

Detección de cambios patológicos presentes en el glaucoma a través de la segmentación de fondo de ojo

GLORIA JUDITH MONCADA CARMONA

Trabajo Dirigido

Tutor

Oscar Julián Perdomo Charry, PhD



**Universidad del
Rosario**



**ESCUELA
COLOMBIANA
DE INGENIERÍA
JULIO GARAVITO**

**UNIVERSIDAD DEL ROSARIO
ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO
PROGRAMA DE INGENIERÍA BIOMÉDICA
BOGOTÁ D.C
2023**

AGRADECIMIENTOS

Agradezco a Dios todas las oportunidades que me ha brindado en la vida, en este caso permitiendo culminar mi formación de pregrado. A mis padres y familia quienes con su apoyo incondicional me acompañaron durante toda la carrera formándome como persona en especial a mi padre Jairo Moncada Gelves, agradezco todo su esfuerzo y anhelo en verme salir adelante. Así mismo quiero agradecer a todas aquellas personas que hicieron posible que hoy en día lograra realizar este trabajo dirigido, el cual representa el camino que ha transcurrido durante estos últimos años en mi proceso de formación académica.

Agradezco a mi compañero de vida Gabriel Mauricio García Márquez y a su familia por todo su apoyo, por haber creído en mí y darme la posibilidad de llegar hasta este punto para culminar con éxito mi carrera profesional, agradezco la confianza depositada en mí y en mi trabajo, por sus valiosos consejos, compañía, ayuda, tiempo y motivación de continuar con todo el esfuerzo y dedicación que representó este reto de realizar el proyecto, trabajar en paralelo y afrontar un tiempo de crisis y cambio.

Finalmente agradezco las bendiciones recibidas y la posibilidad de poder estudiar, logrando culminar esta etapa de mi vida, formándome exitosamente como Ingeniera Biomédica dispuesto a contribuir, ayudar e impactar a la sociedad y a mi país. Doy gracias a mi tutor Oscar Julián Perdomo y profesores de la Universidad del Rosario y la Escuela Colombiana de Ingeniería Julio Garavito, por sus valiosos consejos y comentarios los cuales me ayudaron a culminar este proyecto.

TABLA DE CONTENIDO

| | |
|--|----|
| Resumen | 6 |
| 1. INTRODUCCIÓN..... | 7 |
| 1.1 El Ojo humano..... | 7 |
| 1.2 Enfermedad de interés | 8 |
| 1.2.1 Relación Disco/Copa | 8 |
| 1.2.2 Presión intraocular (PIO) | 8 |
| 1.2.3 Diferentes tipos de glaucoma..... | 9 |
| 1.3 Diagnostico del glaucoma | 11 |
| 1.4 Modelos de segmentación..... | 13 |
| 1.4.1 UNET | 13 |
| 1.4.2 Linknet..... | 14 |
| 1.4.3 FPN | 15 |
| 1.4.4 PSPNet..... | 16 |
| 1.5 Redes neuronales convolucionales (CNN) | 17 |
| 1.6 Métrica de desempeño | 18 |
| 1.6.1 Índice de Jaccard (JI)..... | 18 |
| 1.6.2 Índice de Sorensen Dice | 19 |
| 1.7 Medidas estadísticas del error..... | 19 |
| 1.7.1 Matriz de confusión | 20 |
| 1.7.2 Exactitud..... | 20 |
| 1.7.3 Sensibilidad | 20 |
| 1.7.4 Especificidad..... | 21 |
| 1.7.5 Kappa Cohen..... | 21 |
| 2. OBJETIVOS | 22 |
| 2.1. General..... | 22 |
| 2.2. Específicos | 22 |
| 3. METODOLOGÍA | 23 |
| 3.1 Procedimiento | 23 |
| 3.1.1 Exploración sistemática..... | 23 |
| 3.1.2 Validación de modelos | 24 |
| 4. RESULTADOS | 25 |
| 5. DISCUSIÓN..... | 38 |
| 6. RECOMENDACIONES Y TRABAJOS FUTUROS..... | 40 |

| | |
|-----------------------|----|
| 7. CONCLUSIONES | 41 |
| REFERENCIAS | 42 |

LISTADO DE FIGURAS

| | |
|---|----|
| Figura 1. Ojo humano y sus partes. | 7 |
| Figura 2. Diferentes relaciones disco-copa para el diagnóstico de Glaucoma. | 8 |
| Figura 3. Proceso de circulación del humor acuoso dentro del ojo. | 9 |
| Figura 4. El glaucoma de ángulo abierto se ve el canal de Schlemm normal | 10 |
| Figura 5. El glaucoma de ángulo cerrado se muestra el canal de Schlemm totalmente cerrado. | 10 |
| Figura 6. Dispositivo óptico conocido como lámpara de hendidura o biomicroscopio. | 12 |
| Figura 7. Tonómetro para medir presión intraocular. | 12 |
| Figura 8. Campimetría visual. | 13 |
| Figura 9. Red Neuronal Convolucional UNET. | 14 |
| Figura 10. Red Neuronal Convolucional LinkNet | 15 |
| Figura 11. Red Neuronal Convolucional FPN. | 16 |
| Figura 12. Red Neuronal Convolucional PSPNet. | 17 |
| Figura 13. Ejemplo de arquitectura vgg16. | 18 |
| Figura 14. Distribución de clases para la exploración sistemática. | 24 |

Resumen

Introducción: Según estudios de la Organización Mundial de la Salud (OMS) el glaucoma es la segunda causa de ceguera en todo el mundo [1], esta enfermedad puede causar daño progresivo en el nervio óptico llevando a la pérdida irreversible de la visión si no se trata correctamente. En el 2013 se estimó que 64,3 millones de personas de 40 a 80 años padecían glaucoma y se espera que esta cifra aumente a 76 millones para 2020 y 111,8 millones para 2040 [3]. En este contexto las técnicas de procesamiento de imágenes oculares se han convertido en una herramienta esencial para la detección temprana y prevención del glaucoma.

Objetivos: En el desarrollo de este proyecto, se busca encontrar cambios patológicos en la copa y el disco óptico mediante modelos de segmentación usando un procesamiento que realza información importante y extrae características las cuales permiten evaluar el desempeño de la segmentación por medio del índice de Jaccard e índice de Dice y la eficiencia del modelo.

Metodología: La realización del proyecto se organizó mediante un cronograma de tareas y actividades específicas cada una con seis fases delimitadas. La duración total del proyecto fue de 13 semanas, en las cuales se realizaron 5 actividades cada una con fases de planificación, segmentación, entrenamiento, índice de Jaccard, Índice de Dice, validación y visualización de datos, que permitieron lograr el cumplimiento de los objetivos propuestos, definiendo la problemática, ejecutando las ideas desarrolladas y culminando con los requisitos mínimos óptimos y funcionales para el proyecto.

Resultados: En los resultados del presente documento se recopilan datos de los promedios de índices de Dice y Jaccard para determinando el mejor modelo capaz de segmentar imágenes de fondo de ojo y calculando así las métricas correspondientes.

Conclusiones: Se logró realizar una versión primaria de un modelo de segmentación a partir de imágenes de fondo de ojo capaz de clasificar dichas imágenes en glaucomatosas o sanas.

1. INTRODUCCIÓN

1.1 El Ojo humano

Las enfermedades y afecciones presentes en el ojo humano son muy importantes en el campo de la ingeniería biomédica porque contribuye en el desarrollo de equipos y tecnologías que permiten evaluar con precisión estructuras y funciones del ojo humano, por esto es conveniente entender desde un principio la anatomía básica del ojo [6].

El ojo humano es un órgano sensorial que permite visualizar información del entorno, este se compone de varias estructuras (ver Figura 1) [4]. Estas estructuras hacen posible que podamos cumplir labores diarias como lo es ver una película, manejar un automóvil, leer el periódico, entre otras. Al igual que otros órganos del cuerpo, puede verse afectado por enfermedades o lesiones que si no son tratadas a tiempo pueden desencadenar en la pérdida total de la visión por eso la detección temprana y el tratamiento adecuado de estas condiciones son esenciales para preservar la visión y prevenir problemas más graves.

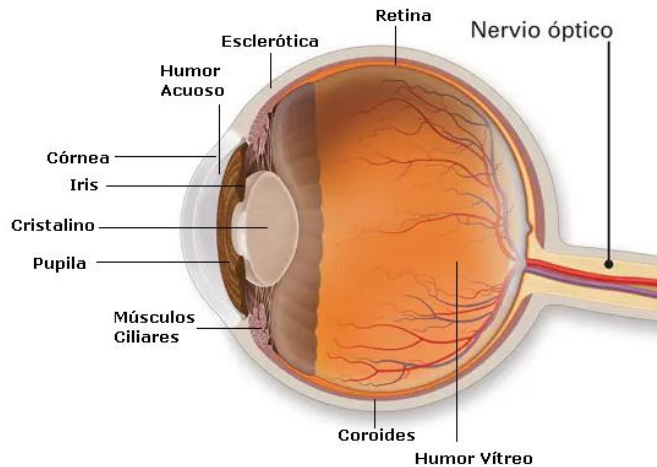


Figura 1. Ojo humano y sus partes. Tomada de [7].

El primer informe mundial publicado por la OMS (Organización mundial de la salud) señala que más de mil millones de personas en todo el mundo padecen trastornos visuales debido a la falta de atención médica adecuada para tratar afecciones como la miopía, la hipermetropía, el glaucoma y las cataratas. De manera global, se estima que al menos 2200 millones de individuos presentan problemas de visión o ceguera, de los cuales al menos 1000 millones tienen una deficiencia visual que podría haber sido prevenida o aún no ha sido tratada. Este panorama refleja la importancia de prestar una atención médica oportuna para detectar y tratar estos trastornos visuales, y así prevenir complicaciones mayores [8].

1.2 Enfermedad de interés

Una de las afecciones más importantes es el glaucoma, ya que de acuerdo con la Organización Mundial de la Salud (OMS), esta enfermedad ocupa el segundo lugar como causa habitual de ceguera, afectando a 4,5 millones de personas en la actualidad. La OMS proyecta que para el año 2020, la cantidad de afectados aumentará a 11 millones [1]. Por este motivo se enfocó este proyecto en detectar el glaucoma, buscando prevenir daños irreversibles causados por la enfermedad mediante una detección temprana que permita tomar decisiones a tiempo.

Esta enfermedad tiene un impacto directo en dos componentes clave del ojo, específicamente en el disco óptico y la copa óptica. El glaucoma está asociado con niveles elevados de presión intraocular (PIO), y este factor ha sido destacado en investigaciones nacionales. Sin embargo, en la actualidad, se considera que la detección del glaucoma en campañas de salud pública basada únicamente en la medición de la presión intraocular elevada no sería lo suficientemente confiable [9] ya que el daño visual puede estar presente sin aumento de la PIO [45].

1.2.1 Relación Disco/Copa

La forma del disco óptico puede variar entre ser plano o mostrar cierta cantidad de excavación considerada normal. Conforme la copa óptica aumenta en tamaño en comparación con el disco óptico, también progresa la excavación del nervio óptico. Este proceso se debe al aumento de la presión en el ojo y/o la reducción del flujo sanguíneo al nervio óptico, lo que provoca la degeneración de las fibras nerviosas y un aumento en la excavación ocupando la mayor parte del área del disco óptico (ver Figura 2) [12][45].

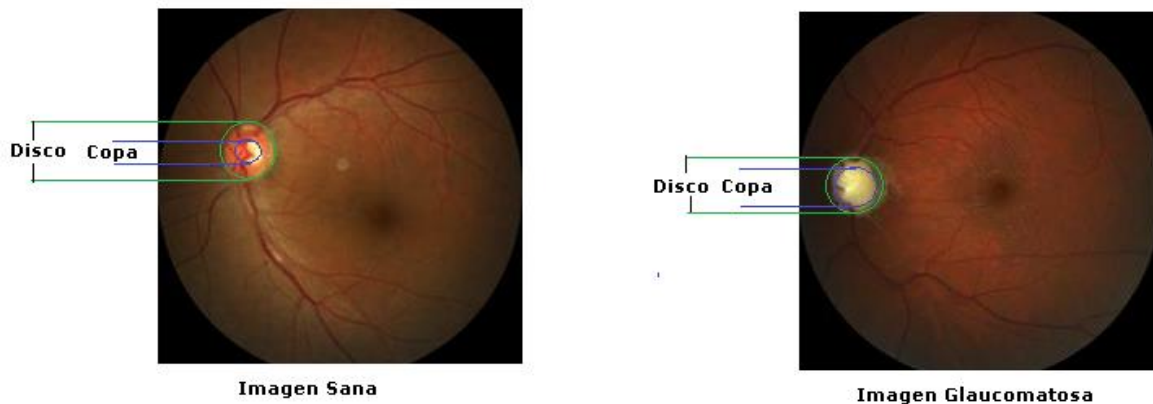


Figura 2. Diferentes relaciones disco-copa para el diagnóstico de Glaucoma. Tomada de [45].

