

Evaluation of a portable retinograph assisted by convolutional neural networks for the early diagnosis and classification of diabetic retinopathy

Pachas Pazos, Adrian Alexander¹; Huaman Albornoz, Paola Andrea²; Hermoza Paz, Luis Ricardo³

^{1,2,3} Universidad Tecnológica del Perú, Perú, u17205692@utp.edu.pe, u20207972@utp.edu.pe, c24362@utp.edu.pe

Abstract– Diabetic retinopathy (DR) is one of the leading causes of vision loss worldwide, especially in regions with limited access to ophthalmological services. The aim of this study was to develop and evaluate a low-cost portable retinograph assisted by Deep Learning for early detection and severity classification of DR. The system integrates a portable device based on indirect ophthalmoscopy, 3D printing, and smartphone image acquisition, together with convolutional neural networks (CNNs) for automated analysis. A combined dataset from the Messidor, APTOS, and EyePACS databases, comprising approximately 49,430 retinal images, was used for training and validation. Image preprocessing, class balancing, and transfer learning techniques were applied using MobileNetV2 and DenseNet121 architectures, evaluated in binary and multiclass classification tasks. The results demonstrated that the prototype captured fundus images with sufficient quality for clinical visualization. In conclusion, the proposed system represents a feasible and accessible solution for DR screening in telemedicine and resource-limited settings.

Keywords-- CNN, diabetic retinopathy, deep learning, fundus imaging.

Evaluación de un retinógrafo portátil asistido por redes neuronales convolucionales para el diagnóstico temprano y la clasificación de la retinopatía diabética

Pachas Pazos, Adrian Alexander¹; Huaman Alborno, Paola Andrea²; Hermoza Paz, Luis Ricardo³

^{1,2,3} Universidad Tecnológica del Perú, Perú, u17205692@utp.edu.pe, u20207972@utp.edu.pe, c24362@utp.edu.pe

Resumen– La retinopatía diabética (RD) es una de las principales causas de pérdida visual a nivel mundial, especialmente en regiones con acceso limitado a servicios oftalmológicos. El objetivo de este estudio fue desarrollar y evaluar un retinógrafo portátil de bajo costo asistido por Deep Learning para la detección temprana y clasificación del grado de severidad de la RD. El sistema integra un dispositivo portátil basado en oftalmoscopia indirecta, impresión 3D y captura de imágenes mediante smartphone, junto con redes neuronales convolucionales (CNN) para el análisis automatizado. Para el entrenamiento y validación se empleó un conjunto de datos combinado de las bases Messidor, APTOS y EyePACS, con aproximadamente 49 430 imágenes retinianas. Se aplicaron técnicas de preprocesamiento, balanceo de clases y Transfer Learning utilizando las arquitecturas MobileNetV2 y DenseNet121, evaluadas en tareas de clasificación binaria y multiclase. Los resultados demostraron que el prototipo capturó imágenes con calidad suficiente para la visualización clínica. En conclusión, el sistema propuesto representa una solución viable y accesible para el tamizaje de la RD en telemedicina y entornos con recursos limitados.

Palabras clave– CNN, retinopatía diabética, aprendizaje profundo, fondo de ojo.

I. INTRODUCCIÓN

La retinopatía diabética (RD) es una de las principales causas de ceguera en personas con diabetes, resultado del daño progresivo de los vasos sanguíneos de la retina por niveles elevados y prolongados de glucosa en sangre, especialmente en pacientes con diabetes tipo 2 [1]. En 2024, se estimó que 537 millones de personas entre 20 y 79 años padecían diabetes mellitus (DM), con una prevalencia del 6,4 % en Perú, equivalente a más de 1,3 millones de casos [2]. Cerca del 31,5 % de los diabéticos presenta signos de RD, aunque menos del 30 % accede a un examen oftalmológico anual [3,4].

El diagnóstico de la RD se realiza mediante retinografía, que consiste en la captura de imágenes del fondo de ojo mediante cámaras especializadas, ya sean midriáticas, las cuales requieren dilatar la pupila con gotas oftálmicas, o no midriáticas, las cuales no requieren dilatar la pupila. Estas últimas son preferidas por su rapidez y facilidad de uso. Sin embargo, su interpretación depende del criterio del especialista, lo que puede generar variaciones diagnósticas y limitar su aplicación en zonas con escasos recursos [5].

Aunque los equipos estacionarios ofrecen alta calidad de imagen, su tamaño y costo restringen su uso en áreas rurales,

donde la prevalencia de RD es mayor (29,1 %) frente a las zonas urbanas (18,1 %) [6]. Esta brecha se debe al menor acceso a exámenes oftalmológicos y al costo de los servicios médicos. Por ello, estrategias como la teleoftalmología, el uso de inteligencia artificial (IA) y la capacitación del personal sanitario son alternativas clave para mejorar la detección temprana y reducir las desigualdades en atención visual [7].

En el contexto regional, se han demostrado los beneficios de la teleoftalmología. Un estudio en La Libertad evidenció que la implementación de una red con 3 cámaras retinianas no midriáticas (modelos estacionarios como Topcon TRC-NW8 y Canon CR-series) resultó en un aumento del 138,1 % en la detección de casos de RD [8]. Sin embargo, el alto costo de estos equipos y la necesidad de interpretación manual por especialistas limitan su aplicación.

La literatura muestra diversos esfuerzos para abordar esta problemática, demostrando la viabilidad del uso de distintos algoritmos de IA para el cribado de RD.

En este contexto de innovación, diversos estudios han aplicado estas metodologías con resultados prometedores. Por ejemplo, en Ecuador evaluaron comparativamente arquitecturas de redes neuronales convolucionales (CNN) contrastando dos modelos residuales (ResNet-18 y ResNet-50) con una CNN personalizada (no preentrenada), los resultados mostraron una variabilidad notable en la precisión: ResNet-18 alcanzó el 26%, ResNet-50 logró el 83%, y la CNN personalizada obtuvo la mayor precisión con un 91%. A pesar de estos altos rendimientos, el estudio demostró que el éxito de los modelos depende intrínsecamente de la calidad de la base de datos, la cual exige imágenes capturadas con cámaras de alta gama. Esta dependencia, a su vez, encarece la solución y dificulta significativamente su implementación en entornos con recursos limitados [9].

En esa misma línea, en México desarrollaron un modelo basado en Deep Learning para la detección de RD, validado con 400 imágenes (200 con RD y 200 sin RD), se obtuvo una tasa de acierto promedio (o exactitud) del 83.23% en las imágenes con RD y del 78.86% en las imágenes sin RD, aunque este hallazgo es positivo para la viabilidad de la detección binaria mediante Deep Learning, se requiere que el modelo demuestre la capacidad de clasificar los distintos grados de severidad de la RD para su aplicación clínica [10].

