuir el material modificado.

**No hay restricciones adicionales** — No puede aplicar términos legales ni [medidas tecnológicas](#) que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia.

### Avisos:

No tiene que cumplir con la licencia para elementos del material en el dominio público o cuando su uso esté permitido por una [excepción o limitación](#) aplicable.

No se dan garantías. La licencia podría no darle todos los permisos que necesita para el uso que tenga previsto. Por ejemplo, otros derechos como [publicidad, privacidad, o derechos morales](#) pueden limitar la forma en que utilice el material.



**Universidad Autónoma de  
Querétaro**

**Facultad de Ingeniería**

**Maestría en Ciencias en Inteligencia  
Artificial**

**Modelo de aprendizaje profundo para detección de  
anomalías en vasos sanguíneos asociadas al  
glaucoma en imágenes de fondo de ojo**

**Tesis**

Que como parte de los requisitos para obtener el Grado de

**Maestro en Ciencias en Inteligencia Artificial**

Presenta

**Jorge Antonio Nava González**

Dirigido por

**Dr. Gendry Alfonso Francia**

Dr. Gendry Alfonso Francia  
Presidente

Dr. Saúl Tovar Arriaga  
Secretario

Dr. Jesús Carlos Pedraza Ortega  
Vocal

Dr. Andras Takacs  
Suplente

Dr. Efrén Gorrostieta Hurtado  
Suplente

Centro Universitario, Querétaro, Qro. México  
Fecha de aprobación por el Consejo Universitario (Septiembre 2025)

Este trabajo está dedicado a la maestra Elia De Asís Valle Díaz, mi Mamá Elia, una mujer fuerte, valiente y luchadora, cuyo amor y ejemplo inspirador han marcado mi vida. Su fortaleza y cariño permanecerán siempre como pilar e inspiración en mi camino.

# Agradecimientos

Primero quiero agradecer a mis padres, Jorge Antonio Nava de los Reyes y Elia Natividad González Valle, por su guía, amor, confianza y apoyo, siendo pilares y ejemplo fundamental en mi formación personal y académica. A mis hermanas, Elia Beatriz y Socorro Natividad, por su paciencia, cariño y por apoyarme incondicionalmente en todo momento, brindándome su acompañamiento y respaldo en cada etapa de este camino.

Un agradecimiento especial a mi madrina, Betty, y a mi padrino, Albino, quienes han sido personas de gran importancia en mi vida, siempre presentes para brindarme orientación, apoyo y enseñanzas de valor incalculable. De igual forma, a sus hijos, Héctor Albino y Juan Pablo, quienes para mí son como hermanos, por ser un modelo de constancia, disciplina y éxito que me impulsa a seguir creciendo.

A mis abuelos, tíos y primos, por la dicha de formar parte de una familia unida, llena de amor, apoyo y enseñanzas que me acompañan siempre.

Mi más sincero agradecimiento a mi director de tesis, el Dr. Gendry Alfonso Francia, por su valiosa orientación, paciencia y apoyo constante en este proceso. Sus enseñanzas no solo enriquecieron esta investigación, sino que su ejemplo de compromiso y excelencia ha sido una inspiración determinante en mi crecimiento profesional y personal.

Agradezco profundamente a los miembros de mi comité de tesis, el Dr. Saúl Tovar, el Dr. Jesús Pedraza, el Dr. Andras Takacs y el Dr. Efrén Gorrostieta, así como a todos los profesores que formaron parte de esta etapa, por su tiempo, dedicación y enseñanzas, que fueron fundamentales para la realización y culminación de esta investigación.

Agradezco a la Universidad Autónoma de Querétaro por brindarme los recursos y el espacio necesarios para la realización de esta investigación, así como a los médicos del Instituto Mexicano de Oftalmología por su apoyo y colaboración durante el desarrollo de este trabajo.

Mi sincero agradecimiento al SECIHTI por el apoyo económico brindado a través de la beca, que me permitió concentrarme totalmente en el desarrollo de esta investigación.

# Resumen

El glaucoma es un grupo de enfermedades neurodegenerativas que causan un daño progresivo a las células ganglionares de la retina y sus axones, lo que conduce a una pérdida irreversible de la visión. Una de las principales herramientas de diagnóstico es la fotografía del fondo de ojo, que permite identificar alteraciones en estructuras anatómicas clave. Una de estas alteraciones es la nasalización de los vasos sanguíneos centrales, cuya presencia está estrechamente relacionada con la progresión del deterioro del campo visual. A pesar de su relevancia clínica, la nasalización ha sido un biomarcador relativamente poco explorado, ya que se considera indirecto. En este trabajo se presenta, por primera vez, la aplicación de la inteligencia artificial para segmentar y clasificar el grado de nasalización en imágenes del fondo de ojo. Se seleccionó la arquitectura YOLOv8 para la tarea de segmentación debido a su velocidad y facilidad de implementación, lo que arrojó excelentes resultados en la detección de las regiones nasal y temporal del disco óptico. Estas regiones se consideran de interés, ya que forman parte de la fórmula propuesta para cuantificar el grado de desplazamiento nasal. Posteriormente, se utilizó ResNet50 como modelo de clasificación debido a su uso generalizado y su robustez en el procesamiento de imágenes médicas, lo que permitió clasificar las imágenes como nasalizadas y no nasalizadas con una precisión constante. Otra contribución destacable fue la creación de una base de datos con imágenes nasalizadas y no nasalizadas, así como máscaras correspondientes a las regiones nasal y temporal del disco óptico. Este enfoque sirve como punto de partida para futuros trabajos sobre la nasalización. Tiene el potencial de convertirse en una herramienta valiosa para evaluar la eficacia del tratamiento, ya que permite una observación objetiva de la progresión de la enfermedad.

**Palabras clave:** Glaucoma, Nasalización, Fondo de Ojo, Inteligencia Artificial, Aprendizaje Profundo.

# Abstract

Glaucoma is a group of neurodegenerative diseases that cause progressive damage to the retinal ganglion cells and their axons, leading to irreversible vision loss. One of the primary diagnostic tools is fundus photography, which enables the identification of alterations in key anatomical structures. One of these alterations is the nasalization of the central blood vessels, whose presence is closely related to the progression of visual field deterioration. Despite its clinical relevance, nasalization has been a relatively underexplored biomarker, as it is considered indirect. This paper presents, for the first time, the application of artificial intelligence for segmenting and classifying the degree of nasalization in fundus images. The YOLOv8 architecture was selected for the segmentation task due to its speed and ease of implementation, yielding excellent results in detecting the nasal and temporal regions of the optic disc. These regions are considered of interest as they form part of the proposed formula for quantifying the degree of nasal displacement. Subsequently, ResNet50 was utilized as a classification model due to its widespread use and robustness in medical image processing, which enabled images to be categorized as nasalized and non-nasalized with consistent accuracy. Another notable contribution was the construction of a database containing photos of nasalized and non-nasalized images, as well as masks corresponding to the nasal and temporal regions of the optic disc. This approach serves as a starting point for future work on nasalization. It has the potential to become a valuable tool for evaluating treatment effectiveness, allowing for an objective observation of disease progression.

**Keywords:** Glaucoma, Nasalization, Fundus Photography, Artificial Intelligence, Deep Learning

# Contenido

Índice de Figuras .....	1
Índice de Tablas.....	2
1. Introducción .....	3
1.1. Glaucoma .....	3
1.2. Prevalencia Global y en México .....	3
1.3. Diagnóstico .....	4
1.4. Nasalización de los Vasos Sanguíneos .....	5
1.5. Inteligencia Artificial en Oftalmología .....	7
1.6. Justificación.....	7
1.7. Planteamiento del Problema.....	9
2. Antecedentes .....	10
3. Marco Teórico .....	15
3.1. Tipos de Glaucoma.....	15
3.2. El Ojo Humano .....	16
3.3. Retina .....	16
3.3.1. Disco Óptico o Papila Óptica.....	17
3.3.2. Vasos Sanguíneos Retinianos.....	18
3.4. Desplazamiento Nasal e Índice de Nasalización.....	18
3.5. Fotografía de Fondo de Ojo .....	20
3.6. Aprendizaje Automático.....	20
3.7. Aprendizaje Profundo .....	21
3.7.1. Redes Neuronales Convolucionales.....	22
3.7.2. Redes Neuronales Residuales .....	23
3.8. Modelos de Detección de Objetos y Segmentación.....	23
3.8.1. Detección de Objetos .....	23
3.8.2. Segmentación Semántica .....	23
3.8.3. Segmentación de Instancias .....	24
3.9. YOLO.....	24



3.9.1.	YOLOv8.....	24
3.10.	Modelos de Clasificación.....	25
3.10.1.	ResNet50.....	25
3.11.	Técnicas de Preprocesamiento .....	25
3.11.1.	Ecualización de Histograma Adaptativo con Contraste Limitado .....	25
3.11.2.	Filtro B-COSFIRE .....	26
3.12.	Métricas de Evaluación .....	26
3.12.1.	Matriz de Confusión.....	26
3.12.2.	Exactitud ( <i>accuracy</i> ) .....	27
3.12.3.	Precisión.....	27
3.12.4.	Sensibilidad ( <i>Recall</i> ) .....	27
3.12.5.	Especificidad .....	27
3.12.6.	Puntuación F1 ( <i>F1-Score</i> ).....	28
4.	Hipótesis.....	29
5.	Objetivos .....	30
5.1.	Objetivo General .....	30
5.2.	Objetivos Específicos.....	30
6.	Material y Metodología.....	31
6.1.	Python .....	31
6.2.	Google Colaboratory .....	31
6.3.	PyTorch .....	31
6.4.	OpenCV .....	32
6.5.	Roboflow.....	32
6.6.	Hardware .....	33
6.7.	Metodología .....	33
6.8.	Investigación del Estado del Arte.....	33
6.9.	Búsqueda y selección de Base de Datos .....	34
6.10.	Preprocesamiento .....	34
6.10.1.	Canal Verde .....	34

6.10.2.	Filtro CLAHE .....	35
6.10.3.	Filtro B-COSFIRE .....	35
6.10.4.	Extracción de Región de Interés .....	35
6.10.5.	Combinación de Imágenes .....	36
6.11.	Etiquetado .....	37
6.12.	Segmentación .....	38
6.13.	Base de Datos de Nasalización .....	39
6.13.1.	Índice Nasalización Propuesto .....	40
6.13.2.	Formación de Base de Datos.....	41
6.14.	Clasificación.....	42
7.	Resultados y Discusión .....	43
7.1.	YOLOv8.....	43
7.2.	ResNet50 .....	44
7.3.	Discusión General .....	48
8.	Conclusiones .....	50
	Referencias Bibliográficas .....	52
	Anexo I .....	57
	Anexo II .....	59

# Índice de Figuras

Figura 1. Imágenes de fondo de ojo, muestran la diferencia entre un ojo normal (sano) y uno con desplazamiento nasal de los VSC.....	6
Figura 2. Representación esquemática del desplazamiento nasal del vaso retiniano central sobre el nervio óptico en pacientes con glaucoma [12]. ....	6
Figura 3. Visualización de la retina y sus principales estructuras en el fondo del ojo. ....	17
Figura 4. Imagen de fondo de ojo que muestra las distancias necesarias para el cálculo del NI y SI propuestos por Kilhwan Shon. ....	19
Figura 5. Diagrama de la metodología realizada. ....	33
Figura 6. Combinación para la imagen final. ....	36
Figura 7. Interfaz de Roboflow para etiquetado.....	37
Figura 8. Criterios de inclusión-exclusión. ....	38
Figura 9. Criterios de exclusión. ....	39
Figura 10. Correlación de SI (índice de desplazamiento nasal). ....	41
Figura 11. Base de datos de nasalización, cuenta con tres clases. ....	42
Figura 12. Matriz de confusión de la configuración óptima (1800 imágenes). ....	46
Figura 13. Gráficas de pérdida y validación (1800 imágenes).....	47
Figura 14. Matriz de confusión con los mismos hiperparámetros de la configuración óptima (2100 imágenes). ....	48
Figura 15. Gráficas de pérdida y validación (2100 imágenes).....	48

## Índice de Tablas

Tabla 1. Pruebas clínicas para el diagnóstico y seguimiento en oftalmología .....	4
Tabla 2. Estado del Arte .....	11
Tabla 3. Estado del Arte sobre segmentación de vasos sanguíneos de la retina. ....	13
Tabla 4. Resultado de YOLOv8 para segmentación de instancias.....	43
Tabla 5. Métricas 80/10/10.....	45
Tabla 6. Métricas 60/20/20.....	45
Tabla 7. Métricas del conjunto de datos ampliado.....	46

# 1. Introducción

## 1.1. Glaucoma

El glaucoma abarca un conjunto de neuropatías ópticas que se caracteriza por la progresiva degeneración de las células ganglionares de la retina (CGR) y sus axones, en muchos casos debido al incremento de la presión intraocular. El nervio óptico está formado por los axones de las neuronas CGR, cuyos cuerpos se localizan en la retina. La pérdida de las CGR y sus axones conduce a cambios estructurales característicos en la apariencia de la cabeza del nervio óptico, y consecuentemente, en una pérdida irreversible del campo visual, que normalmente inicia en la periferia [1]. Conocido como “el ladrón silencioso de la vista” por su naturaleza insidiosa, el glaucoma suele ser asintomático en etapas tempranas, de modo que cuando el paciente percibe alteraciones en el campo visual, el daño al nervio óptico ya es significativo [2].

Uno de los principales factores de riesgo es la presión intraocular (PIO). Este aumento de la PIO se da por el mal drenaje del humor acuoso, ejerciendo un estrés mecánico sobre la cabeza del nervio óptico. Sin embargo, el glaucoma puede desarrollarse y progresar con niveles de PIO normales (10 y 21 mmHg) [3]. Esto resalta la importancia de investigar biomarcadores independientes de la PIO, como la relación copa-disco, muescas en el anillo neuroretiniano (ANR), defectos en la capa de fibras nerviosas de la retina (CFNR), atrofas peripapilares (APP), exposición de la lámina cribosa y las anomalías vasculares, para mejorar la comprensión de la enfermedad y optimizar el diagnóstico.

## 1.2. Prevalencia Global y en México

A nivel global, el glaucoma es la causa más frecuente de pérdida irreversible de la vista, con más de 70 millones de personas afectadas [4]. Se estima que para el año 2040 más de 112 millones de personas padezcan glaucoma [5]. En México, se calcula que 1.5 millones de personas sufren glaucoma [6]. La prevalencia de la enfermedad aumenta significativamente con la edad [7].

La prevalencia varía significativamente entre los distintos grupos étnicos. El tipo más frecuente de glaucoma, el glaucoma primario de ángulo abierto, se presenta con mayor frecuencia en personas de ascendencia africana. El riesgo de glaucoma primario de ángulo cerrado presenta un mayor riesgo en mujeres y aumenta con la edad, con una prevalencia superior en personas de origen asiático [7]. Por último, el glaucoma de tensión normal es más frecuente en la población japonesa y está presente en poblaciones indígenas americanas y nativas de Alaska [8].

### 1.3. Diagnóstico

El diagnóstico del glaucoma se fundamenta en la identificación de una neuropatía óptica progresiva, que se caracteriza por la presencia de anomalías en el campo visual y alteraciones estructurales en el disco óptico y la capa de fibras nerviosas de la retina. Aunque con frecuencia se observa una presión intraocular elevada, no siempre está presente. No hay una sola prueba definitiva, por lo que el diagnóstico se apoya en una exploración clínica e integración de múltiples hallazgos [9].

Debido a que los cambios funcionales en la visión por glaucoma tienden a aparecer en fases avanzadas, la detección temprana del daño glaucomatoso puede ser un desafío si solo se recurre a la evaluación del campo visual. Se vuelve imprescindible, por tanto, la aplicación de múltiples pruebas para identificar alteraciones tanto estructurales como funcionales del nervio óptico. Los métodos actuales para el diagnóstico de glaucoma se muestran en la Tabla 1. [7], [10], [11].

*Tabla 1. Pruebas clínicas para el diagnóstico y seguimiento en oftalmología*

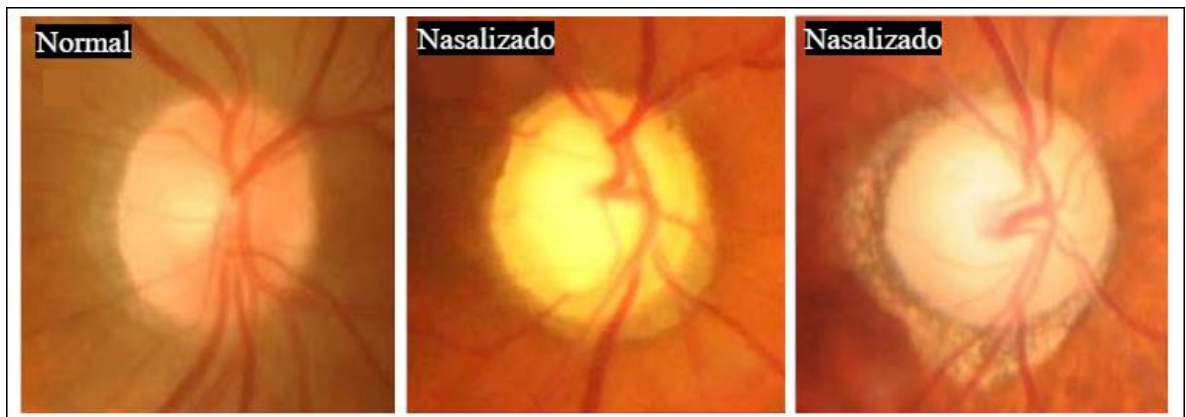
<b>Método</b>	<b>Descripción</b>	<b>Utilidad Principal</b>
Tonometría	Mide la presión intraocular.	Detección de PIO elevada.
Oftalmoscopia /Fondo de Ojo	Permite observar el nervio óptico, el ANR y la CFNR.	Evaluación cualitativa del daño estructural del nervio óptico.