1.2.2 Presión intraocular (PIO)

La presión intraocular del ojo puede variar en diferentes momentos del día, pero normalmente se mantiene dentro de un rango que el ojo puede tolerar [18]. En un ojo sano, el humor acuoso se renueva continuamente en pequeñas cantidades, mientras que la misma cantidad se drena fuera del ojo. La mayor parte de este líquido fluye hacia afuera del ojo a través del ángulo de drenaje ubicado frente al iris, lo que garantiza una presión estable al mantener el flujo de entrada y salida en equilibrio [23].

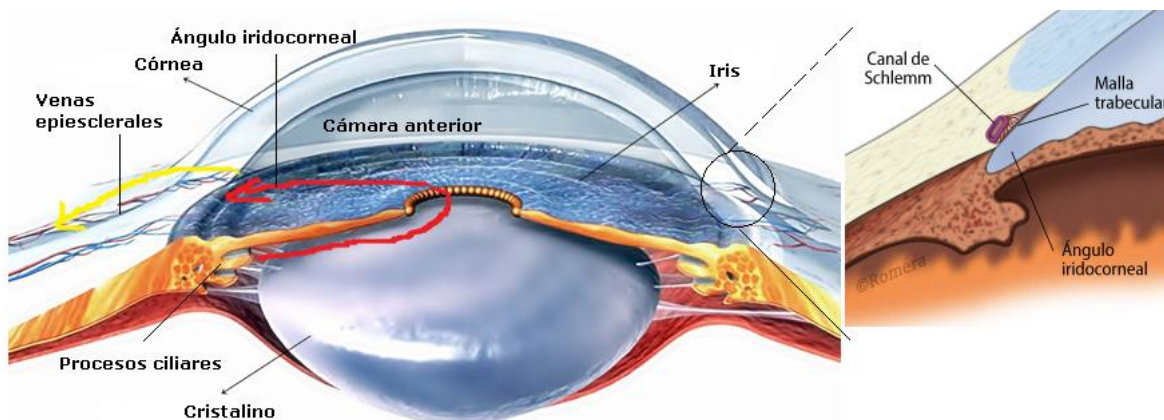


Figura 3. Este esquema ilustra el proceso de circulación del humor acuoso dentro del ojo. La producción de humor acuoso se lleva a cabo en los procesos ciliares y fluye a través de la pupila hacia la cámara anterior. Desde allí, es eliminado mediante la malla trabecular y el canal de Schlemm, lo que permite que el líquido sea llevado a la circulación general. Tomado de [24] [25].

La pérdida de la visión se produce cuando hay daños en el nervio óptico y el líquido intraocular o humor acuoso producido por el cuerpo ciliar, detrás del iris no se drena correctamente a través de la malla trabecular en la parte frontal del ojo, estas anomalías en este sistema de drenaje, resulta en una acumulación de presión intraocular (PIO) que puede dañar el nervio óptico y afectar la visión. En otros casos, el glaucoma puede ocurrir aunque la PIO sea relativamente normal, debido a una debilidad del ojo para soportar el estrés mecánico en las fibras nerviosas o por una deficiente irrigación sanguínea [22].

1.2.3 Diferentes tipos de glaucoma

Existen varios tipos de glaucoma. Los dos tipos principales son el de ángulo abierto y el de ángulo cerrado.

1.2.3.1 Glaucoma de ángulo abierto (GPA)

El glaucoma de ángulo abierto, también conocido como glaucoma primario o crónico, es el tipo más común de glaucoma, donde el ángulo iridocorneal tiene un ancho y una apertura normales. La enfermedad se desarrolla gradualmente a medida que los canales de drenaje del ojo se obstruyen con el tiempo, lo que causa un aumento en la presión intraocular. A diferencia de otros tipos de glaucoma, no hay signos tempranos de advertencia para detectar el glaucoma de ángulo abierto, y la pérdida de la visión puede pasar desapercibida

durante muchos años. Por esta razón, someterse a exámenes oculares regulares es crucial para detectar y tratar la enfermedad a tiempo [18].

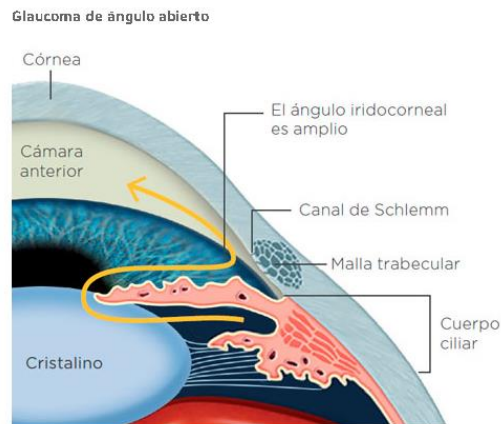


Figura 4. El glaucoma de ángulo abierto se ve el canal de Schlemm normal. Tomado de [27].

1.2.3.2 Glaucoma de ángulo cerrado

El glaucoma de ángulo cerrado es un tipo de glaucoma que se produce cuando el ángulo entre el iris y la córnea se estrecha, lo que hace que el borde externo del iris se pegue a los canales de drenaje. Aunque este problema estructural ocurre solo en un pequeño porcentaje de los ojos, puede causar un ataque de dolor en el ojo, enrojecimiento ocular, dolor de cabeza, náuseas, visión de un arcoíris alrededor de las luces nocturnas o visión borrosa. Es importante buscar atención médica de inmediato si experimenta alguno de estos síntomas para evitar daños graves en el ojo [18].

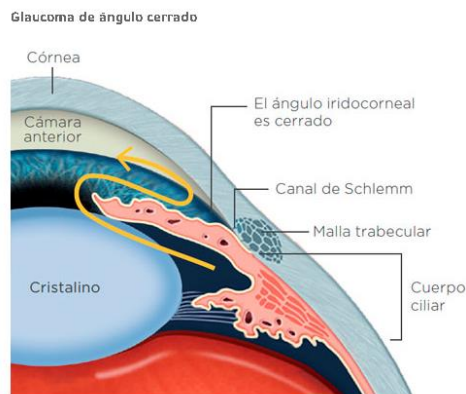


Figura 5. El glaucoma de ángulo cerrado se muestra el canal de Schlemm totalmente cerrado. Tomado de [27].

1.2.3.3 Otros tipos de glaucoma

La mayoría de los otros tipos de glaucoma son versiones modificadas de los tipos de ángulo abierto o cerrado. Estos tipos pueden manifestarse en uno o en ambos ojos [18].

- **Glaucoma de tensión normal**
Se produce daño en el nervio óptico, aunque la presión intraocular no sea alta y los médicos desconocen las razones detrás de esto. Aquellas personas con antecedentes familiares de glaucoma de tensión normal, origen japonés o problemas de corazón sistémicos, como latidos irregulares, tienen un mayor riesgo de desarrollar esta forma de glaucoma [18].
- **Glaucoma congénito**
Se trata de una enfermedad poco común que se presenta en bebés y puede ser hereditaria. Su causa es el desarrollo incorrecto o incompleto de los canales de drenaje del ojo durante la gestación. En casos no complicados, se puede corregir mediante microcirugía para reparar los defectos estructurales. En casos más graves, se utilizan medicamentos y cirugía para tratar la enfermedad [18].
- **Glaucoma neovascular**
Esta enfermedad se produce cuando hay una producción excesiva de vasos sanguíneos que cubren la parte del ojo encargada de drenar el líquido. Por lo general, esta afección se asocia con otras enfermedades como diabetes o presión arterial alta. Los síntomas comunes de este tipo de glaucoma incluyen enrojecimiento o dolor en el ojo y pérdida de visión. La dificultad en el tratamiento radica en tratar la causa subyacente junto con la reducción de la presión ocular generada [30].
- **Glaucoma pigmentario**
Esta enfermedad se presenta cuando los diminutos gránulos de pigmento ubicados en la parte trasera del iris, ingresan al humor acuoso que se produce en el interior del ojo. Estos gránulos de pigmento obstruyen lentamente los canales de drenaje del ojo, provocando un aumento en la presión intraocular. Por lo general, esta forma se trata con medicamentos, cirugía láser o cirugía filtrante [18].
- **Glaucoma de exfoliación**
El glaucoma de exfoliación, también conocido como pseudoexfoliación, es un tipo de glaucoma de ángulo abierto que se produce en algunas personas con síndrome de exfoliación. Esta afección causa la acumulación excesiva de material en ciertas partes del ojo, lo que bloquea el drenaje del líquido y aumenta la presión intraocular. Según estudios este tipo de glaucoma progresa más rápido y la genética puede ser un factor de riesgo [30].

1.3 Diagnostico del glaucoma

El diagnóstico del glaucoma solo puede realizarse a través de un examen oftalmológico completo que incluya la evaluación de múltiples factores como la medición de la presión intraocular, el ángulo de drenaje del ojo, inspección del nervio óptico y evaluación de la visión periférica del paciente [31].

Existen diferentes tecnologías capaces de descartar otras causas y encontrar hallazgos de daños del nervio óptico [26], estas tecnologías son:

La lámpara de hendidura: Es un instrumento utilizado en el examen oftalmológico en el que el paciente se sienta en una silla, coloca la barbilla y la frente en un microscopio. Con una luz de gran aumento, se pueden examinar detalladamente todas las estructuras del ojo y determinar las diferentes afecciones presentes en la misma [32].



Figura 6. *Dispositivo óptico conocido como lámpara de hendidura o biomicroscopio. Tomado de [33].*

Tonometría: La tonometría es una técnica para medir la presión intraocular utilizando un colorante con anestésico en el ojo del paciente y un dispositivo especial, un cono con luz azul, para medir la presión del ojo de forma indolora. Este procedimiento se realiza en la misma lámpara de hendidura utilizada para examinar las estructuras oculares [31].

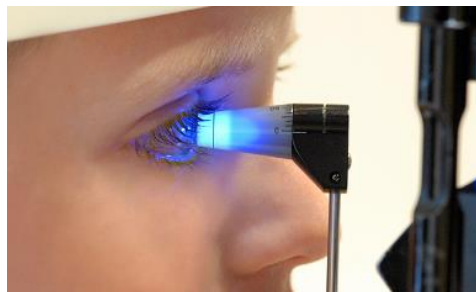


Figura 7. *Tonómetro para medir presión intraocular. Tomado de [34].*

Campimetría: La campimetría es una prueba que se realiza en un instrumento especializado, en la que el paciente presiona un botón cuando ve luces periféricas mientras mantiene la mirada fija en una luz central de referencia. Esta prueba es útil para evaluar la visión periférica del paciente y suele ser supervisada por un optometrista [31].

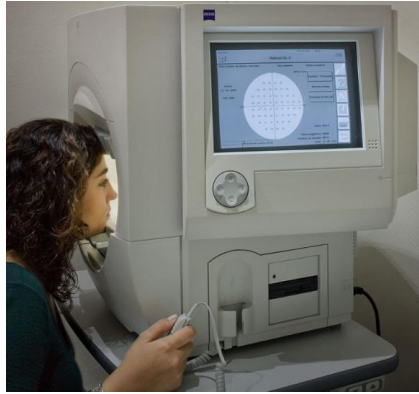


Figura 8. *campimetría visual. Tomado de [35].*

1.4 Modelos de segmentación

La visión artificial es una subcategoría de la inteligencia artificial que busca que los ordenadores puedan entender y procesar imágenes y vídeos de forma automática. El Deep Learning ha permitido grandes avances en este campo y ahora las máquinas pueden competir con la visión humana en algunas situaciones. La visión artificial puede realizar tareas diferentes como la clasificación de imágenes, la detección de objetos, la segmentación semántica y la segmentación de instancias, que permiten al ordenador comprender el contenido de las imágenes de manera cada vez más precisa [36].

