

**Detección, segmentación y clasificación de la retinopatía
diabética en imágenes oculares**

**Daniela Echeverri Martinez
Juliana Andrea Molina Velasco**

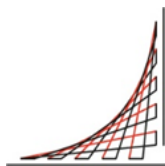
Trabajo Dirigido

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

Queremos agradecer a Dios en primer lugar por permitirnos culminar esta etapa tan importante en nuestras vidas, a nuestros padres por su gran esfuerzo y dedicación que han puesto para vernos crecer profesionalmente, a nuestras familias y seres queridos que nos han acompañado incondicionalmente en todo el proceso. A nuestro director de proyecto, el ingeniero Oscar Perdomo por todo el apoyo que nos brindó durante el desarrollo de este proyecto, por sus conocimientos que compartió y puso a nuestra disposición de manera incondicional, por su tiempo invertido en la enseñanza que sin duda ha sido muy valioso en nuestro crecimiento profesional y por último, queremos agradecer a las instituciones que durante todo este tiempo fueron nuestro segundo hogar, la Escuela Colombiana de Ingeniería Julio Garavito y la Universidad del Rosario por darnos la oportunidad de crecer como personas.

Tabla de contenido

1. INTRODUCCIÓN.....	6
2. OBJETIVOS.....	8
2.1. General.....	8
2.2. Específicos.....	8
3. Justificación.....	9
4. ESTADO DEL ARTE.....	10
4.1. Época.....	10
4.2. Batch size.....	10
4.3. Validation split.....	10
4.4. Coeficiente de similitud de Jaccard.....	10
4.5. coeficiente de similitud de DICE.....	11
4.6. Métricas de clasificación.....	11
4.7. Modelos de backbone.....	12
4.8. MobileNet.....	12
5. METODOLOGÍA.....	13
5.1 Preprocesamiento.....	13
5.2. Entrenamiento.....	16
5.2.1. Segmentación y Clasificación.....	17
5.3 Validación.....	19
6. RESULTADOS.....	19
7. DISCUSIÓN.....	25
8. RECOMENDACIONES Y TRABAJOS FUTUROS.....	26
9. CONCLUSIONES.....	27
10. REFERENCIAS.....	28

Tabla de figuras

Figura 1. Imágenes Originales con color de fondo.....	13
Figura 2. Imágenes de fondo RD.....	14
Figura 3. Imágenes de fondo EDM.....	14
Figura 4. Imágen original Microaneurisma.....	17
Figura 5 . Imágenes con los parches para pacientes no-sanos Microaneurisma.....	18
Figura 6. Modelos de segmentación explorados [33].....	18
Figura 7. Relación entre el índice de Jaccard y el coeficiente de Dice.....	19
Figura 8. Resultados obtenidos en la predicción de HE en conjunto de entrenamiento.....	20
Figura 9. Resultados obtenidos en la predicción de HE en conjunto de prueba.....	21
Figura 10. Resultados para la segmentación de Hemorragias, Microaneurismas, exudados blandos, exudados duros reportados en la literatura [37].....	27
Figura 11. Resultados para la clasificación de la Retinopatía diabética y Edema Macular Diabetico reportados en la literatura [37].....	28

Lista de tablas

Tabla 1. Número total de imágenes en la base de datos.....	15
Tabla 2. Cantidad de imágenes totales.....	15
Tabla 3. Cantidad de imágenes para la clasificación.....	15
Tabla 4. Clasificación grado de la enfermedad RD y EMD training.....	16
Tabla 5. Métrica de Jaccard obtenida en Exudados Blandos.....	22
Tabla 6. Métrica de Jaccard obtenida en Exudados Duros.....	22
Tabla 7. Métrica de Jaccard obtenida de Hemorragias.....	22
Tabla 8. Métrica de Jaccard obtenida de Disco Óptico.....	22
Tabla 9. Métrica de Jaccard obtenida de Microaneurismas.....	23
Tabla 10. Métrica de Dice obtenida en Exudados Blandos.....	23
Tabla 11. Métrica de Dice obtenida en Exudados Duros.....	23
Tabla 12. Métrica de Dice obtenida en Hemorragias.....	24
Tabla 13. Métrica de Dice obtenida en Disco Óptico.....	24
Tabla 14. Métrica de Dice obtenida en Microaneurismas.....	24
Tabla 15. Métricas para clasificación de RD y EMD.....	24

1. INTRODUCCIÓN

Las imágenes de fondo de ojo a color representan para el médico una oportunidad única para evaluar y examinar al paciente, ya que éstas permiten complementar el estudio sistémico dando la posibilidad de evaluar el árbol vascular de la retina, el nervio óptico y la retina en toda su extensión. Es importante mencionar que uno de los datos importantes que deben tenerse en cuenta al momento de observar una retinografía es identificar qué ojo está visualizando, si el derecho o el izquierdo, y esto se puede lograr debido a la ubicación del nervio óptico, también llamado papila óptica, el cual se encuentra situado en el lado nasal de la imagen. De esta manera, si se observa una retinografía en la que el nervio óptico está ubicado a la derecha de la imagen, se está examinando un ojo derecho, y viceversa[1][2].

Actualmente, la exploración del fondo de ojo se realiza mediante oftalmoscopia directa o indirecta, además de la exploración con lámpara de hendidura ayudada por lentes de contacto (tipo Goldmann) o no contacto (de 90 dioptrías). Estas técnicas tienen una gran desventaja, ya que no permiten dejar un registro histórico de imágenes que puedan consultarse tiempo después, por esta razón, para el diagnóstico de enfermedades de la retina como lo es la retinopatía diabética (RD) y su clasificación, se recurre a técnicas de imagen fotográficas. La más común es la retinografía, una prueba diagnóstica que permite obtener una imagen de la retina o del fondo del ojo sin necesidad de contacto, para este examen se utiliza una cámara de fotografía adaptada para tal fin o retinógrafo para tomar una fotografía digital manteniendo el color de la retina y del fondo del ojo. La prueba dura muy poco tiempo, no genera ningún dolor y lo único que nota el paciente es el flash de la cámara; esta técnica empezó a utilizarse en la década de 1980 y en la actualidad sigue siendo de preferencia para el cribado y diagnóstico, sobre todo de la retinopatía diabética y la hipertensiva arterial[1][2][3].

La retinopatía diabética es una de las complicaciones microvasculares de la Diabetes Mellitus (DM), la cual puede empezar con una disfunción endotelial en los capilares retinianos y finalmente llevando a la aparición de un edema macular diabético (EMD). El EMD consiste en la inflamación y acumulación de líquido en la mácula, lo cual requiere un abordaje terapéutico multidisciplinar, un buen control metabólico para mantener las concentraciones de hemoglobina glucosilada por debajo del 7.5%, gestión de los factores de riesgo y de una evolución oftalmológica completa con la realización de exámenes complementarios[4].

Según los resultados de varios estudios realizados en el año 2018 por la doctora Lilian Jeannette Prada Ramirez en la Universidad del Rosario escuela de medicina y ciencias de la salud y la Universidad CES facultad de medicina especialización en epidemiología clínica, la prevalencia de retinopatía diabética en Colombia varía entre el 7,3% y el 18,8%, dependiendo de la población estudiada. Un estudio encontró que la prevalencia de retinopatía diabética en pacientes con diabetes tipo 2 atendidos en una entidad hospitalaria de Bogotá fue del 7,3%. Otro estudio encontró que la prevalencia de retinopatía diabética en una población de pacientes con diabetes en Colombia fue de aproximadamente 20 individuos por cada 100.000 habitantes. Además, un estudio reciente que incluyó a 1.203 pacientes con diabetes en Colombia encontró que aproximadamente 1 de cada 3 personas

con diabetes tenía retinopatía diabética y 1 de cada 10 tenía retinopatía diabética proliferativa o edema macular diabético[5][6][7].

La probabilidad de desarrollar RD está relacionada con la duración de la diabetes. La diabetes tipo 2 tiene un inicio insidioso y puede pasar desapercibida durante muchos años. Por lo tanto, es posible que el paciente ya tenga RD cuando se le diagnostica. El riesgo de desarrollar retinopatía aumenta después de la pubertad. Veinte años después de ser diagnosticados con diabetes, el 80% de las personas con diabetes tipo 2 y la mayoría de las personas con diabetes tipo 1 tienen algunos signos de retinopatía[8].

Se presenta en dos etapas según el grado de degeneración microvascular y el daño isquémico asociado: retinopatía diabética no proliferativa (NPDR) la cual se puede dividir en i) NPDR benigna, ii) NPDR moderado y iii) NPDR grave. El otro tipo de retinopatía diabética es la retinopatía diabética de proliferación progresiva (PDR). Por otro lado, la presencia de exudados en la retina está asociado con EMD, lo que requiere una pronta detección de esta lesión, para realizar un tratamiento pertinente que mejore la calidad de vida de las personas que la padecen[9].

La Inteligencia Artificial (IA) se ha venido implementando en procedimientos médicos contribuyendo en la prevención y seguimiento de algunos casos clínicos, se ha utilizado exhaustivamente en el procesamiento de imágenes médicas ya que demuestra ser bastante eficiente revelando enfermedades o cuerpos extraños en poco tiempo. Por lo cual, actualmente se considera un técnica válida puesto que facilita la realización de un juicio médico rápido y eficiente, así como reduce costos, tiempo y algunos errores médicos permitiendo potenciar el uso de los recursos humanos en las ramas médicas con mayores requerimientos. Dentro de la IA se incluye también los algoritmos de aprendizaje profundo, basados en redes neuronales artificiales, es decir, modelos computacionales compuestos por capas de nodos interconectados, y es ahí donde se encuentra el aspecto diferencial por lo cual lleva el título “profundo”, ya que estos algoritmos se diferencian por tener múltiples capas ocultas entre la capa de entrada y salida, lo cual permite que la red neuronal aprenda representaciones de alto nivel como identificar patrones en las imágenes médicas y así detectar enfermedades específicas con alta precisión [10].