Por su lado, en Colombia aplicaron la técnica de *Transfer Learning* para la detección y clasificación de RD, abordando ambos objetivos por separados, para la detección, lograron una precisión de 98.15% con una sensibilidad (Se) de 95.67% y una especificidad (Sp) de 98.14%, mientras que para la clasificación, obtuvieron una precisión de 82.68% con una Se de 83.52% [11].

Asimismo, en Rusia y Tailandia, evaluaron la eficacia de modelos basados en *Deep Learning* para la detección de RD: en la primera, con una muestra de prueba de 1186 fotografías de fondo de ojo de 593 pacientes, demostró como resultado una alta precisión del 95.9%, junto a una Se y una Sp del 95%, sin embargo, se basó en imágenes obtenidas con equipos no portátiles en entornos hospitalarios, lo que limita la aplicabilidad del modelo en áreas rurales o con recursos limitados, donde no se dispone de estos dispositivos; en la segunda, se analizó en 4852 pacientes diabéticos, utilizando cámaras retinianas no midriáticas, las cuales ser fijas y no portátiles, como resultado, se señaló que el algoritmo presentó mejores resultados que los evaluadores humanos (método manual realizado por oftalmólogos) en términos de Se y Sp [12,13].

De forma análoga, investigadores en Brasil y Estados Unidos emplearon modelos de aprendizaje profundo para la detección de RD utilizando diversos tipos de cámaras retinianas, entre ellas Centervue DRS, Optovue iCam, Canon CR1/DGi/CR2, Topcon NW y la cámara no midriática Topcon TRC-NW6. En Brasil, se analizaron 15816 imágenes de 4590 pacientes capturadas con equipos estándar en hospitales, obteniéndose una Se del 93,5 % y una Sp del 94,6 % frente al diagnóstico manual de oftalmólogos. En Estados Unidos, el modelo se aplicó a 11711 imágenes de 5 871 pacientes, logrando Se entre 96,1 % y 97,5 % y Sp entre 93,4 % y 93,9 %, lo que evidencia un desempeño igualmente elevado [14,15].

No obstante, las mejoras en la precisión algorítmica y la dependencia de equipos de alta gama, la principal barrera para un cribado masivo y sostenible reside en la infraestructura y la portabilidad de los dispositivos de captura. Por ello, otro conjunto de estudios se ha centrado en evaluar la viabilidad y el rendimiento de los retinógrafos portátiles.

En relación con lo mencionado, en España, la evaluación de dos dispositivos portátiles mediante IA mostró un desempeño sobresaliente, con solo 7,98 % de imágenes que no pudieron ser evaluadas, un F1 de 0,99 y una Se de 0,99 para detectar cualquier grado de RD [16]. Mientras, en Brasil se evaluó a 350 pacientes mediante un programa de telemedicina con retinógrafos portátiles, identificando RD leve (12,23 %), moderada (6,31 %) y proliferativa (2,58 %), y reportando una discrepancia mínima del 7,8 %, lo que equivale a más del 92 % de coincidencia con la clasificación estandarizada [17]. Estos resultados demuestran que los retinógrafos portátiles,

son necesarios para la detección de la RD, además demuestran su efectividad en el campo de la telemedicina. Sin embargo el costo de estos equipos son elevados, por ejemplo, equipos no midriáticos portátiles pueden costar entre los US \$15000 a \$20000 según el proveedor y estos equipos no trabajan con una IA asistida por lo que aún requieren el diagnóstico de un retinólogo, esto evidencia la limitación práctica debido a su alto precio.

Esto refuerza su validez del aprendizaje profundo como una herramienta diagnóstica confiable y eficaz para la detección automatizada de RD; no obstante, resalta la necesidad de su integración en un sistema portátil de bajo costo que permita aprovechar el potencial de la IA para mejorar el acceso al diagnóstico de RD en comunidades desatendidas.

Así, la integración de un retinógrafo portátil con un sistema de análisis basado en IA representa un enfoque prometedor que podría transformar significativamente el cribado de la RD, en especial en comunidades con acceso limitado a servicios oftalmológicos especializados. Este tipo de solución permitiría la ejecución de evaluaciones del fondo de ojo en la atención primaria de salud y, mediante el uso de algoritmos de aprendizaje, facilita diagnósticos automatizados en tiempo real [18]. Además, al reducir la dependencia de equipos costosos y personal altamente especializado, se optimizaría el uso de los recursos disponibles y se acortaría el tiempo de respuesta diagnóstica.

El objetivo principal del presente estudio es evaluar la aplicación de un sistema basado en *Deep Learning*, integrado a un retinógrafo portátil para la captura de imágenes del fondo de ojo, implementando modelos de aprendizaje para su análisis y clasificación del nivel de severidad de la RD, los cuales se evaluarán previamente para identificar las estrategias y métodos más adecuados, con el fin de contribuir al diagnóstico temprano de la enfermedad.

II. MATERIALES Y MÉTODOS

La presente investigación se compone por dos etapas principales e interdependientes, hardware y software: 1) fase de adquisición de datos oftálmicos (retinógrafo portátil), 2) fase de análisis y clasificación automática (análisis automático para el diagnóstico de RD mediante una CNN), cada una con sus subfases. La Fig. 1 muestra un diagrama de bloques del prototipo desarrollado en este estudio, desde la adquisición de la imagen de fondo de ojo del paciente hasta la obtención del resultado diagnóstico. En esta sección, se describen todos los pasos en detalle.

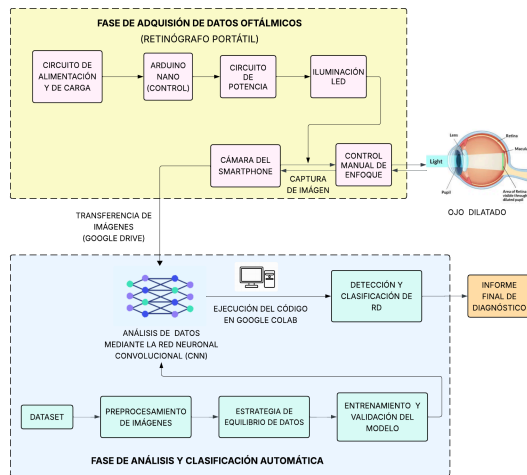


Fig. 1 Diagrama de bloques del flujo de trabajo para el diagnóstico y clasificación de RD

A. Fase de Adquisición de Datos Oftálmicos

Esta etapa comprende el diseño e implementación del prototipo de retinógrafo portátil, el cual se desarrolló en base a la oftalmoscopia indirecta. El sistema se dividió en dos partes: la parte electrónica, encargada de proporcionar la iluminación y el suministro de energía requeridos para realizar el estudio, conformada por 1) el circuito de alimentación y de carga, 2) el Arduino Nano (control), 3) el circuito de potencia y 4) la iluminación LED; y una estructura que integra todos los componentes ópticos necesarios para generar y capturar la imagen del fondo de ojo, compuesta por 5) la cámara del smartphone y 6) el control manual de enfoque.