Actualmente podemos encontrar diferentes métodos que nos permiten resolver problemas de segmentación semántica, estos métodos continúan con los enfoques tradicionales de detectar puntos, líneas o bordes, aunque también busca detectar grupos de píxeles o morfologías [36].

1.4.1 U-NET

La red neuronal U-NET es un modelo convolucional utilizado para la segmentación de imágenes. Esta red consta de dos vías, la primera es de contracción o codificador, que su función es captura el contexto de una imagen mediante capas de convolución y max pooling, la segunda es de expansión simétrica o decodificador, que permite una localización precisa mediante la convolución transpuesta [36].

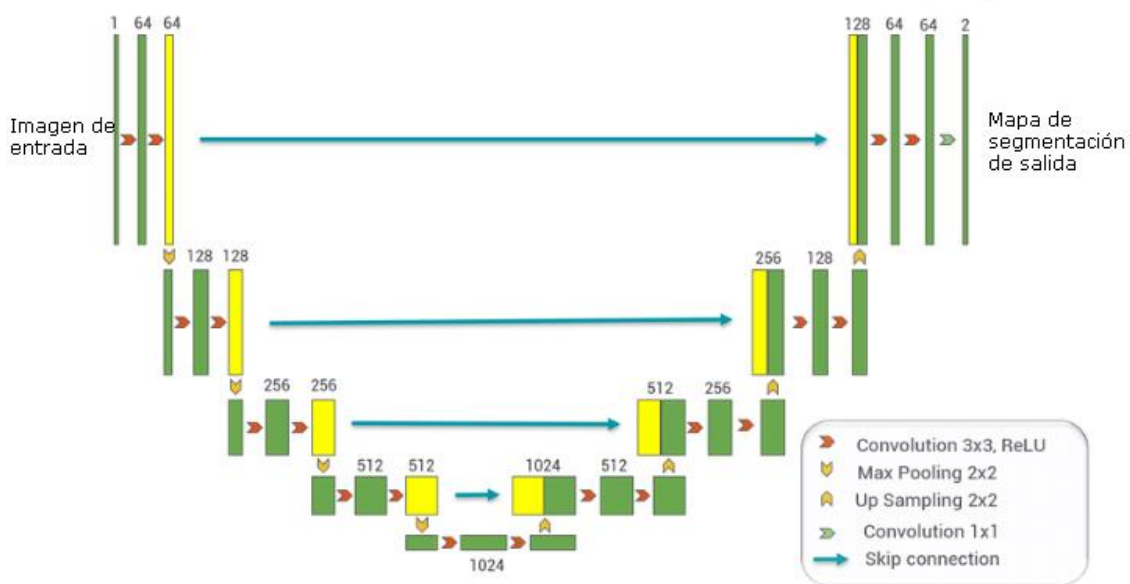


Figura 9. Red Neuronal Convolutacional UNET. Tomado de [14]

U-NET se utiliza para resolver problemas de clasificación de imágenes que requieren grandes cantidades de datos etiquetados. A diferencia de otras arquitecturas U-NET no comprime la entrada de forma lineal, lo que evita la pérdida de características. Este modelo es eficaz incluso con una serie de datos limitada y ofrece una mayor precisión que los modelos convencionales, por lo que es una herramienta útil en campos especializados como el diagnóstico médico. En general, U-NET permite solucionar problemas de clasificación de imágenes con una mayor precisión y con una serie de datos más reducida [36].

1.4.2 Linknet

Linknet es una red neuronal convolutacional para la segmentación semántica de imágenes médicas. La motivación detrás de Linknet es que los enfoques actuales de segmentación semántica tienen limitaciones en términos de precisión, eficiencia y generalización. Linknet se basa en la arquitectura de la red neuronal U-Net, que ha demostrado ser efectiva en la segmentación semántica, y mejora su eficiencia y precisión al incorporar capas de enlace en la red. Las capas de enlace mejoran la propagación de información entre las diferentes capas de la red, lo que permite una mejor representación de las características de las imágenes. Linknet también utiliza técnicas de entrenamiento como la normalización de lotes y la tasa de abandono para evitar el sobreajuste y mejorar la generalización. Los resultados experimentales demuestran que Linknet supera a los enfoques de vanguardia en términos de precisión y eficiencia en la segmentación semántica de imágenes médicas [37].

Su arquitectura consta de dos partes: un encoder y un decoder. El encoder es similar al de U-Net, con capas convolucionales que reducen la dimensionalidad de la entrada y extraen características relevantes. Sin embargo, Linknet utiliza una técnica llamada "skip connection" que conecta las capas del encoder directamente con las del decoder,

permitiendo que la información se transmita de manera más eficiente y conservando detalles de la imagen original en las capas del decodificador. El decodificador, por su parte, utiliza capas deconconvolucionales para aumentar la resolución de la salida y generar una máscara de segmentación precisa. En conjunto, la arquitectura de Linknet permite una segmentación de imágenes más precisa y eficiente que otros modelos convolucionales [37].

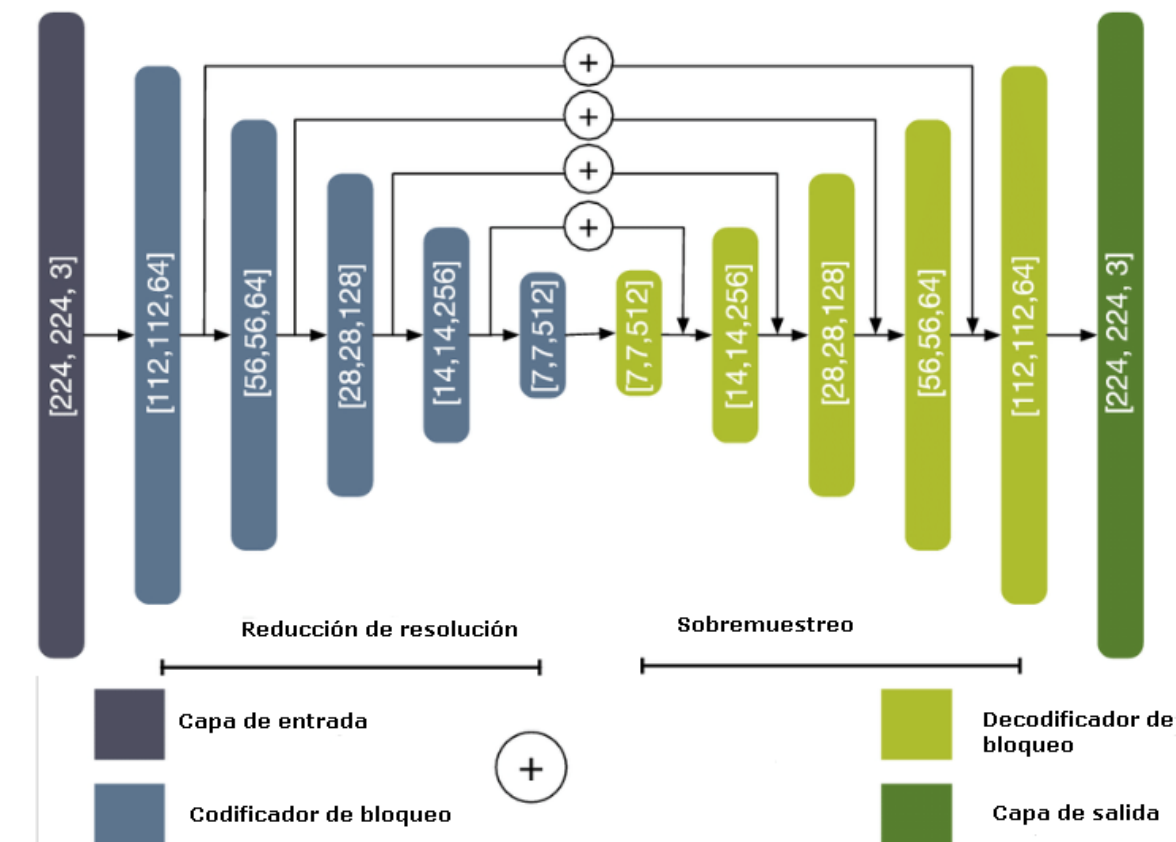


Figura 10. Red Neuronal Convolutional LinkNet. Tomado de [50]

1.4.3 FPN

El Feature Pyramid Network (FPN) es un extractor de características que toma una imagen de escala única de tamaño arbitrario como entrada y produce mapas de características de tamaños proporcionales en múltiples niveles de manera completamente convolucional. Actúa como una solución genérica que permite construir pirámides de características dentro de las redes convolucionales profundas para su uso en tareas como la detección de objetos [38].

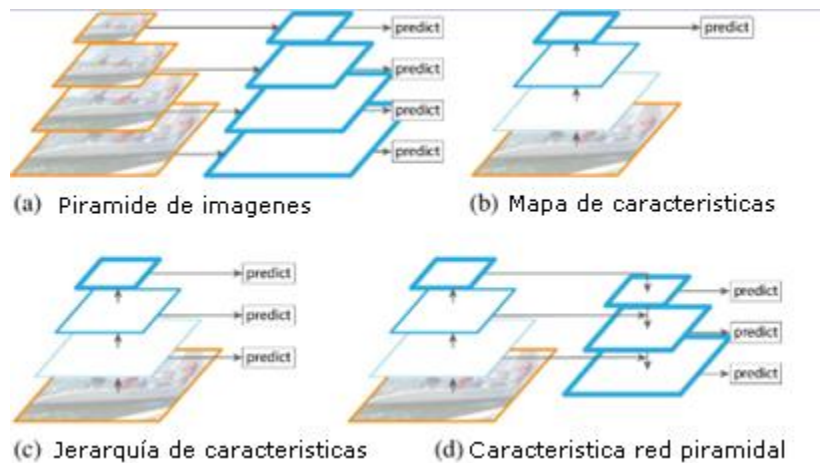


Figura 11. Red Neuronal Convolutiva FPN. Tomado de [38]

Para construir la pirámide, se emplea un camino ascendente y un camino descendente. El camino ascendente es el cálculo feedforward de la ConvNet principal, que genera una jerarquía de características conformado por distintos mapas de características en varias escalas. Se define un nivel de la pirámide de características para cada etapa y se utiliza la salida de la última capa de cada etapa como un conjunto de mapas de características de referencia [38].

El camino descendente imagina características de resolución más alta mediante el muestreo de mapas de características de niveles superiores de la pirámide. Estas características se mejoran con características del camino ascendente a través de conexiones laterales, cada conexión lateral fusiona mapas de características del mismo tamaño espacial del camino ascendente y el camino descendente [38].

1.4.4 PSPNet

PSPNet es un modelo de segmentación semántica que emplea un módulo de análisis piramidal para aprovechar la información contextual global al agregar contexto de diferentes regiones. Al combinar pistas locales y globales, se mejora la confiabilidad de la predicción final. Para una imagen de entrada, PSPNet utiliza una red neuronal convolutiva (CNN) preentrenada con una estrategia de red dilatada para extraer un mapa de características. Este mapa tiene un tamaño final que es $1/8$ del tamaño de la imagen original. Luego, se aplica el módulo de agrupación piramidal para recopilar información contextual [39].

Mediante una pirámide de 4 niveles, los núcleos de agrupación cubren la imagen completa, la mitad y pequeñas porciones de la imagen. Finalmente, se concatenan los resultados del módulo de agrupación piramidal con el mapa de características original y se utiliza una capa de convolución para generar el mapa de predicción final [39].