Finalmente, conociendo que las dos enfermedades mencionadas anteriormente pueden desarrollarse de manera silenciosa, sin ser excluyentes y pueden progresar a etapas avanzadas sin producir ningún síntoma inmediato en el paciente, es allí donde la detección temprana de estas enfermedades es esencial para establecer un tratamiento temprano y evitar pérdida parcial o total de visión. Es así que con el uso de algoritmos de aprendizaje profundo se posibilita llegar a resultados con mayor precisión en la detección y clasificación de enfermedades oculares de forma temprana, incluidos cambios pequeños pero significativos en la función visual en las primeras etapas de la enfermedad. De esta manera y consecuentes a esta problemática, este trabajo de tesis busca generar algoritmos para la detección de lesiones oculares y la clasificación de enfermedades oculares usando imágenes de fondo de ojo en su color original, buscando ayudar en la toma de decisiones médicas.

2. OBJETIVOS

2.1. General

Implementar un algoritmo que permita detectar y segmentar lesiones oculares apoyando la clasificación de Retinopatía Diabética (RD) y Edema Macular Diabético (EMD) en imágenes de fondo de ojo.

2.2. Específicos

- a. Segmentar automáticamente imágenes de fondo del ojo para encontrar microaneurismas, hemorragias, exudados duros, exudados blandos y disco óptico.
- b. Clasificar los diferentes estadios de RD y EMD en imágenes de fondo del ojo.
- c. Validar sistemáticamente los algoritmos de segmentación de lesiones oculares relacionados con RD y EMD.

3. Justificación

La Sociedad Americana estima que la retinopatía diabética es la tercera causa de pérdidas visuales e irreversibles en el mundo, y es la primera causa de países en desarrollo entre las edades de 16 y 64 años. En Colombia, las cifras de RD no son nada alentadoras y se estiman que alrededor de 20 de cada 100,000 habitantes presentan RD, con la problemática que más de la mitad nunca acceden a diagnósticos y tratamientos en estadios leves de la enfermedad[11].

Aunque, se cuenta con la escala para el estudio de tratamiento temprano de la retinopatía diabética y el edema macular diabético (abreviada en inglés como ETDRS) para la estimación del estadio, el tamizaje ocular debería ser realizado al menos 2 veces por año en personas diabéticas tipo I y 1 vez al año en personas con diabetes tipo II [12]. Sin embargo, en Colombia contamos con limitado acceso a servicios de salud ocular y una baja tasa de profesionales en oftalmología que se estima alrededor de 2 especialistas en oftalmología por cada 100.000 habitantes en Colombia [13].

Con este proyecto se busca permitirles a los centros oftalmológicos un mejor proceso de análisis de imágenes del ojo para obtener diagnóstico de la retinopatía diabética, esto debido a que se estima que aproximadamente un tercio de pacientes diabéticos tipo 1 y un tercio de pacientes diabéticos tipo 2 desarrollan retinopatía diabética a lo largo de sus vidas. En los países de bajos y medianos ingresos, hasta el 50 % de los pacientes diabéticos pueden sufrir retinopatía diabética[14][15]. Por tanto, el desarrollo de herramientas que permitan apoyar el diagnóstico con la detección, segmentación y clasificación de hallazgos y patologías oculares como los tratados en este trabajo de grado son la mayor justificación para el desarrollo del mismo.

4. ESTADO DEL ARTE

4.1. Época

Las épocas de entrenamiento en un modelo se refiere al número de veces que dicho modelo debe entrenar en el conjunto de datos de entrenamiento , es decir, cada época implica que el modelo ha pasado por todo el conjunto de datos una vez, éste es un hiper parámetro que puede modificarse para obtener mejores resultados, este número puede ser diferente de acuerdo a la segmentación que se vaya a realizar. El número de épocas de entrenamiento es un factor importante que afecta el rendimiento del modelo, y se debe ajustar cuidadosamente para obtener el mejor rendimiento[16][17][18].

4.2. Batch size

Se refiere a la cantidad de ejemplos de capacitación utilizados en una iteración, este es uno de los hiper parámetros más importantes ya que especifica cuántas muestras se deben procesar antes de que se actualicen los parámetros del modelo interno. Este puede ser de 3 tipos: modo lote, en el cual el batch size es igual al conjunto de datos total, por lo que los valores de iteración y época serán equivalentes, también existe el modo de mini lote, dónde el batch size es mayor a 1 pero menor al tamaño del conjunto de datos total y por último, en el modo estocástico el batch size es igual a 1, y el gradiente y los parámetros de la red se actualizan después de cada muestra. El batch size juega un papel importante en el aprendizaje del modelo, por lo que éste debe ajustarse en función del rendimiento del modelo[19][20][21].

4.3. Validation split

Validation split es una técnica utilizada en el aprendizaje automático para evaluar el rendimiento de un modelo, consiste en dividir el conjunto de datos en tres subconjuntos: el de entrenamiento, utilizado tal como su nombre lo indica para entrenar el modelo, el de validación, que permite ajustar el modelo evitando el sobreajuste y el de prueba, se usa para evaluar el rendimiento final del modelo con datos que no se han utilizado. Es necesario determinar la división óptima que se adapte a las necesidades del conjunto de datos/modelo[22][23].

4.4. Coeficiente de similitud de Jaccard

Este coeficiente mide la similitud de entre dos conjuntos de muestras, comparando los elementos de cada conjunto y determinando cuáles se comparten y cuáles son diferentes; por esto, se define como la relación entre el tamaño de la intersección de ambos conjuntos y el tamaño de la unión, de la siguiente manera:

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

Su valor se mueve en un rango de 0 a 1, cuanto mayor sea el índice, más similares son los dos conjuntos de datos. Algunas de las ventajas de esta medida es su facilidad para

calcular y entender, es útil para conjuntos de datos binarios o categóricos y que tengan elementos repetidos; por el contrario, sus desventajas consisten en que no tiene en cuenta la magnitud de las diferencias entre los elementos de los conjuntos, no resulta útil para conjuntos de datos continuos o numéricos y con elementos faltantes[24].

4.5. Coeficiente de similitud de DICE

Se define como el doble del número de elementos comunes entre las imágenes a comparar dividido por la suma de los elementos de cada imagen, cómo se muestra en la ecuación 2[25].

$$sim_D(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (2)$$

4.6. Métricas de clasificación

Las métricas de clasificación son valores que miden el rendimiento de un modelo de aprendizaje automático, las métricas de evaluación de clasificación más comunes son las siguientes:

- **Exactitud (Accuracy):** Hace referencia a la capacidad del modelo para diferenciar correctamente los casos de pacientes y sanos

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

- **Sensibilidad (Sensitivity):** evalúa la capacidad del modelo para diferenciar correctamente los casos positivos para la enfermedad

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

- **Especificidad (Specificity):** examina la capacidad para determinar correctamente los casos sanos

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

- **Coeficiente de Kappa:** se usa para evaluar la concordancia de los datos respecto a lo que puede suceder por azar.

$$k = \frac{P_o - P_e}{1 - P_e} \quad (6)$$

En las ecuaciones 3, 4, 5 y 6 se encuentra TP como el número de casos identificados correctamente como positivo para la enfermedad, TN el número de casos identificados correctamente como sanos, FP el número de casos identificados incorrectamente como positivo para la enfermedad, FN el número de casos identificados incorrectamente como sanos, p_o la proporción de acuerdos observados y p_e la proporción de acuerdos esperados[26][27].

4.7. Modelos de backbone

El backbone es una red conocida que anteriormente fue entrenada en muchas otras tareas y demuestra su eficacia, es la columna vertebral del aprendizaje automático, es responsable de extraer características de los datos y realizar la clasificación o predicción. Algunos tipos más populares son el VGG, ResNet, DenseNet y AlexNet. Las redes neuronales convolucionales (CNN) se utilizan principalmente para clasificar imágenes, localizar objetos y extraer características de las imágenes, como bordes o esquinas; su éxito radica en que éstas pueden procesar grandes cantidades de datos, como imágenes, videos y texto[28].

Los modelos de backbone son redes neuronales convolucionales utilizados para extraer características de las imágenes y reducir la complejidad de los datos de entrada. Estos modelos son diferentes entre ellos ya que tienen características que los diferencian unos con otros, tales como:

- **Arquitectura:** Cada backbone tiene una arquitectura diferente, es decir, tienen diferentes números de capas, filtros, neuronas y conexiones.
- **Profundidad:** La profundidad hace referencia al número de capas que tiene. Los modelos más profundos suelen tener un mejor rendimiento en tareas de clasificación de imágenes, pero también son más difíciles de entrenar y requieren más recursos computacionales.
- **Precisión:** La precisión se entiende como la capacidad para clasificar correctamente las imágenes, hay modelos con precisión más baja que otros, lo cual depende también del conjunto de datos y de la tarea específica.
- **Velocidad:** La velocidad del backbone se refiere a la rapidez con la que puede procesar las imágenes. Sin embargo, los modelos más rápidos suelen tener menos capas y filtros, por lo tanto una menor precisión.
- **Tamaño:** El tamaño hace referencia a la cantidad de memoria que ocupa en el disco duro. Los modelos más grandes suelen tener una mayor precisión, pero también requieren más recursos computacionales y tardan más en entrenarse[29][30].

4.8. MobileNet

Es la descomposición de los núcleos de convolución. La convolución se utiliza para combinar las salidas de las capas de convolución en profundidad. De esta manera, N núcleos de convolución estándar pueden ser reemplazados por M núcleos de convolución en profundidad y N núcleos de convolución en puntos. Un filtro convolucional estándar combina las entradas en un nuevo conjunto de salidas, mientras que la convolución separable en profundidad divide las entradas en dos capas, una para filtrar y otra para fusionar[31].

