Artículo Científico / Scientific Paper

UNIVERSIDAD POLITÉCNICA SALESIANA ECUADOR
INGENIUS
Revista de Ciencia y Tecnología
https://doi.org/10.17163/ings.n33.2025.08
pISSN: 1390-650X / eISSN: 1390-860X

Redes neuronales convolucionales para detección de retinopatía diabetica

Convolutional neural networks for diabetic retinopathy detection

Darwin Patiño-Pérez[1,*], Luis Armijos-Valarezo[1], Luis Chóez-Acosta[1], Freddy Burgos-Robalino[1]

Recibido: 15-05-2024, Recibido tras revisión: 13-11-2024, Aceptado: 29-11-2024, Publicado: 01-01-2025

Resumen

La detección temprana de la retinopatía diabética representa un desafío crítico en el diagnóstico médico, donde el aprendizaje profundo dentro del campo de la inteligencia artificial emerge como una herramienta prometedora para optimizar la identificación de patrones patológicos en imágenes retinales. Este estudio evaluó comparativamente tres arquitecturas de redes neuronales convolucionales ResNet-18, ResNet-50 y una CNN personalizada o no-preentrenada para clasificar imágenes de retinopatía diabética en un conjunto de datos de imágenes agrupadas en cinco categorías, revelando diferencias significativas en su capacidad para aprender y generalizar. Los resultados demostraron que la arquitectura de red neuronal convolucional no-preentrenada superó consistentemente a los modelos preentrenados basados en ResNet-18 y ResNet-50, alcanzando una precisión del 91 % y una notable estabilidad en la clasificación. Mientras ResNet-18 mostró limitaciones severas, degradándose de un 70 % a un 26 % de precisión, y ResNet-50 requirió ajustes para mejorar su rendimiento, la CNN no preentrenada exhibió una capacidad sobresaliente para manejar el desbalance de clases y capturar patrones diagnósticos complejos. El estudio subraya la importancia de diseñar arquitecturas específicamente adaptadas a problemas médicos, destacando que la complejidad no garantiza necesariamente un mejor desempeño, y que un diseño cuidadoso puede superar modelos preentrenados en tareas de diagnóstico por imagen cuando la cantidad de datos con que se cuenta es limitada.

Palabras clave: retinopatía diabética, ceguera, detección, inteligencia artificial, redes neuronales convolucionales, imágenes oculares

Abstract

The early detection of diabetic retinopathy remains a critical challenge in medical diagnostics, with deep learning techniques in artificial intelligence offering promising solutions for identifying pathological patterns in retinal images. This study evaluates and compares the performance of three convolutional neural network (CNN) architectures ResNet-18, ResNet-50, and a custom, non-pretrained CNN using a dataset of retinal images classified into five categories. The findings reveal significant differences in the models' ability to learn and generalize. The non-pretrained CNN consistently outperformed the pretrained ResNet-18 and ResNet-50 models, achieving an accuracy of 91 % and demonstrating notable classification stability. In contrast, ResNet-18 suffered severe performance degradation, with accuracy dropping from 70 % to 26 %, while ResNet-50 required extensive tuning to improve its outcomes. The non-pretrained CNN excelled in handling class imbalances and capturing complex diagnostic patterns, emphasizing the potential of tailored architectures for medical imaging tasks. These results underscore the importance of designing domain-specific architectures, demonstrating that model complexity does not necessarily guarantee better performance. Particularly in scenarios with limited datasets, well-designed custom models can surpass pre-trained architectures in diagnostic imaging applications.

Keywords: diabetic retinopathy, blindness, detection, artificial intelligence, convolutional neural networks, image analysis

[1,*]Universidad de Guayaquil, Guayaquil, Ecuador. Autor para correspondencia ✉: darwin.patinop@ug.edu.ec.

Forma sugerida de citación: Patiño-Pérez, D., Armijos-Valarezo, L., Chóez-Acosta, L. y Burgos-Robalino, F. "Redes neuronales convolucionales para detección de retinopatía diabetica," *Ingenius, Revista de Ciencia y Tecnología*, N.° 33, pp. 91-101, 2025. DOI: https://doi.org/10.17163/ings.n33.2025.08.

1. Introducción

La retina, ubicada en la parte posterior del ojo, es una capa vital de células sensibles a la luz, esencial para la visión. Lamentablemente, es susceptible a diversas enfermedades, entre las cuales la retinopatía diabética (RD) se destaca como una de las condiciones más comunes y graves (Figura 1). La RD es una complicación ocular de la diabetes, caracterizada por el daño a los vasos sanguíneos de la retina [1]. Este daño vascular puede conducir a varios problemas patológicos, incluyendo:

- **Obstrucción del flujo sanguíneo.** Los vasos sanguíneos bloqueados dificultan el suministro adecuado de sangre a la retina, lo que puede provocar la muerte de las células retinianas y la consecuente pérdida de visión.
- **Fugas de sangre.** Los vasos sanguíneos dañados pueden filtrar sangre y otros fluidos hacia la retina, provocando hinchazón y visión borrosa.
- **Crecimiento de vasos sanguíneos anormales.** Como respuesta a la falta de oxígeno, la retina puede desarrollar nuevos vasos sanguíneos anormales que pueden ser frágiles y propensos a sangrar.

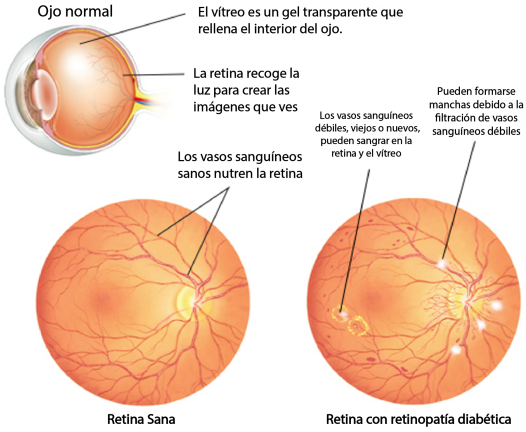

Figura 1. . Retinopatía diabética

La retinopatía diabética (RD) es más prevalente entre las personas con diabetes tipo 1 y tipo 2, especialmente en aquellas que no logran mantener un control adecuado de los niveles de azúcar en sangre [2]. Factores de riesgo adicionales incluyen hipertensión, hipercolesterolemia, tabaquismo, sobrepeso u obesidad y embarazo. En sus etapas iniciales, la RD a menudo se presenta sin síntomas perceptibles.