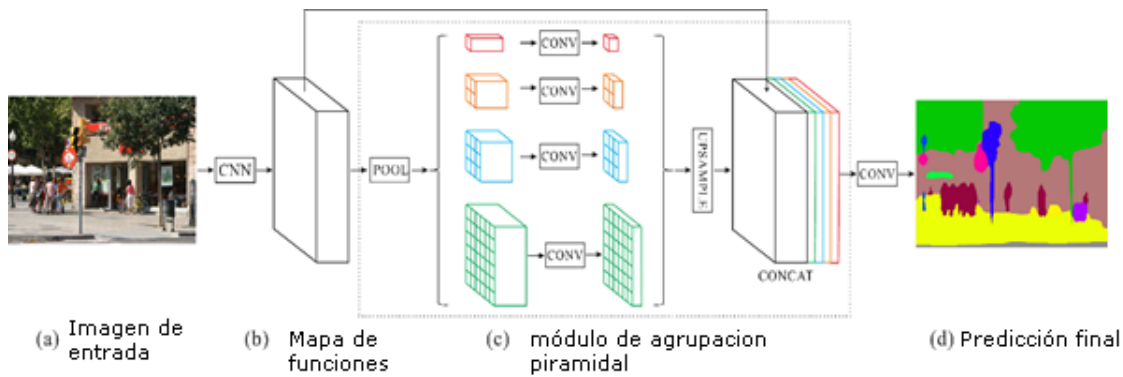


Figura 12. Red Neuronal Convolutiva PSPNet. Tomado de [39]

1.5 Redes neuronales convolucionales (CNN)

Las redes neuronales convolucionales (CNN) son un tipo de algoritmo de aprendizaje profundo diseñado para procesar datos que tienen una estructura de cuadrícula o matricial, ya sea espacial o temporal. Las CNN son similares a otras redes neuronales, pero presentan una complejidad adicional debido a su utilización de capas convolucionales. Estas capas convolucionales son un componente fundamental de las CNN [47].

Las arquitecturas son filtros convolucionales de suma importancia a la hora realizar tareas de clasificación y reconocimiento de imágenes, son altamente utilizados para capturar información de una imagen como formas, texturas y demás detalles a profundidad [48].

A nivel general las arquitecturas son una estructura repetitiva y profunda por distintas capas, en la que cada una de ellas contiene filtros de tamaño reducido a distintas profundidades que le permite extraer una mayor cantidad de información reconociendo partes de las imágenes y logrando una óptima representación de las mismas [48].

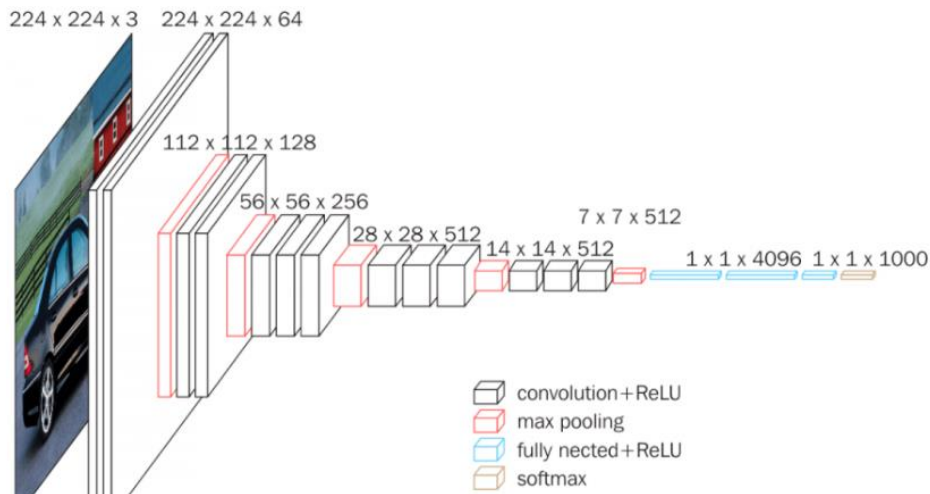


Figura 13. Ejemplo de arquitectura vgg16. Tomado de [48].

A nivel explicativo lo que busca estas arquitecturas es ir adentrándose cada vez más en los detalles de la imagen y por medio de sus filtros en distintas capas extraer características para después procesarla en sus capas de salida [48].

Los modelos utilizados en redes neuronales convolucionales fueron Resnet152, vgg16 y vgg19. Cada modelo cuenta con la precisión top-1 y top-5 las cuales se refiere al desempeño del modelo en el conjunto de datos de validación de ImageNet [51].

La profundidad se refiere a la estructura enredada de la red, lo cual incluye capas de activación, capas de normalización por lotes, entre otros componentes [51].

El tiempo por paso de inferencia se calcula como el promedio de 30 lotes repetidos 10 veces [51].

| Model | Size (MB) | Top-1 Accuracy | Top-5 Accuracy | Parameters | Depth | Time (ms) per inference step (CPU) | Time (ms) per inference step (GPU) |
|-----------|-----------|----------------|----------------|------------|-------|------------------------------------|------------------------------------|
| ResNet152 | 232 | 76.6% | 93.1% | 60.4M | 311 | 127.4 | 6.5 |
| VGG16 | 528 | 71.3% | 90.1% | 138.4M | 16 | 69.5 | 4.2 |
| VGG19 | 549 | 71.3% | 90.0% | 143.7M | 19 | 84.8 | 4.4 |

Tabla 19. Parámetros de los modelos utilizado. Tomado de [51]

1.6 Métricas de desempeño

1.6.1 Índice de Jaccard (JI)

El índice de Jaccard es una medida estadística utilizada para comparar la similitud y la diversidad de conjuntos. En el contexto de la segmentación de imágenes, se utiliza para

evaluar la superposición entre la segmentación de la imagen predicha y la segmentación de la imagen de referencia. El índice de Jaccard se define como la relación entre la intersección de las dos regiones dividida por la unión de las mismas. En términos más simples, mide la cantidad de píxeles comunes entre las dos segmentaciones dividida por la cantidad de píxeles totales de ambas segmentaciones. Cuanto más cercano a 1 sea el índice de Jaccard, mayor será la similitud entre las dos segmentaciones [43].

El índice de Jaccard se calcula dividiendo el tamaño de la intersección de los dos conjuntos entre el tamaño de la unión de los mismos conjuntos. La fórmula matemática para el índice de Jaccard se muestra en la ecuación (1) [43].

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Donde A y B son los dos conjuntos de datos que se están comparando, |A| es el tamaño de A y |B| es el tamaño de B [43].

1.6.2 Índice de Sorensen Dice

El índice de Dice, también conocido como índice de Sorensen (S), se emplea en imágenes binarias para evaluar las discrepancias entre ellas. Se determina comparando los objetos de interés segmentados en la imagen U con los de referencia en la imagen V a través de la intersección de ambas imágenes. Esta intersección se divide entre la cantidad de píxeles correspondientes a los objetos de interés en ambas imágenes. Una segmentación ideal se alcanza cuando el índice S tiende a "1", mientras que una segmentación deficiente se refleja en valores cercanos a "0" para S [20].

$$S = \frac{2|U \cap V|}{|U| + |V|} \quad (2)$$

La medida de distancia se obtiene de manera similar al índice de Jaccard, restando 1 al valor de la similitud. Puedes calcular el índice de Jaccard (4) a partir del coeficiente de Dice (3) y viceversa utilizando las siguientes ecuaciones [21].

$$D = \frac{2J}{(1 + J)} \quad (3)$$

$$J = \frac{D}{(2 - D)} \quad (4)$$

1.7 Medidas estadísticas del error

Es importante utilizar una variedad de métricas de evaluación para mejorar la capacidad predictiva de un modelo de aprendizaje automático. Estas métricas desempeñan un papel fundamental en proyectos de ciencia de datos al permitir evaluar el rendimiento del modelo. Su objetivo principal es estimar qué tan precisa será la generalización del modelo en datos futuros, ya que este podría estar limitado a memorizar en lugar de aprender. Esta limitación

dificulta la capacidad del modelo para adaptarse a situaciones con datos nuevos y desconocidos [40].

1.7.1 Matriz de confusión

La matriz de confusión es una forma matricial de representar los resultados de las predicciones en pruebas binarias, y se utiliza con frecuencia para describir el rendimiento de un modelo de clasificación o clasificador en un conjunto de datos de prueba cuyos valores reales son conocidos. Esta matriz nos permite conocer con qué frecuencia las predicciones son clasificadas de forma incorrecta, y para ello se utilizan cuatro variables de evaluación que forman parte de esta medida [40].

A continuación, se describen cuatro variables de evaluación que se utilizan para medir la precisión de la clasificación de las predicciones en una imagen [40].

Verdadero Positivo (TP): predicción y realidad coinciden y son verdaderos.

Verdadero Negativo (TN): predicción y realidad coinciden y son falsos.

Falso Positivo (FP): predicción verdadera, pero la realidad es falsa.

Falso Negativo (FN): predicción falsa, pero la realidad es verdadera [40].

1.7.2 Exactitud

La precisión es una medida de evaluación utilizada comúnmente para problemas de clasificación en los que las clases tienen tamaños aproximadamente iguales. Esta métrica proporciona la proporción de predicciones correctas realizadas en comparación con el total de predicciones. Por otro lado, la tasa de error es una métrica que nos indica con qué frecuencia nuestro modelo clasifica de manera incorrecta. Esta tasa se puede calcular restando la precisión de uno. En otras palabras, la tasa de error es el porcentaje en que nuestro modelo se equivoca [40].

$$\text{Exactitud} = \frac{(VP + VN)}{Total} \quad (5)$$

$$\text{Tasa de clasificación errónea} = \frac{(FP + FN)}{Total} \quad (6)$$

1.7.3 Sensibilidad

Se puede definir la sensibilidad como la proporción de que una persona que tiene una enfermedad tenga un resultado de prueba positivo. Cuando la sensibilidad de una prueba aumenta, la probabilidad de que la prueba arroje un resultado negativo falso para personas que realmente tienen la enfermedad disminuye [42].

$$\text{Sensibilidad} = \frac{(VP)}{(VP + FN)} \quad (5)$$

1.7.4 Especificidad

La especificidad es la proporción de verdaderos negativos a todos los que deberían haber sido clasificados como negativos. Ambas, la especificidad y la sensibilidad, se consideran en conjunto para analizar la matriz de confusión completa [40].

$$\text{Especificidad} = \frac{VN}{(VN + FP)} \quad (7)$$

1.7.5 Kappa Cohen

La medida estadística de Kappa de Cohen se emplea para evaluar el grado de concordancia entre evaluadores o jueces que categorizan elementos en categorías mutuamente excluyentes. La medida en que los evaluadores asignan la misma puntuación a una variable se llama fiabilidad entre evaluadores. Aunque hay varios métodos para medir la fiabilidad entre evaluadores, la medida tradicional se calcula como el porcentaje de acuerdo, es decir, el número de puntajes acordados dividido por el número total de puntajes [41].

$$K = \frac{(P_0 - P_e)}{(1 - P_e)} \quad (8)$$

dónde:

P_0 : Acuerdo relativo observado entre evaluadores
 P_e : Probabilidad hipotética de concordancia al azar [41].

| Kappa (K) | Grado de acuerdo |
|------------------|-------------------------|
| <0.00 | Sin acuerdo |
| 0.00-0.20 | Insignificante |
| 0.21-0.40 | Mediano |
| 0.41-0.60 | Moderado |
| 0.61-0.80 | Sustancial |
| 0.81-1.00 | Casi perfecto |

Tabla 1. Grado de acuerdo para diferentes valores de Kappa. Tomado de [46]

2. OBJETIVOS

2.1. General

Detectar cambios patológicos presentes en el disco y la copa óptica utilizando modelos de segmentación y clasificación aplicados en imágenes de fondo de ojo.

2.2. Específicos

- a) Segmentar automáticamente el disco óptico y copa fisiológica en imágenes de fondo de ojo.
- b) Extraer características para clasificar imágenes de fondo de ojo sanas y glaucomatosas
- c) Validar la habilidad de los modelos desarrollados en conjuntos de prueba.

3. METODOLOGÍA

3.1 Procedimiento

En enero de 2023 se inició el proceso de investigación para la elaboración del trabajo dirigido en el campo de ingeniería biomédica junto con el Ingeniero Oscar Julián Perdomo Charry. En primera instancia, se realizó una revisión bibliográfica sobre los distintos modelos de redes neuronales más aplicadas en el área de la medicina, más específicamente en el procesamiento de imágenes médicas que en nuestro caso son imágenes de fondo de ojo, así como también se estudió la teoría de la patología del glaucoma buscando que partes del ojo interno es importante analizar para detectar esta patología.