1) Circuito de alimentación y carga

El circuito de alimentación está compuesto por dos baterías de litio de 3.7 V conectadas en serie, cuya función es suministrar energía a los LED de potencia y al circuito de control. Estas baterías se cargan mediante un módulo BMS 2S 18650, que permite una gestión adecuada de carga y protección.

2) Arduino Nano (control)

En la parte de control se utilizó un Arduino Nano, y se utilizó un código de modulación por ancho de pulso (PWM) para controlar la intensidad del led, en este código se lee un valor analogico en el pin A0, lo convierte al rango de 0 a 255 y lo utiliza para generar una señal PWM en el pin 9, la cual varía el brillo del LED haciendo que este se encienda y apague rápidamente a una frecuencia fija de 1 kHz, dando la apariencia de un brillo continuo proporcional al valor leído.

3) Circuito de potencia

Debido a que el arduino nano entrega en el pin 9 una corriente de 20mA se implementó un circuito de potencia, ver Fig. 2. Este circuito regula la corriente de los LEDs usando un MOSFET IRF530 y un BC547 que supervisa la caída de tensión en la resistencia R que se encuentra conectada al

surtidor del mosfet. Al encenderse, el MOSFET conduce porque su gate recibe tensión a través de la resistencia de 4,7 k Ω , permitiendo el paso de corriente a los LEDs. Esa corriente genera una caída de voltaje en R y, cuando supera unos 0,6 V, el BC547 se activa y reduce la tensión del gate del MOSFET, limitando su conducción. Así, el circuito mantiene la corriente del LED estable, determinada principalmente por el valor de la resistencia R.

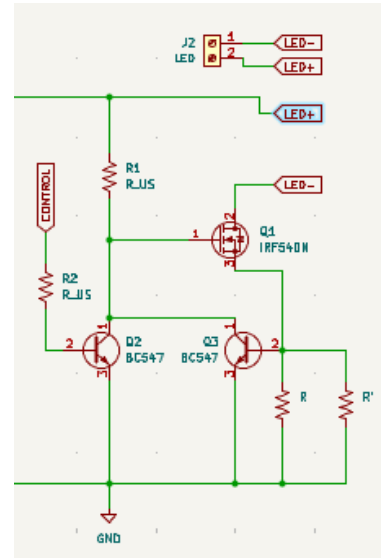


Fig. 2 Circuito de Potencia para la iluminación de los leds

4) Iluminación LED

Para la iluminación se empleó un LED blanco de 5w con disipación de 6500k. Con el fin de asegurar que la luz emitida no fuera perjudicial para el paciente, se realizó una comparación de la intensidad lumínica con un luxómetro, midiendo los lúmenes emitidos por un oftalmoscopio indirecto de la marca RIESTER modelo al pupil y los emitidos por nuestro prototipo. En nuestro prototipo se obtuvo 600 lux y en el oftalmoscopio riester 1000 lux estos valores son con la intensidad máxima, pero cabe resaltar que en los exámenes que se realizan no se usa el 100% de la intensidad lumínica sin embargo el tener el control de la intensidad ayuda mucho ya que hay muchos pacientes que son fotosensibles y en estos casos se puede regular la intensidad y realizar mejor el examen.

5) Cámara del smartphone

Para la captura de las imágenes del fondo de ojo se utilizaron cámaras de distintos smartphones (Xiaomi, Samsung e iPhone). El smartphone no solo cumple la función de capturar las imágenes, sino que trabaja en conjunto con un lente oftálmico de 20 dioptrías, el cual permite visualizar el fondo de ojo del paciente, el cual proporciona un campo de visión aproximado de 46°. Este lente fue acoplado al sistema mediante un adaptador impreso en 3D, diseñado

específicamente para permitir su enroscado seguro en el prototipo, así como un fácil desacople para su posterior limpieza o mantenimiento cuando sea necesario. Este conjunto óptico se integra en la estructura del prototipo, asegurando la correcta alineación entre la cámara del celular, la fuente de iluminación y el lente oftálmico, lo que garantiza una captura nítida y centrada del fondo de ojo, como se muestra en la Fig. 3.

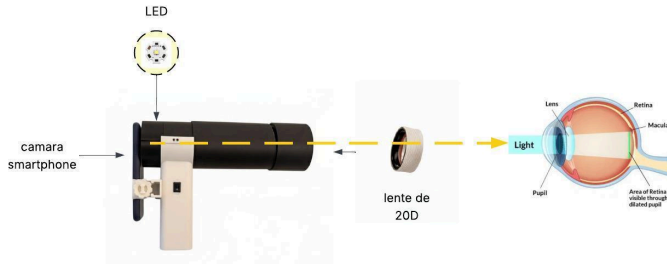


Fig. 3 Sistema óptico para la captura del fondo de ojo

6) Control manual de enfoque

Esta sección describe la construcción de la estructura del dispositivo. El diseño se realizó en Fusion 360, elaborándose inicialmente tres prototipos, de los cuales el tercero fue el más eficiente (Fig. 4). Una vez elegido el diseño final, se procedió a su fabricación mediante impresión 3D: el mango del equipo se imprimió con filamento ABS por su resistencia, mientras que la zona destinada a sostener el lente oftálmico de 20 dioptrías se fabricó con filamento PLA-CF, un material reforzado con fibra de carbono que brinda mayor rigidez y estabilidad.

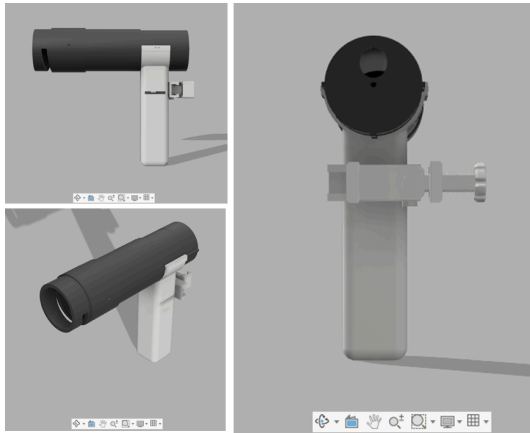


Fig. 4 Diseño final del Retinógrafo Portátil en fusion 360

B. Fase de Análisis y Clasificación Automática

Esta etapa comprende el desarrollo e implementación del modelo de CNN, el cual está alojado en un entorno de ejecución en la nube (Google Colab) y operación sobre las imágenes oftálmicas adquiridas en la Fase de Hardware.

El proceso incluye la organización del conjunto de datos, el acondicionamiento de las imágenes, la aplicación de estrategias para el equilibrio de clases y el entrenamiento y validación del modelo, con el fin de obtener un sistema capaz de identificar la presencia de la enfermedad y determinar su nivel de gravedad, generando finalmente un informe automático de diagnóstico.