Sin embargo, a medida que la enfermedad progresa, pueden manifestarse síntomas, incluyendo visión borrosa, manchas oscuras o flotantes, dificultad para ver de noche, visión distorsionada e incluso pérdida de visión. La detección temprana y el tratamiento oportuno son esenciales para prevenir el deterioro permanente de la visión. Por ello, se recomienda encarecidamente realizar exámenes oculares regulares a las personas con diabetes, especialmente a aquellas con factores de riesgo adicionales, para facilitar la intervención y el manejo tempranos.

La informática se centra en el diseño y desarrollo de sistemas y algoritmos capaces de realizar tareas que típicamente requieren inteligencia humana, como el aprendizaje, la percepción, el razonamiento y la resolución de problemas. Esta disciplina constituye la base de la inteligencia artificial (IA) [3]. La IA integra técnicas de informática, estadística, lógica y matemáticas para crear sistemas que puedan aprender de manera autónoma a partir de datos y mejorar su rendimiento en tiempo real.

La inteligencia artificial (IA) ha surgido como una herramienta prometedora para la detección de la retinopatía diabética (RD). Los algoritmos de aprendizaje automático pueden analizar imágenes retinianas e identificar patrones sutiles indicativos de la enfermedad. Esta tecnología tiene un potencial significativo para mejorar la precisión y la eficiencia del diagnóstico de la RD, permitiendo una detección más temprana y facilitando intervenciones oportunas.

La retinopatía diabética (RD) es una complicación grave de la diabetes que puede llevar a la pérdida de visión si no se trata. La detección temprana y la intervención oportuna son fundamentales para prevenir la progresión de la enfermedad y mitigar su impacto. La inteligencia artificial (IA) ha surgido como una herramienta prometedora para mejorar la detección de la RD, ofreciendo capacidades de diagnóstico más precisas y eficientes, y contribuyendo a la preservación de la salud visual en personas con diabetes.

Los modelos predictivos, que proporcionan pronósticos para resultados dicotómicos (resultados distintos, pero complementarios), se utilizan ampliamente en aplicaciones médicas. La Figura 2 ilustra una evaluación de los modelos más relevantes empleados en este campo [4]. El aprendizaje profundo, un campo prominente dentro de la inteligencia artificial, permite a las máquinas o computadoras aprender y analizar datos de manera similar a la inteligencia humana [5]. Este estudio examina el comportamiento de varios modelos basados en aprendizaje profundo, destacando su capacidad para aprovechar múltiples capas de procesamiento y facilitar el aprendizaje a partir de representaciones de datos en diversos niveles de abstracción [6].

Existen numerosas implementaciones preentrenadas de ResNet disponibles en diversos marcos de aprendizaje automático, incluidos TensorFlow, PyTorch, Keras y MXNet. Cada marco ofrece sus propias variantes y optimizaciones específicas, lo que hace que la selección de un modelo preentrenado de ResNet adecuado sea crucial para abordar una tarea particular.

Los factores clave a considerar al seleccionar un modelo incluyen el tamaño y la complejidad del conjunto de datos, la naturaleza de la tarea (por ejemplo, clasificación, detección de objetos o segmentación) y los recursos computacionales disponibles.

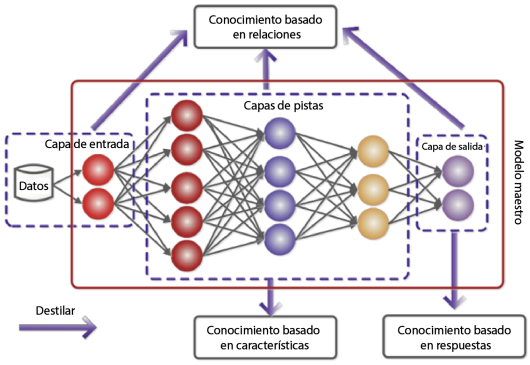

Figura 2. Modelo de aprendizaje profundo

La variedad de arquitecturas de ResNet preentrenadas disponibles es extensa, ofreciendo una gama de opciones adaptadas a diferentes tareas y requisitos. La selección del modelo preentrenado de ResNet adecuado depende de los objetivos específicos del proyecto y las características del problema a abordar. Para este estudio, que se centra en el reconocimiento de enfermedades mediante el análisis de imágenes oculares utilizando aprendizaje supervisado en un marco de clasificación, se eligieron modelos ResNet debido a su rendimiento aceptable demostrado en estudios previos que involucraban otros tipos de imágenes. Esta investigación evalúa el rendimiento de los modelos de redes neuronales artificiales, específicamente las redes neuronales convolucionales (CNN) preentrenadas como ResNet-18 y ResNet-50, como se muestra en la Figura 3.

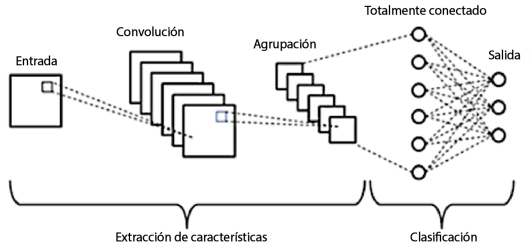

Figura 3. Modelo CNN

2. Materiales y métodos

2.1. Metodología

El aprendizaje profundo (DL), una subdisciplina del aprendizaje automático (ML), representa más que una simple técnica de análisis (Figura 4). Es una metodología integral que abarca toda la cadena de procesamiento de datos, incluyendo la recopilación, preparación, exploración, modelado y evaluación de los datos. Este enfoque permite la identificación de patrones, la generación de predicciones y la toma de decisiones informadas. A diferencia de los métodos estadísticos tradicionales, que dependen de reglas predefinidas y modelos estáticos, el aprendizaje automático emplea algoritmos que aprenden directamente de los datos. Estos algoritmos se adaptan a la complejidad de los datos y evolucionan con el tiempo, mejorando su rendimiento a medida que se exponen a conjuntos de datos más grandes y diversos. Esta adaptabilidad es particularmente evidente en los modelos basados en redes neuronales artificiales, que comprenden numerosas neuronas interconectadas organizadas en capas. Estas redes siguen una estructura jerárquica, como se muestra en la Figura 5.