La primera parte consistió en la recopilación y visualización de las imágenes de la base de datos brindada por la organización IEEE en el reto REFUGE (retinal fundus glaucoma challenge) planteado en el año 2019 [45]. Esta base de datos está conformada por 3 grupos de imágenes identificados como TEST, VALIDATION y TRAIN, cada uno de estos grupos se conforma por 400 imágenes de fondo de ojo con sus respectivas anotaciones permitiéndonos clasificar las imágenes como patológicas o sanas con el fin de validar la eficiencia de los modelos de segmentación.

3.1.1 Exploración sistemática

Para iniciar esta parte del proyecto se agrupo la base de datos en dos segmentos de estudio importantes para la detección del glaucoma, como ya vimos anteriormente estos segmentos son el disco y la copa óptica. Con el fin de encontrar el mejor modelo convolucional para cada uno de estos grupos, se plateó un set de pruebas con cada combinación posible entre las arquitecturas y los modelos planteados en la introducción y se analizó los valores promediados del índice de Jaccard en cada caso.

Las combinaciones de modelos y arquitecturas usadas se presentan en la tabla a continuación.

| Arquitecturas Modelo | Vgg16 | Vgg19 | Resnet152 |
|---------------------------------------|--------------|--------------|------------------|
| Unet | x | x | x |
| Linknet | x | x | x |
| FPN | x | x | x |
| PSPNet | x | x | x |

Tabla 2. Modelos y arquitecturas obtenidas de la exploración sistemática

Aparte del análisis realizado con el índice de Jaccard, se realizó una grilla de imágenes en la que se visualiza la imagen original, la imagen anotada por un especialista y la imagen

obtenida por el modelo, con el fin de observar gráficamente el resultado del modelo (Ver anexos).

3.1.2 Validación de modelos

Luego de finalizar esta etapa de exploración sistemática se entrenó una red CNN MobileNet para la clasificación de imágenes, usando imágenes RGB de fondo de ojo.

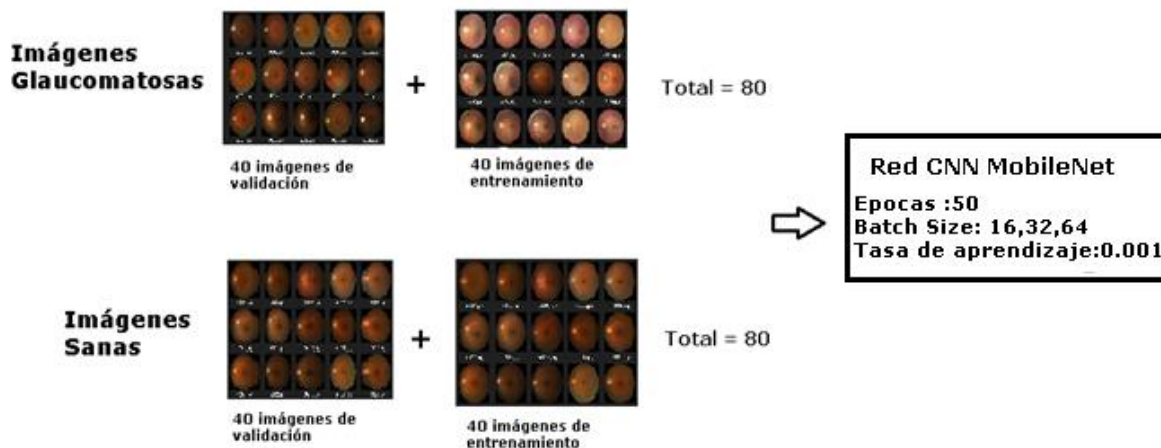


Figura 14. Distribución de clases para la exploración sistemática.

Con el fin de validar los resultados del entrenamiento de la red MobileNet se plantearon 27 pruebas distintas con la base de datos de imágenes de fondo de ojo. Se dividieron las pruebas en tres grupos, en cada uno de ellos aumentando el parámetro de “Batch Size” con el fin de analizar con cuál de estas configuraciones obteníamos mejores resultados.

En cada grupo contábamos con 720 imágenes correspondientes a pacientes sanos, estas fueron divididas en 9 sub grupos de 80 imágenes donde cada sub grupo correspondía a un experimento diferente.

Cada caso de prueba consistía en entrenar la red MobileNet con 80 imágenes glaucomatosas y las 80 imágenes sanas correspondientes a cada experimento, analizando por medio de las métricas dadas (Matriz de confusión, especificidad, exactitud, sensibilidad y valor de Kappa) cual modelo nos daba un mejor resultado.

4. RESULTADOS

En la fase 1a y 1b (expuesta más adelante) muestran los resultados que se obtuvieron en el desarrollo de la metodología correspondientes a la fase de exploración sistemática. A continuación, en la tabla 3, 4, 8 y 9 se presentan los promedios del índice de Jaccard e índices de Dice para cada segmento de estudio en imágenes de entrenamiento e imágenes de test arrojadas por cada modelo de red convolucional utilizado.

En las tablas 5, 6 y 7 se puede observar los resultados obtenidos correspondientes a las grillas de imágenes de cada una de las arquitecturas utilizadas en disco óptico de un paciente, en ellas se puede comparar visualmente los resultados de cada arquitectura con los distintos modelos de entrenamiento.

Al igual que en la sección 1a, en las tablas 10, 11 y 12 se presenta la grilla de imágenes de las distintas arquitecturas utilizadas, pero en esta ocasión se visualizan para copa óptica del mismo paciente.

1.a Resultados de exploración sistemática para disco óptico

| PROMEDIO DE ÍNDICES DE JACCARD PARA DISCO ÓPTICO | | | |
|--|------------------|------------------|------------------|
| Arquitecturas Modelos | Vgg16 | Vgg19 | resnet152 |
| Unet | Traning: 0.6316 | Traning : 0.5683 | Traning : 0.6222 |
| | Test: 0.8068 | Test : 0.7702 | Test : 0.5830 |
| Linknet | Traning : 0.5864 | Traning: 0.5678 | Traning : 0.6283 |
| | Test: 0.7569 | Test: 0.7848 | Test : 0.5952 |
| FPN | Traning : 0.5269 | Traning :0.4966 | Traning : 0.6075 |
| | Test : 0.5301 | Test : 0.6489 | Test : 0.6680 |
| PSPNet | Traning : 0.5746 | Traning: 0.5973 | Traning : 0.633 |
| | Test:0.7676 | Test:0.6977 | Test: 0.6140 |

Tabla 3. Promedios de índices de Jaccard para segmento de disco óptico

| PROMEDIO DE ÍNDICES DE DICE PARA DISCO ÓPTICO | | | |
|---|------------------|------------------|------------------|
| Arquitecturas Modelos | Vgg16 | Vgg19 | resnet152 |
| Unet | Training: 0,7175 | Training: 0,6746 | Training: 0,7281 |
| | Test: 0,8846 | Test: 0,8511 | Test: 0,6994 |
| Linknet | Training: 0,6959 | Training: 0,671 | Training: 0,7323 |
| | Test: 0,8458 | Test: 0,8649 | Test: 0,7157 |
| FPN | Training: 0,6643 | Training: 0,6354 | Training: 0,7238 |
| | Test: 0,6531 | Test: 0,7775 | Test: 0,7723 |
| PSPNet | Training: 0,6781 | Training: 0,6958 | Training: 0,7307 |
| | Test: 0,8587 | Test: 0,8067 | Test: 0,727 |

Tabla 4. Promedios de índices de Dice para segmento de disco óptico

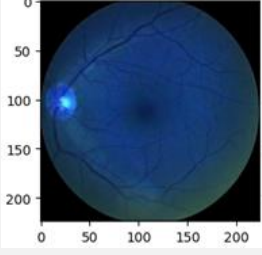
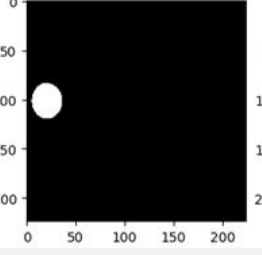
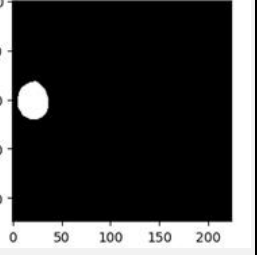
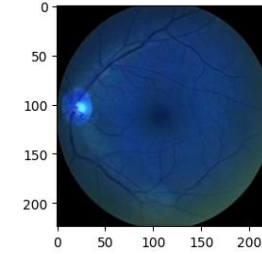
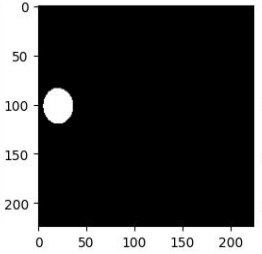
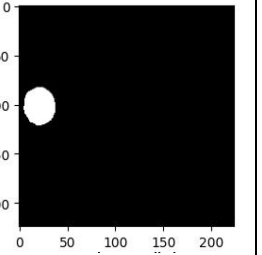
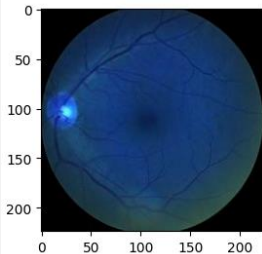
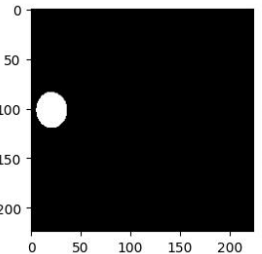
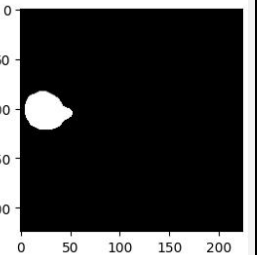
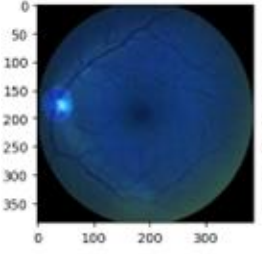
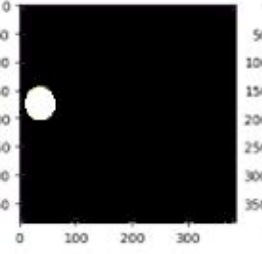
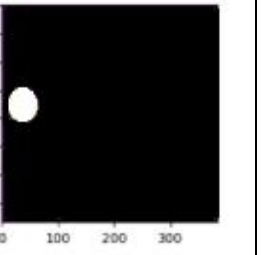
| | | Imagen Original | Imagen Anotada | Imagen Predictiva |
|--------------|---------|---|--|---|
| ARQUITECTURA | MODELO | | | |
| VGG16 | UNET |  |  |  |
| | LinkNet |  |  |  |
| | FPN |  |  |  |
| | PSPNet |  |  |  |

Tabla 5. Arquitectura vgg16 para disco óptico

Imagen Original Imagen Anotada Imagen Predictiva

| ARQUITECTURA | MODELO | |
|--------------|---------|---|
| VGG19 | UNET | <p>The UNET model results show three images side-by-side. The first is the original fundus image with a blue spot on the optic disc. The second is the annotated image with a white circle on the disc. The third is the predicted image, which shows a white circle on the disc, slightly larger and more centered than the annotated one.</p> |
| | LinkNet | <p>The LinkNet model results show three images side-by-side. The first is the original fundus image. The second is the annotated image with a white circle on the disc. The third is the predicted image, which shows a white circle on the disc, very similar in position and size to the annotated one.</p> |
| | FPN | <p>The FPN model results show three images side-by-side. The first is the original fundus image. The second is the annotated image with a white circle on the disc. The third is the predicted image, which shows a white circle on the disc, very similar to the annotated one.</p> |
| | PSPNet | <p>The PSPNet model results show three images side-by-side. The first is the original fundus image. The second is the annotated image with a white circle on the disc. The third is the predicted image, which shows a white circle on the disc, very similar to the annotated one.</p> |

Tabla 6. Arquitectura vgg19 para disco óptico

Imagen Original Imagen Anotada Imagen Predictiva

| ARQUITECTURA | MODELO | |
|--------------|---------|--|
| RESNET152 | UNET | |
| | LinkNet | |
| | FPN | |
| | PSPNet | |