Perimetría	Mide la sensibilidad visual a la luz en diferentes puntos del campo visual.	Evaluación funcional del daño glaucomatoso.
Gonioscopia	Visualiza el ángulo iridocorneal con una lente especial.	Diferencia entre GPAA y GPAC.
Paquimetría	Mide el grosor corneal central.	Ajusta la interpretación de la PIO.
Tomografía de Coherencia Óptica (OCT)	Técnica de imagen no invasiva que cuantifica estructuras como la CFNR, ANR y el complejo de células ganglionares.	Evaluación estructural objetiva del daño glaucomatoso.

Pese a la disponibilidad de diversas herramientas de diagnóstico, la detección de glaucoma representa un reto clínico considerable. La interpretación de la oftalmoscopia depende significativamente del criterio individual del oftalmólogo, lo que conlleva una alta variabilidad interobservador [11]. La disponibilidad limitada de tecnologías diagnósticas como la OCT y los perímetros computarizados en áreas rurales o de bajos recursos representa una barrera importante para la detección oportuna y el monitoreo adecuado del glaucoma.

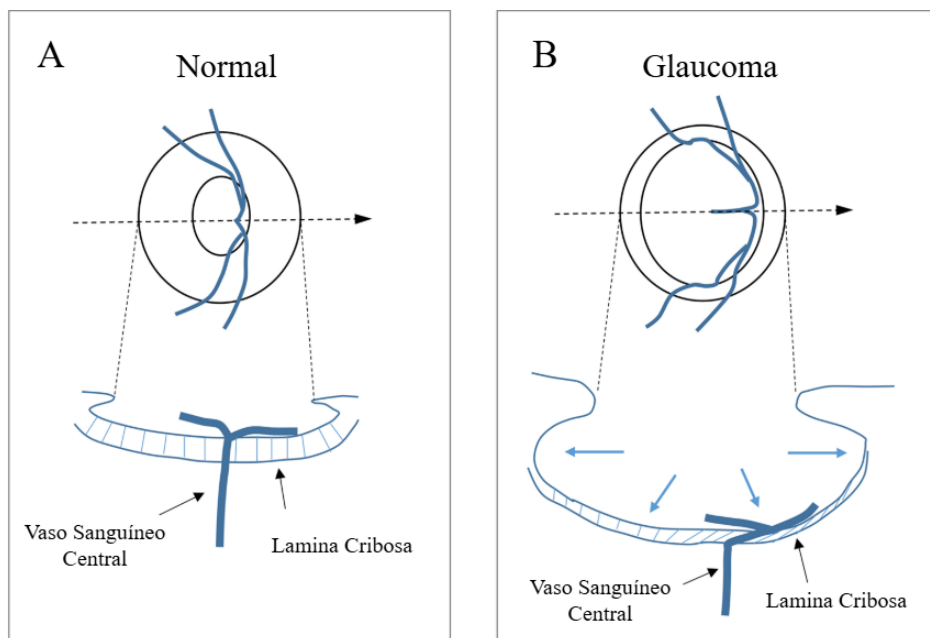
#### 1.4. Nasalización de los Vasos Sanguíneos

La nasalización se define como el desplazamiento de los vasos sanguíneos centrales (VSC) de la retina hacia el lado nasal de la superficie del disco óptico por una deformación en la lámina cribosa. La evidencia sugiere que esta deformación es resultado del estrés biomecánico prolongado generado por una PIO elevada [12].



*Figura 1. Imágenes de fondo de ojo, muestran la diferencia entre un ojo normal (sano) y uno con desplazamiento nasal de los VSC.*

La nasalización de los VSC se considera un biomarcador indirecto del glaucoma, ya que no existe suficiente evidencia que confirme su desplazamiento nasal ni su relevancia diagnóstica. Sin embargo, la nasalización está altamente asociada con la gravedad de los defectos en el campo visual y se manifiesta en fases avanzadas de la enfermedad. Esto sugiere que la nasalización no es un hallazgo incidental, sino un indicador de la magnitud del daño funcional [12].



*Figura 2. Representación esquemática del desplazamiento nasal del vaso retiniano central sobre el nervio óptico en pacientes con glaucoma [12].*



## 1.5. Inteligencia Artificial en Oftalmología

El campo de la informática conocido como Inteligencia Artificial (IA) se dedica a crear máquinas capaces de realizar operaciones como el aprendizaje, la resolución de problemas, el reconocimiento de patrones y la toma de decisiones, que a menudo requieren el intelecto humano [13]. Tanto la IA como el aprendizaje profundo (*Deep Learning*) están emergiendo como tecnologías poderosas con el potencial de transformar el diagnóstico y manejo de diversas enfermedades, incluidas las patologías oftálmicas. Dado su potencial para analizar grandes cantidades de imágenes y reconocer patrones difíciles de identificar por el ojo humano, estas técnicas son idóneas para su aplicación en oftalmología, donde la imagenología es una herramienta diagnóstica fundamental.

Las aplicaciones de la IA en oftalmología se han centrado principalmente en enfermedades de alta prevalencia, en las que la detección temprana y el manejo oportuno son fundamentales. Estas afecciones incluyen glaucoma, degeneración macular relacionada con la edad y retinopatía diabética [14]. En el caso del glaucoma, su curso progresivo y asintomático en etapas tempranas ha motivado un interés considerable en su estudio, dada la importancia de una detección temprana para evitar daños visuales irreversibles.

El aprendizaje profundo se ha consolidado como una técnica eficaz para identificar, segmentar y clasificar biomarcadores relacionados con el glaucoma en imágenes obtenidas mediante fondo de ojo y OCT. La relación copa-disco, las atrofas peripapilares, las muescas, las hemorragias y cambios en la capa de fibras nerviosas de la retina constituyen biomarcadores clave en el diagnóstico del glaucoma. La identificación de estos signos mediante técnicas de aprendizaje profundo ha permitido una clasificación eficaz entre ojos sanos y glaucomatosos [3].

## 1.6. Justificación

El glaucoma es la principal causa de pérdida de visión permanente en el mundo, se proyecta que para el año 2040 esta enfermedad afectará aproximadamente a más de 112 millones de personas, representando un grave problema socioeconómico [5]. En México,

aproximadamente el cuatro por ciento de la población sufre de glaucoma, y debido a la falta de síntomas, la mitad de estas personas desconoce que lo tiene [15].

Teniendo en cuenta que el glaucoma progresa lentamente y que puede transcurrir hasta 15 años antes de que se descubra daño en el nervio óptico, es fundamental que las personas a partir de los 35 años se sometan a revisiones anuales con el oftalmólogo para una detección temprana. Sobre todo, aquellas personas que tienen familiares directos que padecen glaucoma o miopía [15].

La Clínica de Glaucoma del Hospital General de México diagnostica más de 1500 casos al año. En el IMSS, los servicios de oftalmología están a cargo de médicos capacitados para detectar oportunamente esta enfermedad, a través de diversos estudios, como el monitoreo de la presión intraocular y prueba de perimetría automatizada, entre otras [15].

El diagnóstico del glaucoma suele requerir tiempo, ya que la interpretación de los biomarcadores en las imágenes de fondo de ojo puede variar entre especialistas. Por esta razón, se recurre a un conjunto de pruebas diagnósticas que permiten caracterizar la morfología ocular y evidenciar posibles alteraciones compatibles con la enfermedad.

La aplicación de técnicas de IA en el campo de la oftalmología se presenta como una herramienta útil y de gran beneficio para detectar anomalías relacionadas con el glaucoma en imágenes de fondo de ojo. Otros estudios han demostrado que el uso de estas técnicas ayuda a obtener diagnósticos precisos, reduciendo las posibles disparidades en la interpretación médica y ahorrando tiempo. El diagnóstico temprano permite iniciar un tratamiento oportuno, previniendo el deterioro progresivo de la enfermedad y protegiendo al paciente de una pérdida visual permanente.

La nasalización de los vasos sanguíneos en el disco óptico se ha asociado con etapas avanzadas de glaucoma y se correlaciona significativamente con la severidad del daño glaucomatoso y la pérdida del campo visual, lo que subraya su importancia clínica. No obstante, su evaluación sigue siendo subjetiva y no estandarizada. El desarrollo de un sistema automatizado de aprendizaje profundo para la detección y cuantificación de esta alteración tiene el potencial de ofrecer contribuciones significativas al manejo del glaucoma.

## 1.7. Planteamiento del Problema

El glaucoma es un trastorno ocular progresivo que causa daño al nervio óptico, entre sus complicaciones se encuentra la nasalización de los VSC de la retina, indicativa de un control insuficiente de la enfermedad. Es fundamental evaluar estas anomalías a tiempo para evitar que el daño en la vista sea mayor.

La valoración de la nasalización vascular en la práctica clínica habitual se basa en la observación directa del fondo de ojo por un oftalmólogo. Sin embargo, este enfoque presenta una alta variabilidad entre observadores, ya que la interpretación de la nasalización de los VSC y su grado depende mucho de la experiencia, el entrenamiento y la percepción del profesional. El aprendizaje profundo puede ofrecer una alternativa prometedora al permitir una evaluación automatizada, objetiva y con mayor precisión.

La mayoría de los trabajos disponibles sobre el uso del aprendizaje profundo en la identificación de glaucoma en imágenes de fondo de ojo no consideran características como el desplazamiento nasal de los VSC, en su lugar, se centran en la relación copa-disco, las muescas en el ANR, cambios morfológicos en la CFNR, hemorragias en disco y APP. Debido a esta falta de consideración, existe una escasez de información disponible, e incluso la nasalización de los vasos sanguíneos no se consideraba como un biomarcador (indirecto) asociado al glaucoma hasta hace poco.

La mayoría de los métodos de aprendizaje profundo desarrollados para la segmentación de vasos sanguíneos se han empleado principalmente en diagnósticos de enfermedades diferentes al glaucoma. Estos métodos han demostrado ser útiles en la segmentación tanto de vasos sanguíneos gruesos como de los vasos sanguíneos más delgados. A pesar de estos avances, la identificación precisa de todos los vasos sanguíneos en el disco óptico sigue siendo un desafío para la mayoría de los métodos existentes; la precisión en la identificación de estos es de particular importancia, ya que sus anomalías están estrechamente asociadas al glaucoma y constituyen un foco de interés en este estudio.

La nasalización de los VSC es un biomarcador indirecto del glaucoma, su desplazamiento está fuertemente vinculado al grado de pérdida de visión y progreso de la enfermedad. Sin

embargo, a pesar de su relevancia, ningún trabajo revisado en el estado del arte ha abordado las características del desplazamiento de los vasos sanguíneos hacia la parte nasal para desarrollar un método de identificación específico. Esta ausencia puede atribuirse a que los estudios sobre nasalización se han retomado recientemente y a la falta de bases de datos de imágenes de fondo de ojo etiquetadas, donde sea claro un patrón de desplazamiento de dichos vasos.

## 2. Antecedentes

En los últimos años se han publicado un amplio número de trabajos donde se hace uso de técnicas de IA en el área médica, destacando especialmente el uso del aprendizaje profundo para el análisis de imágenes médicas, con aplicaciones relevantes en oftalmología.

Los métodos de aprendizaje profundo tienen la capacidad de aprender, de manera autónoma, características jerárquicas a partir de grandes conjuntos de imágenes, permitiendo la automatización de tareas complejas como la detección, clasificación y segmentación de lesiones o alteraciones patológicas. Esta identificación de patrones glaucomatosos facilita la clasificación entre ojos sanos y enfermos en imágenes de fondo de ojo. La mayoría de estos patrones están vinculados a biomarcadores como la relación copa-disco (RCD), la atrofia de la CFNR, APP y las muescas en el ANR.

La aplicación de la IA en la oftalmología no solo incrementa la precisión diagnóstica, sino que también contribuye a un mayor entendimiento de la progresión de la enfermedad. Los estudios más recientes sobre modelos de aprendizaje profundo aplicados a imágenes del fondo de ojo evidencian una adopción cada vez mayor en la identificación automatizada de anomalías patológicas. En la tabla 2 se presentan estudios que utilizan IA para la identificación del glaucoma en imágenes de fondo de ojo. Estos trabajos se enfocan en patrones derivados de las características de los defectos desarrollados en el nervio óptico debido al glaucoma.

*Tabla 2. Estado del Arte*

<b>Autor/ Año</b>	<b>Biomarcador</b>	<b>Método</b>	<b>Base de Datos</b>	<b>Métricas</b>
Zhixi Li et al. 2018	RCD Atrofia de la CFNR	CNN Inception-V3	LabelMe 48,116	AUC = 0.986 Presición = 0.929 Recall = 0.956 EP = 0.92
Naoto Shibata et al. 2018	RCD Muecas y adelgazamiento del ANR	ResNet	Propia 3,132	AUC = 0.965
Sonia Phene et al. 2019	RCD Adelgazamiento del ANR Defectos de la CFNR	CNN Inception-v3	EyePACS Inoveon AREDS UK Biobank 86,618	AUC = 0.945 0.855 0.881
Ajitha S. et al. 2021	RCD Atrofia de la CFNR	CNN	Públicos 1113	Recall = 0.8542 EP = 1 Precisión = 0.938
Aaron Brown et al. 2023	Hemorragias en disco	Detección de objetos (YOLOv4)	NYEE-MSSM 955 GONE 40 OHTS 206 TES 21 MEE 449	Recall = 70% EP = 94.3%

A diferencia de otros estudios en su tiempo, Zhixi Li et al. (2018) emplearon una amplia cantidad de imágenes de fondo de ojo para entrenar y validar su algoritmo (aprox. 39,745). Utilizó la arquitectura Inception-v3, una red neuronal convolucional que permite el aprendizaje de características en imágenes debidamente etiquetadas de manera predictiva. Cabe mencionar que no está claro cuáles fueron las características específicas que el algoritmo identificó y aprendió en relación con los cambios morfológicos en el disco óptico y la CFNR [16].

Naoto Shibata et al. (2018) desarrollaron un algoritmo de aprendizaje profundo residual (Deep Residual Learning) para el reconocimiento de imágenes, conocido como ResNet. Este algoritmo identificó características en imágenes glaucomatosas, tales como la relación copa-disco, atrofia peripapilar, muecas en el ANR y hemorragias [17].

Sonia Phene et al. (2019) utilizaron una red neuronal convolucional con la arquitectura Inception-v3. El algoritmo fue desarrollado y entrenado en TensorFlow, logrando predecir diferentes anomalías en la cabeza del nervio óptico, como la relación copa-disco, atrofia en la CFNR, muescas en el ANR y vasos circumlineares expuestos [18].

Ajitha S. et al. (2021) desarrollaron un algoritmo empleando redes neuronales convolucionales para la detección de distintas características. Una notable ventaja de su estudio radica en el uso de bases de datos públicas y el entorno de desarrollo de Google Colab, evitando así la necesidad de instalar paquetes adicionales. El algoritmo logró alcanzar una especificidad del 100%, esto significa que detecta correctamente ojos sanos (resultado verdadero negativo) [19].

Aaron B. et al. (2023) desarrollaron un sistema automatizado de aprendizaje profundo (Deep Learning) para encontrar hemorragias en disco óptico en imágenes de fondo de ojo. Este destaca que un modelo de detección de objetos demostró un mejor rendimiento en comparación con un modelo de clasificación binaria de imágenes. Para la implementación del modelo de detección de objetos, se optó por la arquitectura YOLOv4 debido a su gran desempeño y velocidad en la detección de objetos [20].

Una etapa fundamental de este trabajo fue el preprocesamiento de las imágenes de fondo de ojo. Se realizó una investigación exhaustiva de las mejores técnicas de preprocesamiento y segmentación de los VSC de la retina, ya que estas estructuras son de gran interés clínico y anatómico. Los VSC dividen el fondo de ojo en dos regiones clave, la región nasal y la temporal. Los estudios abordados aplican métodos de aprendizaje profundo para realizar la segmentación de estos vasos, haciendo posible un análisis más detallado de estos. En la Tabla 3 se presentan algunos de estos trabajos, detallando el preprocesamiento utilizado para incrementar la calidad visual de las imágenes y resaltar las estructuras anatómicas de interés.