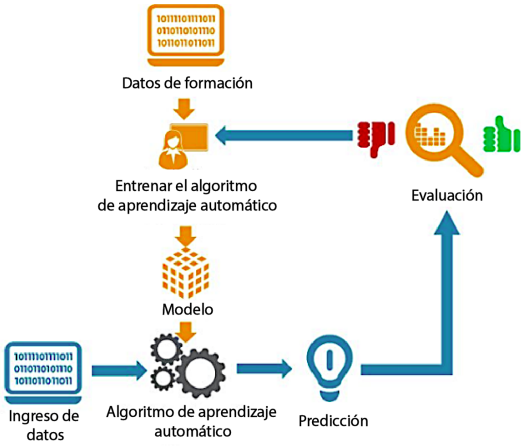

Figura 4. Técnica de análisis – ML

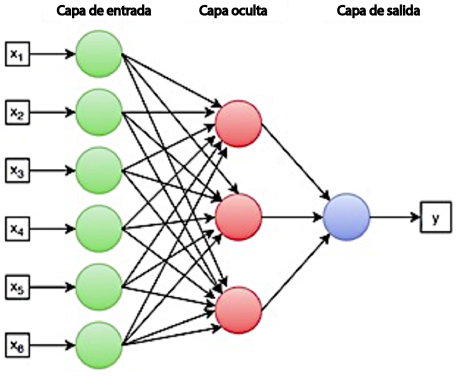

Figura 5. Red neuronal artificial

2.2. El conjunto de datos

El conjunto de datos utilizado en esta investigación consta de 3662 imágenes retinianas obtenidas de la

comunidad en línea Kaggle APTOS 2019 Blindness Detection (BD). Estas imágenes están clasificadas según la gravedad de la retinopatía diabética, categorizadas como sin retinopatía diabética, leve o grave.

La Tabla 1 proporciona una visión general del conjunto de datos, que consta de 3662 imágenes médicas obtenidas de la comunidad en línea de Kaggle.

Tabla 1. Segmentación de imágenes médicas

Clase	Nombre	Número de imagenes	Porcentaje
0	No DR	1805	49.29%
1	Leve	370	10.10%
2	Moderado	999	27.28%
3	Grave	193	5.27%
4	Proliferativa	295	8.06%
Total		3662	100%

2.3. Tratamiento y ajuste de imágenes

Debido a las variaciones en las condiciones de adquisición y el equipo, muchas de las imágenes en el conjunto de datos muestran diferencias en la alineación y calidad de la retina. Para abordar estas inconsistencias y permitir que los modelos aprendan las propiedades de la red de manera más eficiente, se implementó un método de procesamiento de imágenes utilizando la biblioteca OpenCV en Python.

Los pasos de preprocesamiento incluyeron el desenfoque gaussiano y el recorte circular. Se dibujó un contorno alrededor de cada imagen, seguido de la aplicación de un filtro gaussiano. Este proceso reduce los componentes de alta frecuencia, mejorando la claridad de las características clave de cada imagen y su idoneidad para el análisis.

2.4. Descripción de las variables

La Figura 6 ilustra los diferentes niveles de retinopatía diabética (RD), que se categorizan de la siguiente manera:

- **Nivel 0 (Sin RD).** Este nivel indica un estado no patológico, lo que significa la ausencia de retinopatía diabética.
- **Nivel 1 (leve).** Este estadio se caracteriza por una retinopatía diabética leve no patológica, donde están presentes los microaneurismas (manchas rojas). Estos microaneurismas son la fuente de exudados duros, que aparecen como manchas amarillas de alto contraste.
- **Nivel 2 (moderado).** En este estadio, puede ocurrir distorsión e hinchazón de los vasos sanguíneos, lo que podría comprometer su capacidad para transportar sangre de manera efectiva.
- **Nivel 3 (grave).** Este estadio se caracteriza por un bloqueo significativo de los vasos sanguíneos, lo que conduce a un suministro sanguíneo comprometido hacia la retina.
- **Nivel 4 (proliferativo).** Este es el estadio más avanzado, caracterizado por la secreción de factores de crecimiento por parte de la retina, lo que estimula la proliferación de nuevos vasos sanguíneos. Estos vasos anormales crecen dentro de la retina y se extienden hacia el gel vítreo, llenando el ojo.

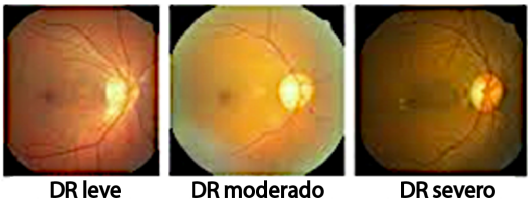

Figura 6. Tipos de retinopatías

Cada estadio de la retinopatía diabética tiene características y propiedades distintas. Sin embargo, durante el análisis, los clínicos pueden pasar por alto ciertos detalles, lo que podría aumentar la probabilidad de un diagnóstico incorrecto.

2.5. Análisis de datos

Inicialmente, las imágenes se descargaron y se subieron a Google Drive. Posteriormente, se organizaron en directorios correspondientes a cada nivel de retinopatía diabética, asegurando una diferenciación precisa. El factor crítico para verificar la corrección de los resultados radica en los datos asociados con cada imagen retinal, los cuales están almacenados de manera segura en la nube.

El conjunto de datos almacenado en la nube se utilizó para permitir que los algoritmos en Google Colab accedieran a la información necesaria para el entrenamiento. La corrección de los resultados se determina al comparar la salida del modelo o algoritmo clasificador con la información disponible en la nube. Si los resultados coinciden, se puede concluir que la clasificación es precisa; de lo contrario, el resultado se considera incorrecto. Una vez completado el entrenamiento de cada algoritmo y obtenidos los resultados, se efectúa un análisis comparativo para evaluar su rendimiento e identificar el algoritmo más eficiente para resolver el problema propuesto.