En muchas tareas, principalmente aquellas tareas sobre las predicciones intensivas como la segmentación semántica de imágenes y la estimación de flujo óptico, es necesario predecir el valor de cada píxel de la imagen de entrada, y el valor de cada píxel de salida necesita un gran campo receptivo para retener información importante[31]

5. METODOLOGÍA

En esta sección se detalla el paso a paso que se realizó en el desarrollo para la implementación de un algoritmo que permita realizar un diagnóstico más rápido de la retinopatía diabética con la utilización de la base de datos de Indian Diabetic Retinopathy Image Dataset (IDRiD), se realizó mediante unos pasos a seguir los cuales son los siguientes, preprocesamiento, segmentación utilizando diferentes modelos como el FPN, PSPNet, Unet y Linknet con algunos tipos como el VGG16, VGG19, ResNet152 y mobilenet, esto para cada una de las patologías mencionadas anteriormente, utilización de una Red Neuronal Convolutiva con la capacidad de implementarse en un dispositivo móvil llamada MobileNet donde se realiza la clasificación de las enfermedades generando 2 modelos, uno para Retinopatía Diabética y otro para el Edema Macular Diabético.

5.1 Preprocesamiento

Se realizó en un primer lugar el análisis de la base de datos utilizada la cual se encuentra dividida en subcarpetas donde se cuenta con un conjunto de entrenamiento y uno de prueba correspondientes al 80% y 20% del total de las imágenes respectivamente con imágenes para cada una de las patologías (MA, HE, SE, EX y OD). Dicha división se proporciona en archivos CSV dónde se encuentra información como el número de imagen, nivel de gravedad de RD en el rango de 0 (sana sin RD aparente), 1 (RD leve), 2 (RD moderada), 3 (RD severa) y 4 (RD proliferativa o muy grave) y el nivel de gravedad del edema macular en el rango de 0 (sana o sin presencia de EMD), 1 (EMD leve) y 2 (EMD presente y grave)[32].

Las imágenes de ojo se encuentran en su color original como por ejemplo las que se pueden observar en la figura 1.

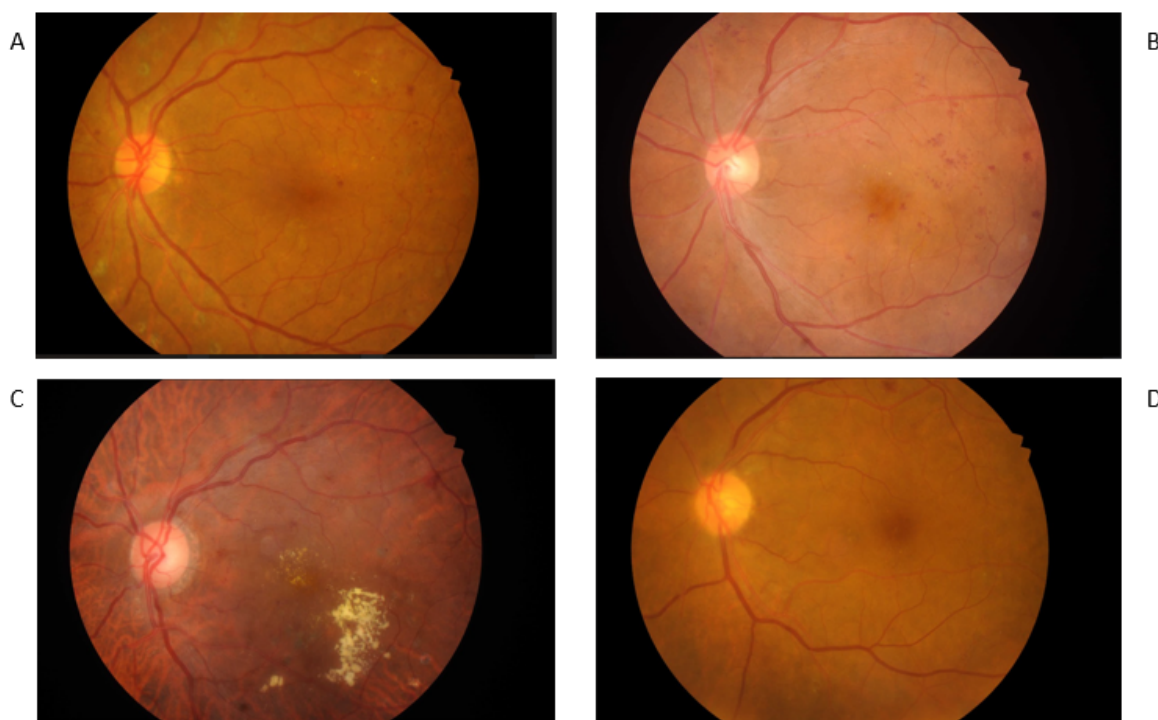


Figura 1. Imágenes Originales con color de fondo

En la ilustración 1, la imagen A y B corresponden a imágenes originales de training, la imagen C y D corresponden a imágenes originales de test, estas imágenes hacen parte de la carpeta de segmentación de la base de datos, estas imágenes se muestran debido a que fueron capturadas con cámaras diferentes y en centros diferentes logrando generar una motivación debido a la complejidad que puede surgir al momento de implementarlas en algoritmos, un ejemplo de este tipo de imágenes se puede observar en la figura 1 la cual corresponde a una imagen original de training.

En otra carpeta de la base de datos encontramos la graduación de la enfermedad, en esta carpeta encontramos las imágenes del ojo en color original para Retinopatía Diabética y para Edema Macular Diabetico, un ejemplo de estas imágenes pueden ser observadas en la figura 2 y la figura 3.

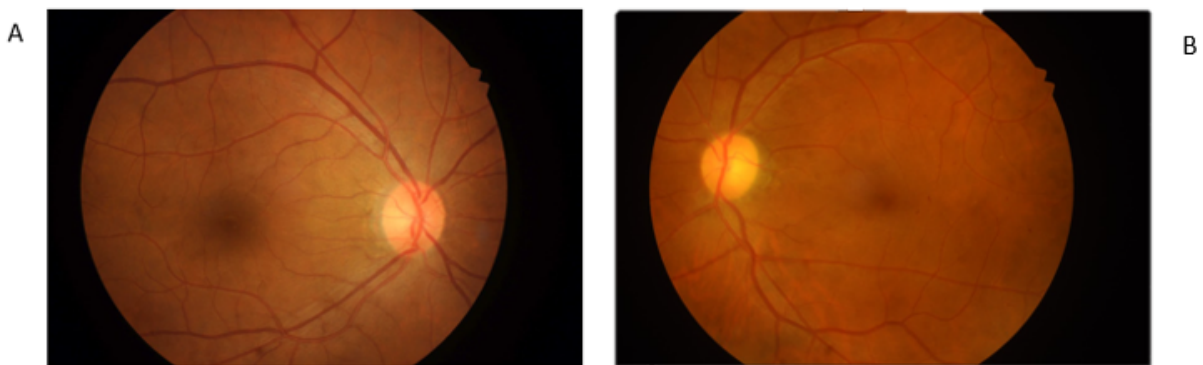


Figura 2. Imágenes de fondo RD

En la figura número 2, se puede observar que la imagen A corresponde a una persona sana de Retinopatía diabética en imágenes de training y la imagen B corresponde a una persona con retinopatía diabética de grado 2.

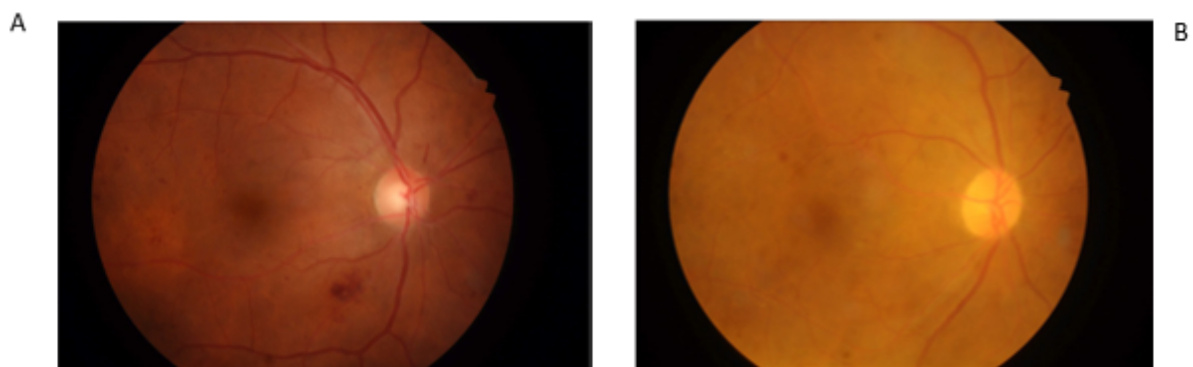


Figura 3. Imágenes de fondo EDM

En la figura número 3, se puede observar que la imagen A corresponde a una persona sana de Edema Macular Diabetico en imágenes de training y la imagen B corresponde a una persona con Edema Macular Diabetico de grado 1

	Número de imágenes en Training	Número de imágenes en Testing
Segmentación	195	276
Método basado en parches	50	50

Tabla 1. Número total de imágenes en la base de datos

En la tabla 1 se puede observar el número total de imágenes que se encuentran en la carpeta de la base de datos, en el caso de microaneurismas se tomaron todas las imágenes.

Patologías	Número de imágenes de Training	Número de imágenes de Testing
Microaneurisma (Basado en parches)	Sanos: 50 No sanos: 50	60
Exudados duros	54	27
Exudados blandos	26	14
Disco óptico	54	27
Hemorragias	53	27

Tabla 2. Cantidad de imágenes totales

En la tabla 2 se puede observar el número total de imágenes que se encuentran en la carpeta de segmentación, en el caso de microaneurismas se encuentra la clasificación de dos formas en el caso de training, esta clasificación se da como sanos y no sanos. Se decide realizar el algoritmo en este programa por la facilidad, organización y además la fácil vinculación de la cuenta drive o subir el archivo con solo las imágenes logrando de esta forma la carga segura del archivo para su respectiva segmentación, localización y graduación.