Tabla 3. Estado del Arte sobre segmentación de vasos sanguíneos de la retina.

Autor	Preprocesamiento	Método	Base de Datos	Precisión	Sensibilidad	Especificidad
Zhexin Jiang et al. 2017	Remoción de fondo negro Método de Otsu Umbralización	Transformación <i>Top-Hat</i>	DRIVE STARE	0.9596 0.9579	0.8375 0.7787	0.9694 0.9705
Zhexin Jiang et al. 2018	Remoción de fondo negro Método Otsu Filtro Gaussiano	<i>Fully Convolutional Network</i>	DRIVE STARE CHASE_DB1 HRF	0.9624 0.9734 0.9668 0.9650	0.7540 0.8352 0.8640 0.8010	0.9825 0.9846 0.9745 0.8010
Zengginan g Yan et al. 2018	Umbralización Método “ <i>Skeletonization</i> ”	<i>Deep Learning Framework (Caffe)</i>	DRIVE STARE CHASE_DB1	0.9538 0.9638 0.9607	0.7631 0.7735 0.7641	0.9820 0.9857 0.9806
Xing Yang et al. 2021	Umbralización CLAHE	<i>Convolutional Neural Network (CNN) Cascade Forest</i>	DRIVE STARE CHASE_DB1	0.9531 0.9611 0.9559	0.8206 0.8762 0.8235	0.9728 0.9713 0.9711
Shuang Xu et al. 2021	Umbralización Filtro Gaussiano Ecuación de Histograma Local Corrección Gamma	ResNet-CNN	DRIVE STARE	0.9590 0.9688	0.8320 0.8295	0.9885 0.9785
Wenjing Li et al. 2022	Extracción canal verde CLAHE Umbralización Canal L CIELab	B-COSFIRE	DRIVE STARE	0.9604 0.9558	0.7339 0.8003	0.9847 0.9705

Zhexin Jiang et al. (2017) propusieron un sistema de segmentación de vasos sanguíneos aplicando el método de procesamiento morfológico y el algoritmo de filtro adaptado, mencionando la importancia de considerar que incluso el vaso sanguíneo más pequeño puede marcar la diferencia en un diagnóstico médico. Por lo tanto, la extracción precisa de los vasos sanguíneos es fundamental. En este sentido, el sistema propuesto segmenta de forma separada las vénulas y los capilares, para más tarde superponer las imágenes y eliminar el ruido de estas [21].

Zhexin Jiang et al. (2018) propusieron un método supervisado para la segmentación de vasos sanguíneos retinales utilizando una red totalmente convolucional preentrenada mediante aprendizaje por transferencia. Subrayaron la importancia de una extracción de vasos

extremadamente precisa para el diagnóstico de enfermedades relacionadas con la retina, ya que hasta el vaso más delgado y diminuto puede influir en la lista de diagnósticos diferenciales. En este sentido, su método simplificó el problema de segmentación de imágenes completas al reconocimiento regional de elementos vasculares y la posterior fusión de los resultados. Adicionalmente, se aplicaron técnicas de postprocesamiento de imágenes no supervisadas para refinar el resultado final de la segmentación [22].

Tanto los vasos sanguíneos gruesos como los más delgados aportan información importante para el diagnóstico médico, Zengqiang Yan et al. (2018) propusieron un método de aprendizaje profundo que consta de tres etapas: la segmentación de los vasos sanguíneos gruesos, la segmentación de los vasos sanguíneos delgados y, finalmente, la unión de los resultados obtenidos en las dos etapas anteriores [23]. Xing Yang et al. (2021) desarrollaron un algoritmo para la segmentación de vasos sanguíneos fusionando redes neuronales convolucionales con el método de “*Cascade Forest*”. La elección de la red neuronal convolucional se basó en su capacidad sobresaliente para extraer automáticamente características, mientras que el método “*Cascade Forest*” demostró ser altamente eficaz al trabajar con bases de datos pequeñas. En el preprocesamiento de las imágenes, se aplicó inicialmente el método CLAHE (*Contrast Limited Adaptive Histogram Equalization*), seguido por el filtrado morfológico, con el objetivo de mejorar el procesamiento [24].

Shuang Xu et al. (2021) propusieron un algoritmo que combina redes neuronales residuales (ResNet) y redes neuronales convolucionales (CNN), utilizó el módulo de atención residual mejorado y el módulo de supervisión profunda para prevenir posibles problemas durante el entrenamiento de la CNN. Su algoritmo demostró una efectividad notable al extraer los vasos sanguíneos de las imágenes de fondo de ojo, incluyendo los vasos sanguíneos más pequeños y delgados (capilares) [25].

Wenjing Li et al. (2022) propusieron un método eficiente para la segmentación de vasos sanguíneos en imágenes de fondo de ojo basado en filtros B-COSFIRE. Destacaron que la extracción de vasos retinales juega un papel importante en el diagnóstico de enfermedades como el glaucoma y la retinopatía diabética, y que la precisión de esta segmentación afecta en gran medida el efecto diagnóstico. Su enfoque abordó los desafíos de mejorar



simultáneamente el rendimiento de la segmentación y la eficiencia en el tiempo, así como la dificultad para distinguir los vasos delgados del ruido similar a los vasos. Para lograrlo, el método aplicó CLAHE para mejorar el contraste, un conjunto de filtros B-COSFIRE para detectar vasos, filtros morfológicos para eliminar el ruido, y un postprocesamiento basado en dominios conectados para eliminar los píxeles no vasculares no conectados y así obtener la imagen vascular final, preservando los vasos delgados conectados [26].

El papel de los métodos de IA, con énfasis en el aprendizaje profundo, ha experimentado un visible crecimiento en el campo de la oftalmología. Los avances significativos resaltados en los estudios revisados destacan la importancia y eficacia de estas técnicas en la detección de biomarcadores, tanto los evidentes como los más discretos. La diversidad de enfoques, desde el uso de redes neuronales convolucionales hasta la combinación de diferentes arquitecturas, refleja el continuo avance y la adaptabilidad de las técnicas de aprendizaje profundo en este campo.

## 3. Marco Teórico

### 3.1. Tipos de Glaucoma

El glaucoma se clasifica en varios tipos, que se dividen principalmente en primarios y secundarios. El glaucoma primario de ángulo abierto es el tipo más común de glaucoma. Se caracteriza por un daño lento, progresivo e indoloro al nervio óptico, causado por un problema en el drenaje del humor acuoso en la cámara anterior del ojo, lo que provoca un aumento de la presión intraocular [27].

El glaucoma de tensión normal se caracteriza por la pérdida del campo visual periférico sin el aumento de la presión intraocular. Este tipo no tiene causa aparente. Entre las posibles explicaciones se encuentran la mayor susceptibilidad del nervio óptico a sufrir daños incluso con presión intraocular normal, o que exista un problema en el riego de sangre al nervio, posiblemente debido a una falla en los mecanismos que regulan ese flujo [27], [28].

El glaucoma primario de ángulo cerrado es una urgencia oftalmológica debido al aumento brusco y notable de la presión intraocular causado por el bloqueo del drenaje del humor

acuoso; si no se trata a tiempo, puede llevar a la pérdida de visión en un corto periodo. En ocasiones, es asintomático y la persona afectada puede no darse cuenta hasta que experimenta ceguera. Las mujeres tienen una mayor predisposición a este tipo de glaucoma que los hombres, y esta predisposición aumenta en el caso de aquellas de ascendencia asiática oriental [27], [28].

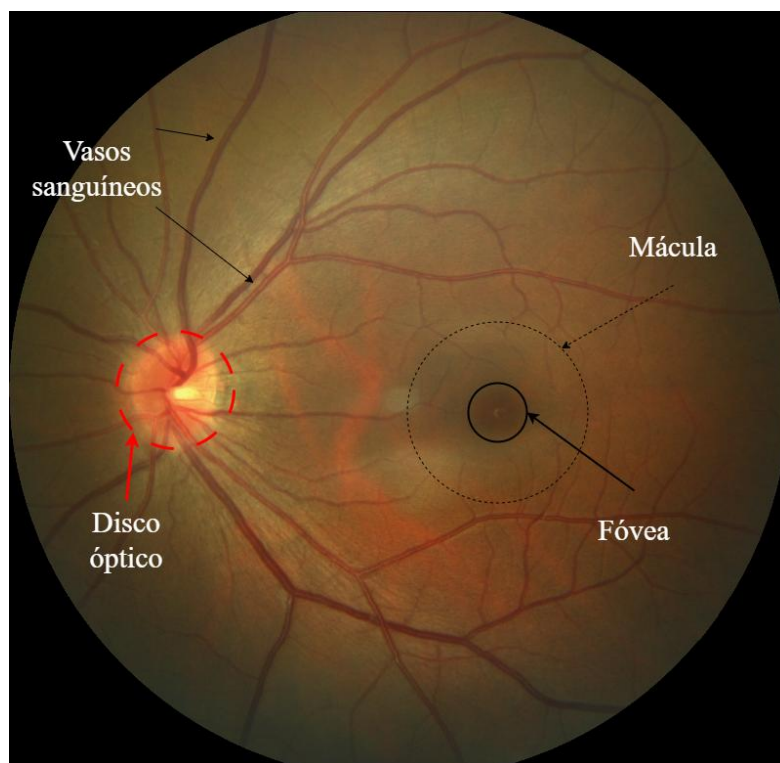
### 3.2. El Ojo Humano

El ojo humano es un órgano sensorial complejo y fundamental, encargado de nuestra percepción visual. Está ubicado dentro de la cavidad orbitaria junto con sus anexos (párpados, cejas, músculos extraoculares). Su función principal es la vista, que se inicia con la refracción de la luz a través de los medios transparentes del ojo y finaliza con la activación de las células fotorreceptoras en la retina [29].

El ojo está conformado por tres capas: externa, media e interna. La capa externa es una capa fibrosa conformada por la esclera y la córnea; su función es la protección del globo ocular. La capa media es una capa principalmente vascular, conocida como úvea, que constituye el iris, el cuerpo ciliar y la coroides. Por último, la capa interna conocida comúnmente como retina [30].

### 3.3. Retina

La retina es un tejido nervioso delgado y complejo, formado por varias capas altamente organizadas que abarcan toda la parte interna del globo ocular. La retina sensorial cumple un papel fundamental en la visión al encargarse de detectar estímulos luminosos, procesar las señales visuales y transmitir dichas señales al cerebro a través del nervio óptico, proceso conocido como fototransducción [29], [30].



*Figura 3. Visualización de la retina y sus principales estructuras en el fondo del ojo.*

En la retina se encuentran estructuras anatómicas de gran interés: el disco óptico (papila óptica), la mácula, la fóvea, la fóveola y los vasos sanguíneos.

### 3.3.1. Disco Óptico o Papila Óptica

El disco óptico es una estructura circular u ovalada de aproximadamente 1.5 mm diámetro donde los axones de las células ganglionares convergen y abandonan el ojo para formar el nervio óptico. El disco óptico presenta un color amarillento y bordes generalmente bien definidos. En su centro se encuentra la copa óptica, la cual es una depresión de tamaño variable por donde emergen la arteria y la vena central de la retina. La evaluación de la forma, el color, la nitidez de los bordes y el tamaño de la excavación (relación copa-disco) es importante en el diagnóstico de enfermedades como el glaucoma, en el que un aumento de la excavación constituye un biomarcador característico [31], [32].

### 3.3.2. Vasos Sanguíneos Retinianos

La arteria central de la retina es la responsable del suministro sanguíneo a la retina interna, emerge a través del disco óptico y se ramifica en divisiones superior e inferior para distribuirse a lo largo del tejido retiniano. Previo a su salida del nervio óptico, las ramas principales superior e inferior se ramifican nuevamente en divisiones nasales y temporales, se ramifican progresivamente para irrigar los cuatro cuadrantes de la retina. Las anomalías vasculares, como la neovascularización, hemorragias, microaneurismas, tortuosidad vascular y nasalización, son signos característicos de diversas enfermedades en la retina [32].

### 3.4. Desplazamiento Nasal e Índice de Nasalización

La nasalización es un desplazamiento perceptible de los vasos sanguíneos que emergen de la cabeza del nervio óptico hacia el lado nasal. En la mayoría de los casos en condiciones normales, los VSC se distribuyen de una manera relativamente simétrica alrededor del disco óptico. Sin embargo, en ojos con glaucoma, esta disposición de los vasos tiende a agruparse o desviarse hacia el sector nasal del disco óptico. Esta alteración es más notable en el tronco vascular retiniano central (TVRC), que es el punto donde la arteria y la vena centrales de la retina emergen. Este desplazamiento de los VSC se da por la pérdida progresiva del ANR en la región temporal del disco óptico y cambios en la lámina cribosa, ambas alteraciones estructurales son consecuencia del proceso glaucomatoso [12].

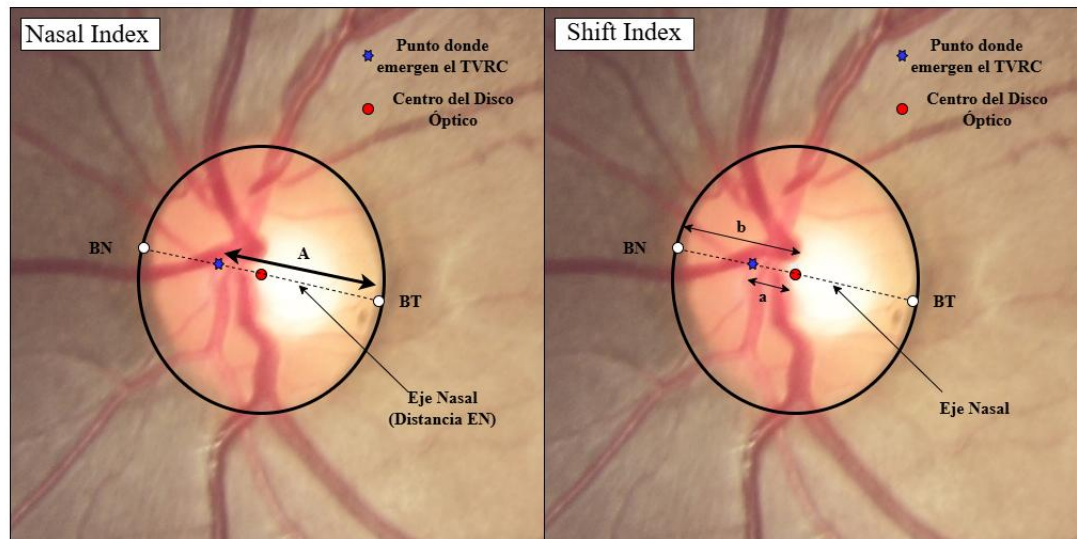
Para cuantificar el grado de nasalización, Kilhwan Shon desarrolló el índice de nasalización (*Nasalization Index, NI*), que es una métrica cuantitativa diseñada específicamente para evaluar el grado de desplazamiento hacia el sector nasal del TVRC dentro del disco óptico [33].

La metodología para calcular el índice es la siguiente:

- Localización del TVRC. Con ayuda de OCT se localiza el punto donde emerge el TVRC y se marca dicho punto en la imagen de fondo de ojo.
- Centro del disco óptico. Se dibuja una elipse alrededor del disco óptico (sobrepuesta a este) y se localiza el centro de la elipse.

- Eje nasal. Se establece un eje de referencia para medir el desplazamiento del TVRC trazando una línea que pase por los puntos del centro de la elipse y el punto donde emerge el TVRC. El eje conecta directamente el borde nasal (punto BN) y temporal (punto BT) del disco óptico, atravesando su región central.
- Distancia A y EN. La distancia A abarca del punto donde emerge el TVRC al punto BT. EN es la distancia del eje nasal que va desde el punto BN al punto BT.
- Cálculo de NI. El valor de NI está dado por la distancia A dividida por la distancia EN.

$$NI = \frac{A}{EN} \quad (1)$$



*Figura 4. Imagen de fondo de ojo que muestra las distancias necesarias para el cálculo del NI y SI propuestos por Kilhwan Shon.*

De acuerdo con los resultados del estudio, tras calcular el NI, los pacientes se agruparon en dos categorías: alta y baja nasalización. El grupo con alta nasalización mostró un valor promedio de NI de 0.84, mientras que el grupo con baja nasalización registró un promedio de 0.52 [33].

En su trabajo, Shon propone un índice adicional denominado “*Shift Index*” (SI), el cual está relacionado con NI, pero su procedimiento de cálculo es diferente. Este índice se calcula dividiendo la distancia “a” (desde el centro de la elipse hasta el punto de emergencia del TVRC) entre la distancia “b” (desde el centro de la elipse hasta el punto BN). Un valor de  $SI \geq 0.5$  indica alta nasalización; en cambio, un  $SI < 0.5$  corresponde a baja nasalización [33].

