



# Aprendizaje Profundo para la Clasificación de Retinopatía Diabética e Hipertensiva

| Deep Learning for Classification of Diabetic and Hypertensive Retinopathy |

 María Carolina Niño Rivera

maria.nino02@usa.edu.co

Universidad Sergio Arboleda

Colombia

Recibido: propuesta

Aceptado: propuesta

**Resumen:** La detección y diagnóstico de enfermedades crónicas presenta un escenario desafiante, especialmente en sus etapas iniciales. Actualmente, las enfermedades de más rápido crecimiento son la diabetes y la hipertensión. Ambos provocan daños en múltiples órganos, incluidos los ojos, porque son degenerativos y tienen graves consecuencias. Para observar a fondo estas alteraciones en la investigación, fue necesario utilizar fondos de imágenes retinianas. Los conjuntos de datos de libre acceso utilizados fueron de Kaggle. A través de redes neuronales convolucionales, como técnica de visión por computadora, se realizó el preprocesamiento y clasificación de imágenes. Los resultados de las redes convolucionales utilizando ambos conjuntos de datos fueron aproximadamente un 70 % precisos. Después de revisar los diferentes modelos, existe un sesgo hacia la no enfermedad debido a la mayor cantidad de datos seleccionados para esta clase. Además, el uso de la CPU en lugar de la GPU presenta un largo tiempo de ejecución; por lo tanto, no se puede ejecutar el número deseado de épocas, lo que da un porcentaje de precisión más bajo en comparación con el modelo original.

**Palabras Clave:** Retinopatía, Redes Convolucionales.

**Abstract:** The detection and diagnosis of chronic diseases presents a challenging scenario, especially in their early stages. Currently, the fastest-growing diseases are diabetes and hypertension. Both cause damage to multiple organs, including the eyes, because they are degenerative and have serious consequences. To observe these alterations thoroughly in the investigation, it was necessary to use retinal image backgrounds. The free access data sets used were from Kaggle. Through convolutional neural networks, as a computer vision technique, it was performed image preprocessing and classification. The results of the convolutional networks using both data sets were approximately 70 % accurate. After reviewing the different models, there is a bias towards non disease due to the larger amount of data selected for this class. In addition, the use of the CPU instead of the GPU presents a long time for execution; therefore, the desired number of epochs cannot be run, thus giving a lower percentage of accuracy compared to the original model.

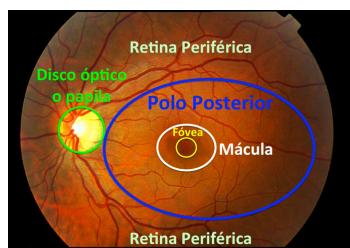
**Keywords:** Retinopathy, Convolutional Networks.

## 1. Introducción

---

Las enfermedades crónicas de acuerdo a la Organización Mundial de la Salud pueden iniciar a temprana edad, pero las manifestaciones clínicas en muchos casos se observan varios años después [1], por esta razón son consideradas como enfermedades silenciosas. En Colombia según el reporte del Ministerio de Salud, la cifra oficial de las estadísticas del 2020 indica que 3 de cada 100 colombianos sufre de diabetes mellitus, aunque también se alerta que la realidad puede ser mucho peor debido a que algunos pacientes aún no saben que la padecen. La cifra real podría llegar a ser una de cada 10 personas [2]. Los número para el caso de la hipertensión, son similares aproximadamente 4 de cada 10 adultos la presenta aunque cerca de un 60 % no lo saben o no han recibido un diagnóstico [3]. Estas dos enfermedades, generan un gran impacto en las funciones de los vasos sanguíneos, debido a esto los pacientes tienen más riesgo de presentar complicaciones en la retina.