Patología	Training Enfermos	Testing Enfermos	Training Sanos	Testing Sanos
Retinopatía Diabética	279	69	134	34
Edema Macular Diabetico	236	58	177	45

Tabla 3. Cantidad de imágenes para la clasificación

En la tabla número 3 podemos observar el número total de imágenes de la clasificación realizada de manera binaria para las patologías de la retinopatía diabética y el edema

macular diabetico, se evidencia un problema de desbalance en el número de imágenes dado que las diferencias entre las cantidades para cada grado son bastante grandes.

En la base de datos encontramos las tablas informativas con el grado de la enfermedad que presenta el paciente para las imagenes de training y testing, esta información se puede observar en la tabla 4 donde se encuentra el número de la imagen el grado de la enfermedad que presenta para retinopatía diabética y edema macular diabetico.

A continuación se muestra parte de la información encontrada en las tablas anteriormente mencionadas que nos permite ver que las enfermedades son no excluyentes entre sí.

Nombre de la imagen	Grado de retinopatía	Grado de EMD
IDRID_001	3	2
IDRID_002	3	2
IDRID_003	2	2
IDRID_004	3	2
IDRID_005	4	0
IDRID_006	4	1
IDRID_007	4	0
IDRID_008	4	2
IDRID_009	3	2
IDRID_010	4	1
IDRID_011	3	1
IDRID_012	3	2
IDRID_013	3	0
IDRID_014	4	2
IDRID_015	4	2

Tabla 4. Clasificación grado de la enfermedad RD y EMD training

Con estas tablas y siguiendo recomendaciones clínicas de oftalmólogos de la Universidad del Rosario se realizó una clasificación de manera binaria para determinar qué imagen corresponde a paciente con RD o EMD y qué imagen a una de un paciente sano. La clasificación se determinó de 1 para pacientes con RD o EMD y de 0 para pacientes sanos creando de esta forma las etiquetas a ser usadas por el modelo binario para la clasificación de la enfermedad.

5.2. Entrenamiento

Las imágenes encontradas en la base de datos presentan un tamaño original de 4.288 x 2.848 y se realiza el cambio de tamaño a uno de 224 x 224 debido a que este tamaño es compatible con modelos de preentrenamiento como lo son el VGG16 o ResNet, el entrenamiento de la red neuronal se hizo con 100 épocas, un batch size igual a 2 y el validation split de 0.1, al finalizar el entrenamiento se calcula el coeficiente de similitud de Jaccard y Dice para cada modelo permitiendo así evaluar las predicciones resultantes comparado con la segmentación que provee la base de datos. Estos resultados se observan en la tabla 5 a la tabla 14 en las cuales se puede concluir el mejor modelo para cada patología.

5.2.1. Segmentación y Clasificación

Se ejecutó una etapa de exploración para cada patología implementando todas las imágenes, esto se realizó con el fin de poder identificar en un primer plano la enfermedad. En el caso de microaneurisma como se puede observar en la la figura 4, su segmentación tuvo que ser realizada primero haciendo una división en parches, esta división se realizó en las imágenes correspondiente para pacientes sanos y en las imagenes para pacientes patológicos ésto debido al tamaño de la afección. Una vez obtenida la segmentación de cada uno de los parches, se implementó con los diferentes backbones y realizando 1000 épocas buscando de esta forma obtener un modelo con mejores resultados.

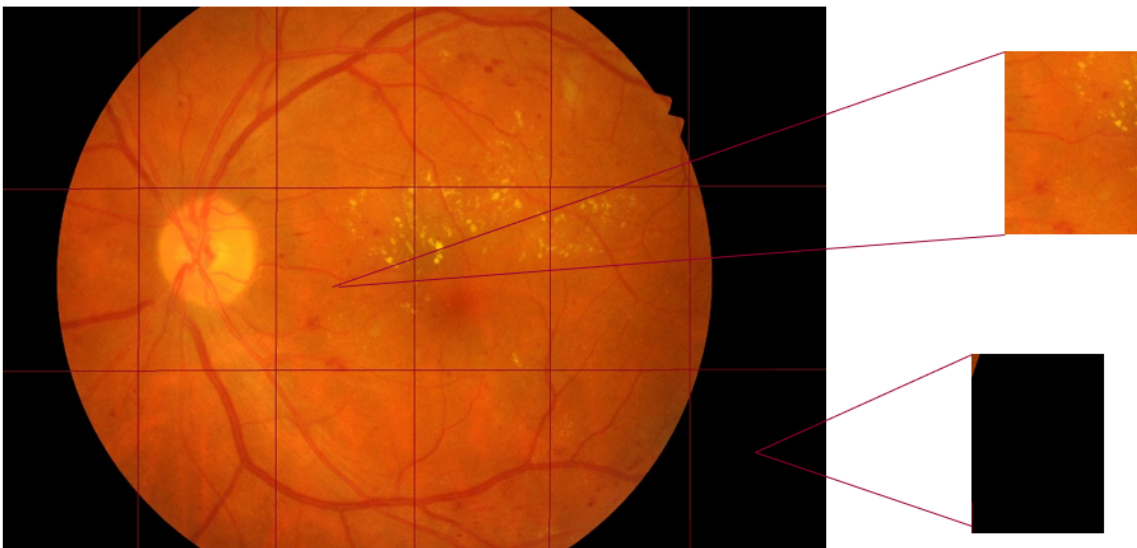


Figura 4. Imagen original Microaneurisma

En la figura 5 se evidencia parte de los parches realizados en las imágenes originales de los microaneurisma, estos parches se realizaron para las imágenes de pacientes patológicos como para pacientes sanos en las dos categorías Training y testing.

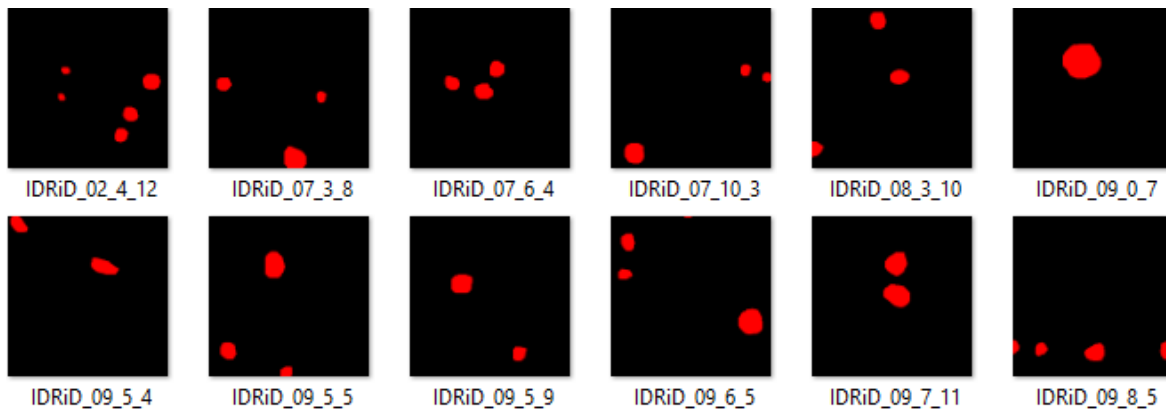


Figura 5 . Imágenes con los parches para pacientes no-sanos Microaneurisma

Se realizaron modelos de segmentación utilizando 4 modelos de backbone, los cuales se muestran en la figura número 6, estos backbone se implementaron con 4 tipos (VGG16, VGG19, ResNet 152 y mobilenet), teniendo como resultado un total de 80 modelos de segmentación, esto con el fin de tener un buen margen de comparación y así identificar el mejor modelo de segmentación para cada patología.

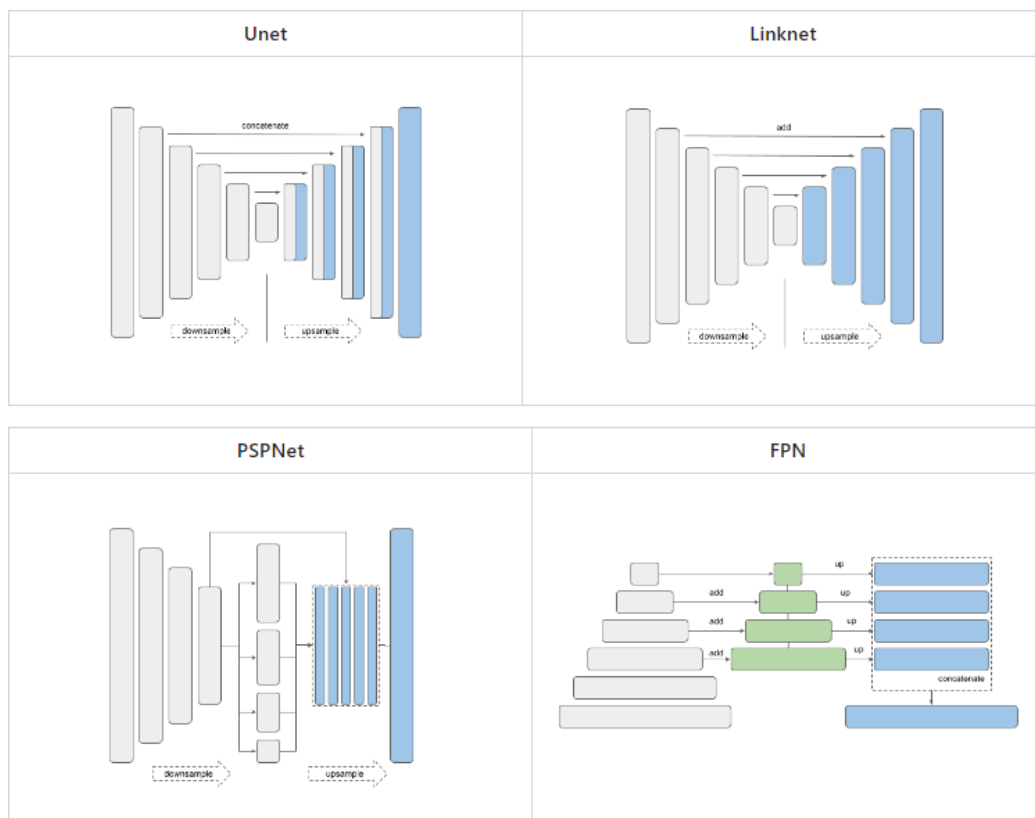


Figura 6. Modelos de segmentación explorados [33]

En cuanto a la clasificación de la enfermedad se realizaron 2 modelos, para el desarrollo de cada uno se clasificaron las imágenes en 2 clases pacientes sanos y pacientes con RD o EMD dependiendo el caso, esto se hizo con ayuda de las anotaciones proporcionadas por la base de datos y tomando el 0 como paciente sano y cualquier número diferente como

paciente patológico. Basado en lo anterior se realizaron diferentes experimentos con la exploración sistemática de hiper-parámetros, con la premisa de que debían ser modelos ligeros para que se puedan ejecutar en un dispositivo móvil o una tarjeta electrónica, y así posiblemente llegar a zonas alejadas donde el acceso a internet es limitado y en algunos casos completamente imposible, por esta razón se utiliza la Red Neuronal Convolutiva de MobileNet.