### 3.5. Fotografía de Fondo de Ojo

La fotografía del fondo de ojo (retinografía) es una técnica diagnóstica que consiste en capturar una imagen de la superficie interna del ojo, que contiene a la retina, el nervio óptico, la mácula, la fovea, los vasos sanguíneos y el polo posterior [34].

En la práctica oftalmológica, la fotografía de fondo de ojo representa una herramienta clave para la detección temprana de enfermedades oculares. Además, permite un seguimiento clínico más preciso y genera un registro visual permanente para futuras comparaciones.

La capacidad de las imágenes de fondo de ojo para permitir la detección temprana de patologías amplifica significativamente la utilidad de la IA en este campo. Los modelos de aprendizaje profundo pueden identificar patrones sutiles que podrían pasar desapercibidos por el ojo humano.

### 3.6. Aprendizaje Automático

El aprendizaje profundo (*machine learning*) es un subcampo de la IA que facilita la detección de patrones a través del análisis de grandes conjuntos de datos, utilizando técnicas estadísticas y algoritmos matemáticos que mejoran el aprendizaje a través de la experiencia, es decir, puede aprender de sus errores y realizar predicciones precisas [35].

Existen tres categorías de aprendizaje automático: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo.

En el aprendizaje supervisado, se le ayuda al algoritmo a realizar tareas a través de entradas y salidas conocidas (etiquetadas), tiene como objetivo principal aprender patrones o reglas generales que le permitan asociar las entradas con las salidas [36].

En el aprendizaje no supervisado, no se proporciona información previa al algoritmo (los datos no están etiquetados), este se centra en aprender pocas características de los datos y busca automáticamente patrones o tendencias, separando la información de manera autónoma [36].

El aprendizaje por refuerzo implica que el algoritmo interactúa con su entorno obteniendo recompensas o castigos según sus actos, aprendiendo de estas experiencias y perfeccionando su desempeño. Este tipo de aprendizaje puede recibir tanto datos etiquetados como no etiquetados [36].

### 3.7. Aprendizaje Profundo

El aprendizaje profundo (Deep Learning) es un subcampo del aprendizaje automático; consiste en modelos compuestos por múltiples capas de unidades neuronales que permiten la extracción automática de características y la identificación de patrones complejos directamente desde los datos en bruto. Estas redes neuronales profundas se inspiran, en cierta medida, en la estructura y funcionamiento del cerebro humano, particularmente en la interconexión entre sus neuronas. Una gran ventaja del aprendizaje profundo es su habilidad para adquirir automáticamente representaciones jerárquicas: desde patrones sencillos en las capas iniciales hasta conceptos progresivamente más complejos y abstractos en las capas más avanzadas [37].

La capacidad del aprendizaje profundo para identificar automáticamente patrones significativos a partir de los datos ha sido esencial para su éxito en dominios exigentes como el procesamiento del lenguaje natural, la visión por computadora y el reconocimiento de voz [38].

Entre las arquitecturas de aprendizaje profundo más utilizadas en el estado del arte para el campo de la visión por computadora se encuentran las redes neuronales convolucionales (*Convolutional Neural Networks*, CNN) y las redes neuronales residuales (*Residual Neural Networks*, ResNet).

### 3.7.1. Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNN) constituyen una categoría de redes neuronales profundas que destacan por su efectividad en el análisis de imágenes médicas. Diseñadas para procesar datos en matrices múltiples, su objetivo principal es identificar diversas características visuales mediante el uso de capas convolucionales, las cuales aplican filtros para detectar patrones locales en las entradas [39].

La arquitectura característica de una CNN se estructura en múltiples capas especializadas que procesan la información jerárquicamente, tal como se ilustra en la figura. Sus componentes esenciales son capas convolucionales, capas de agrupación (pooling) y capas completamente conectadas.

- Capas convolucionales (*convolutional layers*). Su función principal es aprender y extraer características jerárquicas de los datos de entrada. Esta extracción se logra mediante la operación de convolución entre la entrada y un conjunto de filtros (*kernels*). Cada filtro se desliza a través de la imagen de entrada y calcula el producto punto entre los pesos del filtro y la porción correspondiente de la imagen, generando así un "mapa de activación" que resalta la presencia de la característica identificada por el filtro en diferentes ubicaciones espaciales. La función de activación no lineal, como ReLU, se emplea al final de la convolución para dotar al modelo de la capacidad de aprender representaciones más abstractas y complejas [40].
- Capas de agrupación (*pooling layers*). Estas capas tienen como propósito disminuir gradualmente las dimensiones espaciales (ancho y alto) de los mapas de características, lo que a su vez disminuye la cantidad de parámetros y la complejidad computacional de la red. La reducción de dimensionalidad también actúa como una forma de regularización, ayudando a controlar el sobreajuste [40].
- Capas completamente conectadas. Los mapas de características de alto nivel que se producen tras una serie de capas convolucionales y de agrupación suelen aplanarse en un vector unidimensional y, a continuación, introducirse en una o varias capas totalmente conectadas. En una capa completamente conectada, cada neurona está



conectada a todas las activaciones de la capa anterior, similar a las redes neuronales tradicionales. Estas capas realizan la tarea final de clasificación o regresión [40].

### 3.7.2. Redes Neuronales Residuales

Las redes neuronales residuales (ResNet) son una arquitectura de aprendizaje profundo diseñada para resolver el problema de degradación que ocurre al aumentar la profundidad de la red. En redes neuronales convencionales, añadir más capas puede reducir el rendimiento en lugar de mejorarlo. ResNet soluciona este problema mediante el uso de conexiones residuales [41].

El elemento distintivo de la arquitectura ResNet es las conexiones de salto, las cuales implementan un mapeo de identidad, es decir, transfieren la entrada directamente a la salida sin modificaciones ni costos computacionales adicionales. Su presencia es fundamental para permitir el entrenamiento eficaz de redes neuronales muy profundas [41].

La arquitectura ResNet ha demostrado versatilidad y rendimientos excepcionales, teniendo éxito en tareas de detección de objetos, segmentación y clasificación de imágenes.

## 3.8. Modelos de Detección de Objetos y Segmentación

### 3.8.1. Detección de Objetos

La detección de objetos es una técnica que permite identificar y localizar objetos de ciertas clases dentro de una imagen, utilizando cuadros delimitadores (*bounding boxes*). Esta técnica integra dos tareas: la localización del objeto y su clasificación. Es especialmente efectiva cuando los objetos ocupan una porción moderada de la imagen o cuando es necesario identificar agrupaciones como una única entidad [42].

### 3.8.2. Segmentación Semántica

El objetivo de la segmentación semántica es asignar una etiqueta a cada píxel de la imagen, generando un mapa de segmentación en el que todos los píxeles que pertenecen a la misma clase comparten la misma etiqueta. Esta técnica no distingue entre diferentes instancias de una misma categoría, tratándolas como una sola entidad [43].

### 3.8.3. Segmentación de Instancias

A diferencia de la segmentación semántica, la segmentación de instancias no solo clasifica cada píxel, sino que también distingue e individualiza cada objeto perteneciente a la misma clase. Para ello, produce máscaras de segmentación a nivel de píxel que delinean con exactitud la forma y extensión de cada instancia [43].

## 3.9. YOLO

YOLO (*You Only Look Once*) es una familia de algoritmos de detección de objetos en una sola etapa. YOLO cambió el paradigma de la detección de objetos al predecir tanto las ubicaciones como las clases de los objetos directamente desde la imagen completa, todo en una sola operación. Este diseño eficiente permite alcanzar detección en tiempo real sin comprometer significativamente la precisión [42], [44].

La arquitectura general de YOLO se compone comúnmente de tres partes principales:

- *Backbone*. Es la responsable de extraer jerarquías de la imagen de entrada. Por lo general, se recurre a arquitecturas de CNN preentrenadas con grandes volúmenes de datos, como ImageNet, para facilitar la tarea de clasificación y detección de imágenes [45].
- *Neck*. Su función es fusionar las características extraídas por el *backbone* a diferentes escalas.
- *Head*. Es el componente final de la red, responsable de generar las predicciones. Para la detección de objetos, esto incluye los cuadros delimitadores y las probabilidades de clase, mientras que en segmentación de instancias también se encarga de producir las máscaras de segmentación a nivel de píxel [45].

### 3.9.1. YOLOv8

Desarrollado por Ultralytics en 2023, YOLOv8 constituye una evolución significativa de la serie YOLO, al integrar mejoras en precisión y velocidad, y extender sus aplicaciones más allá de la detección de objetos hacia tareas como la segmentación de instancias. Su arquitectura rediseñada incorpora módulos escalables y procesos de entrenamiento

optimizados, lo que permite abordar distintos problemas de visión por computadora con mayor eficiencia [46].

La arquitectura YOLOv8 para segmentación de instancias se apoya en gran medida en el modelo de detección de objetos, lo que la diferencia son las modificaciones para generar las máscaras a nivel de píxel para cada instancia detectada.

### 3.10. Modelos de Clasificación

#### 3.10.1. ResNet50

ResNet50 es una variante de la arquitectura ResNet que se caracteriza por tener 50 capas, las cuales contienen 48 capas convolucionales, una capa de entrada (*MaxPool*) y otra capa totalmente conectada (*AveragePool*). La arquitectura ResNet50 ha sido pre-entrenada con el conjunto de datos “ImageNet” lo que la convierte en una opción valiosa y robusta para tareas de clasificación por la variedad de patrones visuales aprendidos [47].

### 3.11. Técnicas de Preprocesamiento

Las técnicas de preprocesamiento tienen como objetivo principal optimizar la calidad de los datos de la imagen. Al aplicar la técnica correcta, podemos resaltar las características deseadas y/o eliminar las que no necesitamos (por ejemplo, el ruido), facilitando así las subsiguientes tareas de procesamiento y análisis [48].

#### 3.11.1. Ecualización de Histograma Adaptativo con Contraste Limitado

El filtro CLAHE (*Contrast Limited Adaptive Histogram Equalization*) es una técnica avanzada de mejora del contraste. A diferencia de otros métodos que aplican una única transformación a toda la imagen, CLAHE opera de forma adaptativa en pequeñas regiones locales (o "tiles"). Cada una de estas secciones se mejora individualmente, lo que permite que el contraste de los píxeles de fondo y los niveles de contraste ajustados resalten eficazmente incluso las áreas con niveles bajos o ruidosos, donde la mejora es más necesaria [49].

### 3.11.2. Filtro B-COSFIRE

El filtro B-COSFIRE (*Bar-selective Combination Of Shifted Filter Responses*) es una herramienta propuesta para la detección de estructuras en forma de barra, considerada un método novedoso y válido para la segmentación automática de vasos sanguíneos.

El principio de B-COSFIRE radica en la combinación de las respuestas de múltiples filtros base, generalmente filtros DoG (Diferencia de Gaussianas). Un filtro DoG se crea restando dos versiones de la misma imagen y cada una suavizada con kernels Gaussianos de diferente desviación estándar. Este proceso da como resultado un filtro pasa-banda que realza regiones con frecuencias espaciales específicas, como bordes o líneas, lo que permite al B-COSFIRE identificar las estructuras deseadas [26].

## 3.12. Métricas de Evaluación

Las métricas de evaluación son cruciales para medir el rendimiento de un modelo y facilitar la comparación entre distintas arquitecturas. De esta manera, es posible identificar la más adecuada para la tarea específica y el conjunto de datos empleado. Además, juegan un papel clave en la optimización de hiperparámetros, al servir como criterio de referencia para mejorar el rendimiento.

### 3.12.1. Matriz de Confusión

La matriz de confusión es una tabla que resume el rendimiento de un modelo de clasificación comparando las clases reales con las que anticipaba el modelo (*ground truth*) para un conjunto de datos de prueba [50].

La matriz de confusión está compuesta por 4 términos fundamentales:

- Verdaderos Positivos (TP): Representa la cantidad de instancias que se clasificaron correctamente como pertenecientes a la clase positiva.
- Falsos Positivos (FP): Representa la cantidad de instancias que se clasificaron incorrectamente como pertenecientes a la clase positiva.
- Verdaderos Negativos (TN): Representa la cantidad de instancias que se clasificaron correctamente como pertenecientes a la clase negativa.

- Falsos Negativos (FN): Representa la cantidad de instancias que se clasificaron incorrectamente como pertenecientes a la clase negativa.

Estos términos se utilizan para calcular métricas de rendimiento del modelo como la precisión, la sensibilidad, la especificidad y la puntuación F1 (*F1-Score*).

### 3.12.2. Exactitud (*accuracy*)

La exactitud cuantifica la efectividad general del modelo al medir la proporción de predicciones correctas con respecto al número total de predicciones realizadas [50].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

### 3.12.3. Precisión

La precisión cuantifica cuántas de las instancias que el modelo clasificó como positivas son realmente positivas [50].

$$Presición = \frac{TP}{TP + FP} \quad (3)$$

### 3.12.4. Sensibilidad (*Recall*)

La sensibilidad, a veces denominada tasa de verdaderos positivos, evalúa la capacidad del modelo para detectar instancias positivas correctamente [50].

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

### 3.12.5. Especificidad

La especificidad (tasa de verdaderos negativos) es importante cuando es fundamental evitar falsos positivos, ya que mide la proporción de casos realmente negativos que el modelo clasifica correctamente como negativos [50].

$$Especificidad = \frac{TN}{TN + FP} \quad (5)$$

### 3.12.6. Puntuación F1 (*F1-Score*)

La puntuación F1 es la medida armónica de precisión y sensibilidad. Ofrece una única métrica que logra un equilibrio entre ambas. Es especialmente útil cuando se presenta un desequilibrio en las clases o cuando tanto los falsos negativos como los falsos positivos tienen un impacto similar [50].

$$F1\ Score = 2 * \frac{Presición * Recall}{Presición + Recall} \quad (6)$$

## 4. Hipótesis

Un modelo de aprendizaje profundo aplicado al análisis de imágenes de fondo de ojo puede detectar con alta precisión la nasalización de los vasos sanguíneos del disco óptico, sirviendo como un biomarcador indirecto crucial para evaluar el daño glaucomatoso establecido y la progresión de la enfermedad en el campo visual.

## 5. Objetivos

### 5.1. Objetivo General

Desarrollar un modelo de aprendizaje profundo para detectar y clasificar anomalías en los vasos sanguíneos asociadas al glaucoma en el disco óptico en imágenes de fondo de ojo.

### 5.2. Objetivos Específicos

- Seleccionar una base de datos de imágenes de fondo de ojo que muestre de manera clara la nasalización de los vasos sanguíneos.
- Realizar un preprocesamiento que delimite la región del nervio óptico y que destaque tanto los vasos sanguíneos centrales gruesos como los más delgados.
- Adaptar técnicas de preprocesamiento específicamente diseñadas para la segmentación de vasos sanguíneos en la región del anillo neuroretiniano.
- Implementar un algoritmo con métodos de aprendizaje profundo para la detección de características asociadas a la nasalización de los vasos sanguíneos.



## 6. Material y Metodología

### 6.1. Python

Python es un lenguaje de programación interpretado, orientado a objetos y de alto nivel, ampliamente adoptado para la programación de propósito general y, de manera particular, en la comunidad científica y en el área de investigación en inteligencia artificial.

Fue diseñado pensando en que el código sea fácil de leer y que su sintaxis permita escribir ideas de forma simple. Python es ampliamente reconocido por su facilidad de aprendizaje y uso, lo que lo hace accesible tanto para principiantes como para desarrolladores con experiencia [51].

Para este trabajo se utilizó la versión de Python 3.11.9.

### 6.2. Google Colaboratory

*Google Colaboratory* (*Google Colab*) es una plataforma gratuita en la nube desarrollada por Google que permite la escritura y ejecución de código Python directamente desde el navegador. Funciona como un cuaderno interactivo al estilo de Jupyter Notebooks, y ofrece acceso a recursos potentes como GPUs (*Graphics Processing Unit*) y TPUs (*Tensor Processing Unit*) en la nube [52].

Si bien el plan gratuito de Google Colab ya incluye acceso a recursos como GPUs y TPUs, existen versiones de pago (Colab Pro y Pro+) pensadas para usuarios que requieren mayor capacidad. Estas versiones ofrecen ventajas como más velocidad, más memoria y mayor tiempo de ejecución [52].

### 6.3. PyTorch

PyTorch es una biblioteca de código abierto que permite construir modelos de aprendizaje profundo en Python, ampliamente aplicada en tareas de visión por computadora y procesamiento del lenguaje natural. Fue desarrollado por Facebook AI Research (ahora Meta AI).

PyTorch combina la flexibilidad del lenguaje Python con potentes herramientas para crear redes neuronales, desde modelos simples hasta arquitecturas complejas como Transformers. Su enfoque en la ejecución en tiempo real facilita la depuración y la creación rápida de prototipos, lo que lo convierte en una herramienta clave en investigación y desarrollo [53].