Existen cuatro grupos de problemas visuales entre las cuales se incluyen la retinopatía ya sea diabética o hipertensiva [4], en ambos casos, los vasos sanguíneos se ven afectados provocando diferentes grados de clasificación, para el caso de la retinopatía diabética se puede clasificar en cuatro grupos. Clase 0: sin evidencia de ninguna anormalidad, clase 1 etapa más temprana, Clase 2 etapa moderada, Clase 3 etapa grave, Clase 4 etapa proliferativa, es la más grave de todas [5]. De la misma manera la retinopatía hipertensiva se puede clasificar en cuatro grados, grado 0 donde no se tiene evidencia de la enfermedad, Grado 1 o hipertensión benigna, Grado 2 más marcada, Grado 3 angiospástica moderada y Grado 4 maligna [6]. El procedimiento más común para detectarla es a través de la oftalmoscopía directa la cual se realiza por medio del uso de dilatación de la pupila con un midriático, el profesional toma una fotografía del fondo de la retina [7]. Las partes más importantes en estas imágenes de fondos se detallan en la figura 1. Estas partes son muy importantes para realizar el proceso de segmentación, debido a que aportan información sobre el avance de la enfermedad al detectar las anomalías presentadas en los vasos sanguíneos o la formación de hemorragias. Los vasos sanguíneos se pueden conceptualizar anatómicamente como una red intrincada, o una estructura en forma de árbol (o vasculatura), con tubos huecos de diferentes tamaños y composiciones que incluyen arterias, arteriolas, capilares, vénulas y venas [8].



**Figura 1:** Imagen de un Fondo de Retina donde se resaltan las partes más importantes afectadas por la retinopatía

[9]

Para la clasificación de la retinopatía a través de la visión por computador es necesario comenzar con un preprocesamiento de imagen el objetivo es homogenizar los datos. Dentro de las técnicas más usadas para la mejora están la ecualización de iluminación para eliminar las variaciones intra e interimágenes además de la utilización de segmentación mediante técnicas morfológicas [10]. Estas técnicas se utilizan para destacar los vasos sanguíneos y encontrar las características particulares de cada retinopatía. Una vez terminado este proceso se suele utilizar diferentes tipos de redes para hacer predicciones y clasificaciones. Una de estas es la red ConvNet las neuronas no se encuentran implementadas de forma lineal o vectorial, sino en una estructura bidimensional a manera de matriz. Debido a ello, la imagen de entrada también se considera una retícula de neuronas, donde cada neurona corresponde a un pixel de la imagen [11].

En la actualidad existen diferentes estrategias para mejorar el desempeño de los modelos, este es

el caso de la transferencia de aprendizaje. El aprendizaje automático tradicional utiliza métodos de entrenamiento y prueba de datos con una distribución de datos y una entrada de datos similares. A raíz de la diferencia generada en la distribución de datos entre el conjunto de entrenamiento y el de prueba, el resultado o la predicción pueden deteriorarse o mejorarse [12]. En algunos casos, adquirir datos de entrenamiento que se ajusten al conjunto de características de entrada de los datos de prueba, así como el resultado anticipado de las características de los datos de distribución, puede ser bastante desafiante y muy costoso. TL se enfoca en amplios dominios, tareas y patrones en conjuntos de datos de entrenamiento y prueba [12].

El objetivo principal al realizar esta investigación es aplicar los métodos de visión por computador desde el preprocesamiento de las imágenes hasta el uso de red ConvNet para el análisis y clasificación de la retinopatía diabética e hipertensiva, aportando de manera positiva mediante la difusión de las enfermedades que hoy por hoy se siguen considerando silenciosas como son la retinopatía diabética e hipertensiva.