1) Transferencia y carga de imágenes

Las imágenes retinianas capturadas por el retinógrafo portátil son transferidas a la plataforma Google Colab mediante almacenamiento en la nube, permitiendo su acceso y gestión para el procesamiento computacional

2) Dataset

Para garantizar la solidez, generalización y validez clínica del modelo, se empleó un conjunto de datos combinado de grandes bases de referencia a nivel mundial: Messidor, APTOS (Asia Pacific Tele-Ophthalmology Society) y EyePACS. La combinación de estas fuentes es fundamental, ya que permite exponer a la CNN a una mayor variabilidad en la calidad de imagen, los equipos de adquisición y las poblaciones de pacientes. De esta manera, se minimiza la dependencia del modelo respecto a los parámetros de adquisición de las imágenes, permitiendo que el desempeño diagnóstico sea consistente al ser aplicado en un entorno clínico [19].

El conjunto de datos combinado totaliza aproximadamente 49430 imágenes oftálmicas. En la Tabla I, se detallan las características de las bases utilizadas:

TABLA I
DESCRIPCIÓN DEL CONJUNTO DE DATOS

Base de Datos	Origen geográfico	Formato	Cantidad
MESSIDOR	Francia	TIFF	1200
APTOS 2019	India	JPEG	3662
EyePACS	EEUU	JPEG	44568
Total de imágenes			49430

El conjunto de datos inicial se sometió a un riguroso proceso de filtrado. Este proceso fue crítico para excluir imágenes de técnicas no deseadas, como la Tomografía de Coherencia Óptica (OCT) y la Angiografía, conservando exclusivamente las fotografías de fondo de ojo relevantes para el diagnóstico de RD.

Todas las imágenes filtradas y conservadas fueron etiquetadas y clasificadas en tres datasets independientes para clasificación binaria, asimismo de tres y cinco niveles de RD, siguiendo la escala estandarizada de la Asociación Americana de Oftalmología (AAO), con el fin de entrenar a la CNN. Los cinco niveles son: Grado 0 (Sin RD), Grado 1 (RD Leve), Grado 2 (RD Moderada), Grado 3 (RD Severa), y Grado 4

(RD Proliferativa), mientras que para el modelo de tres niveles se clasificó en Grado 0 (Sin RD), Grado 1 (RD Leve), Grado 2 (RD Grave).

3) Preprocesamiento de imágenes

El preprocesamiento de las imágenes retinianas se realizó con el fin de estandarizar el formato, el tamaño y las características visuales del fondo de ojo previo al entrenamiento de la CNN. La Fig. 5 muestra de manera secuencial el flujo metodológico aplicado, iniciando con la carga de las imágenes mediante la librería OpenCV, las cuales se visualizan inicialmente en el espacio de color BGR, lo que explica la tonalidad azulada observada. Posteriormente, se realizó la conversión al espacio de color RGB y la extracción de la región de interés mediante la eliminación de los bordes negros, conservando únicamente la zona retinal relevante.

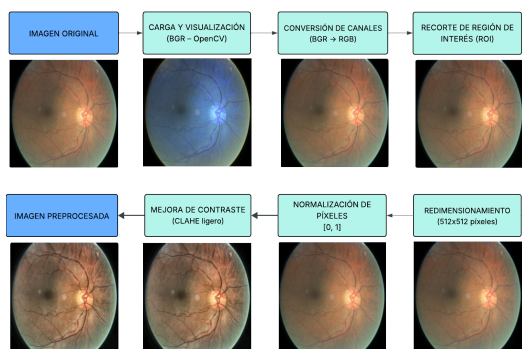


Fig. 5 Diagrama del proceso de preprocesamiento de imágenes de fondo de ojo

Las imágenes procesadas fueron redimensionadas a 512×512 píxeles, normalizadas al rango $[0,1]$ y sometidas a una mejora de contraste ligera mediante CLAHE. Este procedimiento permitió obtener imágenes estandarizadas y optimizadas, favoreciendo la estabilidad del entrenamiento y el desempeño del modelo en la clasificación de retinopatía diabética

4) Entrenamiento y validación del modelo

Fase fundamental donde se evaluaron las arquitecturas: MobileNetV2 y DenseNet121, basadas en deep learning, con el objetivo de determinar el modelo con mejor desempeño para la detección y clasificación de la RD. Se estableció un procedimiento técnico orientado a reducir el sobreajuste y mejorar la capacidad de generalización del modelo, mediante el control del proceso de aprendizaje y la selección de la arquitectura más adecuada.

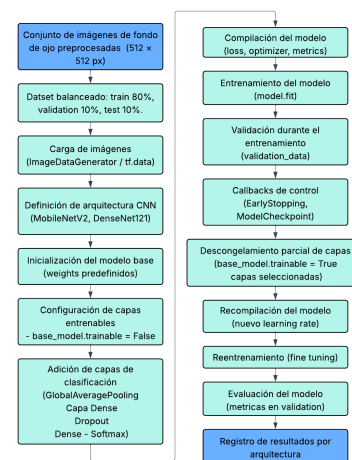


Fig. 6 Diagrama del proceso de entrenamiento de modelos CNN

La Fig. 6 muestra el flujo técnico implementado para el entrenamiento y validación de las arquitecturas CNN, el cual abarca desde la utilización de imágenes de fondo de ojo previamente preprocesadas hasta la evaluación del desempeño del modelo. Asimismo, se incluyen etapas clave como el congelamiento de capas convolucionales, la adición de capas de clasificación, el ajuste fino del modelo y la aplicación de mecanismos de control durante el entrenamiento, permitiendo el registro y comparación objetiva de los resultados obtenidos.

5) Análisis de datos mediante la CNN

En esta etapa se emplea la CNN previamente entrenada y validada para procesar las imágenes retinianas de entrada. El análisis se realiza mediante la ejecución del modelo en el entorno de cómputo, generando para cada imagen un conjunto de salidas asociadas a las clases definidas para la RD.

Dichas salidas corresponden a las probabilidades estimadas por el modelo y constituyen la base para la evaluación del comportamiento del sistema frente a los distintos niveles de severidad de la enfermedad. Los resultados obtenidos en esta etapa permiten verificar la consistencia del modelo y analizar su respuesta ante imágenes no utilizadas durante el entrenamiento, sirviendo como insumo para la posterior etapa de detección y clasificación automática de la RD.

III. RESULTADOS Y DISCUSIÓN

A. Validación funcional del prototipo en pruebas de campo

El prototipo de retinógrafo portátil mostrado, de acuerdo a la Fig. 9, fue diseñado mediante el software Fusion 360 e impreso en 3D. Este dispositivo fue empleado durante la etapa de prueba de campo, con el objetivo de validar su ergonomía, alineación óptica así como la intensidad de luz y capacidad de acoplamiento a un smartphone para la captura de imágenes del fondo de ojo.