Los problemas de clasificación en aprendizaje automático se dividen ampliamente en dos categorías principales: problemas binarios y problemas de múltiples clases. La principal distinción radica en la cantidad de clases que el modelo debe identificar dentro de los datos. En el caso de los problemas binarios, el modelo

distingue entre solo dos clases. Estos problemas se caracterizan por su simplicidad, ya que los modelos binarios generalmente son más fáciles de entrenar e interpretar debido al número limitado de clases involucradas. Comprender las decisiones del modelo también es más sencillo, ya que solo existen dos resultados posibles. En contraste, la clasificación de múltiples clases implica distinguir entre más de dos, lo que aumenta la complejidad de la tarea. Estos problemas son más desafiantes de entrenar e interpretar debido al mayor número de clases y las relaciones complejas entre ellas. Las redes neuronales convolucionales (CNN) son especialmente adecuadas para resolver problemas de múltiples clases, pero interpretar las decisiones de dichos modelos puede ser más difícil, dada la mayor cantidad de resultados potenciales.

2.6. Métricas de validación

Matriz de confusión

La matriz de confusión desempeña un papel crucial en la identificación de errores, permitiendo evaluaciones tanto descriptivas como analíticas de los modelos de clasificación. Muestra las diversas asignaciones correctas e incorrectas realizadas por el modelo [7]. Usando los valores proporcionados por la matriz de confusión, se pueden calcular métricas clave de evaluación para evaluar el rendimiento del modelo, como se ilustra en la Figura 7.

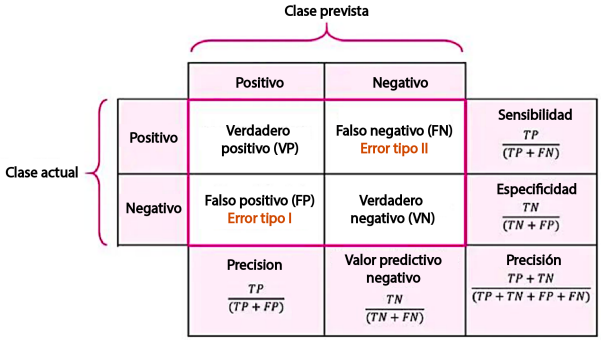

Figura 7. Matriz de confusión

La matriz de confusión es una herramienta esencial para validar redes neuronales, especialmente en tareas de clasificación. Ofrece una visión detallada del rendimiento del modelo al cuantificar el número de predicciones correctas e incorrectas para cada clase.

Entre las métricas derivadas de la matriz de confusión y comúnmente aplicadas a tareas de clasificación en redes neuronales convolucionales (CNN) se encuentran la precisión y la pérdida. Estas métricas de rendimiento son ampliamente utilizadas para evaluar modelos de clasificación de imágenes, tanto en redes neuronales convolucionales preentrenadas (CNN preentrenadas) como en modelos desarrollados con scikit-learn (sklearn). Sin embargo, es crucial comprender sus limitaciones y utilizarlas junto con otras métricas para una evaluación más completa del rendimiento del modelo.

Precisión

Representa la proporción de predicciones correctas realizadas por el modelo, calculada como el número de predicciones correctas dividido entre el número total de predicciones. Es una métrica intuitiva y sencilla de interpretar; un valor alto de precisión indica que el modelo está realizando predicciones generalmente acertadas. La precisión también es una métrica útil para comparar rápida y fácilmente diferentes modelos. Sin embargo, la precisión es sensible a la distribución de las clases. Si una clase domina el conjunto de datos, el modelo puede lograr una alta precisión al predecir predominantemente la clase mayoritaria, incluso si su rendimiento en otras clases es deficiente. Esta limitación hace que la precisión sea menos confiable en presencia de un desbalance de clases.

Pérdida

Representa el error promedio de las predicciones del modelo y se calcula como la suma de los errores individuales de cada predicción. Proporciona información sobre la magnitud del error, donde un valor de pérdida más bajo indica que el modelo está realizando predicciones con errores globales menores. La pérdida desempeña un papel crucial en la optimización del modelo. Durante el entrenamiento, se utiliza para ajustar los pesos de la red neuronal con el fin de minimizar el error y mejorar el rendimiento. La interpretación y la escala de la pérdida dependen de la función de pérdida específica utilizada, ya que diferentes funciones de pérdida pueden tener significados y escalas variables. Sin embargo, la pérdida puede verse influenciada por el desbalance de clases, lo cual debe ser considerado cuidadosamente al evaluar el rendimiento del modelo.

Para las redes neuronales convolucionales (CNN) preentrenadas, la efectividad de las métricas de precisión y pérdida depende de la calidad del modelo preentrenado y su adecuación para la tarea de clasificación específica. La selección cuidadosa de la red preentrenada, junto con el ajuste adecuado de los hiperparámetros, es esencial para optimizar el rendimiento y garantizar una evaluación precisa. En los modelos desarrollados utilizando sklearn, las métricas de precisión y pérdida son directamente aplicables a tareas de clasificación. Sin embargo, es crucial tener en cuenta las características específicas del modelo y del problema de clasificación al seleccionar las métricas y técnicas de evaluación adecuadas.

La efectividad y confiabilidad de las métricas de precisión y pérdida dependen de varios factores, incluyendo la complejidad del problema, la calidad de

los datos, la arquitectura del modelo y las métricas adicionales utilizadas. Es esencial comprender las limitaciones de estas métricas y utilizarlas de manera responsable junto con otros métodos de evaluación para garantizar una evaluación integral y robusta de los modelos de clasificación de imágenes.

2.7. Modelos de aprendizaje profundo utilizados

2.7.1. Modelos preentrenados

Las redes neuronales preentrenadas utilizadas en este estudio se basan en la arquitectura de red residual (ResNet), que aborda el problema de la degradación del gradiente mediante la incorporación de bloques residuales. Un bloque residual sirve como una unidad fundamental de construcción en las ResNets y consta de dos rutas dentro de la red.

1. **Ruta principal.** Incluye las capas convolucionales o totalmente conectadas típicas de una red neuronal profunda.
2. **Ruta directa.** Esta es una conexión que elude las capas en la ruta principal, sumando su salida directamente a la salida de la ruta principal.

Esta estructura de doble ruta permite que la información se propague a través de la red sin ser distorsionada por las transformaciones aplicadas en la ruta principal. En consecuencia, simplifica el proceso de aprendizaje y facilita el entrenamiento de redes neuronales mucho más profundas en comparación con las arquitecturas tradicionales. Para este estudio, se utilizaron dos variantes de ResNet: ResNet-18 y ResNet-50.