## 2. Metodología

---

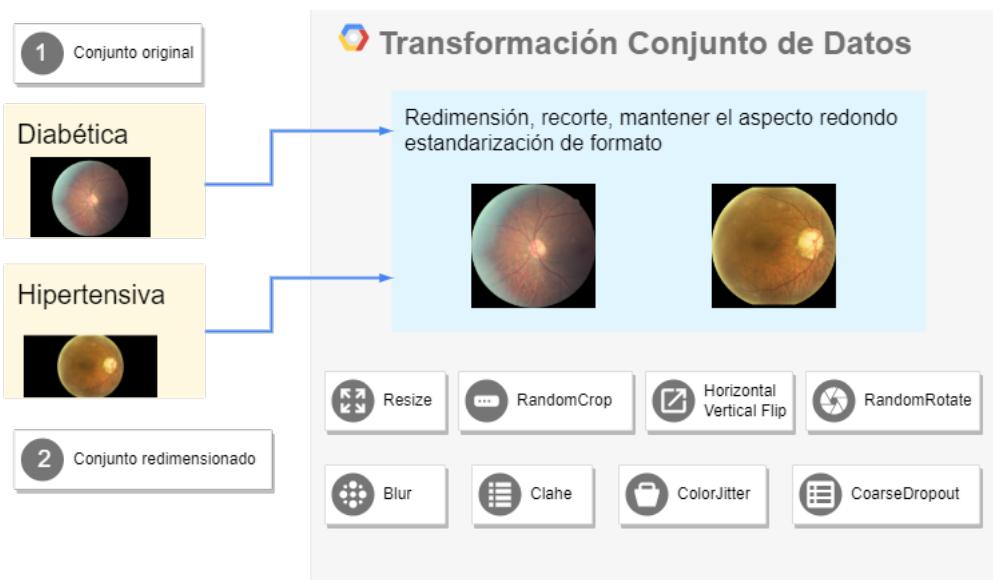
La metodología propuesta para la detección y clasificación de la retinopatía diabética e hipertensiva está dividida en cuatro etapas principales. La primera consiste en una adquisición de los dataset, la segunda un preprocesamiento de estas imágenes, la tercera parte en la selección de las diferentes herramientas de aprendizaje profundo y la cuarta en la validación de los resultados, explicadas a continuación.

### 2.1. Conjunto de Datos

Para esta investigación se utilizó la fuente de datos de Kaggle [13], en la parte de la retinopatía diabética, en este se encuentra 35.126 imágenes digitales de fondo de retina para realizar el entrenamiento y 53.576 para realizar las pruebas, todas estas en formato "JPEG". Esta fuente es la que tiene el conjunto de datos más grande disponible de manera gratuita. También contiene un archivo en formato CSV que contenía las etiquetas con la información del ojo para determinar si es el derecho o izquierdo y la clasificación de la enfermedad de 0 a 4. Para el estudio de la retinopatía hipertensiva se utilizó el dataset de Kaggle [14], es un conjunto de datos público el cual contenía dos carpetas la primera con un total de 2.674 imágenes de entrenamiento y otra con 190 imágenes para prueba, dentro de cada una de estas carpetas se encontraban divididos los datos entre fondos con presencia de la enfermedad y en otra carpeta los fondos sin ninguna enfermedad.

#### 2.1.1. Preprocesamiento

Debido a que existen pocos estándares para la captura de las fotografías de retina, lo primero que se realizó fue un pre-procesamiento, el cual consistió en redimensionar las imágenes para poder comprobar cuál generaba mejores resultados. Ambos conjunto de datos fueron redimensionados en grupos de 150, 250, 350, 450, 550, 650 y 1000 pixeles. Estos cambios fueron realizados mediante la elaboración de un algoritmo que utilizaba las librerías de Pillow y CV2. Algunas de las imágenes contaban con un marco grueso negro. Para evitar errores se creó una función que recorría el arreglo de la imagen para identificar y minimizar estos bordes. Además de esto, para generar más imágenes, se utilizaron varias técnicas las cuales consisten en generar recortes aleatorios, giros verticales y horizontales, utilizar el efecto de desenfoque, eliminación de rectángulos aleatorios mediante la utilización de CoarseDropout, en especial para prevenir un sobreajuste en el modelo [15]. Este aumento de imágenes se adicionó como una capa más al modelo establecido. En la figura 2 se observa una diagrama de los



**Figura 2:** Diagrama transformación conjunto de datos

pasos principales utilizados en la transformación del conjunto de datos En el proceso realizado en este proyecto debido a la cantidad tan desigual de imágenes de los conjuntos de datos, se utilizó la distribución mostrada en la tabla 1:

**Tabla 1:** Distribución conjunto de datos utilizados en el proyecto

| Conjunto     | Tipo          | Cantidad |
|--------------|---------------|----------|
| Diabética    | Entrenamiento | 1251     |
| Hipertensión | Entrenamiento | 2674     |
|              |               | 3925     |
| Diabética    | Prueba        | 280      |
| Hipertensión | Prueba        | 190      |
|              |               | 470      |

La diferencia presentada en los conjuntos de datos se da debido a que se utilizaron las imágenes sin ningún nivel de la enfermedad para que de esta manera el conjunto estuviera equilibrado entre presentar y no la enfermedad. En la tabla 2 se presenta la distribución de los diferentes tipos de enfermedad para los conjuntos de entrenamiento y validación

**Tabla 2:** Distribución conjunto de datos utilizados en entrenamiento

| Nivel                        | Entrenamiento | Validación |
|------------------------------|---------------|------------|
| Sin Enfermedad (0)           | 1300          | 325        |
| Retinopatía Hipertensiva (1) | 839           | 210        |
| Retinopatía Diabética (2)    | 1041          | 210        |
| Total                        | 3180          | 740        |

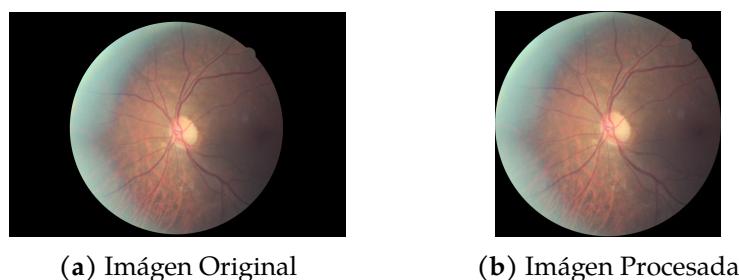
Además se realizó un archivo unificado en csv que contenía los datos de ambos conjuntos etiquetando las imágenes sin enfermedad como nivel 0, con retinopatía hipertensiva como 1 y con un 2 para las imágenes de retinopatía diabética.

## 2.2. Red Convolucional

La clasificación de la retinopatía se realizó utilizando el modelo de aprendizaje automático introducido por Aladdin Persson [16], es una red neuronal convolucional (ConvNet) en un bucle, el cual permite realizar una clasificación de los 3 tipos de niveles propuestos, 0 Normal, 1 Hipertensiva y 2 Diabética. Estas redes se diseñaron para aprender de forma automática y adaptativa, utilizando múltiples bloques de construcción como capas convolucionales, capas de agrupación y capas totalmente conectadas [17]. Los algoritmos de aprendizaje profundo tienen la capacidad de aprender funciones útiles para la tarea de clasificación automáticamente. Las redes neuronales convolucionales (CNN) son variantes de estos algoritmos de aprendizaje [17]. Las capas ocultas generalmente consisten en capas convolucionales, capas ReLU, capas de agrupación y capas completamente conectadas [17]. La red original solo estaba planteada para la clasificación de la retinopatía diabética. La propuesta trabajada hace uso de la arquitectura EfficientNet, la cual utiliza el aprendizaje por transferencia para ahorrar tiempo y poder de cómputo. Por un lado, los Modelos EfficientNet se basan en métodos de escalado compuestos simples y altamente efectivos. Este método permite escalar una ConvNet de línea de base a cualquier restricción de recursos de destino mientras se mantiene la eficiencia del modelo, que se utiliza para transferir conjuntos de datos de aprendizaje [18]. Para el proyecto se utilizó el modelo preentrenado de ImageNet efficientnet-b3. En la arquitectura planteada se comienza con una capa "Linear", o capa lineal, se denomina capa totalmente conectada o capa densa en Keras. Esta capa ayuda a cambiar la dimensionalidad de la salida de la capa anterior para que el modelo pueda definir fácilmente la relación entre los valores de los datos en los que trabaja el modelo [19]. Las capas lineales utilizan la multiplicación de matrices para transformar sus entidades de entrada en entidades de salida mediante una matriz de peso. Las características de entrada son recibidas por una capa lineal, se pasan en forma de un tensor unidimensional aplanado y luego se multiplican por la matriz de peso [19]. Para el modelo se escogió la función de activación Adam. El optimizador se utiliza para disminuir las tasas de error durante el entrenamiento de las redes neuronales se usa como un optimizador de reemplazo para el gradiente en descenso [20]. Al finalizar esta etapa de entrenamiento se generaron el archivo CSV con la clasificación de los datos de prueba en los tres niveles definidos.