## 6.4. OpenCV

OpenCV (*Open Computer Vision*) es una biblioteca de código abierto diseñada para tareas de visión por computadora e IA. Fue lanzada en 1999 por Gary Bradski [54]. Entre las funciones de OpenCV para visión por computadora se encuentran el procesamiento básico de imágenes, transformaciones geométricas, suavizado de imágenes, dibujo, umbralización, detección de bordes y contornos, entre otras.

## 6.5. Roboflow

Roboflow es una plataforma integral diseñada a fin de facilitar y acelerar el flujo de trabajo completo asociado a la creación, entrenamiento y despliegue de modelos de visión por computadora. Se propone como una solución que ofrece un conjunto de herramientas tanto para desarrolladores individuales como para empresas, simplificando el proceso de un proyecto de visión artificial [55].

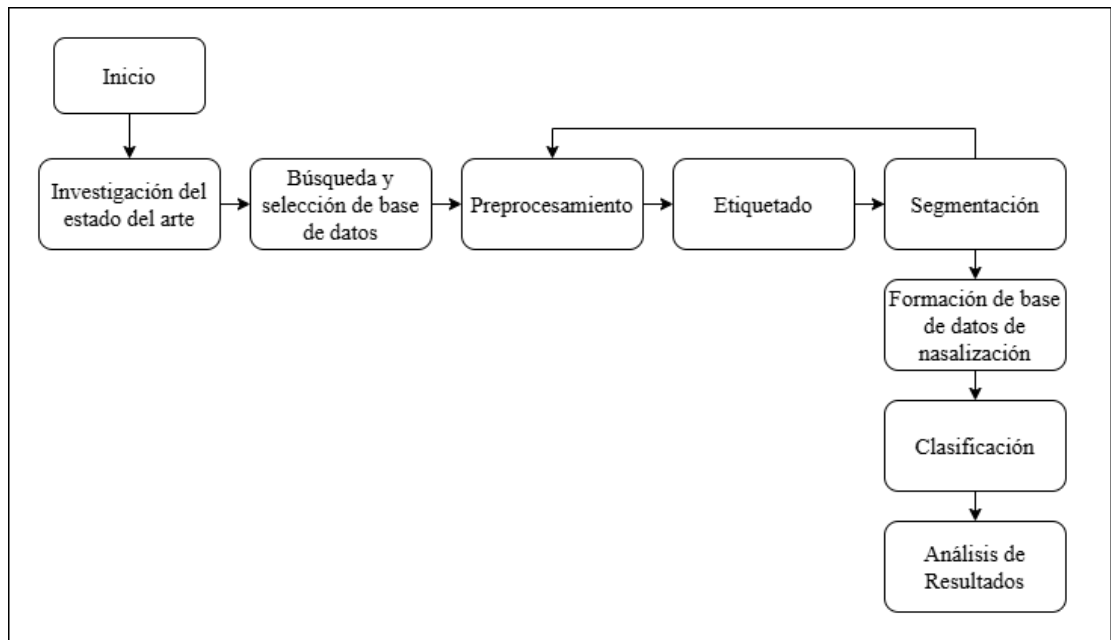
Roboflow cuenta con una extensa variedad de herramientas para la implementación de modelos de visión por computadora. Facilita el proceso de anotación y etiquetado al proporcionar herramientas de anotación en línea que soportan diversos tipos de tareas, incluyendo la detección de objetos mediante cuadros delimitadores (*bounding boxes*), la segmentación de instancias a través de polígonos y la clasificación de imágenes. Para agilizar este proceso laborioso, Roboflow incluye capacidades de etiquetado asistido por inteligencia artificial (*Label Assist*), el cual permite utilizar modelos previamente entrenados por el usuario, sugiriendo anotaciones y acelerando el trabajo de etiquetado [55].

Otra de las grandes virtudes de Roboflow es la importación y exportación de anotaciones en una amplia variedad de formatos. Su versión de pago ofrece la aplicación de técnicas de preprocesamiento, aumento de datos y entrenamiento de modelos. Además, cuenta con una interfaz para la validación y visualización de modelos.

## 6.6. Hardware

- Equipo: PC con procesador Intel(R) Core (TM) i7-9750H CPU @ 2.60GHz, 30 GB de RAM y tarjeta de video NVIDIA GeForce GTX 1660 Ti.
- Google Colab: Suscripción de pago de Google Colab Pro, GPU NVIDIA A100, 40 GB de VRAM.

## 6.7. Metodología



*Figura 5. Diagrama de la metodología realizada.*

## 6.8. Investigación del Estado del Arte

El punto de partida de este trabajo consistió en una revisión profunda del estado del arte en la aplicación de algoritmos de aprendizaje profundo en el área de la oftalmología, centrándose específicamente en la detección de biomarcadores asociados al glaucoma en imágenes de fondo de ojo.

Se analizaron diversas técnicas de preprocesamiento de imágenes de fondo de ojo, con un énfasis particular en aquellas destinadas a la segmentación y el realce de los vasos

sanguíneos, ya que son cruciales para el estudio. Esto abarcó desde la corrección de brillo y contraste hasta filtros para detectar bordes y contornos, todo para asegurar que los vasos se visualizaran nítidamente.

Esta etapa posibilitó la identificación de los modelos de aprendizaje profundo más empleados en tareas de segmentación y clasificación. La mayoría de estos trabajos utilizan tanto bases de datos públicas como privadas, lo que facilitó una primera interacción en la búsqueda de recursos de datos disponibles y la posterior evaluación de su idoneidad para los objetivos de este estudio.

Un hallazgo de esta investigación fue la ausencia de estudios previos que aborden la detección de la nasalización en imágenes de fondo de ojo mediante el uso de modelos de IA, lo que representa una gran oportunidad de contribución.

## 6.9. Búsqueda y selección de Base de Datos

Se realizó una búsqueda de bases de datos públicas entre las cuales destacaron ORIGA, G1020, PAPILA, Drishti-GS y REFUGE. El objetivo fue encontrar un conjunto de datos que cumpliera con requisitos de representatividad, diversidad y alta calidad. De las opciones evaluadas, REFUGE (*Retinal Fundus Glaucoma Challenge*) se eligió por su robustez, al incluir 1200 imágenes de fondo de ojo con resoluciones de 2124 x 2056 y 1634 x 1634. Además de incluir imágenes glaucomatosas y no glaucomatosas, presenta un escenario clínico más realista al contener casos con retinopatía diabética, macrodiscos y miopía. Un aspecto fundamental de su elección es que las 1200 imágenes están acompañadas de máscaras segmentadas del disco y la copa óptica, realizadas por oftalmólogos experimentados [56].

## 6.10. Preprocesamiento

La etapa de preprocesamiento tuvo como objetivo la aplicación de técnicas para mejorar la calidad de la imagen en donde los vasos sanguíneos y el disco óptico resaltarán.

### 6.10.1. Canal Verde

El proceso de preprocesamiento se inició con la conversión de las imágenes a escala de grises utilizando OpenCV. Posteriormente, se procedió a la separación de los canales RGB, optando

por conservar exclusivamente el canal verde para la aplicación de los pasos siguientes. Esta decisión se fundamenta en que el canal verde es reconocido por su superior contraste para la visualización de la vasculatura retiniana y las estructuras relacionadas con la sangre, además de ofrecer una clara delimitación del disco óptico [57].

#### 6.10.2. Filtro CLAHE

Como segundo paso, se aplicó el filtro CLAHE para mejorar el contraste local de la imagen de fondo de ojo. Esto tuvo como objetivo principal hacer que los vasos sanguíneos fueran más visibles. Esta mejora en la visualización de la vasculatura facilitó significativamente la posterior aplicación del filtro B-COSFIRE.

El filtro CLAHE opera dividiendo la imagen en pequeñas regiones, conocidas como "*tiles*". Con la ayuda de OpenCV, se implementó una cuadrícula de 8x8, resultando en 64 regiones individuales. Esta división permitió aplicar el CLAHE a cada región de forma individual, asegurando una mejora de contraste local más efectiva y uniforme en toda la imagen.

#### 6.10.3. Filtro B-COSFIRE

Para este punto del preprocesamiento, se parte de una imagen de entrada a la que ya se le aplicó el filtro CLAHE, lo que resulta en un realce del contraste de los vasos sanguíneos. Sobre esta imagen, se implementó el filtro B-COSFIRE con el objetivo de detectar y acentuar la estructura vascular retiniana. Este filtro se seleccionó por su probada capacidad para responder selectivamente a los patrones característicos de los vasos sanguíneos. La sinergia entre CLAHE y B-COSFIRE fue clave, ya que el CLAHE optimizó los detalles finos, asegurando que estos fueran óptimamente detectados por B-COSFIRE, lo que maximiza la visibilidad de la vasculatura.

#### 6.10.4. Extracción de Región de Interés

La región de interés (ROI) para este estudio se estableció en el disco óptico y su entorno inmediato, dada su relevancia directa para la detección de la nasalización de los VSC. Este proceso de extracción de la ROI se realizó de forma automática. Inicialmente, se aprovecharon las máscaras que, al contener únicamente el disco y la copa óptica representados en blanco, negro y gris, facilitaron la binarización. Una vez binarizada, la

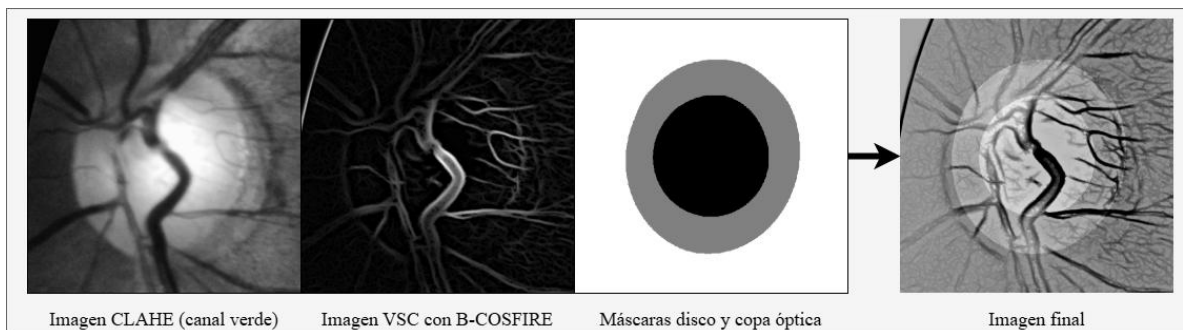
región del disco óptico se representó en blanco sobre un fondo negro, lo que permitió el trazado de sus bordes y la obtención de su centroide. A partir de este punto central, se delimitó una ventana de 500 x 500 píxeles, cuyas coordenadas se registraron para aplicarlas de manera uniforme a las imágenes que ya habían sido procesadas con los filtros CLAHE y B-COSFIRE.

#### 6.10.5. Combinación de Imágenes

Hasta el momento tenemos las siguientes imágenes:

- Imagen CLAHE: El filtro CLAHE se aplicó al canal verde permitiendo obtener un contraste entre los vasos sanguíneos y el fondo.
- Imagen VSC: Esta imagen binaria muestra los vasos sanguíneos segmentados gracias a la aplicación del filtro B-COSFIRE.
- Imagen máscaras: Esta imagen corresponde a las máscaras del disco y copa óptica anotadas manualmente por especialistas.

Cada imagen resultante de los pasos anteriores delimita dos estructuras de interés: el disco óptico y los VSC. Para realzar aún más los VSC, se procedió a combinar las tres imágenes preprocesadas en un solo canal utilizando la función “*cv2.merge*”. El resultado fue una imagen final donde los VSC contrastan notablemente, funcionando como un límite claro entre las regiones nasal y temporal del disco óptico.



*Figura 6. Combinación para la imagen final.*

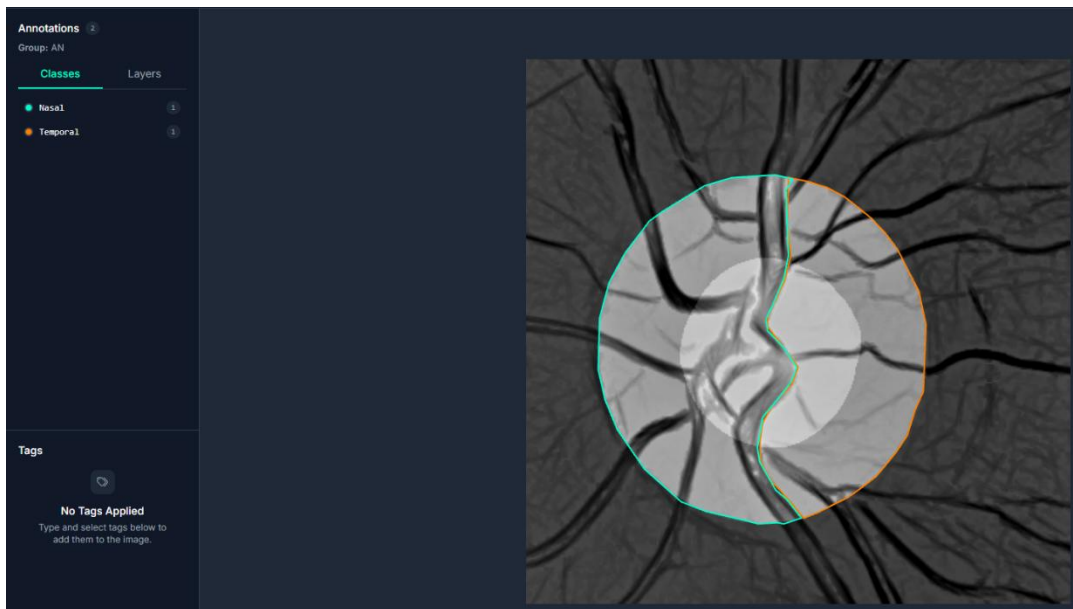
Este preprocesamiento no solo destaca los vasos sanguíneos más gruesos, sino también aquellos más delgados que se encuentran en la región del ANR. Además, el uso de las

máscaras facilita la visualización del ANR, lo que permitirá observar con mayor precisión la relación copa-disco (o excavación del ANR) en las imágenes de fondo de ojo.

### 6.11. Etiquetado

Tras el preprocesamiento de las 1200 imágenes de la base de datos REFUGE, se seleccionaron 400 imágenes correspondientes a la carpeta "*test*" para la fase de etiquetado. Este proceso se llevó a cabo manualmente utilizando la herramienta Roboflow, creando un proyecto de segmentación de instancias.

El criterio para el etiquetado se basó en la definición del TVRC como la frontera entre la región nasal y temporal del disco óptico. De acuerdo con esta definición, se procedió a etiquetar estas áreas específicas, con el área del TVRC abarcando la porción nasal, como se visualiza en la Figura 6. Los etiquetados fueron validados por expertos del Instituto Mexicano de Oftalmología (IMO).



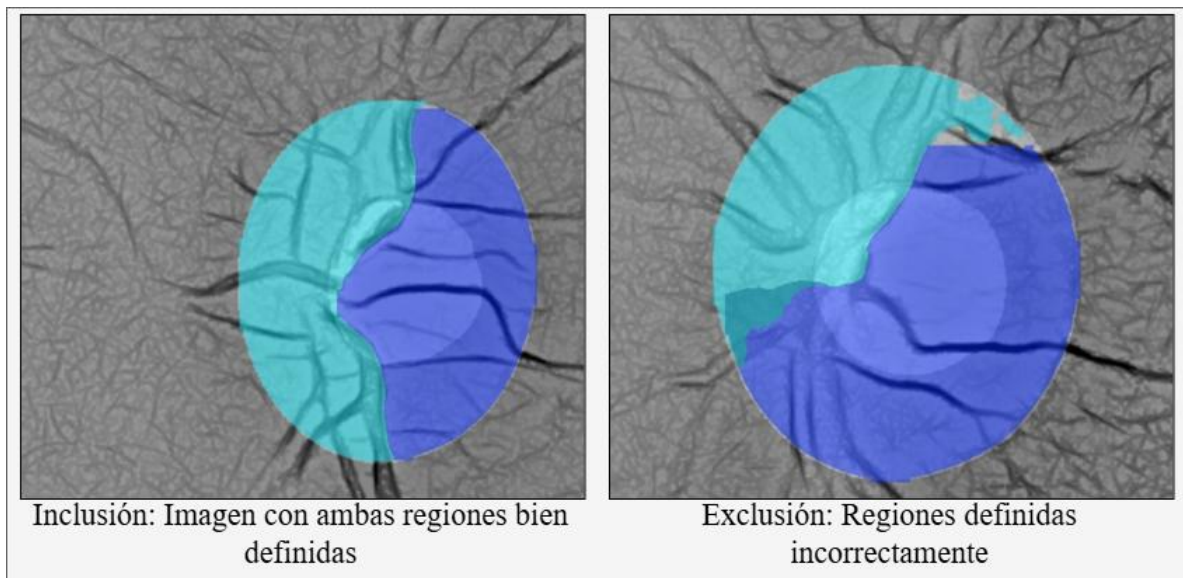
*Figura 7. Interfaz de Roboflow para etiquetado.*

## 6.12. Segmentación

Para agilizar el proceso de obtención de máscaras para las 800 imágenes restantes de la base de datos, se decidió utilizar la arquitectura YOLOv8 con la tarea de segmentación de instancias. Roboflow facilitó la exportación de anotaciones en un formato compatible con las arquitecturas YOLO. La razón principal para elegir YOLOv8 para esta tarea fue su velocidad, eficiencia y la disponibilidad de una variedad de modelos preentrenados para segmentación.