ResNet-18

Es una red neuronal convolucional de 18 capas de profundidad, como se ilustra en la Figura 8. Debido a su arquitectura relativamente superficial, ResNet-18 puede retener eficazmente características a baja escala, lo que la hace apropiada para servir como un extractor de características (codificador). La arquitectura de ResNet-18 consta de 16 capas convolucionales, 2 capas de muestreo descendente y varias capas completamente conectadas [8].

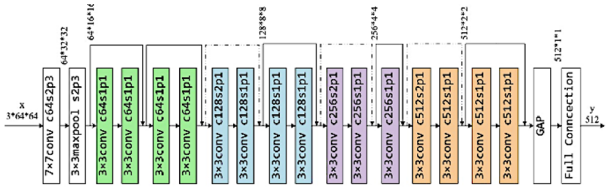

Figura 8. Modelo ResNet-18

ResNet-50

Es una red neuronal artificial convolucional con una profundidad de 50 capas, como se muestra en la Figura 9. Puede utilizar una versión preentrenada entrenada con más de un millón de imágenes de la base de datos ImageNet [9]. La arquitectura de ResNet-50 consta de 48 capas convolucionales, una capa de Max Pooling y una capa de average pooling. Requiere aproximadamente 3.8×109 operaciones de punto flotante.

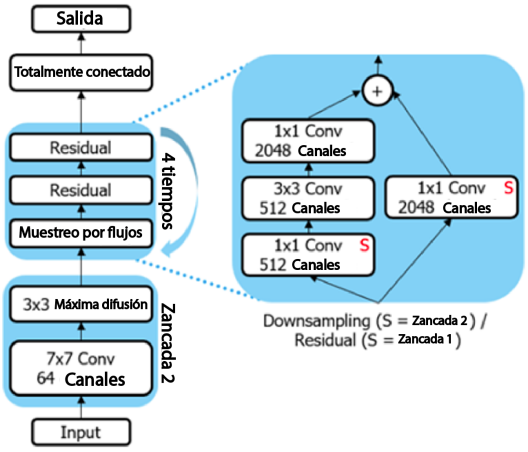

Figura 9. Modelo ResNet-50

2.7.2. Modelos no preentrenados

La arquitectura de la red neuronal convolucional (CNN) no preentrenada utilizada en este estudio consta de tres capas convolucionales 2D con 8, 16 y 32 filtros, respectivamente. Cada filtro tiene un tamaño de 3×3, asegurando que cada operación de convolución procese una región de 3 píxeles × 3 píxeles de la entrada. La red también incluye tres capas de agrupación, tres capas densas con 64, 32 y 3 neuronas, respectivamente, y dos capas de abandono, cada una con una tasa de abandono del 15 %.

Las redes neuronales convolucionales (CNN) son un tipo de algoritmo de inteligencia artificial basado en redes neuronales multilayer. Estas redes están diseñadas para aprender y extraer características relevantes de las imágenes, como se ilustra en la Figura 10. Las CNN son capaces de realizar diversas tareas, incluyendo clasificación de objetos, detección y segmentación [10]. Son un componente fundamental en el campo del aprendizaje profundo [11].

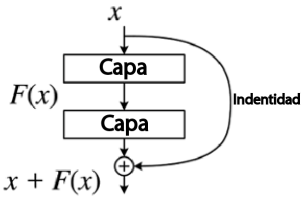

Figura 10. CNN

El modelo de análisis de componentes principales (PCA) también fue utilizado en este estudio. PCA es una técnica estadística altamente efectiva, ampliamente aplicada en campos como el reconocimiento facial y la compresión de imágenes. Se utiliza comúnmente para identificar patrones en datos de alta dimensión [12].

La función de activación ReLU (unidad lineal rectificada) fue empleada en las redes neuronales convolucionales (CNN) utilizadas en este estudio. Su función principal es mejorar las propiedades de activación no lineales de la red sin alterar los campos receptivos de las capas convolucionales [13].

Convoluciones

Una convolución en una imagen es una transformación píxel por píxel lograda mediante la aplicación de una operación específica definida por un conjunto de pesos, comúnmente llamados filtros. La capa convolucional en una red neuronal consiste en una colección de filtros aprendibles. Cada filtro es espacialmente pequeño en términos de ancho y alto, pero se extiende a través de toda la profundidad del volumen de entrada [14].

Submapeo

La capa de agrupación, también conocida como capa de submuestreo, tiene como función reducir progresivamente las dimensiones espaciales de la representación, como se ilustra en la Figura 11. Esta reducción minimiza el número de parámetros y la complejidad computacional dentro de la red [14].

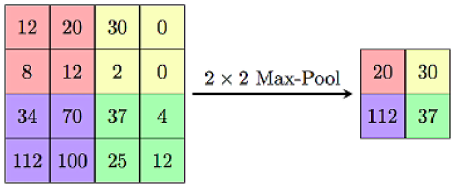

Figura 11. Submuestreo

Capa de agrupación

La capa de agrupación se utiliza para reducir las dimensiones de los mapas de características, con el objetivo principal de disminuir los tiempos de procesamiento mientras se preserva la información más crítica. Esta reducción de la dimensionalidad ayuda a mitigar el sobreajuste en la red e introduce un grado de invariancia ante traslaciones [15].

Retinografías

La retinografía es un procedimiento diagnóstico que captura una imagen en color no invasiva y sin dolor del fondo de ojo [16].

Como funcionan las CNN

Las redes neuronales convolucionales (CNN) operan a través del aprendizaje automático [17] y el aprendizaje supervisado [18], aprovechando varios componentes clave que funcionan de manera integrada. El núcleo de las CNN se encuentra en sus capas convolucionales, que realizan operaciones de convolución para analizar las imágenes de entrada utilizando pequeños filtros (kernels). Estos filtros extraen características relevantes, como bordes, texturas y patrones, mediante multiplicación de matrices. Al deslizarse por la imagen, los filtros generan mapas de características convolucionales [19].

Funciones de activación

Después de la operación de convolución, se aplica una función de activación no lineal, como la Unidad Lineal Rectificada (ReLU). Esto introduce no linealidad en el modelo, permitiéndole capturar y extraer características más complejas.