Tabla 7. Arquitectura resnet152 para disco óptico

1.b Exploración sistemática para copa óptica

| PROMEDIO DE ÍNDICES DE JACCARD COPA ÓPTICA | | | |
|--|------------------|------------------|------------------|
| Arquitecturas Modelos | Vgg16 | Vgg19 | resnet152 |
| Unet | Traning: 0.3671 | Traning : 0.3499 | Traning : 0.3888 |
| | Test: 0.5020 | Test : 0.5031 | Test : 0.3283 |
| Linknet | Traning : 0.3032 | Traning: 0.2816 | Traning : 0.4190 |
| | Test: 0.6013 | Test: 0.4963 | Test : 0.4906 |
| FPN | Traning : 0.3728 | Traning : 0.2948 | Traning: 0.3609 |
| | Test : 0.2407 | Test :0.2490 | Test: 0.4506 |
| PSPNet | Traning: 0.3626 | Traning: 0.3513 | Traning: 0.3331 |
| | Test: 0.1830 | Test:0.3697 | Test: 0.3952 |

Tabla 8. Promedios de índices de Jaccard para segmento de copa óptica

| PROMEDIO DE ÍNDICES DE DICE COPA ÓPTICA | | | |
|---|------------------|------------------|------------------|
| Arquitecturas Modelos | Vgg16 | Vgg19 | resnet152 |
| Unet | Training: 0,4864 | Training: 0,4713 | Training: 0,5175 |
| | Test: 0,6342 | Test: 0,6345 | Test: 0,4452 |
| Linknet | Training: 0,4222 | Training: 0,3983 | Training: 0,5389 |
| | Test: 0,7123 | Test: 0,605 | Test: 0,6272 |
| FPN | Training: 0,5106 | Training: 0,4304 | Training: 0,4924 |
| | Test: 0,3323 | Test: 0,3633 | Test: 0,576 |
| PSPNet | Training: 0,4853 | Training: 0,4643 | Training: 0,4652 |
| | Test: 0,2703 | Test: 0,507 | Test: 0,5227 |

Tabla 9. Promedios de índices de Dice para segmento de copa óptica

Imagen Original Imagen Anotada Imagen Predictiva

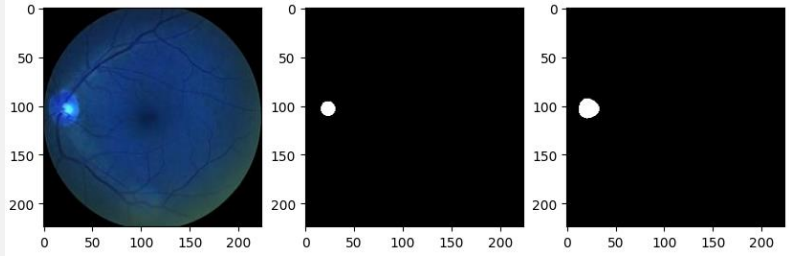
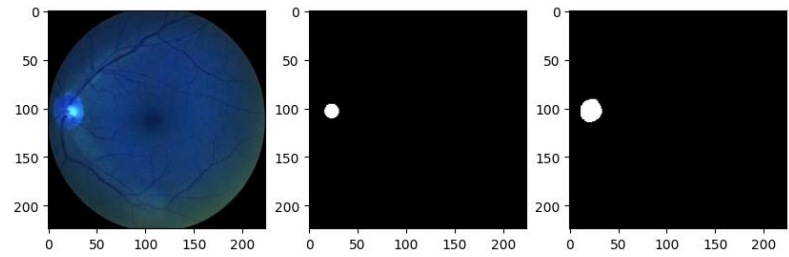
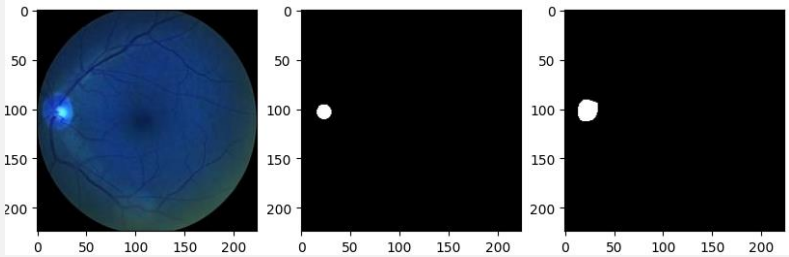
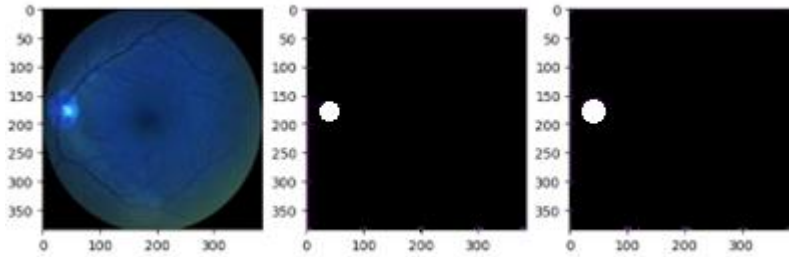
| ARQUITECTURA | MODELO | |
|--------------|---------|--|
| VGG16 | UNET |  |
| | LinkNet |  |
| | FPN |  |
| | PSPNet |  |

Tabla 10. Arquitectura vgg16 para copa óptica

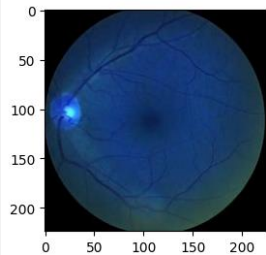
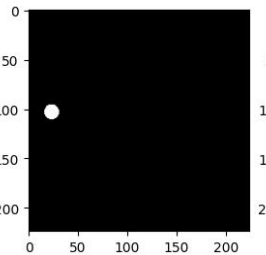
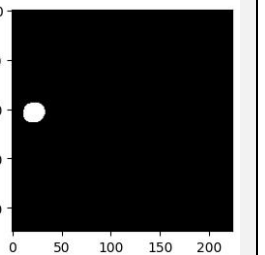
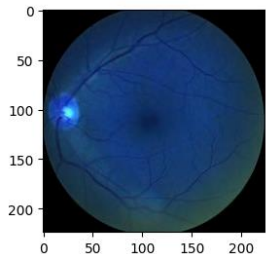
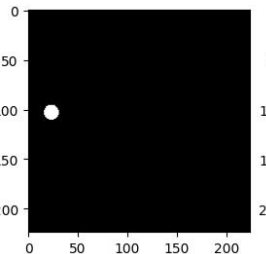
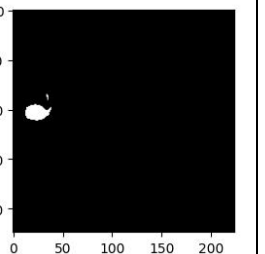
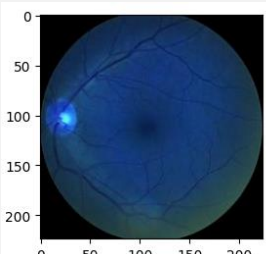
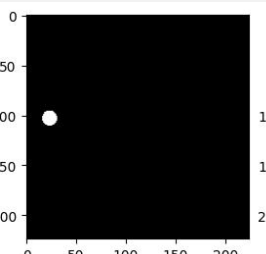
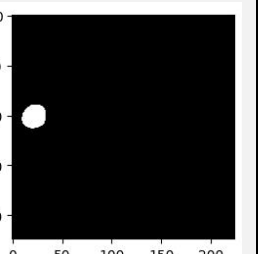
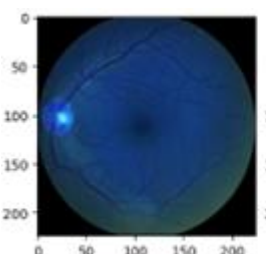
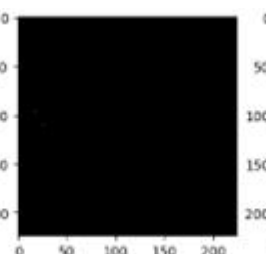
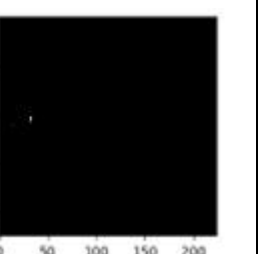
| | | Imagen Original | Imagen Anotada | Imagen Predictiva |
|--------------|---------|---|--|---|
| ARQUITECTURA | MODELO | | | |
| VGG19 | UNET |  |  |  |
| | LinkNet |  |  |  |
| | FPN |  |  |  |
| | PSPNet |  |  |  |

Tabla 11. Arquitectura vgg19 para copa óptica

Imagen Original
Imagen Anotada
Imagen Predictiva

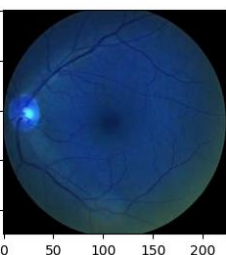
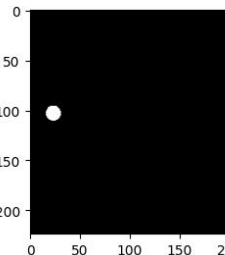
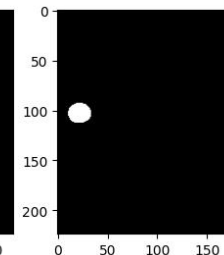
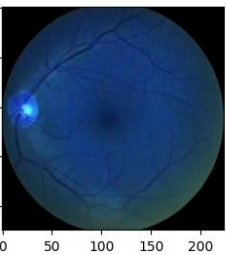
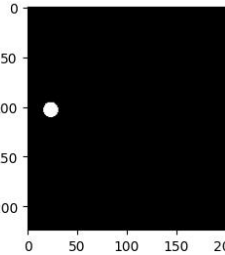
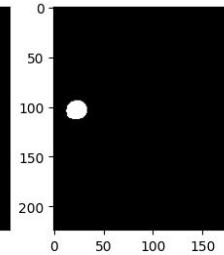
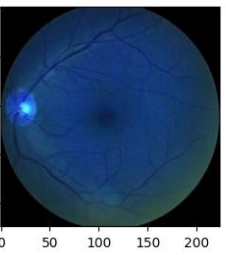
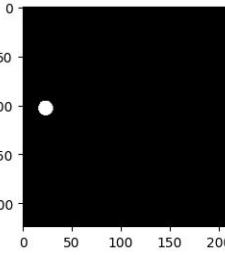
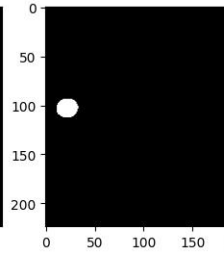
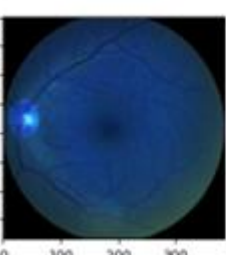
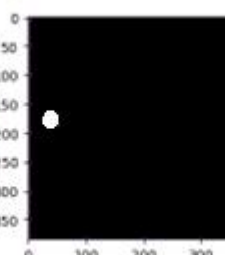
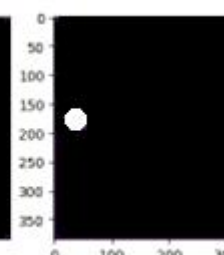
| ARQUITECTURA | MODELO | | | |
|------------------|----------------|---|--|---|
| RESNET152 | UNET |  |  |  |
| | LinkNet |  |  |  |
| | FPN |  |  |  |
| | PSPNet |  |  |  |

Tabla 12. *Arquitectura resnet152 para copa óptica*

2.a Resultados obtenidos en la tarea de clasificación usando la Red Neuronal Convolutiva MobileNet

En esta fase de validación del modelo se obtuvieron medidas estadísticas que nos permitieron analizar los resultados de cada uno de los entrenamientos de la red MobileNet, estos resultados se pueden ver en las tablas 13, 14 y 15, donde en cada una de ellas realizamos una variación en el parámetro de Batch Size para analizar como afectaba este cambio en los resultados de las métricas del modelo.