Se aplicó un proceso de aumento de datos mediante técnicas de volteo y rotación de imágenes, alcanzando un total de 814 imágenes. Estas imágenes fueron distribuidas en una proporción de 80% para entrenamiento, 10% para validación y 10% para prueba.

Ya con el modelo entrenado se procedió a la obtención de las máscaras de la región nasal y la región temporal de las 800 imágenes restantes. La plataforma de Roboflow permitió observar las máscaras obtenidas y así poder realizar criterios de inclusión y exclusión. Aquellas máscaras con una perfecta delimitación de las regiones nasal y temporal se mantuvieron para la siguiente etapa. Las máscaras con una mala definición de las regiones y/o el TVRC recargado hacia la parte temporal fueron excluidas. En la figura 8 se puede observar el criterio de inclusión-exclusión.

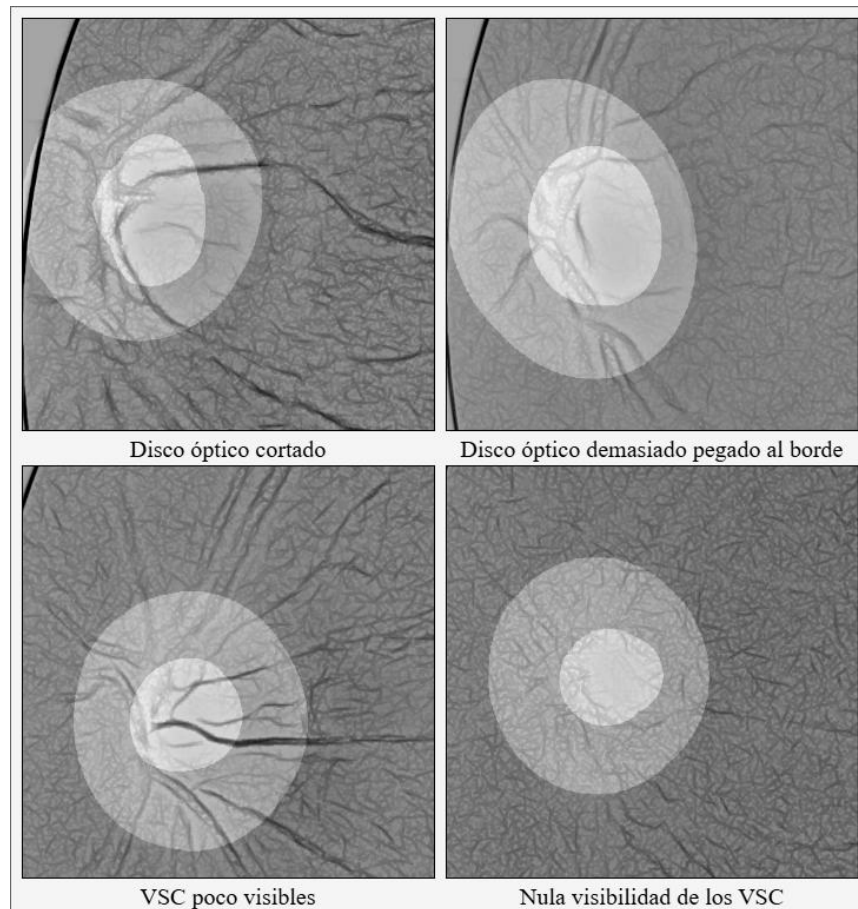


*Figura 8. Criterios de inclusión-exclusión.*



Además de la mala definición de los límites de las regiones, se excluyeron aquellas imágenes de fondo de ojo con las siguientes características:

- Disco óptico cortado.
- Disco óptico pegado al borde por el ángulo de la fotografía.
- VSC poco visibles.
- Imágenes sin VSC visibles.



*Figura 9. Criterios de exclusión.*

### 6.13. Base de Datos de Nasalización

Para la construcción de una base de datos de nasalización, se propuso y aplicó una fórmula como alternativa al índice de desplazamiento propuesto por Kilhwan Shon en su trabajo [33]. El cálculo de este índice se explica a detalle en el marco teórico. El índice de nasalización

propuesto tuvo como objetivo clasificar las imágenes de fondo de ojo según el grado de la nasalización, con el propósito de utilizar esta nueva base de datos para entrenar un modelo de clasificación.

#### 6.13.1. Índice Nasalización Propuesto

En condiciones normales, las áreas de las regiones nasal y temporal del disco óptico son similares. Sin embargo, esta simetría del área puede variar debido a factores como la distribución de los VSC, la presencia de macrodiscos, entre otros.

En imágenes de fondo de ojo que presentan nasalización, se observa un incremento en el área temporal y una reducción en el área nasal. Esto ocurre porque el TVRC se recarga hacia el borde nasal del disco óptico, modificando la proporción entre ambas regiones.

Para calcular esta proporción de áreas se propuso la fórmula siguiente:

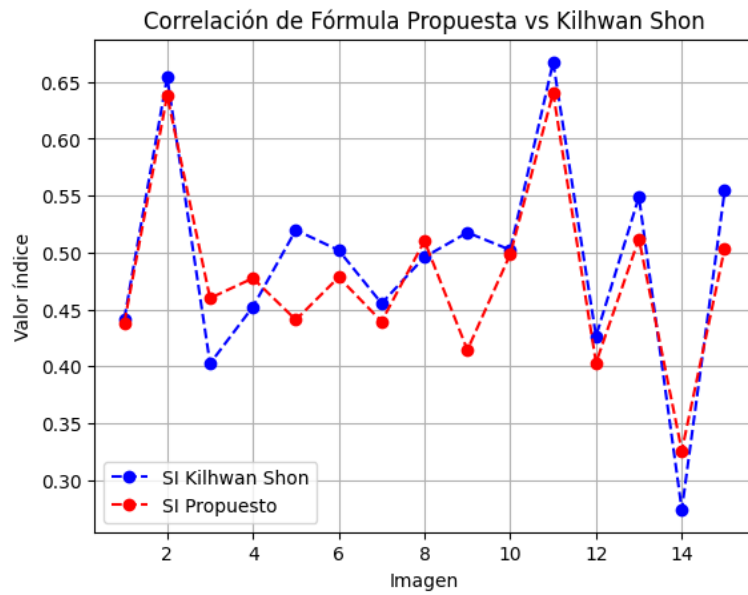
$$SI = \frac{AT}{AT + AN} \quad (6)$$

Donde:

- SI – Índice de Nasalización (desplazamiento)
- AT – Área Temporal
- AN – Área Nasal

Para comprobar la correlación entre la fórmula propuesta y la fórmula de Kilhwan Shon, se seleccionaron 15 imágenes de fondo de ojo de forma aleatoria y se aplicó el procedimiento para el cálculo del SI descrito en el marco teórico. Al igual que el autor, se utilizó el programa “*ImageJ*” para hacer el trazo de la elipse, los ejes y puntos.

Los valores obtenidos con el índice propuesto mostraron una aproximación significativa con los generados por el índice de Kilhwan Shon. Esto se puede observar en la figura 10.



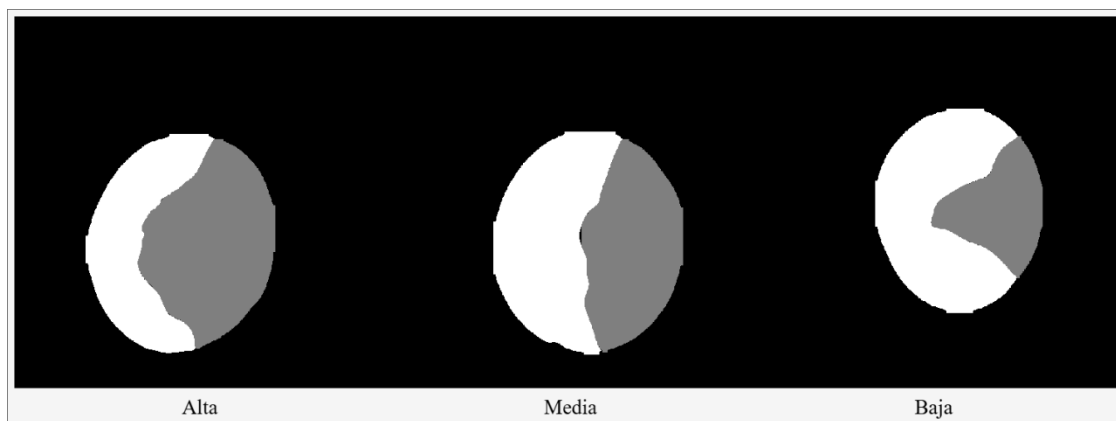
*Figura 10. Correlación de SI (índice de desplazamiento nasal).*

#### 6.13.2. Formación de Base de Datos

Se formó una base de datos con 985 imágenes, distribuidas en tres clases según los siguientes criterios de clasificación:

- Alta:  $SI \geq 0.6$
- Media:  $0.6 > SI > 0.4$
- Baja:  $SI \leq 0.4$

El índice SI, para la distribución de clases, se obtuvo mediante la fórmula propuesta. Para el cálculo de áreas, la región nasal son todos los píxeles blancos en la máscara. Los píxeles grises corresponden a la región temporal. Las clases baja y media agrupan imágenes de fondos de ojo sin nasalización (ojos sanos). La clase alta contiene imágenes de fondo de ojo con sospecha de nasalización de los VSC.



*Figura 11. Base de datos de nasalización, cuenta con tres clases.*

#### 6.14. Clasificación

ResNet50 fue la arquitectura de aprendizaje profunda seleccionada para la tarea de clasificación. Este modelo es ampliamente adoptado y reconocido en el campo de la clasificación de imágenes médicas. Otro factor importante de su selección fue que tiene disponibles los pesos pre-entrenados de ImageNet. Aunque ImageNet no tiene imágenes médicas, estos pesos proporcionan una buena base para la extracción de características generales.

Se ha observado que, a pesar de la diferencia de dominios, las características de bajo nivel aprendidas de ImageNet (como bordes, texturas y formas) son altamente transferibles y beneficiosas para el análisis de imágenes médicas. Esto reduce significativamente los recursos y el tiempo necesarios para el entrenamiento, especialmente en conjuntos de datos médicos que a menudo son más pequeños.

La arquitectura ResNet50 mantuvo sus capas convolucionales y residuales inalteradas, aprovechando su potente capacidad de extracción de características. Sin embargo, fue esencial modificar la capa de salida (la capa de clasificación final). Esto se hizo porque la capa original de ResNet50 está diseñada para las 1000 clases de ImageNet, y necesitábamos reemplazarla con una nueva capa que tuviera el número de neuronas de salida específico para nuestras tres categorías de clasificación.

## 7. Resultados y Discusión

La investigación se realizó en dos fases principales: la segmentación de instancias de las regiones nasal y temporal del disco óptico con YOLOv8, y la posterior clasificación de estas regiones según el índice propuesto con ResNet50. A continuación, se detallan los resultados obtenidos en cada fase y su discusión.

### 7.1. YOLOv8

La segmentación de las regiones nasal y temporal del disco óptico fue un paso clave en este estudio, ya que la detección de nasalización requiere una delimitación precisa de ambas regiones, separadas por el trayecto del TVRC. YOLOv8 fue la arquitectura seleccionada para dicha tarea.

El entrenamiento del modelo YOLOv8 se realizó con un conjunto de 814 imágenes. Para asegurar una evaluación robusta y representativa, el conjunto de datos se dividió en proporciones del 80% para entrenamiento, 10% para validación y 10% para pruebas.

*Tabla 4. Resultado de YOLOv8 para segmentación de instancias.*

Métrica	Resultado
Precisión	99.2
Recall	100
F1-Score	99.59

Los resultados cuantitativos obtenidos para la segmentación de las regiones nasal y temporal del disco óptico con YOLOv8 fueron muy buenos, demostrando un rendimiento sobresaliente con una precisión del 99.2%, un recall del 100% y un F1-score del 99.59%. La alta calidad de la segmentación de estas regiones fue corroborada cualitativamente mediante la inspección visual detallada de las imágenes de prueba y las máscaras generadas en la plataforma de Roboflow.

A pesar del buen rendimiento general que muestra la capacidad del modelo para reconocer las dos áreas de interés, se notaron algunas características inusuales en algunas imágenes de prueba. En estos casos particulares, la máscara generada por el modelo no siempre cubrió por completo el área total de la región detectada, a pesar de la correcta identificación de la etiqueta (nasal o temporal).

Esta diferencia en la delineación precisa de los límites de las máscaras se debe al hecho de que, al crear el conjunto de entrenamiento YOLOv8, no se aplicaron los criterios de inclusión-exclusión para el filtrado de imágenes con ciertas características. Por ejemplo, se encontraron casos donde las imágenes de fondo de ojo no tenían los vasos sanguíneos definidos de manera clara a pesar del preprocesamiento realizado. En estos casos, la localización de las regiones fue precisa, pero la máscara no abarcaba de manera precisa la región detectada.

Es importante destacar que incluso en imágenes donde la visibilidad de los VSC era particularmente limitada, el modelo YOLOv8 demostró una capacidad extraordinaria, logrando segmentar las regiones del disco óptico de manera adecuada. Esto destaca la robustez de YOLOv8 para la detección y segmentación de objetos, incluso en imágenes de calidad escasa. Sin embargo, para futuras mejoras, se enfatiza la importancia de mejorar el filtrado del conjunto de datos de entrenamiento y la mejora del preprocesamiento. Esto con el fin de optimizar la precisión en la delineación de las máscaras, lo cual es un factor crítico para la fase de clasificación basada en el tamaño de las regiones.

## 7.2. ResNet50

La base de datos de nasalización formada está compuesta por 985 imágenes, distribuidas en tres clases distintas de acuerdo con el índice de nasalización propuesto (“Alta”, “Media”, “Baja”). Dado que las clases se encontraban desbalanceadas, se aplicó un aumento de datos con el fin de lograr un balance en la representación de cada categoría. Entre las técnicas de aumento de datos utilizadas se incluyeron volteo (vertical y horizontal), rotación, cambios de brillo y modificaciones de saturación. Tras finalizar este proceso, se obtuvo un conjunto de imágenes balanceado con 600 imágenes cada uno.

Posteriormente, el entrenamiento se realizó con un total de 1800 imágenes. Para evaluar su robustez y optimizar el rendimiento del modelo, se probaron dos particiones del conjunto de datos en subconjuntos para entrenamiento, validación y prueba.

- 80/10/10
- 60/20/20

Para ambos grupos de división, se mantuvieron constantes algunos hiperparámetros clave, como el tamaño de lote (*batch size*) y el optimizador (Adam). Sin embargo, se experimentó con la tasa de aprendizaje (*learning rate*) y el número de épocas (*epochs*). Ajustando estos últimos con el objetivo de observar qué configuración ofrecía las mejores métricas. Las tablas 5 y 6 muestran las métricas obtenidas.

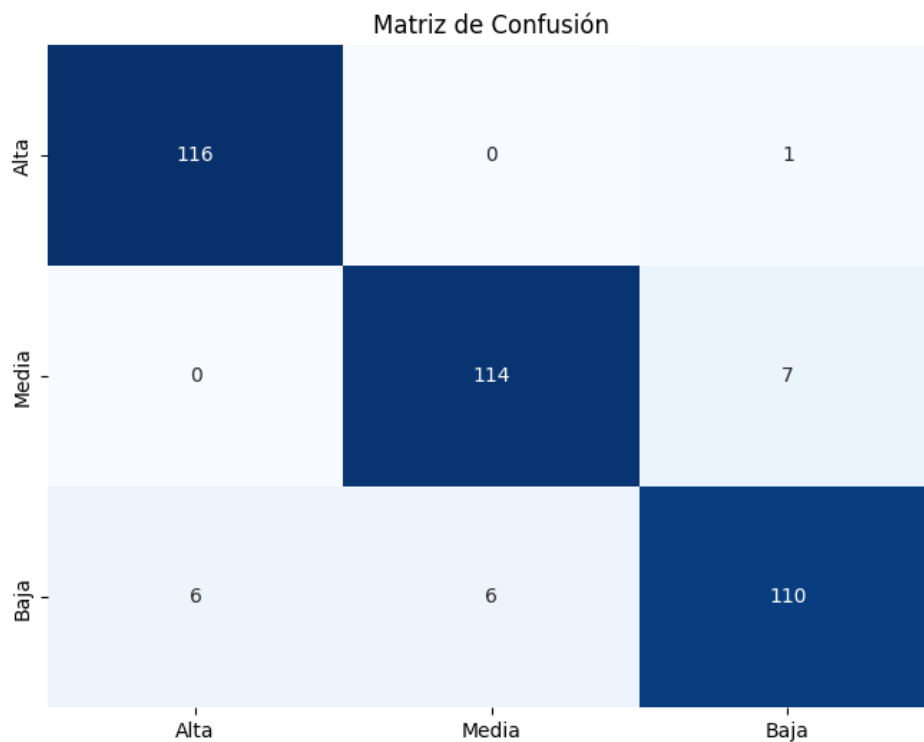
*Tabla 5. Métricas 80/10/10.*

Tasa de aprendizaje	Épocas	Tamaño de lote	Precisión	Especificidad	Sensibilidad
0.001	20	32	0.9328	0.9616	0.9247
0.001	50	32	0.9280	0.9636	0.9340
0.0001	20	32	0.9415	0.9689	0.9383
0.0001	30	32	0.9454	0.9693	0.9422
0.0001	50	32	0.9407	0.9691	0.9402
0.0001	100	32	0.9406	0.9695	0.9411

*Tabla 6. Métricas 60/20/20.*

Tasa de aprendizaje	Épocas	Tamaño de lote	Precisión	Especificidad	Sensibilidad
<b>0.0001</b>	<b>50</b>	<b>32</b>	<b>0.9442</b>	<b>0.9722</b>	<b>0.9451</b>
0.0001	100	32	0.9300	0.9656	0.9286

La configuración que tuvo los mejores resultados se observa resaltada en la tabla 6. Esta configuración resultó ser la óptima para el entrenamiento de 1800 imágenes. Este punto permitió al modelo converger de manera eficaz sin caer en sobreajuste, aprovechando la capacidad de ResNet50. Su arquitectura de bloques residuales permite extraer las características sutiles necesarias para diferenciar los distintos grados de nasalización a partir de las relaciones de tamaño de las áreas segmentadas.