5.3 Validación

Se utilizó el coeficiente de similitud de Jaccard para evaluar y categorizar los mejores modelos para cada patología esto debido a que se utiliza principalmente para conjuntos de datos binarios o categóricos. Por otro lado, también se calculó el coeficiente de Dice, el cual es una métrica relacionada con el índice de *Jaccard*.

Estas métricas mencionadas son utilizadas para medir la similitud entre dos conjuntos de datos, para este caso específico, se utiliza con el fin de evaluar la calidad del modelo para identificar ciertas patologías en imágenes de fondo de ojo. En la ilustración 9 se ejemplifica la relación existente entre el índice de Jaccard y el coeficiente de Dice, dónde se observa que estos valores oscilan entre 0 y 1, siendo 0 un indicador de conjuntos completamente diferentes y 1 completamente idénticos[34].

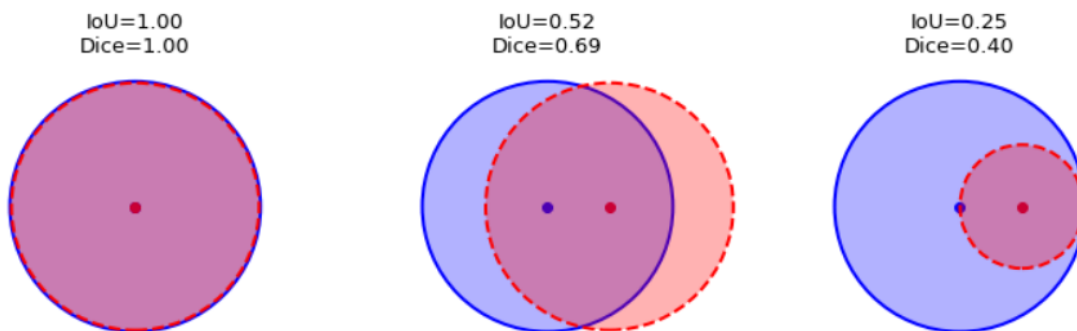


Figura 7. Relación entre el índice de Jaccard y el coeficiente de Dice

El valor reportado, tanto de jaccard como de Dice, para cada patología es el promedio del coeficiente calculado para cada una de las imágenes de entrenamiento y prueba, obteniendo un valor para cada una de estas, se decide implementar el índice Jaccard y el coeficiente Dice debido a que son las más empleadas y en relación al artículo “Optimization for Medical Image Segmentation: Theory and Practice when evaluating with DiceScore or Jaccard Indexse” suelen ser estudiadas las dos, sin embargo, dependiendo del tamaño de los hallazgos se recomienda usar Dice o Jaccard[35].

Para la parte de clasificación, se utilizan 4 métricas para caracterizar y evaluar el modelo resultante, estas son, exactitud (o en inglés “Accuracy”); la sensibilidad (o en inglés “Sensitivity”); la especificidad (en inglés “Specificity”); y por último, el coeficiente de Kappa.

6. RESULTADOS

Con la extracción de características por medio de la segmentación realizada con el procedimiento descrito en la sección 5.2.1. Modelos de segmentación y Clasificación, se obtuvieron diferentes predicciones para cada una de las patologías tanto para training como para testing a continuación se muestra el resultado de training y test para 5 imágenes.

PREDICTIONS IN THE TRAINING SET

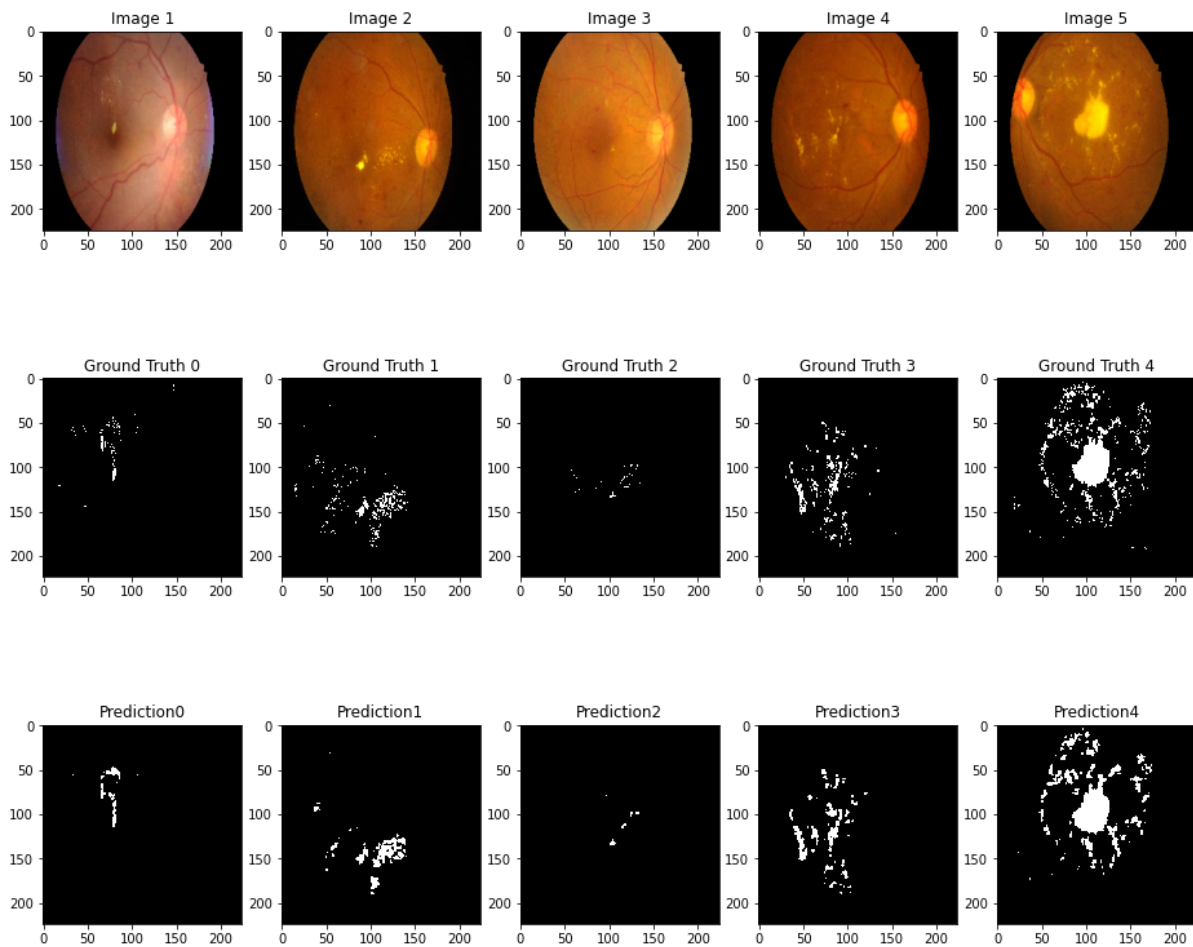


Figura 8. Resultados obtenidos en la predicción de HE en conjunto de entrenamiento