Capas de agrupamiento

Estas capas se utilizan para reducir la dimensionalidad de los mapas de características mediante la reducción de la información extraída por las capas convolucionales. Esta operación se realiza típicamente utilizando técnicas como la agrupación máxima o el promedio de agrupación, que reducen efectivamente el tamaño de las características mientras retienen su información más relevante.

Capas completamente conectadas

Después de pasar por varias capas convolucionales y de agrupamiento, la información extraída se aplana y se introduce en una o más capas densas (totalmente conectadas). Estas capas realizan operaciones de clasificación o regresión para generar la salida final.

Regularización

Para prevenir el sobreajuste y mejorar las capacidades de generalización del modelo, se emplean técnicas de regularización. Estas incluyen métodos como el abandono, que desactiva aleatoriamente neuronas durante el entrenamiento para reducir la dependencia de características específicas, y la normalización por lotes *(batch normalization)*, que normaliza las activaciones de las capas intermedias.

Función de pérdida y optimización

Durante el entrenamiento, se emplea una función de pérdida para cuantificar la discrepancia entre las predicciones del modelo y las etiquetas reales. Los algoritmos de optimización, como el descenso de gradiente estocástico (SGD) y sus variantes, se utilizan luego para minimizar esta pérdida. Al ajustar iterativamente los pesos de la red neuronal, estos algoritmos mejoran el rendimiento y la precisión predictiva del modelo.

3. Resultados y discusión

ResNet (redes residuales) aborda los problemas de degradación en redes neuronales profundas mediante la introducción de bloques residuales. Las principales diferencias entre los modelos ResNet radican en su profundidad, el tamaño de los bloques residuales, la capacidad de aprendizaje y el costo computacional. El proceso de entrenamiento se realizó en dos fases, incorporando tanto los modelos ResNet preentrenados como la CNN no preentrenada. Esto se llevó a cabo utilizando un conjunto de datos con distribuciones de clases desbalanceadas.

Fase-1

Como se muestra en la Tabla 2, durante el entrenamiento del modelo ResNet-18, se observó que la pérdida en el conjunto de entrenamiento fue del 86 %, mientras que la pérdida de validación (val_pérdida) fue significativamente más alta, alcanzando el 194 %. Esto fue acompañado de una precisión del 60 % y una precisión de validación (val_precisión) del 70 %. Estos resultados indican posibles problemas de calibración, detención temprana o configuraciones incorrectas de entrenamiento causadas por factores como subajuste, regularización excesiva, datos no representativos o problemas de muestreo.

Tabla 2. Fase de Entrenamiento y Validación - Fase 1

	ResNet-18	ResNet-50	CNN
Pérdida de formación (Pérdida)	0.86	1.32	0.19
Precisión de entrenamiento (exactitud)	0.6	0.48	0.92
Pérdida en la validación (val_pérdida)	1.94	1.26	0.22
Precisión de validación (val_precisión)	0.7	0.54	0.91

Para el modelo ResNet-50, la pérdida de entrenamiento fue del 132 % y la pérdida de validación fue del 126 %, con una precisión de entrenamiento del 48 % y una precisión de validación del 54 %. Estas métricas sugieren desafíos relacionados con la capacidad de aprendizaje y generalización del modelo, posiblemente debido a su mayor complejidad y requisitos computacionales. En contraste, el CNN no preentrenado demostró un rendimiento superior, logrando una pérdida de entrenamiento del 19 % y una pérdida de validación del 22 %, con precisiones de entrenamiento y validación del 92 % y 91 %, respectivamente. La alineación de las métricas de pérdida y precisión entre los conjuntos de entrenamiento y validación indica que este modelo está generalizando bien y aprendiendo eficazmente de los datos.

Como se muestra en la Figura 12, un número significativo de muestras fue clasificado en la clase 0 (No DR) con un total de 330 muestras. La clase 1 (leve) contenía 19 muestras, la clase 2 (moderada) incluía 87 muestras, la clase 3 (severa) tenía 20 muestras, y la clase 4 (proliferativa) comprendía 35 muestras.

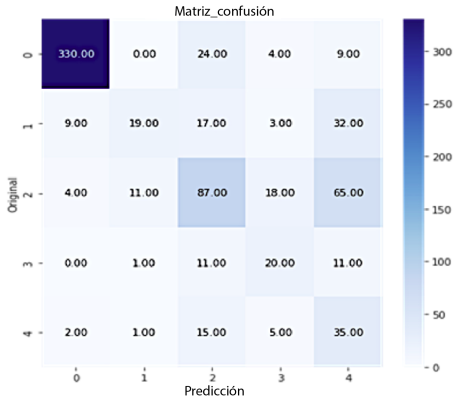

Figura 12. Matriz de Confusión ResNet18

Fase-2

En la Fase-2, se realizaron una serie de ajustes en las configuraciones de los hiperparámetros de los modelos ResNet-18, ResNet-50 y CNN no preentrenado para desarrollar un modelo robusto y consistente. Como se indica en la Tabla 3, los modelos basados en ResNet no mostraron mejoras notables en comparación con los resultados obtenidos en la Fase-1. En contraste, el modelo CNN no preentrenado demostró una mejora significativa en el rendimiento y la precisión, logrando una exactitud del 94 %, una exactitud de validación (val_precisión) del 93 %, una pérdida del 18 % y una pérdida de validación (val_pérdida) del 19 %. Estas métricas indican una generalización efectiva del conocimiento adquirido, con resultados consistentes y confiables. El modelo CNN no preentrenado superó claramente a los modelos basados en ResNet y demostró ser una opción superior y más adecuada para predecir la retinopatía hepática.

Como se muestra en la Figura 13, un gran número de muestras fueron clasificadas en la clase 0 (Sin DR), con un total de 351. La clase 1 (leve) incluyó 8 muestras, la clase 2 (moderada) comprendió 162 muestras,

la clase 3 (severa) tuvo 25 muestras, y la clase 4 (proliferativa) representó 23 muestras.