| Epocas = 50 , Batch Size = 16 , Learning Rate = 0.001 | | | | | | |
|---|----------------|---------------------------------|---------------|--------------|---------------|---------------|
| Experimentos | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 0-80 | [[30 10] [41 319]] | 0.8725 | 0.75 | 0.886 | 0.4731 |
| 2 | 80-160 | [[28 12] [79 281]] | 0.772 | 0.7 | 0.780 | 0.2754 |
| 3 | 160-240 | [[29 11] [44 316]] | 0.862 | 0.72 | 0.877 | 0.4410 |
| 4 | 240-330 | [[28 12] [36 324]] | 0.88 | 0.7 | 0.9 | 0.4736 |
| 5 | 320-400 | [[31 9] [68 292]] | 0.8075 | 0.77 | 0.811 | 0.3540 |
| 6 | 400-480 | [[29 11] [54 306]] | 0.8375 | 0.72 | 0.85 | 0.3890 |
| 7 | 480-560 | [[30 10] [37 323]] | 0.8825 | 0.75 | 0.897 | 0.4978 |
| 8 | 560-640 | [[29 11] [28 332]] | 0.9025 | 0.725 | 0.922 | 0.5443 |
| 9 | 640-720 | [[27 13] [34 326]] | 0.882 | 0.675 | 0.905 | 0.4707 |

Tabla 13. Métricas de la validación de los modelos con batch size de 16

| Epocas = 50 , Batch Size = 32 , Learning Rate = 0.001 | | | | | | |
|---|----------|-------------------------|--------------|--------------|---------------|---------------|
| Experimentos | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 0-80 | [[29 11] [49 311]] | 0.85 | 0.725 | 0.8638 | 0.4140 |
| 2 | 80-160 | [[30 10] [64 296]] | 0.815 | 0.75 | 0.8222 | 0.3576 |
| 3 | 160-240 | [[29 11] [40 320]] | 0.872 | 0.725 | 0.8888 | 0.4642 |
| 4 | 240-330 | [[29 11] [44 316]] | 0.8625 | 0.725 | 0.8777 | 0.4410 |
| 5 | 320-400 | [[33 7] [72 288]] | 0.8025 | 0.825 | 0.8 | 0.3629 |
| 6 | 400-480 | [[28 12] [41 319]] | 0.8675 | 0.7 | 0.8866 | 0.4432 |
| 7 | 480-560 | [[29 11] [41 319]] | 0.87 | 0.725 | 0.8861 | 0.4583 |
| 8 | 560-640 | [[30 10] [36 324]] | 0.885 | 0.75 | 0.9 | 0.5043 |
| 9 | 640-720 | [[28 12] [35 325]] | 0.882 | 0.7 | 0.9027 | 0.4800 |

Tabla 14. Métricas de la validación de los modelos con batch size de 32

| Epocas = 50 , Batch Size = 64 , Learning Rate = 0.001 | | | | | | |
|---|----------------|---------------------------------|---------------|--------------|---------------|---------------|
| Experimentos | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 0-80 | [[30 10] [41 319]] | 0.8725 | 0.75 | 0.886 | 0.4731 |
| 2 | 80-160 | [[28 12] [55 305]] | 0.8325 | 0.7 | 0.8472 | 0.37300 |
| 3 | 160-240 | [[30 10] [39 321]] | 0.8775 | 0.75 | 0.8916 | 0.4852 |
| 4 | 240-330 | [[30 10] [41 319]] | 0.8725 | 0.75 | 0.8861 | 0.4731 |
| 5 | 320-400 | [[30 10] [57 303]] | 0.8325 | 0.75 | 0.8416 | 0.3886 |
| 6 | 400-480 | [[30 10] [67 293]] | 0.8075 | 0.75 | 0.8138 | 0.3452 |
| 7 | 480-560 | [[29 11] [42 318]] | 0.8675 | 0.725 | 0.8833 | 0.4524 |
| 8 | 560-640 | [[30 10] [31 329]] | 0.8975 | 0.75 | 0.9138 | 0.5382 |
| 9 | 640-720 | [[27 13] [29 331]] | 0.895 | 0.675 | 0.9194 | 0.5047 |

Tabla 15. Métricas de la validación de los modelos con batch size de 64

Durante los experimentos realizados para la validación de los modelos se pudo observar que a medida que la cantidad de muestras fuera mayor y tendiera a ser simétrica se aumentaba a exactitud del modelo y mejoraban notablemente los resultados, por lo tanto, se realizó un experimento adicional en el que se buscaba tener una muestra balanceada y

con un mayor número de imágenes, estimando que esto podía llevar a que los resultados obtenidos fueran mejores que los anteriores experimentos realizados.

En las tablas 16, 17 y 18 se presentan los resultados estadísticos de realizar el entrenamiento del modelo con un grupo muestral balanceado de 160 imágenes, con una variación del parámetro de batch size para analizar la afectación de este cambio en los resultados.

| Epocas = 50 , Batch Size = 16 , Learning Rate = 0.001 | | | | | | |
|---|----------|------------------------|-----------|--------------|---------------|--------|
| Experimento | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 160 | [[22 18] [1 359]] | 0.9525 | 0.55 | 0.9977 | 0.6746 |

Tabla 16. Métricas de muestra balanceada con batch size de 16

| Epocas = 50 , Batch Size = 32 , Learning Rate = 0.001 | | | | | | |
|---|----------|------------------------|-----------|--------------|---------------|--------|
| Experimento | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 160 | [[32 8] [52 308]] | 0.85 | 0.8 | 0.8555 | 0.4402 |

Tabla 17. Métricas de muestra balanceada con batch size de 32

| Epocas = 50 , Batch Size = 64 , Learning Rate = 0.001 | | | | | | |
|---|----------|------------------------|-----------|--------------|---------------|--------|
| Experimento | Imágenes | Matriz de confusión | Exactitud | Sensibilidad | Especificidad | Kappa |
| 1 | 160 | [[22 18] [1 359]] | 0.9525 | 0.55 | 0.9972 | 0.6746 |

Tabla 18. Métricas de muestra balanceada con batch size de 64

5. DISCUSIÓN

Como se puede observar en la tabla 3, para los modelos de segmentación de disco óptico se obtuvo los mejores resultados al usar el modelo de red neuronal U-NET con la arquitectura Vgg16 tanto para Training como para test, aunque no es mucha la diferencia entre algunos modelos si se puede observar que para el grupo de imágenes de TEST es el único valor de Jaccard que supera el rango de 0.8, siendo esta combinación la mejor configuración de entrenamiento para el segmento de estudio del disco óptico.

En la tabla 8 podemos observar uno de los modelos que nos dio los resultados más bajos, este es el modelo FPN con la arquitectura VGG16, visualmente se puede notar que el entrenamiento de este modelo nos da como resultado algunas áreas poco uniformes o desarraigadas a lo esperado, la tabla 3 nos confirman que no cumplen con los estándares adecuados para considerarlo uno de los modelos óptimos en nuestro contexto, aunque da valores moderados, para este tipo de imágenes que estamos manejando en el proyecto es más conveniente usar modelos especializados en el ámbito médico que nos puede llegar a dar mejores resultados como la arquitectura de DeepMedic.

Para la sección 1b también podemos analizar que la combinación que nos da los resultados más óptimos es la del modelo LinkNet con la arquitectura Vgg16, aunque se ven valores más bajos que en la tabla 8 podemos atribuir esto a que el área de la copa es menor al área del disco, por lo tanto, es más probable que los resultados del modelo se diferencien en mayor proporción a los resultados esperados. De igual modo los resultados obtenidos para TEST entran dentro del rango de aceptación del modelo, con un valor promedio de Jaccard mayor a 0.6.

Los resultados obtenidos de los índices de Dice en las tablas 4 y 9 son diferentes a los índices de Jaccard esto difiere en que el índice de Dice es más estricto al momento de evaluar las segmentaciones considerando solo los píxeles comunes, mientras que el índice de Jaccard evalúa la superposición en relación con la unión de los conjuntos.

Por este motivo el mejor modelo para disco óptico (vgg16-U-Net) de la tabla 3 obtuvo un Jaccard en test de 0.80 mientras que en Dice en test (ver tabla 4) fue de 0.88 y el mejor modelo para copa óptica (vgg16-LinkNet) de la tabla 8 presentó un Jaccard en test de 0.601 mientras que Dice en test (ver tabla 9) fue de 0.712, siendo estos unos resultados considerados aceptables para la base de datos con la que se contaba.

Dentro de los primeros 27 experimentos realizados en la validación de la red MobileNet es importante destacar que en cada uno de los sub grupos de imágenes siempre obtuvimos una mejor validación del modelo en el experimento 8, aunque este valor es aleatorio debido a la organización o enumeración de los pacientes, podemos concluir que en este grupo de imágenes el modelo pudo clasificar con una mayor facilidad si el paciente era glaucomatoso o sano, indicándonos que si afecta al entrenamiento de los modelos la calidad de las muestras tomadas.

En vista de que los resultados que estábamos obteniendo en todos los experimentos entraban en la clasificación de moderados (según la tabla 1), se decidió probar un ambiente

fuera del mundo real. Actualmente se cree que de cada 10 pacientes que atiende un médico oftalmólogo, aproximadamente solo 1 de ellos requiere atención especializada y que los demás solo tienen patologías leves, por lo tanto, al entrenar un modelo asumiendo que teníamos una muestra balanceada con la misma cantidad de pacientes sanos que glaucomatosos.

Este resultado fue de acuerdo a lo esperado, claramente al tener una muestra más grande y proporcional, el modelo tendrá un mejor entrenamiento, lo cual le permitirá tener una mejor predicción, este resultado lo podemos observar en las tablas 16, 17 y 18, donde ya podemos observar una Kappa mayor a 0.6 y una exactitud de 0.95 corroborando nuestra hipótesis en este ambiente alejado de lo real.

6. RECOMENDACIONES Y TRABAJOS FUTUROS

Se espera que con este proyecto se impulse al desarrollo y la investigación de estudios oftalmológicos que permitan aportar más soluciones desde el área de la ingeniería biomédica, en lo posible que sean soluciones de bajo costo, accesibles a más personas y de más impacto que permitan disminuir los índices de esta enfermedad.

Continuar con este proyecto buscando profundizar en los modelos obtenidos con el fin de mejorar los índices de Jaccard y las métricas estadísticas de la validación de la red convolucional. También sería interesante el trabajar en un software que permita a los especialistas el uso intuitivo de esta herramienta, brindado una mayor facilidad en el uso y una entrega de resultados más fácil de comprender y expresar a los pacientes.

Con el fin de mejorar los resultados obtenidos en este trabajo se propone realizar alianzas con empresas especializadas en la toma de imágenes, buscando ampliar la base de datos trabajada y con esto poder tener un grupo muestral que nos permita llegar a resultados más óptimos.

Esta tesis está abierta a más áreas de investigación referentes a patologías presentes en el disco y/o copa óptica como lo es: Adelgazamiento de la retina, cataratas, drusas, etc.

7. CONCLUSIONES

Los resultados obtenidos durante el proyecto se clasifican en una escala moderada, con lo que podemos entender que, aunque el trabajo esta apto para mejoras, los resultados fueron aceptables para iniciar en este campo de estudio.

Se logro desarrollar un modelo capaz de detectar glaucoma a partir de imágenes de fondo de ojo, si bien es cierto que hay muchos estudios utilizados en este campo con grandes bases de datos, aun así, el modelo logro segmentar con pocas imágenes y predecir métricas de desempeño favorables para ser una primera etapa de estudio.

Los modelos lograron segmentar las características relevantes para la detección del glaucoma como lo son el disco y la copa óptica por medio de un CNN adecuado para las imágenes utilizadas, ya que sin unas características extraídas correctamente es muy difícil llegar a una buena predicción.

Se determino que para este campo de estudio el mejor modelo capaz de predecir glaucoma a partir de imágenes de fondo de ojo con buenos índices de Jaccard y Dice fueron el vgg16-UNET y el vgg16-LinkNet.

El batch size en una muestra es importante y más cuando se quiere segmentar estructuras pequeñas donde los detalles son importantes para obtener buenos resultados en nuestro experimento la métrica de validación con mejor kappa fue con un batch size de 16, aun así, cuando las muestras no eran balanceadas.

Este tipo de modelos ayudan a facilitar el trabajo evitando realizar un alistamiento de la imagen para poder procesarla, lo que causa una mayor optimización en el tiempo invertido en segmentar y los recursos utilizados.