PREDICTIONS IN THE TEST SET

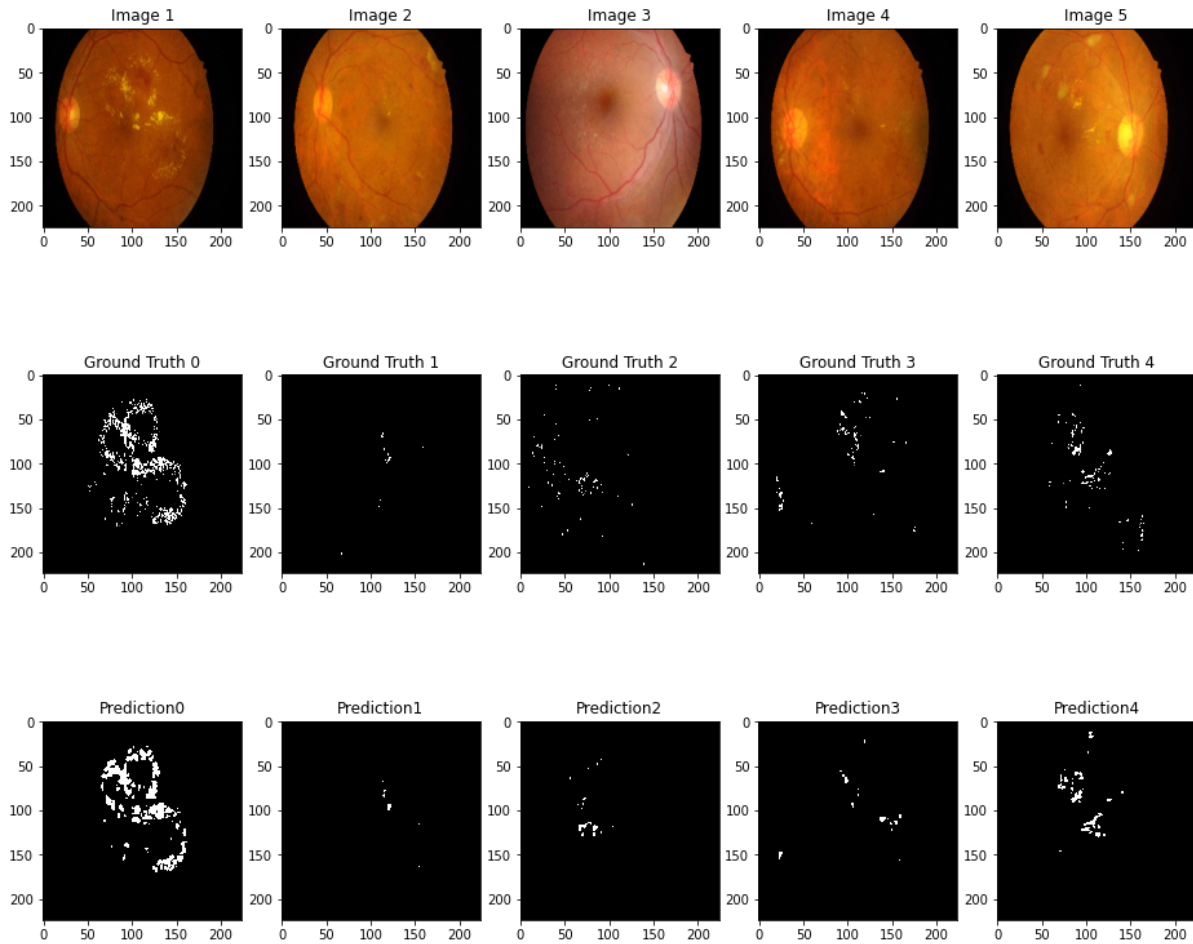


Figura 9. Resultados obtenidos en la predicción de HE en conjunto de prueba

Con el objetivo de encontrar el mejor modelo para cada patología, se calculó el índice de jaccard, el cual mide el grado de similitud entre la predicción arrojada por la red neuronal y la realidad, de esta forma se encontraron modelos con mayor precisión que otros. Los resultados fueron organizados en tablas discriminadas por patología y modelos, también fueron resaltados los 2 modelos que obtuvieron el jaccard más alto para cada patología.

SOFT EXUDATES								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0,6427	0,63461	0,6465	0,6465076	0,2785	0,2785468	0,7860	0,7859581
LINKNET	0,3465	0,3465138	0,5367	0,5367085	0,6590	0,6589521	0,8194	0,8193609
PSPNET	0,3997283	0,3997283	0,5436	0,5435616	0,0000	0,0000	0,4640	0,463997
FPN	0,1086199	0,1069	0,0000	0	0,0000	0,0000	0,7250	0,7249929

Tabla 5. Métrica de Jaccard obtenida en Exudados Blandos.

En la tabla 5 se presentan los resultados de Jaccard para los exudados blandos, se evidencia que los 2 mejores modelos son utilizando el backbone de Mobilenet con arquitecturas como Unet y Linknet con valores de 0,786 y 0,8194 respectivamente, tanto en entrenamiento como en prueba.

HARD EXUDATES								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0,2383	0,2891	0,0000	0,0000	0,2244	0,2941	0,3380	0,4017
LINKNET	0,3343	0,3854	0,0018	0,0035	0,2639	0,3124	0,2994	0,3618
PSPNET	0,1764	0,2291	0,0000	0,0000	0,1822	0,2302	0,1498	0,1995
FPN	0,1861	0,2511	0,2215	0,2838	0,3878	0,4506	0,4252	0,4890

Tabla 6. Métrica de Jaccard obtenida en Exudados Duros.

La tabla 6 muestra los valores de Jaccard al segmentar exudados duros, dónde los mejores resultados se obtuvieron con Mobilenet (Unet y FPN) en entrenamiento y prueba.

HEMORRAGIAS								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0,0447	0,0513	0,2095	0,2047	0,1238	0,1272	0,4116	0,4317
LINKNET	0,0105	0,0104	0,1598	0,1651	0,1595	0,1628	0,4116	0,4317
PSPNET	0,0000	0,0000	0,0000	0,0000	0,1540	0,1365	0,2538	0,2722
FPN	0,0000	0,0000	0,0000	0,0000	0,3298	0,3624	0,3866	0,4177

Tabla 7. Métrica de Jaccard obtenida de Hemorragias.

Durante la exploración de hiperparametros para la segmentación de hemorragias se encontraron algunos resultados muy malos, cuyo coeficiente de jaccard fue igual a 0, sin embargo, los mejores valores se encontraron con Mobilenet, esto se evidencia en la tabla 7

DISCO ÓPTICO								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0,0039	0,0068	0,9239	0,9269	0,8818	0,8652	0,9761	0,9825
LINKNET	0,9419	0,9508	0,9330	0,9419	0,9312	0,9387	0,9606	0,9669
PSPNET	0,8686	0,8722	0,9133	0,9139	0,8741	0,8350	0,9432	0,9465
FPN	0,8692	0,8657	0,8478	0,8385	0,7896	0,7358	0,9511	0,9533

Tabla 8. Métrica de Jaccard obtenida de Disco Óptico

Respecto al disco óptico es uno de los hallazgos con mejores resultados, estos se presentan en la tabla 8, señalando los 2 valores de jaccard más altos en entrenamiento y prueba. Finalmente, los resultados de Jaccard para microaneurismas se encuentran

registrados en la tabla 9, evidenciando mayor dificultad en la segmentación dado que algunos modelos tienen un desempeño muy bajo.

MICROANEURISMAS								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0,0000	0,0000	0,2014	0,0033	0,2420	0,0017	0,7800	0,0113
LINKNET	0,0000	0,0000	0,7719	0,0103	0,0000	0,0000	0,8144	0,0104
FPN	0,0000	0,0000	0,5179	0,0090	0,0000	0,0001	0,6862	0,0096

Tabla 9. Métrica de Jaccard obtenida de Microaneurismas.

De igual manera que se hizo con el Jaccard, se calculó el valor del coeficiente de Dice y se plasmaron en las siguientes tablas, dando valores para entrenamiento y prueba en cada uno de los modelos utilizados. En primer lugar se encuentran los resultados de exudados blandos en la tabla 10, donde se registran resultados muy buenos para el backbone de Mobilenet junto a Unet y Linknet.

SOFT EXUDATES								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0.7825	0.7765	0.7853	0.7853	0.4357	0.4357	0.8802	0.8802
LINKNET	0.5147	0.5147	0.6985	0.6985	0.7944	0.7944	0.9007	0.9007
PSPNET	0.5712	0.5712	0.7043	0.7043	0.0000	0.0000	0.6339	0.6339
FPN	0.1960	0.1931	0.0000	0.0000	0.0000	0.0000	0.8406	0.8406

Tabla 10. Métrica de Dice obtenida en Exudados Blandos

La tabla 11 presenta los coeficientes de Dice resultantes de la segmentación de exudados duros, luego de haber entrenado 16 modelos, allí se señalan los 2 con rendimiento y por lo tanto precisión más alta.

HARD EXUDATES								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0.3849	0.4486	0.0000	0.0000	0.3665	0.4545	0.5052	0.5731
LINKNET	0.5010	0.5564	0.0037	0.0071	0.4176	0.4761	0.4608	0.5313
PSPNET	0.2999	0.3728	0.0000	0.0000	0.3082	0.3743	0.2606	0.3326
FPN	0.3137	0.4014	0.3626	0.4422	0.5588	0.6213	0.5967	0.6568

Tabla 11. Métrica de Dice obtenida en Exudados Duros

Los valores más altos de Dice para la segmentación de hemorragias se lograron al utilizar Mobilenet, los cuales fueron 0.5832 en entrenamiento y 0.603 en prueba, tanto para Unet como Linknet, dicha evidencia se encuentra organizada en la tabla 12.

HEMORRAGIAS								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0.0855	0.0976	0.3464	0.3398	0.2204	0.2257	0.5832	0.6030
LINKNET	0.0208	0.0205	0.2756	0.2834	0.2751	0.2800	0.5832	0.6030
PSPNET	0.0000	0.0000	0.0000	0.0000	0.2669	0.2401	0.4048	0.4279
FPN	0.0000	0.0000	0.0000	0.0000	0.4961	0.5320	0.5576	0.5893

Tabla 12. Métrica de Dice obtenida en Hemorragias

En el disco óptico, tal como se presentó en Jaccard, tiene los mejores resultados, dando lugar a modelos muy precisos, en la tabla 13 se encuentran los valores de dice para cada experimento realizado

DISCO ÓPTICO								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0.0078	0.0134	0.9604	0.9620	0.9372	0.9277	0.9879	0.9912
LINKNET	0.9701	0.9748	0.9653	0.9701	0.9644	0.9684	0.9799	0.9832
PSPNET	0.9297	0.9317	0.9547	0.9550	0.9328	0.9101	0.9708	0.9725
FPN	0.9301	0.9280	0.9176	0.9121	0.8825	0.8478	0.9749	0.9761

Tabla 13. Métrica de Dice obtenida en Disco Óptico

Por último, la tabla 14 registra los resultados de la segmentación para microaneurismas, de igual manera que en la mayoría de las tablas ya presentadas, los valores más altos fueron obtenidos al utilizar el backbone de Mobilenet.