Fig. 9 Prototipo de retinógrafo portátil impreso en 3D

Las pruebas en campo se realizaron en una clínica, con la asistencia de un retinólogo y siguiendo estrictamente todos los protocolos de seguridad establecidos. Durante dichas pruebas, se aplicó previamente una gota de tropicamida con el fin de inducir la dilatación pupilar y facilitar una mejor visualización del fondo de ojo. Posteriormente, el médico efectuó la evaluación mediante el método tradicional de oftalmología indirecta y, una vez realizado el diagnóstico correspondiente, se procedió a la captura de las imágenes del fondo de ojo utilizando el prototipo desarrollado, tal como se observa en la Figura 7. Estas imágenes fueron empleadas para su posterior análisis mediante la CNN entrenada.

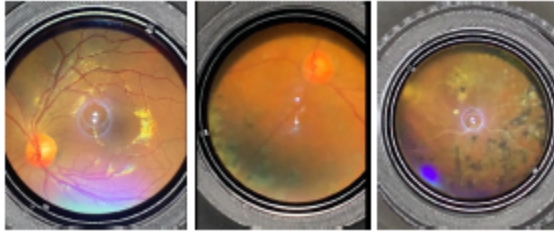


Fig. 8 Imágenes de fondo de ojo de tres pacientes capturadas por el prototipo de retinógrafo portátil, la primera es sin RD, la segunda RD Proliferativo y la tercera es RD Severo.

B. Métricas de Evaluación

En la Tabla II se exponen los resultados obtenidos a partir de la evaluación del desempeño de los tres modelos CNN implementados, métricas de rendimiento que permiten valorar la capacidad de clasificación del modelo.

TABLA II
MÉTRICAS DE DESEMPEÑO PARA CADA MODELO CNN

Métricas de desempeño	MobileNetV2 (2 clases)	DenseNet121 (3 clases)	DenseNet121 (5 clases)
Precisión	67.42%	64.04%	51.01%
Especificidad (Sp)	96.54%	95%	92.36%
Sensibilidad (Se)	93.82%	54.70%	51.56%
F1-score	79.14%	78%	49.48%
AUC	65.02%	97.10%	83%

1) MobileNetV2 (modelo binario)

En la Fig. 9 se evidencia que el modelo logra identificar correctamente la gran mayoría de los casos con RD (1844), presentando un número reducido de falsos negativos (66), lo que evidencia una alta sensibilidad y una adecuada capacidad para la detección de pacientes con la enfermedad.

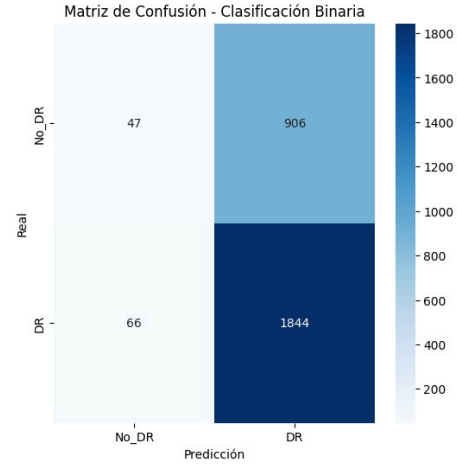


Fig. 9 Matriz de confusión del modelo binario con arquitectura MobileNetV2

En la Fig. 10 se aprecia que la curva ROC alcanza un AUC de 0.650, indicando una capacidad consistente del modelo para diferenciar entre imágenes con y sin RD.

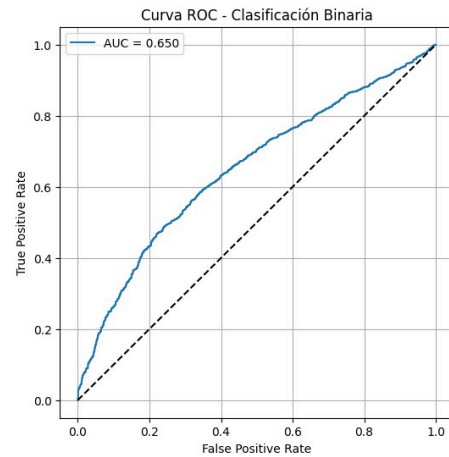


Fig. 10 Curva ROC de clasificación binaria

2) DenseNet121 (modelo de 3 clases)

En la Fig. 11 se evidencia que en el modelo destaca la clasificación correcta de la clase Grave (1809), con un número mínimo de errores, lo que refleja una alta capacidad para reconocer estadios avanzados de la enfermedad. Asimismo, la clase Leve presenta una adecuada proporción de aciertos (1011), mientras que la clase No_DR es reconocida en menor medida, aunque mantiene una tendencia de clasificación hacia estados con presencia de patología.

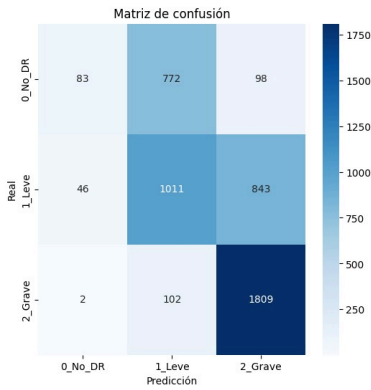


Fig. 11 Matriz de confusión del modelo de tres clases con arquitectura DenseNet121

En la Fig. 12 se aprecian las curvas ROC donde evidencian una alta capacidad de discriminación para la clase Grave (AUC = 0.971 y un buen desempeño para No_DR (AUC = 0.855), mientras que la clase Leve presenta un comportamiento aceptable (AUC = 0.714).

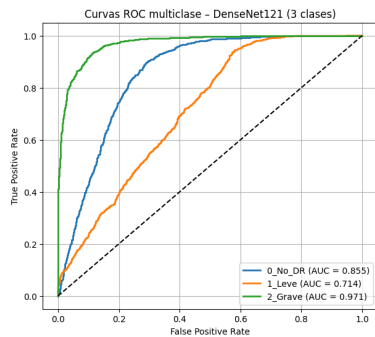


Fig. 12 Curva ROC para el nivel sin RD de clasificación de tres clases

3) DenseNet121 (modelo de 5 clases)

La Fig. 13 evidencia un desempeño consistente del modelo con una alta cantidad de aciertos en las clases Moderate (748) y Severe (611), lo que confirma su capacidad para identificar estadios clínicamente relevantes de la RD. Asimismo, la clasificación de Proliferative muestra una proporción importante de detecciones correctas (379), reflejando un adecuado reconocimiento de los casos más avanzados..

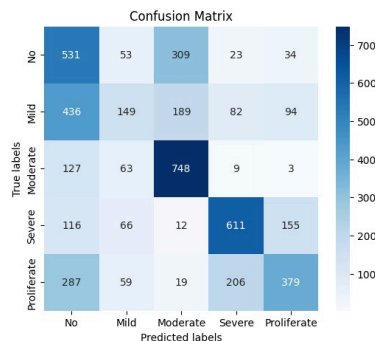


Fig. 13 Matriz de confusión del modelo de cinco clases con arquitectura DenseNet121

En la Fig. 14 las curvas ROC confirman una alta capacidad de discriminación para Moderate y Severe (AUC = 0.91), además de un buen desempeño en Proliferative_DR (AUC = 0.83) y No_DR (AUC = 0.73).