## 3. Resultados

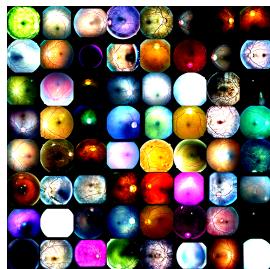
Luego de utilizar la función creada por medio de las librerías de Python OpenCV y Pillow, se puede evidenciar como las imágenes fueron cambiadas para lograr una mejor precisión en el algoritmo en el caso de la Figura 3a se eliminó el borde extra para que esté más centrado y homogéneo. Además, también se cambiaron a diferentes tamaños encontrando mejor porcentaje de precisión en las que se cambiaron a 650 píxeles.



**Figura 3:** Resultado de la transformación para estandarización de formatos y recorte de bordes

La capa de aumento inicial que permitió la generación de diferentes imágenes para ayudar a entrenar el modelo, e incrementar el número de imágenes y de esta forma evitar sobreajustes se puede apreciar en la Figura 4, la misma imagen fue redimensionada a 150 por 150 pixeles, se realizó recortes

aleatorios, giros verticales y horizontales, se añadió diferentes contrastes, y diferentes normalizaciones, forzando también al modelo a no memorizarse los datos. Sin embargo, se evidencia que algunas imágenes son muy diferentes lo que puede aumentar los falsos positivos.



**Figura 4:** Resultado del conjunto de datos generado

Los resultados obtenidos en el proyecto son de aproximadamente 70 % de precisión, se pudo observar que al dejar menos datos de la clase 1 se presentaban mayores errores en esta clase y al utilizar un porcentaje ligeramente mayor de imágenes sin enfermedad también determinaba en una identificación errónea de las demás clases.

## 4. Conclusiones

---

El trabajo realizado propuso la generación de una red convolucional para la detección y clasificación de la retinopatía diabética e hipertensiva. Destacando el papel que cumple los procesos de la visión por computador para mejorar, adaptar y extraer las características de una imagen. Además, se describe el papel que juega el aprendizaje profundo a través de la red propuesta. Se desea hacer visible tanto los avances y aplicaciones de la tecnología, como crear conciencia del crecimiento de estas enfermedades y su impacto en Colombia debido al enorme crecimiento de pacientes con enfermedades crónicas. Debido a que no se pudo ejecutar en un ambiente GPU ya que al realizarlo se agotaba la memoria del equipo, los resultados no logran la precisión deseada las especificaciones técnicas del equipo utilizado son: Sistema operativo Windows 10, procesador Intel i5 7200U CPU con capacidad de 4 CPUs, memoria 16 GB NVIDIA GeForce 940MX. Además, se pudo evidenciar las ventajas de encontrar los pesos del modelo por medio de la transferencia de aprendizaje.

## 5. Bibliografía

---

- [1] Organización Mundial de la Salud *Enfermedades no transmisibles* <https://www.who.int/es/news-room/fact-sheets/detail/noncommunicable-diseases>. Recuperado el 20 de Marzo de 2023.
- [2] Ministerio de Salud *Tres de cada 100 colombianos tienen diabetes,*» Ministerio de Salud y Protección Social <https://www.minsalud.gov.co/Paginas/Tres-de-cada-100-colombianos-tienen-diabetes.aspx>. Recuperado el 20 de Marzo de 2023.
- [3] Ministerio de Salud *Conoce tus números para prevenir la hipertensión arterial,*» Ministerio de Salud y Protección Social <https://www.minsalud.gov.co/Paginas/Conoce-tus-numeros-para-prevenir-la-hipertension-arterial.aspx#:~:text>. Recuperado el 20 de Marzo de 2023.