MICROANEURISMAS								
MODELOS	VGG16		VGG19		RESNET152		MOBILENET	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
UNET	0.0000	0.0000	0.3353	0.0066	0.3896	0.0033	0.8764	0.0223
LINKNET	0.0000	0.0000	0.8713	0.0204	0.0000	0.0000	0.8977	0.0206
FPN	0.0000	0.0000	0.6824	0.0178	0.0000	0.0002	0.8139	0.0190

Tabla 14. Métrica de Dice obtenida en Microaneurismas

En la tabla 15 se presentan los resultados de las métricas obtenidas en los modelos de clasificación de la enfermedad (RD y EMD), aquí se relacionan cuatro métricas (exactitud, sensibilidad, especificidad e índice de Kappa) en el conjunto de entrenamiento y de prueba para cada enfermedad. Es importante mencionar que durante la búsqueda del mejor modelo se variaron el número de épocas (50, 100, 200), el tamaño de lote o batch-size (4, 8, 16, 32, 64, 128) y la tasa de aprendizaje o learning rate (0.1, 0.01, 0.001, 0.0001, 0.00001), dónde los mejores hiperparámetros para la clasificación de RD fue: 50 épocas, batch size de 16 y tasa de aprendizaje de 0.001; y los mejores hiperparámetros para la clasificación de EMD fue: 50 épocas, tamaño de lote de 8 y tasa de aprendizaje de 0.0001, cuyos resultados quedaron registrados a continuación.

CLASIFICACIÓN								
Enfermedad	Exactitud		Sensibilidad		Especificidad		Kappa	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
RD	0.9104	0.7184	0.7388	0.4705	0.9928	0.8405	0.7816	0.3283
EMD	0.9128	0.7669	0.8813	0.6666	0.9364	0.8448	0.8212	0.5192

Tabla 15. Métricas para clasificación de RD y EMD

7. DISCUSIÓN

Teniendo en cuenta los resultados presentados anteriormente, se puede evidenciar que en la segmentación de las patologías hay bastantes diferencias entre los modelos utilizados, sin embargo, la red neuronal convolucional MobileNet presenta los mejores resultados tanto en el conjunto de entrenamiento como en el de prueba en todas las patologías estudiadas. Sin embargo, la segmentación de algunas patologías como exudados blandos y duros tuvieron mejores resultados de manera general en todos los modelos, por lo que se puede afirmar que es mucho más sencillo evidenciarlo debido a su tamaño y color que permiten identificarse con facilidad. En especial, los exudados son un hallazgo vital para el diagnóstico de EMD, lo que puede ayudar al oftalmólogo a emplear los modelos de clasificación de EMD combinado con la salida del modelo de segmentación para un mejor diagnóstico.

En el caso del disco óptico su segmentación es la más cercana al esperado coeficiente de 1 tanto en jaccard como en dice, esto debido a que es una parte del ojo que se percibe con mayor facilidad dado su tamaño, forma y color característico. Los resultados de esta segmentación se muestran en las tablas 7 y 12 correspondientes al valor de Jaccard y Dice respectivamente, allí se puede evidenciar que todos los modelos empleados segmentan de manera muy precisa el disco óptico a excepción del VGG16 con Unet, cuyo resultado es bastante bajo. La segmentación es vital debido a que podemos evidenciar que hallazgos están más cercanos o lejanos (cerca de la mácula y fovea) que pueden estar relacionados con un RD o EMD moderado o severo y que por ende comprometen la visión de los pacientes.

Por otro lado, los microaneurismas fue la patología más difícil de segmentar, por lo que fue necesario utilizar una técnica diferente a la segmentación de las demás, ya que se dividió cada imagen en 18 parches con el fin de entrenar el modelo con imágenes donde fuese mucho más clara la afección, ya que, como se ha mencionado el tamaño de estas protuberancias es muy pequeño dificultando que sean vistas de manera simple. En la actualidad, la clasificación ETDRS no está teniendo en cuenta a los microaneurismas para el diagnóstico de RD debido a que no son fáciles de detectar y por su tamaño no representan un compromiso de visión. De allí, la importancia de contar con modelos para hallazgos de tamaño mayores a 5 micras como las hemorragias.














Lesion	Team name	AUPR	Approach	Ensemble	Input Size (Pixels)	External dataset
Microaneurys	 iFLYTEK	0.5017	Cascaded CNN	✓	320 × 320	×
	 VRT	0.4951	U-Net	×	1280 × 1280	×
	 PATech	0.4740	DenseNet+U-Net	✓	256 × 256	×
	 SDNU	0.4111	Mask R-CNN	×	3584 × 2380	×
Hemorrhages	 VRT	0.6804	U-Net	×	640 × 640	×
	 PATech	0.6490	DenseNet+U-Net	✓	256 × 256	×
	 iFLYTEK	0.5588	Cascaded CNN	✓	320 × 320	×
	 SOONER	0.5395	U-Net	×	380 × 380	×
Soft Exudates	 VRT	0.6995	U-Net	×	640 × 640	×
	 LzyUNCC-I	0.6607	FCN+DLA	×	1024 × 1024	E-ophtha
	 iFLYTEK	0.6588	Cascaded CNN	✓	320 × 320	×
	 LzyUNCC-II	0.6259	FCN+DLA	×	1024 × 1024	E-ophtha
Hard Exudates	 PATech	0.8850	DenseNet+U-Net	✓	256 × 256	×
	 iFLYTEK	0.8741	Cascaded CNN	✓	320 × 320	×
	 SAIHST	0.8582	U-Net	×	512 × 512	×
	 LzyUNCC-I	0.8202	FCN+DLA	×	1024 × 1024	E-ophtha

Figura 10. Resultados para la segmentación de Hemorragias, Microaneurismas, exudados blandos, exudados duros reportados en la literatura [32]

No es posible realizar una comparación directamente debido a que AUPR solo permite tener una medida respecto a la proporción de hallazgos que son encontrados en la base de datos, al querer ofrecer al médico una herramienta que le diga dónde están los hallazgos y que los pinte, se enfoca la tarea de segmentación hacia detección / identificación de hallazgos permitiendo de esta forma ofrecer algo adicional a los médicos con la finalidad de que sean los médicos quienes tomen la decisión sobre lo que nos da el modelo de clasificación y pasando cada imagen por el mejor modelo de cada hallazgo para tener la información de que posibles zonas pueden tener Hemorragias, Microaneurismas, exudados blandos, exudados duros, disco óptico.

Respecto a la clasificación de RD y EMD se registran valores mayores a 0,7 en entrenamiento en las 4 métricas tal como aparece en la tabla 15, los valores de prueba tienden a ser un poco menores lo cual, es normal ya que se utiliza un conjunto de datos diferente al utilizado para entrenar el modelo. Sin embargo, al comparar estos resultados, en especial la precisión de los modelos, con los datos registrados en la literatura y presentados en la figura 11, dónde se evidencian valores de precisión por debajo de los obtenidos en este trabajo, adicional a eso, cabe mencionar que no se utilizaron bases de datos externas para obtener dichos resultados, como si se hizo en el estudio reportado en el artículo “Conjunto de datos de imágenes de retinopatía diabética india (IDRiD): una base de datos para la investigación de detección de retinopatía diabética” [32].







Team Name	Accuracy	Approach	Ensemble	Input Size (Pixels)	External Dataset
 LzyUNCC	0.6311	Resnet + DLA	5	896 × 896	Kaggle
 VRT	0.5534	CNN	10	640 × 640	Kaggle, Messidor
 Mammoth	0.5146	DenseNet	✓	512 × 512	Kaggle
 HarangiM1	0.4757	AlexNet + GoogLeNet	2	224 × 224	Kaggle
 AVSASVA	0.4757	ResNet + DenseNet	DR-8, DME-5	224 × 224	DiaretDB1
 HarangiM2	0.4078	AlexNet + Handcrafted features	2	224 × 224	Kaggle

Figura 11. Resultados para la clasificación de la Retinopatía diabética y Edema Macular Diabetico reportados en la literatura [32]

8. RECOMENDACIONES Y TRABAJOS FUTUROS

En la segmentación de patologías se obtuvo buenos resultados con algunos modelos, sin embargo, respecto a los microaneurismas es importante explorar otros modelos y otras técnicas que permitan una segmentación mucho más precisa debido a que esta patología genera un reto muy grande por su pequeño tamaño y ubicación de dichas protuberancias, por lo que detectarlas por medio del aprendizaje automático requiere de una mayor cantidad de imágenes, más anotaciones y analizar otros modelos que puedan combinar los resultados de segmentación y de clasificación para de esta forma poder ayudar en la toma de decisiones clínicas, resaltando la importancia de un diagnóstico temprano para evitar o retrasar posibles complicaciones de la retinopatía diabética como la ceguera.

Sería interesante contar con bases de datos más grandes y con otras patologías que nos ayuden a tamizar la salud ocular de nuestras ciudades principales como Bogotá, Cali y Medellín. Por otro lado, la necesidad de evaluar estas herramientas desarrolladas en un entorno clínico pueden posibilitar el futuro uso de estos algoritmos y la aceptación por parte de personal clínico como herramienta para el soporte en la toma de decisión médica.

9. CONCLUSIONES

Los algoritmos desarrollados en este trabajo dirigido permiten la detección y segmentación automática de lesiones oculares como microaneurismas, hemorragias, exudados duros, exudados blandos y partes del ojo como el disco óptico. Adicionalmente, la segmentación de estos hallazgos en conjunto con los algoritmos de clasificación de RD y EMD son potencialmente herramientas útiles para la toma de decisiones médicas.

Por un lado, los métodos reportados en este trabajo dirigido fueron validados sistemáticamente y poseen la habilidad de detectar lesiones oculares con diversos tamaños y diferentes características morfométricas, texturas y color. Por otro lado, nuestros métodos poseen un número reducido de parámetros lo que puede posibilitar a futuro la implementación en dispositivos móviles como celulares y tabletas, pero también el desarrollo e implementación en tarjetas embebidas como Raspberry pi, Jetson nano o Google coral.

Los resultados obtenidos en detección y clasificación son considerablemente buenos de acuerdo a lo reportado en la literatura. Las lesiones relacionadas con microaneurismas representaron un desafío mayor debido a que los modelos de segmentación requirieron un enfoque diferencial basado en parches, esto por su variación en cuanto a su tamaño (≤ 100 micras), sus texturas y localización cercana a venas y arterias, lo que hace difícil hasta para los médicos el poder establecer criterios para dar un diagnóstico apropiado.