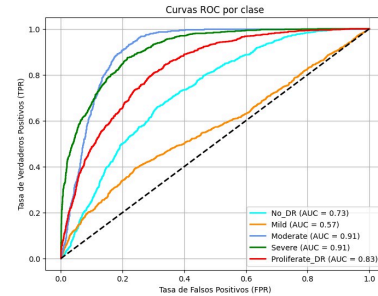


Fig. 14 Curva ROC de clasificación de cinco clases

C. Detección y clasificación de la RD

Con el fin de evaluar el desempeño de los modelos en un entorno clínico real, se realizaron predicciones individuales sobre imágenes de fondo de ojo capturadas con el retinógrafo portátil desarrollado. Estas pruebas evidencian la capacidad del sistema para la detección y clasificación automática de la RD, mostrando las probabilidades asociadas a cada clase y el diagnóstico final generado.

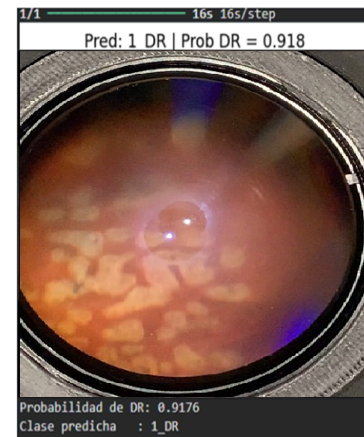


Fig. 15 Predicción de diagnóstico clínico para paciente con RD severo, en el modelo binario clasificado como DR con una probabilidad de 91.80%, analizado por MobilNetV2



Fig. 16 Predicción de diagnóstico clínico para paciente con RD proliferativa, en el modelo de tres clases clasificado como Grave con una probabilidad de 95.01%, analizado por DenseNet121

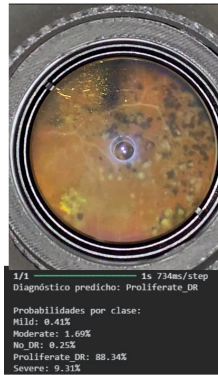


Fig. 17 Predicción de diagnóstico clínico para paciente con RD proliferativa, en el modelo de cinco clases clasificado como Proliferate_DR con una probabilidad de 95.01%, analizado por DenseNet121

D. Discusión

Este estudio abordó el desarrollo de un retinógrafo portátil construido mediante impresión 3D y acoplado a la cámara de un teléfono celular. Con este dispositivo se capturaron imágenes del fondo de ojo de pacientes diabéticos, las cuales fueron evaluadas posteriormente mediante una CNN. Dicho modelo fue el encargado de detectar la presencia de RD y determinar el grado de la enfermedad, clasificando a los pacientes según no presentaran la patología o se encontraran en diferentes estados de severidad.

Los hallazgos principales de este estudio demostraron la viabilidad del retinógrafo portátil desarrollado mediante impresión 3D, el cual, en conjunto con un lente Volk de 20D y un sistema de iluminación LED con control de luminosidad, permitió la captura y visualización efectiva de imágenes del fondo de ojo utilizando un teléfono celular. A pesar de contar con un campo de visión aproximado de 46°, fue posible observar estructuras anatómicas relevantes como la mácula, el disco óptico y la vasculatura retinal, fundamentales para la identificación de signos asociados a la RD.

En cuanto al análisis automatizado, la evaluación de las redes neuronales convolucionales evidenció un comportamiento diferenciado entre los modelos analizados. MobileNetV2 en la clasificación binaria destacó por su alta Se (93,82%) y Sp (96,54%), lo que evidenció una adecuada capacidad para la detección temprana de la enfermedad. Asimismo, alcanzó un F1-score de 79,14%, reflejando un equilibrio aceptable entre precisión y Se en escenarios de tamizaje. Por otro lado, DenseNet121 en la clasificación de tres clases presentó una Sp del 95% y un AUC elevado (97,10%), lo que indicó una mayor capacidad discriminativa global. Sin embargo, su Se del 54,70% reveló limitaciones para detectar todos los casos positivos, especialmente en un contexto clínico orientado a la detección temprana. En la clasificación de cinco clases, el desempeño disminuyó de manera notable, con una precisión de 51,01%, sensibilidad de 51,56% y F1-score de 49,48%, lo que reflejó la mayor complejidad del problema y la influencia del desbalance de datos en tareas multiclase.

Los resultados obtenidos en el presente estudio, resumidos en la Tabla II, evidenciaron que el incremento en el número de clases afectó directamente el rendimiento de los modelos, particularmente en términos de Se y F1-score. Este comportamiento fue consistente con estudios previos, los cuales señalaron que la clasificación multiclase en retinopatía diabética representa un desafío significativo debido a la similitud entre los distintos estadios de la enfermedad, la variabilidad en la calidad de las imágenes y el desbalance de las bases de datos [20,21].

En este contexto, se observó que ambos modelos presentaron características complementarias. MobileNetV2 mostró una elevada Se, lo que favoreció la identificación temprana de casos sospechosos, mientras que DenseNet121 evidenció una mayor capacidad discriminativa y Sp. En consecuencia, los resultados sugirieron que una estrategia combinada podría mejorar el desempeño global del sistema, empleando MobileNetV2 como un primer nivel de tamizaje para la detección inicial de pacientes con posible RD, seguido de DenseNet121 como un segundo nivel de clasificación y confirmación. Este enfoque secuencial permitiría reducir el riesgo de falsos negativos sin incrementar de manera significativa los falsos positivos, optimizando la utilidad clínica del sistema en entornos de atención primaria o con recursos limitados, tal como ha sido propuesto en esquemas de diagnóstico asistido por inteligencia artificial orientados al cribado poblacional [22,23].

Al comparar estos resultados con estudios recientes, se observó que la mayoría de los trabajos reportaron métricas superiores. En particular, los enfoques más recientes alcanzaron valores de exactitud entre 94,0% y 99,4%, junto con F1-scores superiores al 96% y porcentajes de Se cercanas o superiores al 98%, especialmente cuando se emplearon arquitecturas diseñadas específicamente para la detección o

gradación de la enfermedad. Estos niveles de desempeño se asociaron al uso de modelos más complejos y a esquemas de entrenamiento optimizados sobre conjuntos de datos extensos y balanceados [20,21].

Asimismo, los modelos híbridos basados en la fusión de CNN con Transformers y CNN con Vision Mamba reportaron mejoras consistentes en métricas globales, con F1-scores entre 92,4% y 99,4% y Sp que alcanzaron hasta el 99,8% en conjuntos como Messidor y APTOS [22,23]. No obstante, estos enfoques implicaron un incremento considerable en la complejidad computacional, lo que limitó su implementación en sistemas portátiles o de bajo costo, como el planteado en este trabajo.