- [4] A. Casares, N. Rojas, L. Santos y J. Olivera, Artists. "Tecnológicos en Salud Desde el área de la atención Prehospitalaria" **Universidad CES en convenio con la Fundación Universitaria de Ciencias de la Salud FUCS, 2021**
- [5] J. Gallart, Artist. "Diabetic Retinopathy Classification and Interpretation using Deep Learning Techniques" **Universitat Rovira I Virgili, 2019**
- [6] A. Trisán Anoro, H. Menéndez Montes, A. Trisán Alonso, D. Soto Mesa y H. Suárez Cásodo. "Retinopatía Hipertensiva," **Revista Medicina General, 2000**
- [7] G. S y J. R. "Detección de patologías derivadas de las afecciones diabéticas: una revisión del análisis digital de imágenes de retina," **Ingeniería y Desarrollo, vol. 31, pp. 316-338, 2013**
- [8] Zhao, Yitian and Rada, Lavdie and Chen, Ke and Harding, Simon P. and Zheng, Yalin. "Automated Vessel Segmentation Using Infinite Perimeter Active Contour Model with Hybrid Region Information with Application to Retinal Images" **IEEE Transactions on Medical Imaging, vol. 34, num. 9, pp. 1797-1807, 2015**
- [9] Sergio Bonafonte. "Fondo de ojo - Parte II - Polo posterior, mácula y fóvea - Centro de Oftalmología Bonafonte" <https://www.youtube.com/watch?v=FdJeog93ppk&list=LL&index=4>
- [10] D. Ms.D.Durga y M. Mr.K.B. A Survey on Automatic detection of hypertensive and diabetic retinopathy in Fundus Images *International Journal for Research & Development in Technology, vol. 5, n 6.*
- [11] A. Arista-Jalife, G. Calderón-Auza, A. Fierro-Radilla y M. NakanoClasificación de Imágenes Urbanas Aéreas: Comparación entre Descriptores de Bajo Nivel y Aprendizaje Profundo *Información Tecnológica, vol. 28, pp. 209-224, 2017*
- [12] Hosna, A., Merry, E., Gyalmo, J. et al Transfer learning: a friendly introduction *Big Data* 9, 102 (2022) <https://doi.org/10.1186/s40537-022-00652-w>
- [13] Diabetic Retinopathy Detection <https://www.kaggle.com/competitions/diabetic-retinopathy-detection>
- [14] Ocular\_disease\_data <https://www.kaggle.com/datasets/meryemdouider/ocular-disease-data>
- [15] SIIM-ISIC Melanoma Classification <https://www.kaggle.com/c/siim-isic-melanoma-classification/discussion/16972>
- [16] A. Persson, [aladdinpersson/Machine-Learning-Collection](https://github.com/aladdinpersson/Machine-Learning-Collection) <https://github.com/aladdinpersson/Machine-Learning-Collection> [Último acceso: 20 06 2022].
- [17] M. Ech-Chouyyekh, H. Omara y M. LAZAAR, Scientific paper classification using Convolutional Neural Networks *Research Gate, pp. 1-6, 2019*
- [18] G. Marques, D. Agarwal y I. de la Torre Díez, Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network *Applied Soft Computing, vol. 96, p. 106691, 2020*
- [19] Knowledge Transfer, PyTorch Linear Layer (Fully Connected Layer) Explained. <https://androidkt.com/pytorch-linear-layer-fully-connected-layer-explained/> [Último acceso: 24 08 2022]
- [20] B. Kumar, Adam optimizer PyTorch with Examples. <https://pythonguides.com/adam-optimizer-pytorch/> [Último acceso: 24 08 2022]