Por otro lado, los modelos de clasificación usando la red neuronal convolucional MobileNet nos permite realizar la clasificación de RD y EMD con métricas de desempeño sobresalientes, lo que posibilita a futuro dos grandes desafíos: la fácil implementación y aceptación de estos algoritmos por parte de personal clínico; y el desarrollo de trabajos futuros en pro de masificar la validación clínica en diferentes entornos y con diferentes dispositivos de adquisición de imágenes de fondo de ojo.

Finalmente, se espera que este trabajo sea un paso inicial en el desarrollo de telemedicina a nivel local, regional y nacional. A su vez posiciona a la academia frente al sector salud como un aliado para a futuro desarrollar otro tipo de herramientas que apoyen la cobertura de servicios en zonas de difícil acceso. Lo anterior, podría potencialmente reducir la desigualdad y mejorar la calidad de vida de personas que desconocen el estado de la salud ocular y el tipo de tratamiento que tempranamente debería tomar.

10. REFERENCIAS

1. Romero-Aroca, P. (s/f). *La retinopatía diabética e hipertensiva*. Comcordoba.com. Recuperado el 15 de mayo de 2023, de <https://www.comcordoba.com/wp-content/uploads/2018/08/La-retinopat%C3%ADa-diab%C3%A9tica-e-hipertensiva.-AMF-2018.pdf>
2. López-Gálvez, M. I., & García-Campos, J. M. (2012). De la evidencia científica a la práctica clínica: pautas de tratamiento del edema macular diabético. *Archivos de la Sociedad Española de Oftalmología*, 87 Suppl 1, 38–45. [https://doi.org/10.1016/s0365-6691\(12\)70050-3](https://doi.org/10.1016/s0365-6691(12)70050-3)
3. Ocular, S. (s/f). *Rodrigo Abreu González*. Revistadiabetes.org. Recuperado el 28 de mayo de 2023, de <https://www.revistadiabetes.org/wp-content/uploads/Retinografia-y-su-implantacion-con-inteligencia-artificial.pdf>
4. Marano, R. C. (2016). Fisiopatología del edema macular diabético: nuevos conceptos. *Annals d'oftalmologia: òrgan de les Societats d'Oftalmologia de Catalunya, Valencia i Balears*, 24(4), 3. <https://dialnet.unirioja.es/servlet/articulo?codigo=5768287>
5. Asociación Americana de Diabetes. Costos económicos de la diabetes en los EE. UU. en 2017. *Diabetes Care*. 2018;41(5):917-928
6. Ramirez, P., & Jeannette, L. (2018). *Prevalencia de retinopatía diabética en diabéticos tipo II en un centro de referencia 2012 – 2017*. <https://repository.urosario.edu.co/items/72fc4c50-8924-45a8-86cb-600255ffa541>
7. Schmidl D, Garhöfer G, Schmetterer L. La naturaleza compleja de la retinopatía de los pacientes diabéticos e hipertensos: la necesidad de un enfoque multimodal. *Eur J Ophthalmol*. 2018;28(4):359-368.
8. VISLISEL, J.E.S.S.E. y OETTING, T.H.O.M.A.S. (2010) RETINOPATÍA DIABÉTICA, Retinopatía diabética para estudiantes de medicina. Disponible en: <http://www.eyerounds.org/tutorials/diabetic-retinopathy-med-students/> (Consulta: 12 de enero de 2023).
9. W.StittaPerson1Envelope, A. et al. (2015) El progreso en la comprensión y el tratamiento de la retinopatía diabética, Progreso en la investigación ocular y retinal. Pérgamo. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S135094621500066X> (Consultado: 12 de enero de 2023).
10. Rojas, A. M., Aparicio-pico, L.-E., Electrónica, I., José, F., & Bogotá, D. C. (1841). *Técnicas de inteligencia artificial aplicadas al análisis de Artificial intelligence techniques applied to the analysis of diagnostic images . Di*.
11. Bravo, J. D., Correa, A., Bravo, A., Bravo, R., & Villada, O. A. (2022). Retinopatía Diabética y Edema Macular Diabético en Población de Antioquia. Estudio transversal. *IATREIA*, 35(2), 98–107. <https://doi.org/10.17533/udea.-iatreia.125>.
12. López Gálvez, M. I. (2004). Escala internacional de severidad de la retinopatía y del edema macular diabético. *Archivos de la Sociedad Española de Oftalmología*, 79(4), 149–150. <https://doi.org/10.4321/s0365-66912004000400001>
13. (S/f-b). Gov.co. Recuperado el 15 de mayo de 2023, de <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/PP/ENT/asis-salud-visual-colombia-2016.pdf>
14. Organización Mundial de la Salud. (2016b). Informe mundial sobre la diabetes. Organización Mundial de la Salud.

15. Yau JW, Rogers SL, Kawasaki R, et al. Prevalencia mundial y principales factores de riesgo de la retinopatía diabética. *Cuidado de la diabetes*. 2012;35(3):556-564.
16. *Entrenar un modelo de aprendizaje profundo (Image Analyst)*. (s/f). Arcgis.com. Recuperado el 28 de mayo de 2023, de <https://pro.arcgis.com/es/pro-app/3.0/tool-reference/image-analyst/train-deep-learning-model.htm>
17. Cervera, P. C., Rodríguez Vázquez, J., & Sampedro, C. (s/f). *Métodos de aprendizaje profundo para la segmentación semántica de personas TRABAJO FIN DE GRADO PARA LA OBTENCIÓN DEL TÍTULO DE GRADUADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES*. Upm.es. Recuperado el 28 de mayo de 2023, de https://oa.upm.es/64643/1/TFG_EDUARDO_TEMPRANO_COLETO.pdf
18. Wikipedia contributors. (s/f). *Red neuronal artificial*. Wikipedia, The Free Encyclopedia. https://es.wikipedia.org/w/index.php?title=Red_neuronal_artificial&oldid=149306247
19. Varikuti, M. (2022, 1 de mayo). ¿Cuál es el efecto del tamaño del lote en el aprendizaje del modelo? Hacia la IA. <https://pub.towardsai.net/what-is-the-effect-of-batch-size-on-model-learning-196414284add>
20. ¿Qué es el tamaño del lote en la red neuronal? (s/f). Cruz validado. Recuperado el 28 de mayo de 2023, de <https://stats.stackexchange.com/questions/153531/what-is-batch-size-in-neural-network>
21. *Batch Size in a Neural Network explained*. (s/f). Deeplizard.com. Recuperado el 28 de mayo de 2023, de <https://deeplizard.com/learn/video/U4WB9p6ODjM>
22. Kumar, A. (2021, 13 de junio). Aprendizaje automático: conjunto de datos de entrenamiento, validación y prueba. Análisis de datos. <https://vitalflux.com/machine-learning-training-validation-test-data-set/>
23. Baheti, P. (2023, abril 24). Train test validation split: How to & best practices [2023]. *V7labs.com*. <https://www.v7labs.com/blog/train-validation-test-set>
24. Benites, L. (2021, diciembre 21). *Índice Jaccard / Coeficiente de similitud*. Statologos: El sitio web para que aprendas estadística en Stata, R y Phytion. <https://statologos.com/indice-jaccard/>
25. Lorenzetti, C. M. (2011). Caracterización Formal y Análisis Empírico de Mecanismos Incrementales de Búsqueda basados en Contexto. *Universidad Nacional Del Sur*
26. Marano, R. C. (2016). Fisiopatología del edema macular diabético: nuevos conceptos. *Annals d'oftalmologia: òrgan de les Societats d'Ofthalmologia de Catalunya, Valencia i Balears*, 24(4), 3. <https://dialnet.unirioja.es/servlet/articulo?codigo=5768287>
27. Baratloo, A., Hosseini, M., Negida, A., & Ashal, G. El. (2015). *Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity*. 3(2), 48–49. www.jemerg.com
28. Elharrouss, O., Akbari, Y., Almaadeed, N., & Al-Maadeed, S. (2022). Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches. En *arXiv [cs.CV]*. <http://arxiv.org/abs/2206.08016>
29. Pablo, J., & Rodriguez, R. (s/f). *UNIVERSIDAD DE CHILE FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS DEPARTAMENTO DE INGENIERÍA ELÉCTRICA DESTILACIÓN DE MODELO EN REDES CONVOLUCIONALES MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELECTRICO*. Uchile.cl.

- Recuperado el 28 de mayo de 2023, de <https://repositorio.uchile.cl/bitstream/handle/2250/178006/Destilacion-de-modelo-en-redes-convolucionales.pdf?sequence=1>
30. (S/f). Redalyc.org. Recuperado el 28 de mayo de 2023, de <https://www.redalyc.org/journal/5537/553768213002/html/>
 31. Wang, W., Hu, Y., Zou, T., Liu, H., Wang, J. y Wang, X. (2020). Un nuevo enfoque de clasificación de imágenes a través de modelos MobileNet mejorados con expansión del campo receptivo local en capas poco profundas. *Inteligencia Computacional y Neurociencia*, 2020, 8817849. <https://doi.org/10.1155/2020/8817849>
 32. Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V. y Meriaudeau, F. (2018). Conjunto de datos de imágenes de retinopatía diabética india (IDRiD): una base de datos para la investigación de detección de retinopatía diabética. *Datos*, 3(3), 25. <https://doi.org/10.3390/data3030025>
 33. Iakubovskii, P. (s/f). *segmentation_models*: modelos de segmentación con backbones preentrenados. Keras y TensorFlow Keras
 34. Métricas para segmentación semántica - Excursiones en datos. (Dakota del Norte.). Recuperado el 14 de mayo de 2023, de <https://ilmonteux.github.io/2019/05/10/segmentation-metrics.html>
 35. (S/f). Redalyc.org. Recuperado el 28 de mayo de 2023, de <https://www.redalyc.org/journal/5537/553768213002/html/>
 36. Eelbode, T., Bertels, J., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R., & Blaschko, M. B. (2020). Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index. *IEEE Transactions on Medical Imaging*, 39(11), 3679-3690.
 37. *Keras applications*. (s/f). Keras.io. Recuperado el 29 de mayo de 2023, de <https://keras.io/api/applications/>