Por otro lado, estudios que emplearon arquitecturas CNN convencionales con técnicas de preprocesamiento y aumento de datos mostraron una mayor variabilidad en el desempeño. En particular, se reportaron Ses cercanas al 59 % y F1-scores alrededor del 66 % cuando se trabajó con conjuntos de datos multiclase como EyePACS, resultados que fueron comparables a los obtenidos en el presente estudio. Estos hallazgos evidenciaron la influencia del desbalance de clases y la complejidad del problema en el rendimiento final del modelo [24].

Finalmente, trabajos recientes demostraron que la incorporación de mecanismos de atención en arquitecturas CNN tradicionales permitió mejoras significativas en el desempeño. En este sentido, versiones optimizadas de ResNet50 alcanzaron accuracies entre 94,6 % y 96,7 %, con Sp cercanas al 99 % y F1-scores superiores al 96 %, evidenciando el impacto positivo de la atención externa y de modificaciones en las ramas residuales [25].

En conjunto, el desempeño alcanzado por el presente proyecto se encontró en concordancia con el enfoque metodológico adoptado y con las restricciones inherentes a un sistema orientado a la portabilidad, priorizando la viabilidad práctica sin comprometer la capacidad de detección. La alta Se alcanzada por MobileNetV2 reforzó su potencial como herramienta de tamizaje inicial, mientras que la elevada capacidad discriminativa de DenseNet121 estableció una base sólida para futuras mejoras orientadas a optimizar el equilibrio entre Se y Sp.

No obstante, una de las principales limitaciones del presente estudio fue el campo de visión restringido del retinógrafo portátil, el cual, al ser de aproximadamente 46°, fue inferior al de los retinógrafos comerciales convencionales. Esta limitación pudo afectar la detección de lesiones ubicadas en regiones periféricas de la retina, especialmente en estadios avanzados de la RD. Asimismo, el desempeño de los modelos de clasificación estuvo condicionado por el tamaño y la distribución del conjunto de datos, lo que pudo influir en

métricas como la Se y el F1-score, particularmente en escenarios de clasificación multiclase. Adicionalmente, no se incorporaron arquitecturas híbridas ni mecanismos de atención, lo que limitó la comparación directa con modelos más complejos reportados en la literatura reciente; sin embargo, estas limitaciones fueron coherentes con el objetivo del estudio, el cual priorizó la simplicidad del sistema, la portabilidad y el bajo costo computacional.

En este contexto, los resultados obtenidos sugirieron que el sistema desarrollado tuvo potencial como herramienta de tamizaje inicial para la detección de RD, especialmente en entornos con acceso limitado a equipamiento oftalmológico especializado. Como trabajo futuro, se propuso optimizar el diseño óptico del retinógrafo con el fin de ampliar el campo de visión, así como integrar estrategias híbridas o secuenciales que permitan combinar modelos de alta sensibilidad y alta Sp, con el objetivo de mejorar la capacidad diagnóstica sin comprometer la portabilidad del sistema. Asimismo, la validación del dispositivo con un mayor número de pacientes y la utilización de imágenes capturadas directamente con el retinógrafo desarrollado permitirán fortalecer la robustez y aplicabilidad clínica del sistema.

IV. CONCLUSIONES

Se puede apreciar que el prototipo de retinógrafo portátil, fue capaz de capturar imágenes del fondo de ojo con suficiente calidad para la visualización de estructuras anatómicas relevantes, demostrando su viabilidad como dispositivo funcional de bajo costo orientado a aplicaciones portátiles.

La evaluación de los modelos de Deep Learning confirmó que las CNN fueron efectivas para el análisis y clasificación de la RD. En particular, MobileNetV2 presentó una alta Se (93,82 %), adecuada para la detección temprana, mientras que DenseNet121 mostró una elevada Sp (95 %) y AUC (97,10 %), evidenciando una adecuada capacidad de discriminación entre imágenes sanas y patológicas.

En conjunto, el sistema basado en un retinógrafo portátil y modelos de Deep Learning demostró potencial como herramienta de tamizaje inicial para la detección de RD, contribuyendo al diagnóstico temprano de la enfermedad. Estos resultados validan la aplicabilidad del enfoque propuesto y establecen una base sólida para futuras mejoras orientadas a optimizar el desempeño y ampliar su uso en entornos con recursos limitados.