Tabla 3. Entrenamiento y Validación - Fase 1

	ResNet-18	ResNet-50	CNN
Pérdida de formación (Pérdida)	0.83	0.065	0.18
Precisión de entrenamiento (exactitud)	0.68	0.48	0.94
Pérdida en la validación (val_pérdida)	1.88	0.2	0.19
Precisión de validación (val_precisión)	0.26	0.83	0.93

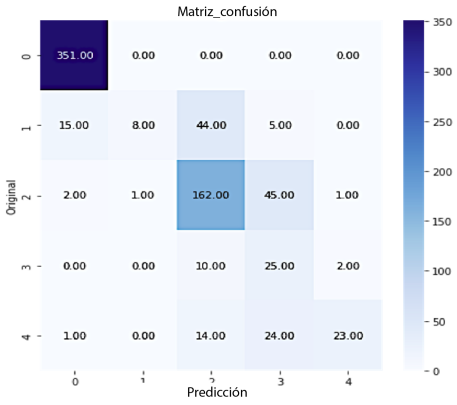

Figura 13. Matriz de Confusión ResNet50

Los resultados obtenidos de los modelos basados en ResNet plantean varios problemas para discusión. La alta pérdida observada tanto en las fases de entrenamiento como de validación puede atribuirse al desbalance de clases en el conjunto de datos. Además, la combinación de valores bajos de precisión y alta pérdida sugiere que los modelos no están aprendiendo de manera efectiva a partir de los datos. Esto podría deberse a la falta de convergencia o configuraciones subóptimas de los hiperparámetros, lo que lleva a un subajuste. Aunque ResNet-50 es inherentemente más potente que ResNet-18 debido a su mayor profundidad y capacidad, puede que no sea adecuadamente adecuado o suficientemente adaptado al problema específico en cuestión.

Los indicadores de pérdida y precisión observados en la Fase-1 y Fase-2 subrayan la efectividad del modelo CNN no preentrenado. La alta precisión tanto en los conjuntos de entrenamiento como de validación sugiere que el modelo captura con éxito los patrones dentro de los datos y generaliza de manera efectiva el conocimiento adquirido. La mínima discrepancia entre la precisión de entrenamiento y la precisión de validación (val_precisión) está dentro de los límites aceptables y puede atribuirse al ruido en los datos o a ligeras variaciones entre los conjuntos de datos de entrenamiento y validación.

El enfoque propuesto para la detección de retinopatía diabética ofrece ventajas significativas a través de su rigurosa evaluación de múltiples arquitecturas de redes neuronales. Este proceso proporciona una comprensión integral de cómo los diversos modelos de inteligencia artificial abordan un problema médico complejo. La metodología es particularmente destacable por su capacidad para resaltar las fortalezas y limitaciones de cada arquitectura, demostrando que una mayor complejidad del modelo no necesariamente se traduce en un rendimiento superior. El CNN no preentrenado emergió como una solución altamente innovadora, logrando una precisión constante superior al 90 %, con sólidas capacidades de generalización y una gestión eficiente del desequilibrio de clases, factores clave en el diagnóstico de enfermedades caracterizadas por presentaciones raras pero potencialmente graves.

A pesar de sus fortalezas, el enfoque propuesto tiene limitaciones notables que deben ser consideradas. La dependencia de una arquitectura específica de red neuronal puede restringir la transferibilidad de la solución a otros contextos médicos, ya que el diseño está altamente adaptado al conjunto de datos utilizado en este estudio. Además, la investigación destacó los desafíos que enfrentaron los modelos preentrenados, como ResNet-18 y ResNet-50, al adaptarse a conjuntos de datos médicos con características complejas e intrincadas. Esto subraya la necesidad de estrategias adicionales, como técnicas avanzadas de remuestreo, funciones de pérdida ponderadas y la ampliación de datos específicos del dominio. Estas complejidades introducen un proceso de desarrollo más laborioso, lo que requiere experiencia especializada tanto en aprendizaje automático como en el dominio médico específico.

4. Conclusiones

Este análisis proporcionó información clave sobre el rendimiento de varios modelos de inteligencia artificial para la detección de retinopatía diabética, destacando una variabilidad significativa entre las arquitecturas de redes neuronales evaluadas [20].

ResNet-18 mostró limitaciones críticas, con una disminución drástica de la precisión, pasando de un 70 % inicial a un 26 % en la fase final, lo que resalta su insuficiencia para manejar la complejidad de la clasificación de imágenes médicas. En contraste, ResNet-50 exhibió una mayor capacidad de aprendizaje, logrando una mejora sustancial y alcanzando una precisión del 83 % en la fase final, lo que subraya la importancia de la afinación y adaptación.

El CNN no preentrenado surgió como la solución más efectiva, manteniendo consistentemente altos niveles de precisión, acercándose al 91 %, a lo largo de ambas fases de entrenamiento y superando significativamente a los modelos preentrenados. Esta arquitec-

tura logró una precisión de entrenamiento del 92 % y una precisión de validación (val_precisión) del 91 % desde el principio. Su estabilidad en los métricos y la baja pérdida de validación (val_pérdida: 0.19 en la Fase 2) demostraron su capacidad para capturar los patrones necesarios para una clasificación precisa de imágenes [21]. Estos resultados destacan que una arquitectura cuidadosamente diseñada y más sencilla puede superar a modelos más complejos en términos de eficiencia y precisión para problemas específicos.

El desequilibrio de clases se identificó como un factor crítico, afectando especialmente el rendimiento de los modelos preentrenados de ResNet. El CNN no preentrenado manejó este desafío de manera notable, lo que sugiere que un diseño arquitectónico reflexivo puede superar las limitaciones estructurales de modelos más complejos. Mientras que el CNN no preentrenado gestionó exitosamente el desequilibrio de clases, ResNet-18 y ResNet-50 enfrentaron dificultades, particularmente durante las primeras fases de entrenamiento. Esto resalta la importancia de implementar estrategias adicionales, como funciones de pérdida ponderadas, aumento de datos o técnicas avanzadas de remuestreo, para mitigar el impacto del desequilibrio y mejorar el rendimiento de modelos más complejos. Asegurar imágenes de retinografía de alta calidad [**?**] también es crucial para evitar inconsistencias durante la fase de entrenamiento.

La investigación futura debe centrarse en estrategias avanzadas para gestionar el desequilibrio de clases en conjuntos de datos médicos, abordando uno de los desafíos más significativos identificados en este estudio. Estos esfuerzos deberían tener como objetivo crear metodologías que aseguren una representación más equilibrada de las diferentes categorías de imágenes, especialmente para las clases minoritarias que son cruciales para el diagnóstico de la retinopatía diabética.