Se ha demostrado que una muestra desequilibrada es más difícil llegara a obtener buenos resultados y que también es vital para el entrenamiento de un modelo contar con muestras considerables que le permitan llegar a tener un buen aprendizaje

REFERENCIAS

- [1]. *Día Mundial del Glaucoma: OMS reporta 4,5 millones de afectados por la enfermedad / Noticias ONU*. (n.d.). Retrieved April 14, 2023, from <https://news.un.org/es/story/2009/03/1158791>
- [2]. *Tipos de glaucoma / glaucoma.org*. (n.d.). Retrieved April 28, 2023, from <https://glaucoma.org/tipos-de-glaucoma/>
- [3]. Orlando, J. I., Fu, H., Breda, J. B., van Keer, K., Bathula, D. R., Diaz-Pinto, A., Fang, R., Heng, P.-A., Kim, J., Lee, J., Lee, J., Li, X., Liu, P., Lu, S., Murugesan, B., Naranjo, V., Phaye, S. S. R., Shankaranarayana, S. M., Sikka, A., ... Bogunović, H. (2019). *REFUGE Challenge: A Unified Framework for Evaluating Automated Methods for Glaucoma Assessment from Fundus Photographs*. <https://doi.org/10.1016/j.media.2019.101570>
- [4]. *El ojo humano*. (n.d.). Retrieved April 28, 2023, from <https://www.zeiss.es/vision-care/mejor-vision/entender-la-vision/el-ojo-humano.html>
- [5]. *Cómo funcionan los ojos / National Eye Institute*. (n.d.). Retrieved April 28, 2023, from <https://www.nei.nih.gov/espanol/aprenda-sobre-la-salud-ocular/vision-saludable/como-funcionan-los-ojos>
- [6]. *Anatomía Del Ojo: Partes Del Ojo Y Cómo Vemos / Instituto de Ojos*. (n.d.). Retrieved May 7, 2023, from <https://institutoojospr.com/anatomia-del-ojo-partes-del-ojo-y-como-vemos/>
- [7]. *Nervio óptico - American Academy of Ophthalmology*. (n.d.). Retrieved May 7, 2023, from <https://www.aao.org/salud-ocular/anatomia/nervio-optico>
- [8]. *La OMS presenta el primer Informe mundial sobre la visión*. (n.d.). Retrieved May 9, 2023, from <https://www.who.int/es/news/item/08-10-2019-who-launches-first-world-report-on-vision>
- [9]. *Aspecto del Nervio Optico en la Detección del Glaucoma*. (n.d.). Retrieved May 10, 2023, from https://sisbib.unmsm.edu.pe/bvrevistas/anales/v59_n2/glaucoma.htm
- [10]. *Esclerótica: MedlinePlus enciclopedia médica*. (n.d.). Retrieved May 8, 2023, from <https://medlineplus.gov/spanish/ency/article/002295.htm>

- [11]. (PDF) *Nuevo Algoritmo para el Cálculo de la Relación Disco Óptico- Excavación Basado en Distancias de Color*. (n.d.). Retrieved January 23, 2023, from https://www.researchgate.net/publication/267860252_Nuevo_Algoritmo_para_el_Calculo_de_la_Relacion_Disco_Optico-Excavacion_Basado_en_Distancias_de_Color
- [12]. *Ventosas del nervio óptico* | glaucoma.org. (n.d.). Retrieved May 10, 2023, from <https://glaucoma.org/optic-nerve-cupping/>
- [13]. *El sentido de la vista: Definición, partes y funcionamiento*. (n.d.). Retrieved May 8, 2023, from <https://www.oftalvist.es/blog/sentido-de-la-vista>
- [14]. *U-Net architecture diagram modified from the original study [27].... | Download Scientific Diagram*. (n.d.). Retrieved May 15, 2023, from https://www.researchgate.net/figure/U-Net-architecture-diagram-modified-from-the-original-study-27-Green-yellow-boxes_fig1_342456648
- [15]. *Cuerpo ciliar* - *Wikipedia, la enciclopedia libre*. (n.d.). Retrieved May 8, 2023, from https://es.wikipedia.org/wiki/Cuerpo_ciliar
- [16]. Rao, A. S., Nguyen, T., Palaniswami, M., & Ngo, T. (2021). Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. *Structural Health Monitoring*, 20(4), 2124–2142. <https://doi.org/10.1177/1475921720965445>
- [17]. ▷ ¿Qué es el HUMOR ACUOSO? - *Área Oftalmológica Avanzada*. (n.d.). Retrieved May 8, 2023, from <https://areaoftalmologica.com/terminos-de-oftalmologia/humor-acuoso/>
- [18]. Del Portillo, C. (n.d.). “*Mi deseo para el futuro es que nadie tenga que vivir con el miedo de perder la vista.*”
- [19]. ▷ ¿Qué es el HUMOR VÍTREO? - *Área Oftalmológica Avanzada*. (n.d.). Retrieved May 8, 2023, from <https://areaoftalmologica.com/terminos-de-oftalmologia/humor-vitreo/>
- [20]. Sørensen–Dice coefficient - *Wikiwand*. (n.d.). Retrieved November 27, 2022, from https://www.wikiwand.com/en/Sørensen–Dice_coefficient#introduction
- [21]. *Métricas de similitud para cadenas de texto. Parte III: Métricas de conjuntos y emparejamiento de caracteres.* | by *Diego Campos Sobrino* | *Medium*. (n.d.). Retrieved May 15, 2023, from <https://medium.com/@diego.campos.sobrino/métricas-de-similitud-para-cadenas-de-texto-83bb09cd8591>

- [22]. *Información de glaucoma ¿Cómo ocurre el daño al nervio óptico? - Información de glaucoma.* (n.d.). Retrieved May 8, 2023, from <https://www.glaucomapatient.org/es/basica/como-ocurre-el-dano-de-glaucoma/>
- [23]. *La presión ocular - American Academy of Ophthalmology.* (n.d.). Retrieved May 8, 2023, from <https://www.ao.org/salud-ocular/anatomia/la-presion-ocular>
- [24]. *Sinequia Ocular - ¿Qué es? - Tipos y Tratamiento.* (n.d.). Retrieved May 8, 2023, from <https://www.martinezdecarneros.com/sinequia-ocular-que-es-tipos-y-tratamiento/>
- [25]. *¿Por dónde drena el líquido intraocular? - Glaucoma - IO·ICO Barcelona.* (n.d.). Retrieved May 8, 2023, from <https://www.icoftalmologia.es/es/noticias/por-donde-drena-el-liquido-intraocular-glaucoma/>
- [26]. *Generalidades sobre el glaucoma - Trastornos oftálmicos - Manual MSD versión para profesionales.* (n.d.). Retrieved May 8, 2023, from <https://www.msmanuals.com/es/professional/trastornos-oftalmicos/glaucoma/generalidades-sobre-el-glaucoma>
- [27]. *Glaucoma | Salud Ocular.* (n.d.). Retrieved May 8, 2023, from <https://salud-ocular.com/glaucoma/>
- [28]. Piñero, R. T., Lora, M., & Andrés, M. I. (2005). Glaucoma. *Offarm*, 24(2), 88–96. <https://www.elsevier.es/es-revista-offarm-4-articulo-glaucoma-13071464>
- [29]. *Glaucoma | Salud Ocular.* (n.d.). Retrieved May 8, 2023, from <https://salud-ocular.com/glaucoma/>
- [30]. *Tipos de glaucoma | National Eye Institute.* (n.d.). Retrieved May 8, 2023, from <https://www.nei.nih.gov/espanol/aprenda-sobre-la-salud-ocular/enfermedades-y-afecciones-de-los-ojos/glaucoma/tipos-de-glaucoma>
- [31]. *Diagnóstico del Glaucoma | PortalCLÍNICA.* (n.d.). Retrieved May 8, 2023, from <https://www.clinicbarcelona.org/asistencia/enfermedades/glaucoma/diagnostico>
- [33]. *Para qué sirve un examen con lámpara de hendidura - Blog de Clínica Baviera.* (n.d.). Retrieved May 8, 2023, from <https://www.clinicabaviera.com/blog/bye-bye-gafasconoce-tus-ojospara-que-sirve-un-examen-con-lampara-hendidura/>
- [34]. *¿Qué es la Tonometría Ocular? - Clínica ISV.* (n.d.). Retrieved May 8, 2023, from <https://www.clinicaisv.cl/que-es-la-tonometria-ocular/>

- [35]. *Campimetria o perimetria computerizada – Laser Optical*. (n.d.). Retrieved May 8, 2023, from <https://www.laseroptical.com.mx/campimetria-o-perimetria-computerizada/>
- [36]. *U-NET : todo lo que tienes que saber sobre la red neuronal de Computer Vision*. (n.d.). Retrieved May 8, 2023, from <https://datascientest.com/es/u-net-lo-que-tienes-que-saber>
- [37]. Chaurasia, A., & Culurciello, E. (n.d.). *LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation*.
- [38]. *FPN Explained | Papers With Code*. (n.d.). Retrieved May 8, 2023, from <https://paperswithcode.com/method/fpn>
- [39]. *PSPNet Explained | Papers With Code*. (n.d.). Retrieved May 8, 2023, from <https://paperswithcode.com/method/pspnet>
- [40]. *Métricas De Evaluación De Modelos En El Aprendizaje Automático*. (n.d.). Retrieved May 8, 2023, from <https://www.datasource.ai/es/data-science-articles/metricas-de-evaluacion-de-modelos-en-el-aprendizaje-automatiko>
- [41]. McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276. <https://doi.org/10.11613/bm.2012.031>
- [42]. *Sensibilidad y especificidad | Cigna*. (n.d.). Retrieved May 8, 2023, from <https://www.cigna.com/es-us/knowledge-center/hw/sensibilidad-y-especificidad-sts14487>
- [43]. *Índice de Jaccard Similitud de atributos binarios asimétricos y Similitud y distancia ponderadas de Jaccard*. (n.d.). Retrieved May 8, 2023, from https://hmong.es/wiki/Jaccard_index
- [44]. *Trastornos del nervio óptico: MedlinePlus en español*. (n.d.). Retrieved May 8, 2023, from <https://medlineplus.gov/spanish/opticnervedisorders.html>
- [45]. Orlando, J. I., Fu, H., Barbossa Breda, J., van Keer, K., Bathula, D. R., Diaz-Pinto, A., Fang, R., Heng, P. A., Kim, J., Lee, J. H., Lee, J., Li, X., Liu, P., Lu, S., Murugesan, B., Naranjo, V., Phaye, S. S. R., Shankaranarayana, S. M., Sikka, A., ... Bogunović, H. (2020). REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical Image Analysis*, 59, 101570. <https://doi.org/10.1016/J.MEDIA.2019.101570>

- [46]. *Fleiss' Kappa and Inter rater agreement interpretation [24] | Download Table.* (n.d.). Retrieved May 8, 2023, from https://www.researchgate.net/figure/Fleiss-Kappa-and-Inter-rater-agreement-interpretation-24_tbl3_281652142
- [47] *Redes neuronales convolucionales (CNN) y aprendizaje profundo.* (n.d.). Retrieved May 11, 2023, from <https://www.intel.es/content/www/es/es/internet-of-things/computer-vision/convolutional-neural-networks.htm>
- [48]. *Arquitectura VGG16 y VGG19 en Deep Learning.* (n.d.). Retrieved May 11, 2023, from <https://keepcoding.io/blog/arquitectura-vgg16-vgg19-deep-learning/>
- [49]. *Architecture of ResNet-152 layer. | Download Scientific Diagram.* (n.d.). Retrieved May 11, 2023, from https://www.researchgate.net/figure/Architecture-of-ResNet-152-layer_fig2_347170147
- [50]. *Linknet architecture. The numbers in squared brackets indicate the... | Download Scientific Diagram.* (n.d.). Retrieved May 15, 2023, from https://www.researchgate.net/figure/Linknet-architecture-The-numbers-in-squared-brackets-indicate-the-outputs-of-the-blocks_fig8_336819290
- [51]. *Keras Applications.* (n.d.). Retrieved May 15, 2023, from <https://keras.io/api/applications/>