REFERENCIAS

- [1] Ministerio de Salud de Chile, "Guía Clínica: Retinopatía Diabética", 2.^a ed., Santiago de Chile: MINSAL, 2010. [En línea]. Disponible en: https://www.superdesalud.gob.cl/difusion/572/articles-657_guia_clinica.pdf
- [2] International Diabetes Federation, "Diabetes in Peru (2024)". [En línea]. Disponible en: <https://idf.org/our-network/regions-and-members/south-and-central-america/members/peru/>
- [3] S. Paredes, A. Quispe, y J. B. Quispe, "Prevalencia y factores de riesgo de retinopatía diabética en Perú: una revisión sistemática", Trabajo de investigación (bachiller), Universidad Peruana Cayetano Heredia, Lima, Perú, 2022. [En línea]. Disponible en: <https://repositorio.upch.edu.pe/handle/20.500.12866/11930>
- [4] A. Barrenechea-Pulache, A. Portocarrero-Bonifaz, A. Hernández-Vásquez, C. Portocarrero-Ramos y J. Moscoso-Carrasco, "Determinants of Eye Care Service Utilization among Peruvian Adults: Evidence from a Nationwide Household Survey", *Ophthalmic Epidemiology*, vol. 29, no. 3, pp. 339–348, jun. 2022. DOI: 10.1080/09286586.2021.1948577
- [5] L. McKenna, C. Chen, D. Barnett, J. K. Ting, L. T. Leong, T. J. Ting, W. K. Ting, W. Hsu, Y. J. Wong, M. Y. Wong, and T. Y. Wong, "Accuracy of trained rural ophthalmologists versus non-medical image graders in the diagnosis of diabetic retinopathy in rural China," *British Journal of Ophthalmology*, vol. 102, no. 12, pp. 1642–1646, Dec. 2018, doi: 10.1136/bjophthalmol-2017-311687
- [6] J. Burmeister, M. N. Pham, F. Bohler y C. North, "Diabetic retinopathy in rural communities: a review of barriers to access of care and potential solutions," *Annals of Medicine*, vol. 57, no. 1, p. 2522323, Jul. 2025, doi: 10.1080/07853890.2025.2522323.
- [7] G. A. Cummings, K. M. Haskett, R. J. Jaffe, and L. J. Boucher, "Screening for diabetic retinopathy in rural areas: the potential of telemedicine," *The Journal of Rural Health*, vol. 17, no. 3, pp. 328–334, 2001, doi: 10.1111/j.1748-0361.2001.tb00964.x
- [8] O. Salamanca, A. Geary, N. Suárez, S. Benavent, y M. Gonzalez, "Implementation of a diabetic retinopathy referral network, Peru," *Bull. World Health Organ.*, vol. 96, no. 10, pp. 674–681, Aug. 2018, doi: 10.2471/BLT.18.212613
- [9] D. Patiño-Pérez, L. Armijos-Valarezo, L. Chóez-Acosta, and F. Burgos-Robalino, "Redes neuronales convolucionales para detección de retinopatía diabética," *Ingenius: Revista de Ciencia y Tecnología*, no. 33, pp. 91–101, Jan.–Jun. 2025, doi: 10.17163/ings.n33.2025.08.
- [10] E. Bernal, E. de la Cruz, J. Montero and R. Hernandez, "Detección automática de retinopatía diabética aplicando visión artificial y redes neuronales convolucionales," 2021. Instituto Tecnológico de Acapulco, Guerrero, México, 2021, [Online]. Available: https://acapulco.tecnm.mx/wp-content/uploads/maestria/repositorio/memoria_congreso/87_Deteccion_automatica_de_retinopatia_diabetica_aplicando_visin_artificial_y_redes_neuronales_convolucionales.pdf
- [11] J. Cuello and C. Barraza, "Construcción y Evaluación de un modelo de Red Neuronal Convolucional (CNN) Para la Detección y Clasificación de la Retinopatía Diabética," 2021. Universidad Autónoma del Caribe, Colombia. Accessed: Nov. 21, 2025. [Online]. Available: <https://repositorio.uac.edu.co/bitstream/handle/11619/4126/582-BARRAZA-CUELLO.pdf?sequence=1&isAllowed=y>
- [12] D. Gognieva, M. Durzhinskaya, I. Vorobyeva, P. Chomakhidze, A. Suvorov, N. Kuznetsova, A. Bektimirova, B. Al-Dwa, M. Abdullaev, Y. Yusef, V. Pavlov, M. Budzinskaya, D. Sychev, L. Moshetova and P. Kopylov, "Diabetic retinopathy diagnosis based on convolutional neural network in the Russian population: A multicenter prospective study," *Curr. Diabetes Rev.*, vol. 20, no. 8, Oct. 2024, doi: <https://doi.org/10.2174/0115733998268034231101091236>
- [13] R. Srinivasan, J. Surya, P. Ruamviboonsuk, P. Chotcomwongse and R. Raman, "Influence of different types of retinal cameras on the performance of deep learning algorithms in diabetic retinopathy screening," *Life*, vol. 12, no. 10, Art. no. 1610, Oct. 2022, doi: <https://doi.org/10.3390/life12101610>
- [14] M. Reis, C. Kūnas, T. Da Silva, J. Schneiders, P. de Azevedo, L. Nakayama, D. Rados, R. Umpierre, O. Berwanger, D. Lavinsky, F. Malerbi, P. Navaux and B. Schaan, "Advancing healthcare with artificial intelligence: Diagnostic accuracy of machine learning algorithm in diagnosis of diabetic retinopathy in the Brazilian population," *Diabetol. Metab. Syndr.*, vol. 16, no. 1, Art. no. 209, 2024, doi: <https://doi.org/10.1186/s13098-024-01447-0>
- [15] V. Gulshan, L. Peng, M. Coram, M. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. Nelson, J. Mega and D. Webster, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016, doi: <https://doi.org/10.1001/jama.2016.17216>
- [16] P. Romero-Aroca, B. Fontoba-Poveda, E. García-Curto, A. Valls, J. Cristiano, M. Llagostera-Serra, C. Morente-Lorenzo, I. Méndez-Marín y M. Baget-Bernaldiz, "Two handheld retinograph devices evaluated by ophthalmologists and an artificial intelligence algorithm," *Journal of Clinical Medicine*, vol. 13, no. 22, art. 6935, 2024, doi: 10.3390/jcm13226935.
- [17] F. V. R. Cyrino, S. Ferronato, S. A. V. de Araújo, V. Giachetto y L. D. Saud, "Applicability of portable retinal cameras and telemedicine as facilitating tools in screening diabetic retinopathy in the COVID-19 pandemic scenario," *Arquivos Brasileiros de Oftalmologia*, vol. 87, no. 2, e2021-0498, 2022, doi: 10.5935/0004-2749.2021-0498.
- [18] S. Araujo, A. Galeno, J. Silva, C. Aragao, C. Henrique and R. Crescencio, "A new strategy for the detection of diabetic retinopathy using a smartphone app and machine learning methods embedded on cloud computer," in *Proc. 2020 IEEE 33rd Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Rochester, MN, USA, 2020, pp. 542–545, doi: <https://doi.org/10.1109/CBMS49503.2020.00108>
- [19] Chen, PN., Lee, CC., Liang, CM., Pao, SI., Huang, KH. and Lin, KF., "Modelo general de aprendizaje profundo para la detección de la retinopatía diabética," *BMC Bioinformatics* 22 (Supl. 5), 84 (2021). <https://doi.org/10.1186/s12859-021-04005-x>
- [20] M. Sushith, A. Sathiy, V. Kalaipoonguzhali, and V. Sathya, "A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images," *Scientific Reports*, vol. 15, art. 15166,
- [21] S. Akhtar, S. Aftab, O. Ali, M. Ahmad, M. A. Khan, S. Abbas, and T. M. Ghazal, "A deep learning-based model for diabetic retinopathy grading," *Scientific Reports*, vol. 15, art. 3763, 2025, <https://doi.org/10.1038/s41598-025-87171-9>
- [22] K. Rezaee and F. Farnami, "Innovative Approach for Diabetic Retinopathy Severity Classification: An AI-Powered Tool using CNN-Transformer Fusion," *Journal of Biomedical Physics & Engineering*, vol. 15, no. 2, pp. 137–158, 2025, doi: <https://doi.org/10.31661/jbpe.v0i0.2408-1811>
- [23] Z. Liu, A. Gao, H. Sheng, and X. L. Wang, "Identification of diabetic retinopathy lesions in fundus images by integrating CNN and vision mamba models," *PLOS ONE*, vol. 20, no. 1, e0318264, 2025, doi: <https://doi.org/10.1371/journal.pone.0318264>
- [24] C. Suedumrong, S. Phongmoo, T. Akarajaka, and K. Leksakul, "Diabetic Retinopathy Detection Using Convolutional Neural Networks with Background Removal and Data Augmentation," *Applied Sciences*, vol. 14, no. 19, 8823, 2024, doi: <https://doi.org/10.3390/app14198823>
- [25] M. Feng, Y. Cai, and S. Yan, "Enhanced ResNet50 for Diabetic Retinopathy Classification: External Attention and Modified Residual Branch," *Mathematics*, vol. 13, no. 10, 1557, 2025, doi: <https://doi.org/10.3390/math13101557>