*Figura 12. Matriz de confusión de la configuración óptima (1800 imágenes).*

Con el objetivo de obtener mejores métricas de rendimiento, se aumentó el volumen del conjunto de datos. En esta ocasión se generaron 700 imágenes por clase, obteniendo un conjunto total de 2100 imágenes. Este nuevo conjunto se entrenó con una división 80/10/10, conservando hiperparámetros óptimos identificados previamente. Sin embargo, no se logró obtener mejores métricas de rendimiento con este conjunto de imágenes ampliado.

*Tabla 7. Métricas del conjunto de datos ampliado.*

80/10/10						
Tasa de aprendizaje	Épocas	Tamaño de lote	Precisión	Especificidad	Sensibilidad	F1-Score
0.001	100	32	0.8944	0.9471	0.8947	0.8908
0.0001	60	32	0.9211	0.9607	0.9214	0.9205
<b>0.0001</b>	<b>100</b>	<b>32</b>	<b>0.9355</b>	<b>0.9679</b>	<b>0.9357</b>	<b>0.9353</b>
0.0001	150	32	0.9250	0.9614	0.9238	0.9230
60/20/20						
0.0001	50	32	0.9318	0.9656	0.9310	0.9311



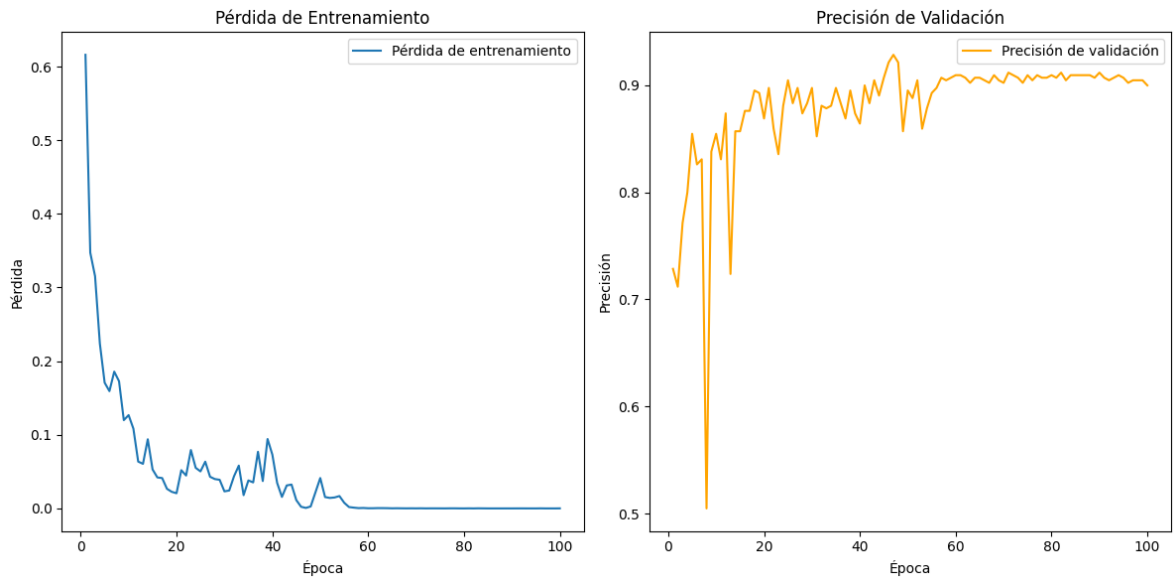
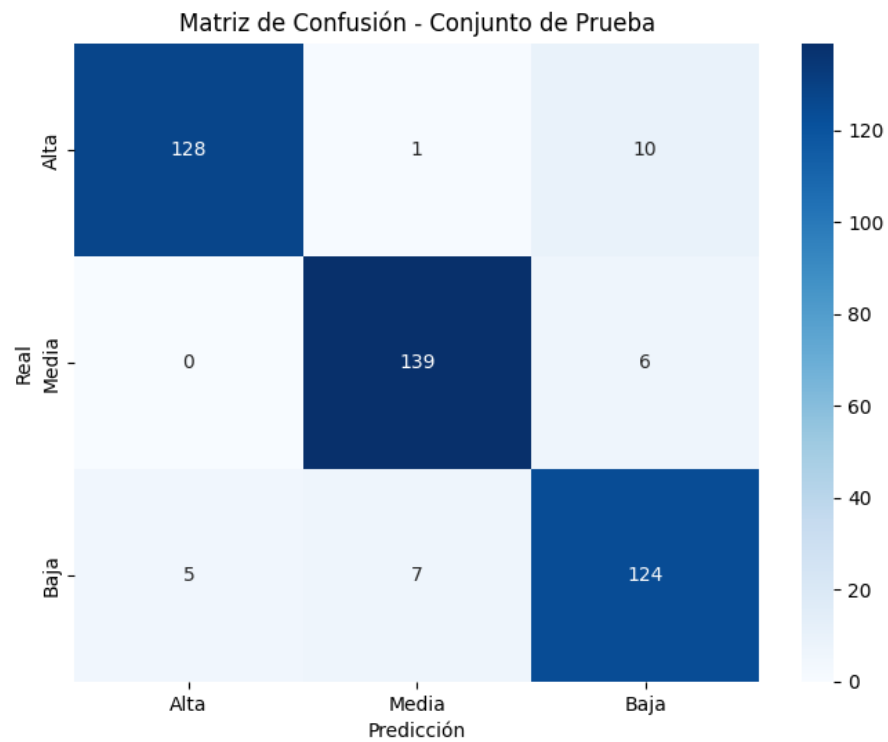
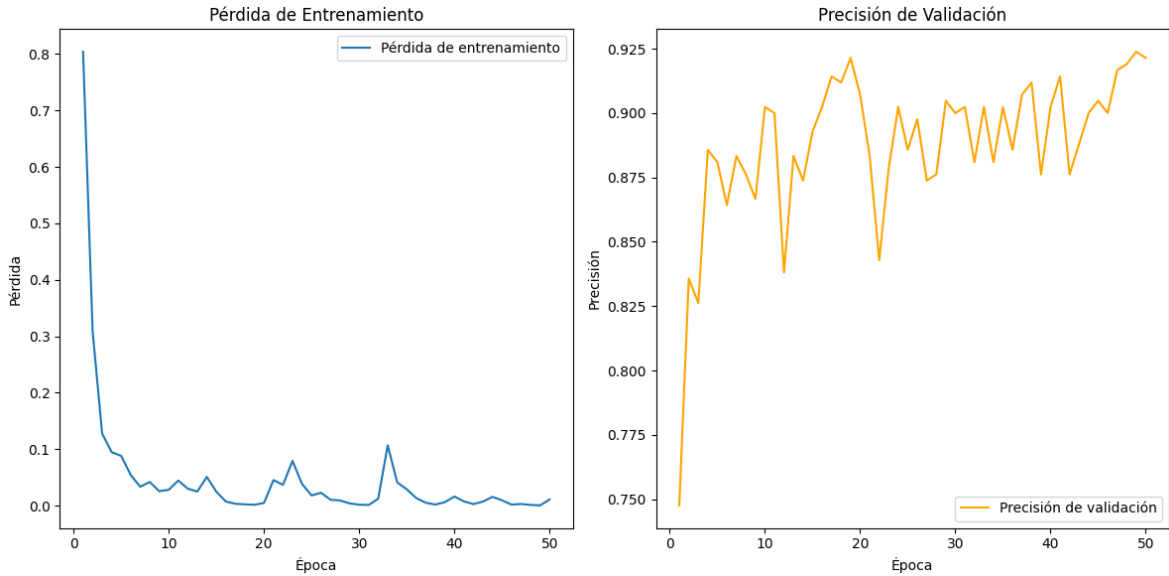


Figura 13. Gráficas de pérdida y validación (1800 imágenes).



*Figura 14. Matriz de confusión con los mismos hiperparámetros de la configuración óptima (2100 imágenes).*



*Figura 15. Gráficas de pérdida y validación (2100 imágenes).*

### 7.3. Discusión General

Los resultados de esta investigación muestran que un enfoque en dos fases puede ser una estrategia viable y efectiva para evaluar la nasalización de los VSC en imágenes de fondo de ojo. En la primera fase, enfocada en la segmentación de instancias de las regiones nasal y temporal del disco óptico, se utilizó el modelo YOLOv8. Este modelo demostró un rendimiento sobresaliente, con una precisión del 99.2%, un recall del 100% y un F1-score del 99.59%. Estos valores se encuentran a la par con el estado del arte en segmentación de estructuras anatómicas en imágenes médicas. Además, oftalmólogos especialistas validaron que, en la mayoría de los casos, las regiones fueron correctamente identificadas, lo que respalda la confiabilidad del modelo para este tipo de tareas. Esta etapa de segmentación es clave, ya que sobre ella se basan los cálculos posteriores relacionados con el tamaño de las regiones.

En la segunda fase, se llevó a cabo la clasificación del grado de nasalización utilizando el modelo ResNet50. Se seleccionó la arquitectura ResNet50 en lugar de VGG16 e InceptionV3 porque en las pruebas iniciales se entrenaron estos tres modelos y, mediante el uso de Grad-

CAM, se evidenció que ResNet50 enfocó su atención en las regiones de interés (nasal y temporal). Por el contrario, VGG16 e InceptionV3 se concentraron en estructuras distintas, como los vasos sanguíneos o el área alrededor del disco óptico.

ResNet50 fue capaz de distinguir con buena precisión las imágenes considerando la relación entre las regiones nasal y temporal. Con un conjunto de 1800 imágenes balanceadas y una configuración de entrenamiento de 60/20/20, una tasa de aprendizaje de  $1e-3$  y 50 épocas, el modelo alcanzó más de 90% de precisión. La matriz de confusión confirmó un buen desempeño, mostrando una baja tasa de errores entre las distintas categorías.

No obstante, se identificaron algunas limitaciones que pueden influir en la segmentación. Por ejemplo, la forma en que se distribuyen los VSC en la retina juega un papel importante. El sistema funcionó bien cuando las imágenes fueron tomadas desde un ángulo adecuado y los vasos presentaban una distribución relativamente compacta. Sin embargo, cuando los VSC estaban más dispersos o mal definidos, YOLOv8 no siempre logró cubrir por completo el área de interés, a pesar de que la localización general de la región fuera correcta. Esto pone en evidencia la sensibilidad del modelo ante variaciones anatómicas y de calidad en las imágenes, lo cual representa un desafío común en aplicaciones clínicas reales.

Por último, al ampliar el conjunto de datos de 1800 a 2100 imágenes, no se observó una mejora significativa en las métricas de desempeño. Esto sugiere que, una vez alcanzado cierto umbral, aumentar el número de muestras no necesariamente mejora el aprendizaje del modelo, especialmente si las nuevas imágenes no aportan suficiente variabilidad o información adicional. Es probable que, con las condiciones actuales de aumento de datos, el modelo haya llegado cerca de su capacidad máxima de generalización, y que futuros avances dependan más de la calidad y diversidad de los datos que de su cantidad.

## 8. Conclusiones

Las conclusiones en cuanto al cumplimiento de los objetivos de este trabajo de investigación son las siguientes.

La combinación de modelos de segmentación y clasificación demostró ser efectiva para la detección y clasificación de anomalías en los vasos sanguíneos, específicamente la nasalización, alcanzando métricas dentro del estado del arte. Tengo que destacar que este es el primer trabajo que aplica técnicas de inteligencia artificial para la detección de nasalización en vasos sanguíneos en imágenes de fondo de ojo, lo que impide una comparación directa con estudios previos.

La base de datos seleccionada se caracterizó por su diversidad de imágenes de fondo de ojo, lo que recrea un escenario clínico real. Se observó que la mayoría de las imágenes glaucomatosas en esta base de datos presentaban nasalización. A partir de esta base, el modelo de segmentación empleado y el índice propuesto, se logró conformar una nueva base de datos de máscaras que categoriza distintos grados de presencia de nasalización, convirtiéndola en una valiosa herramienta para futuros trabajos.

Las técnicas de preprocesamiento aplicadas permitieron el realce de estructuras de gran interés como los vasos sanguíneos y la región del ANR en imágenes de fondo de ojo. No obstante, considero que aún existe un margen significativo para mejorar la imagen final. Se encontraron imágenes donde los vasos sanguíneos no eran visibles, incluso con la aplicación de filtros específicos para el realce de estos (B-COSFIRE). Este realce de vasos es determinante, ya que marca la frontera entre las regiones nasal y temporal en el disco óptico.

Finalmente, el algoritmo ResNet50 se mostró eficiente en la tarea de detección y clasificación de la nasalización de vasos sanguíneos. Sus métricas, superiores al 90%, y su matriz de confusión, validan su excelente desempeño.

Para futuras investigaciones, se propone explorar el uso de otros modelos de segmentación de instancias con un doble paso de procesamiento, como Mask R-CNN. Este tipo de modelo, al combinar la propuesta de regiones con la predicción de máscaras, podría ofrecer una

delineación de contornos aún más precisa y completa para las regiones del disco óptico. Esto sería beneficioso para la exactitud de los cálculos de tamaño y, consecuentemente, para la precisión de la clasificación de la nasalización.

La integración de un modelo de clasificación con pruebas de campo visual tiene el potencial de validar los hallazgos de Kilhwan Shon, demostrando una fuerte correlación entre los defectos del campo visual detectados en las pruebas y la severidad del daño glaucomatoso por nasalización. Esta sinergia abriría la puerta a una herramienta de detección automatizada, mejorando significativamente la cuantificación de este daño.

## Referencias Bibliográficas

- [1] K. Singh *et al.*, “Assessment of optic disk by disk damage likelihood scale staging using slit-lamp biomicroscopy and optical coherence tomography in diagnosing primary open-angle glaucoma”, *Indian J Ophthalmol*, vol. 70, núm. 12, pp. 4152–4157, dic. 2022, doi: 10.4103/ijo.IJO\_1113\_22.
- [2] Y. C. Tsai, H. P. Lee, T. H. Tsung, Y. H. Chen, y D. W. Lu, “Unveiling Novel Structural Biomarkers for the Diagnosis of Glaucoma”, el 1 de junio de 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/biomedicines12061211.
- [3] C. P. Bragança, J. M. Torres, L. O. Macedo, y C. P. de A. Soares, “Advancements in Glaucoma Diagnosis: The Role of AI in Medical Imaging”, el 1 de marzo de 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/diagnostics14050530.
- [4] R. N. Weinreb, T. Aung, y F. A. Medeiros, “The pathophysiology and treatment of glaucoma: A review”, el 14 de mayo de 2014, *American Medical Association*. doi: 10.1001/jama.2014.3192.
- [5] K. Allison, D. Patel, y O. Alabi, “Epidemiology of Glaucoma: The Past, Present, and Predictions for the Future”, *Cureus*, nov. 2020, doi: 10.7759/cureus.11686.
- [6] Secretaría de Salud, “123. Se estima que en México 1.5 millones de personas padecen glaucoma: Secretaría de Salud”, Gobierno de México.
- [7] J. B. Jonas, T. Aung, R. R. Bourne, A. M. Bron, R. Ritch, y S. Panda-Jonas, “Glaucoma”, el 11 de noviembre de 2017, *Lancet Publishing Group*. doi: 10.1016/S0140-6736(17)31469-1.
- [8] Yvonne Ou, “How Glaucoma Affects Different Ethnic Groups”, *National Glaucoma Research*, feb. 2025.
- [9] A. Zaleska-Żmijewska *et al.*, “A new platform designed for glaucoma screening: Identifying the risk of glaucomatous optic neuropathy using fundus photography with deep learning architecture together with intraocular pressure measurements”, *Klin Oczna*, núm. 1, pp. 1–6, 2020, doi: 10.5114/KO.2020.94205.
- [10] Glaucoma Research Foundation, “Testing for Glaucoma”.
- [11] L. F. Machado, R. L. Furlanetto, y C. P. B. Gracitelli, “Anatomy and evaluation of the optic nerve head”, *Arq Bras Oftalmol*, vol. 85, núm. 6, pp. 636–643, 2022, doi: 10.5935/0004-2749.20220080.