Las estrategias propuestas incluyen el desarrollo de técnicas avanzadas de remuestreo, como SMOTE, el diseño de funciones de pérdida personalizadas que ponderen dinámicamente las clases y la creación de métodos de aumento de datos específicamente diseñados para imágenes médicas. Estos enfoques no solo tienen como objetivo mejorar la precisión del modelo, sino también mejorar su capacidad para detectar casos raros, pero clínicamente significativos, lo que representa un avance sustancial en la aplicación de la inteligencia artificial al diagnóstico médico.

La relevancia de este trabajo radica en su potencial para transformar las capacidades de los sistemas de inteligencia artificial para manejar conjuntos de datos complejos y desbalanceados, especialmente en contextos médicos donde la detección temprana y precisa es crucial para un tratamiento efectivo. Esta dirección ofrece vías prometedoras para mejorar la precisión diagnóstica y abordar desafíos críticos en la imagenología médica.

Referencias

[1] OMS, *TADDS* Instrumento para la evaluación de los sistemas de atención a la diabetes y a la retinopatía diabética.* Organización Mundial de la Salud, 2015. [Online]. Available: https://upsalesiana.ec/ing33ar8r1

[2] L. García Ferrer, M. Ramos López, Y. Molina Santana, M. Chang Hernández, E. Perera Miniet, and K. Galindo Reydmond, "Estrategias en el tratamiento de la retinopatía diabética," *Revista Cubana de Oftalmología*, vol. 31, pp. 90–99, 03 2018. [Online]. Available: https://upsalesiana.ec/ing33ar8r2

[3] F. Tablado. (2020) Inteligencia artificial en el trabajo ¿cómo afecta la ia al ámbito laboral de las empresas? Grupo Artico34. [Online]. Available: https://upsalesiana.ec/ing33ar8r3

[4] E. W. Steyerberg, A. J. Vickers, N. R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M. J. Pencina, and M. W. Kattan, "Assessing the performance of prediction models: a framework for some traditional and novel measures," *PubMed Central*, vol. 21, no. 1, pp. 128–138, 2010. [Online]. Available: https://doi.org/10.1097/EDE.0b013e3181c30fb2

[5] L. Rouhiainen, *Inteligencia Artificial: 101 cosas que debes saber hoy sobre nuestro futuro.* Editorial Alienta, 2018. [Online]. Available: https://upsalesiana.ec/ing33ar8r5

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. [Online]. Available: https://doi.org/10.1038/nature14539

[7] J. M. Sanchez Muñoz, "Análisis de calidad cartográfica mediante el estudio de la matriz de confusión," *Pensamiento Matematico*, vol. 6, no. 2, pp. 9–26, 2016. [Online]. Available: https://upsalesiana.ec/ing33ar8r13

[8] X. Ou, P. Yan, Y. Zhang, B. Tu, G. Zhang, J. Wu, and W. Li, "Moving object detection method via resnet-18 with encoder–decoder structure in complex scenes," *IEEE Access*, vol. 7, pp. 108 152–108 160, 2019. [Online]. Available: https://doi.org/10.1109/ACCESS.2019.2931922

[9] G. Celano, "A resnet-50-based convolutional neural network model for language id identification from speech recordings," *ACL Anthology*, 2021. [Online]. Available: https://upsalesiana.ec/ing33ar8r8

[10] E. Todt and B. A. Krinski, *Convolutional Neural Network–CNN*. Universidade Federal do Paraná, 2019. [Online]. Available: https://upsalesiana.ec/ing33ar8r9

[11] J. Torres, *DEEP LEARNING Introducción práctica con Keras*. LuLu, 2018. [Online]. Available: https://upsalesiana.ec/ing33ar8r10

[12] C. Zhu, C. U. Idemudia, and W. Feng, "Improved logistic regression model for diabetes prediction by integrating pca and k-means techniques," *Informatics in Medicine Unlocked*, vol. 17, p. 100179, 2019. [Online]. Available: https://doi.org/10.1016/j.imu.2019.100179

[13] C. Bonilla Carrion, *Redes Convolucionales*. Trabajo Fin de Grado Inédito. Universidad de Sevilla, 2020. [Online]. Available: https://upsalesiana.ec/ing33ar8r12

[14] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, and S. Ridella, "The 'k' in k–fold cross validationthe 'k' in k-fold cross validation," in *European Symposium on Artificial Neural Networks, Computational Intelligenceand Machine Learning*, 2012. [Online]. Available: https://upsalesiana.ec/ing33ar8r15

[15] L. Herguedas Fenoy, *Guía práctica clínica para la realización de una retinografía*. Universidad de Valladolid, 2018. [Online]. Available: https://upsalesiana.ec/ing33ar8r17

[16] J. Chua, C. X. Y. Lim, T. Y. Wong, and C. Sabanayagam, "Diabetic retinopathy in the asia-pacific," *Asia-Pacific Journal of Ophthalmology*, vol. 7, no. 1, pp. 3–16, 2018. [Online]. Available: https://doi.org/10.22608/APO.2017511

[17] G. Stiglic, P. Kocbek, N. Fijacko, M. Zitnik, K. Verbert, and L. Cilar, "Interpretability of machine learning-based prediction models in healthcare," *WIREs Data Mining and Knowledge Discovery*, vol. 10, no. 5, p. e1379, 2020. [Online]. Available: https://doi.org/10.1002/widm.1379

[18] Z.-H. Zhou, "A brief introduction to weakly supervised learning," *National Science Review*, vol. 5, no. 1, pp. 44–53, Jan 2018. [Online]. Available: https://doi.org/10.1093/nsr/nwx106

[19] J. Wu, *Introduction to Convolutional Neural Networks*. LAMDA Group, 2017. [Online]. Available: https://upsalesiana.ec/ing33ar8r22

[20] R. Gonzalez Gouveia, *Diferencias entre Inteligencia Artificial vs Machine Learning vs Deep Learning*. YouTube, 2021. [Online]. Available: https://upsalesiana.ec/ing33ar8r23

[21] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv*, 03 2015. [Online]. Available: https://upsalesiana.ec/ing33ar8r24