- [12] Y. Sawada, M. Araie, H. Shibata, y T. Iwase, “Nasal displacement of retinal vessels on the optic disc in glaucoma associated with a nasally angled passage through lamina cribrosa”, *Sci Rep*, vol. 11, núm. 1, dic. 2021, doi: 10.1038/s41598-021-83720-0.
- [13] M. Nair, S. Tagare, R. Venkatesh, y A. Odayappan, “Artificial intelligence in glaucoma”, *Indian J Ophthalmol*, vol. 70, núm. 5, pp. 1868–1869, may 2022, doi: 10.4103/ijo.IJO\_1015\_22.
- [14] E. Rahimy, “Deep learning applications in ophthalmology”, el 1 de mayo de 2018, *Lippincott Williams and Wilkins*. doi: 10.1097/ICU.0000000000000470.
- [15] Secretaría de Salud, “Glaucoma, enfermedad que causa ceguera irreversible”.
- [16] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, y M. He, “Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs”, *Ophthalmology*, vol. 125, núm. 8, pp. 1199–1206, ago. 2018, doi: 10.1016/j.ophtha.2018.01.023.
- [17] N. Shibata *et al.*, “Development of a deep residual learning algorithm to screen for glaucoma from fundus photography”, *Sci Rep*, vol. 8, núm. 1, dic. 2018, doi: 10.1038/s41598-018-33013-w.
- [18] S. Phene *et al.*, “Deep Learning and Glaucoma Specialists: The Relative Importance of Optic Disc Features to Predict Glaucoma Referral in Fundus Photographs”, *Ophthalmology*, vol. 126, núm. 12, pp. 1627–1639, dic. 2019, doi: 10.1016/j.ophtha.2019.07.024.
- [19] S. Ajitha, J. D. Akkara, y M. V. Judy, “Identification of glaucoma from fundus images using deep learning techniques”, *Indian J Ophthalmol*, vol. 69, núm. 10, pp. 2702–2709, oct. 2021, doi: 10.4103/ijo.IJO\_92\_21.
- [20] A. Brown *et al.*, “Deep Learning for Localized Detection of Optic Disc Hemorrhages”, *Am J Ophthalmol*, vol. 255, pp. 161–169, nov. 2023, doi: 10.1016/j.ajo.2023.07.007.
- [21] Z. Jiang, J. Yezpe, S. An, y S. Ko, “Fast, accurate and robust retinal vessel segmentation system”, *Biocybern Biomed Eng*, vol. 37, núm. 3, pp. 412–421, 2017, doi: 10.1016/j.bbe.2017.04.001.
- [22] Z. Jiang, H. Zhang, Y. Wang, y S. B. Ko, “Retinal blood vessel segmentation using fully convolutional network with transfer learning”, *Computerized Medical Imaging and Graphics*, vol. 68, pp. 1–15, sep. 2018, doi: 10.1016/j.compmedimag.2018.04.005.

- [23] Z. Yan, X. Yang, y K. T. Cheng, “A Three-Stage Deep Learning Model for Accurate Retinal Vessel Segmentation”, *IEEE J Biomed Health Inform*, vol. 23, núm. 4, pp. 1427–1436, jul. 2019, doi: 10.1109/JBHI.2018.2872813.
- [24] X. Yang, Z. Li, Y. Guo, y D. Zhou, “Retinal vessel segmentation based on an improved deep forest”, *Int J Imaging Syst Technol*, vol. 31, núm. 4, pp. 1792–1802, dic. 2021, doi: 10.1002/ima.22610.
- [25] S. Xu, Z. Chen, W. Cao, F. Zhang, y B. Tao, “Retinal Vessel Segmentation Algorithm Based on Residual Convolution Neural Network”, *Front Bioeng Biotechnol*, vol. 9, dic. 2021, doi: 10.3389/fbioe.2021.786425.
- [26] W. Li *et al.*, “Retinal Vessel Segmentation Based on B-COSFIRE Filters in Fundus Images”, *Front Public Health*, vol. 10, sep. 2022, doi: 10.3389/fpubh.2022.914973.
- [27] J. Dietze, K. Blair, M. Zeppieri, y S. J. Havens, “Glaucoma”, en *StatPearls [Internet]*, Treasure Island, FL, USA: StatPearls Publishing, 2025.
- [28] P. Chiaradía, *Introducción a la Oftalmología*, 1ra ed. Medica Panamericana, 2019.
- [29] D. Vaughan y T. Asbury, *Vaughan & Asbury's General Ophthalmology*, 18a ed. New York: McGraw-Hill Medical, 2011.
- [30] E. Graue, *Oftalmología en la Práctica de la Medicina General*, 2a ed. Ciudad de México: McGraw-Hill Interamericana, 2006.
- [31] C. Serran, “Anatomía de la Vía Visual”, Kenhub GmbH .
- [32] F. Sánchez, “El Nervio Óptico y los Trastornos de la Visión”, Barcelona, nov. 2001.
- [33] K. Shon, Y. Hye Jo, J. Won Shin, J. Kwon, D. Jeong, y M. S. Kook, “Nasalization of Central Retinal Vessel Trunk Predicts Rapid Progression of Central Visual Field in Open-Angle Glaucoma”, *Sci Rep*, vol. 10, núm. 1, dic. 2020, doi: 10.1038/s41598-020-60355-1.
- [34] E. Pérez y L. León, “La fotografía de fondo de ojo como método de diagnóstico en el glaucoma”, *Medicentro Electrónica*, vol. 21, núm. 1, pp. 3–10, 2017.
- [35] M. Álvarez Vega, L. M. Quirós Mora, y M. V. Cortés Badilla, “Inteligencia artificial y aprendizaje automático en medicina”, *Revista Medica Sinergia*, vol. 5, núm. 8, p. e557, ago. 2020, doi: 10.31434/rms.v5i8.557.
- [36] B. Mahesh, “Machine Learning Algorithms - A Review”, *International Journal of Science and Research (IJSR)*, vol. 9, núm. 1, pp. 381–386, ene. 2020, doi: 10.21275/art20203995.


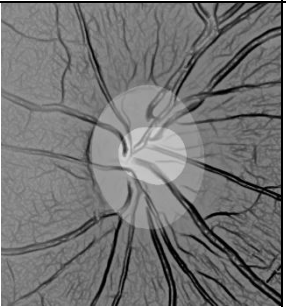
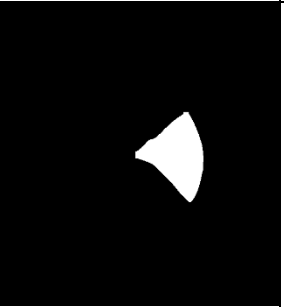
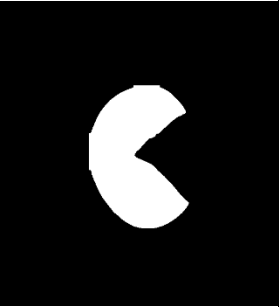

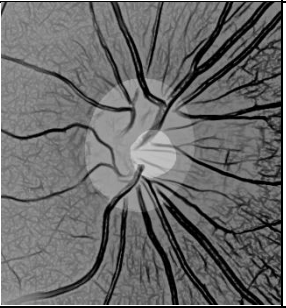
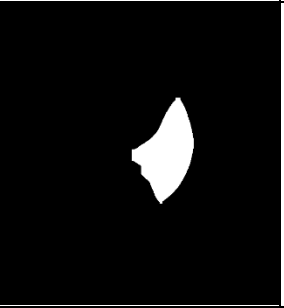


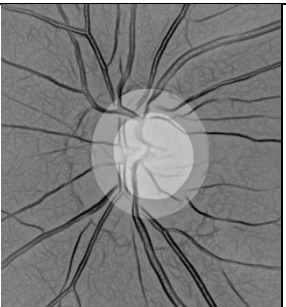
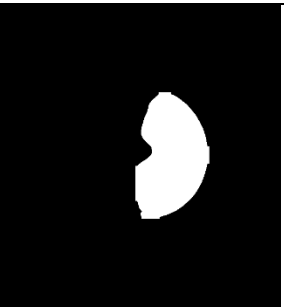
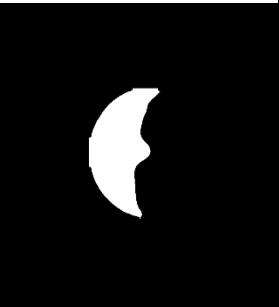
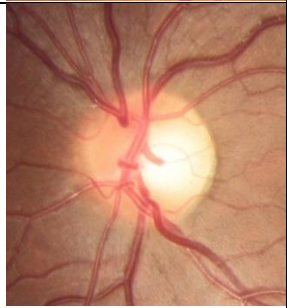
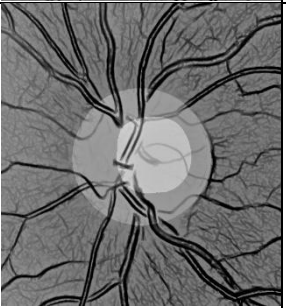
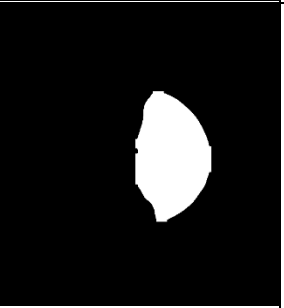



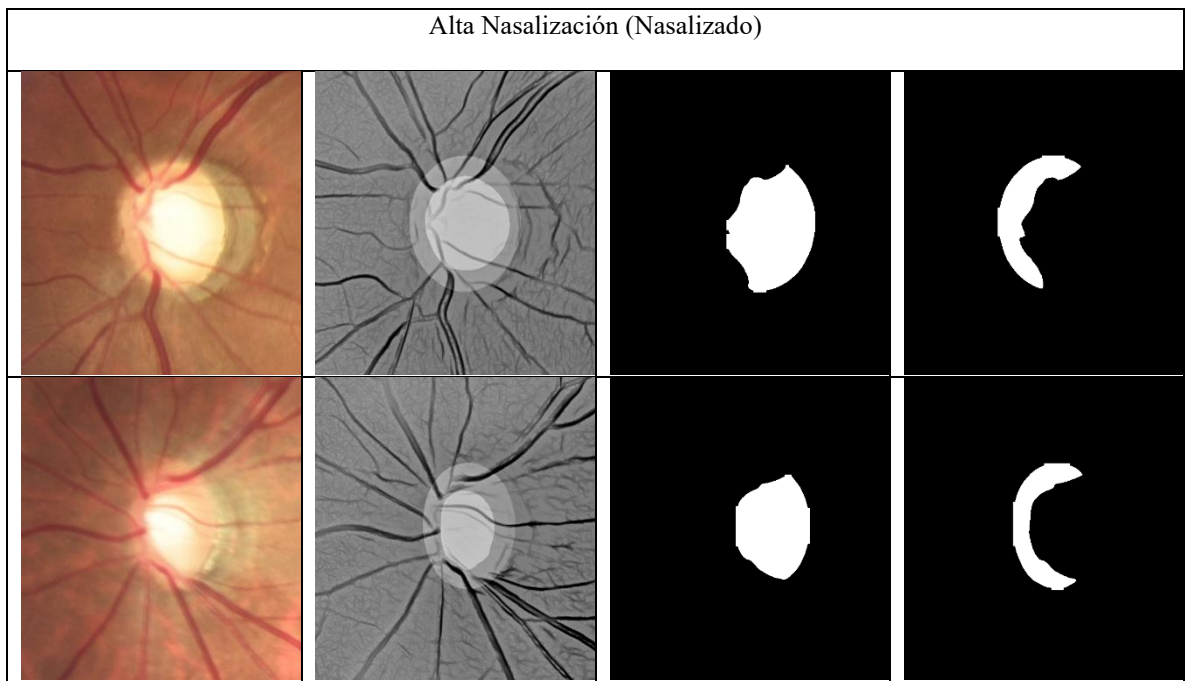
- [37] M. H. M. Noor y A. O. Ige, “A Survey on State-of-the-art Deep Learning Applications and Challenges”, mar. 2024, [En línea]. Disponible en: <http://arxiv.org/abs/2403.17561>
- [38] Y. Lecun, Y. Bengio, y G. Hinton, “Deep learning”, el 27 de mayo de 2015, *Nature Publishing Group*. doi: 10.1038/nature14539.
- [39] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan, y P. Ilono, “Deep Convolutional Neural Networks in Medical Image Analysis: A Review”, el 1 de marzo de 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/info16030195.
- [40] K. O’Shea y R. Nash, “An Introduction to Convolutional Neural Networks”, nov. 2015, [En línea]. Disponible en: <http://arxiv.org/abs/1511.08458>
- [41] K. He, X. Zhang, S. Ren, y J. Sun, “Deep Residual Learning for Image Recognition”, dic. 2015, [En línea]. Disponible en: <http://arxiv.org/abs/1512.03385>
- [42] J. Murel y E. Kavlakoglu, “¿Qué es la detección de objetos?”, IBM Think.
- [43] D. Bergmann y C. Stryker, “¿Qué es la segmentación de instancias?”, IBM Think.
- [44] N. Jegham, C. Y. Koh, M. Abdelatti, y A. Hendawi, “YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLOv11, and Their Previous Versions”, oct. 2024, [En línea]. Disponible en: <http://arxiv.org/abs/2411.00201>
- [45] X. Yuan, A. Chakravarty, L. Gu, Z. Wei, E. Lichtenberg, y T. Chen, “An Empirical Study of Methods for Small Object Detection from Satellite Imagery”, feb. 2025, [En línea]. Disponible en: <http://arxiv.org/abs/2502.03674>
- [46] G. Boesch, “YOLOv8: A Complete Guide”, viso.ai.
- [47] C. E. Mbonu, K. Anigbogu, D. Asogwa, y T. Belonwu, “An Explorative Analysis of SVM Classifier and ResNet50 Architecture on African Food Classification”, may 2025, [En línea]. Disponible en: <http://arxiv.org/abs/2505.13923>
- [48] Great Learning Editorial Team, “Introduction to Image Pre-processing | What is Image Pre-processing?”, Great Learning .
- [49] P. Raja Rajeswari Chandni, J. Justin, y R. Vanithamani, “Fundus Image Enhancement Using EAL-CLAHE Technique”, en *Lecture Notes in Networks and Systems*, Springer Science and Business Media Deutschland GmbH, 2022, pp. 613–624. doi: 10.1007/978-981-16-5689-7\_54.
- [50] B. Klein, “Evaluation Metrics”, python-course.eu.

- [51] K. R. Srinath, “Python-The Fastest Growing Programming Language”, *International Research Journal of Engineering and Technology*, 2017, [En línea]. Disponible en: [www.irjet.net](http://www.irjet.net)
- [52] Google Research, “Preguntas frecuentes sobre Google Colaboratory”, Google Research.
- [53] D. Bergmann y C. Stryker, “¿Qué es PyTorch?”, IBM.
- [54] B. Hangün y Ö. Eyecioğlu, “Performance Comparison Between OpenCV Built in CPU and GPU Functions on Image Processing Operations”, 2017.
- [55] A. Cavin, “Computer Vision Made Simple with ReductStore and Roboflow”, ReductStore Blog.
- [56] J. I. Orlando *et al.*, “REFUGE Challenge: A Unified Framework for Evaluating Automated Methods for Glaucoma Assessment from Fundus Photographs”, oct. 2019, doi: 10.1016/j.media.2019.101570.
- [57] S. Biswas, M. I. A. Khan, M. T. Hossain, A. Biswas, T. Nakai, y J. Rohdin, “Which Color Channel Is Better for Diagnosing Retinal Diseases Automatically in Color Fundus Photographs?”, *Life*, vol. 12, núm. 7, jul. 2022, doi: 10.3390/life12070973.

# Anexo 1

Máscaras obtenidas de las regiones nasal y temporal mediante YOLOv8.

Baja Nasalización (No nasalizado)			
Imagen Original	Imagen Preprocesada	Máscara Temporal	Máscara Nasal
			
			
Media Nasalización (No nasalizado)			
			
			



## Anexo II

La base de datos de nasalización formada en este trabajo se encuentra disponible en el siguiente enlace:

<https://www.kaggle.com/datasets/antonionava/nasal-and-temporal-mask-retinal-fundus-images>