

# Detección automática de retinopatía diabética aplicando visión artificial y redes neuronales convolucionales.

Ing. Eduardo Bernal Catalán  
División de Estudios de Posgrado  
e Investigación.  
TecNM/Instituto Tecnológico de  
Acapulco.  
Guerrero, México.  
ebcatalan18@gmail.com

Dr. Eduardo de la Cruz Gámez  
División de Estudios de Posgrado  
e Investigación.  
TecNM/Instituto Tecnológico de  
Acapulco.  
Guerrero, México.  
gamezeduardo@yahoo.com

Dr. José Antonio Montero Valverde  
División de Estudios de Posgrado  
e Investigación.  
TecNM/Instituto Tecnológico de  
Acapulco.  
Guerrero, México.  
jamontero1@infinitummail.com

M.T.I. Rafael Hernández Reyna  
División de Estudios de Posgrado  
e Investigación.  
TecNM/Instituto Tecnológico de  
Acapulco.  
Guerrero, México.  
rherman7@yahoo.com.mx

**Resumen.** – En este artículo se propone una metodología para el desarrollo de un sistema de visión artificial que mediante redes neuronales convolucionales pueda diagnosticar retinopatía diabética a través del análisis de imágenes digitales de la retina, aplicando una metodología de estandarización para las imágenes recopiladas en diversas bases de datos públicas con diferentes características y ser usadas en un algoritmo de entrenamiento usando Python como lenguaje de programación y utilizando librerías para Deep Learning como Tensorflow y Keras. Para el entrenamiento y validación se ocuparon 500 imágenes seleccionadas de forma aleatoria y clasificadas manualmente con las etiquetas Positivo y Negativo a las cuales se le aplicó una red neuronal convolucional con dos capas convolucionales para la detección de patrones y predecir la etiqueta de imágenes nuevas que no forman parte del entrenamiento obteniendo un grado de certeza del 91%. Durante las pruebas y obtención de resultados se seleccionaron 400 imágenes que no formaron parte del entrenamiento, como nueva información a clasificar divididas en 200 positivas con 174 diagnósticos correctos y 26 errores con un promedio de certeza del 83.23%, por otro lado, las pruebas sobre 200 imágenes negativas el resultado fue de 187 diagnósticos correctos y 13 erróneos con un promedio de certeza de 78.86%.

**Palabras Clave**— Retinopatía Diabética, Retinografía, Detección Automática, Visión Artificial, Redes Neuronales Convolucionales,

## I. Introducción

La medicina es una de las ramas de interés que más pueden beneficiarse con la implementación de tecnologías que automaticen o faciliten alguna tarea específica en alguna rama de la medicina, las altas velocidades de respuesta de una computadora pueden analizar y determinar el resultado de algún examen visual para el humano que puede ser interpretado como información digital por la computadora, en el caso de la visión

artificial se enseña y entrena a la computadora como interpretar datos de medios digitales como lo es una imagen y así poder analizar una radiografía para interpretar la información en algún diagnóstico que a la percepción del ojo humano pueda pasar desapercibida.

Una de las complicaciones más común entre las personas con diabetes suele ser la pérdida de la vista que va de manera progresiva de pequeños puntos hasta la ceguera total de manera irreversible, este padecimiento se le conoce como retinopatía diabética que consiste en la aparición de lesiones en el árbol vascular del ojo como hemorragias, microaneurismas y exudados.

La retinopatía diabética es la complicación crónica que aparece frecuentemente en pacientes con diabetes mellitus, y es la causa más común de ceguera entre la población mundial en edad productiva [1]. El estudio más extenso realizado en la república mexicana durante el año 2005 encontró una prevalencia del 31.5% [2] y 8 años después un nuevo cribado en 2013 demuestra que, de cada 100 pacientes, 48 padecen retinopatía diabética [3].



Figura 1. (a) Retina sana (b) Retina con signos de retinopatía diabética (c) Acercamiento a zona con lesión

Para la detección de estas lesiones en imágenes digitales previamente recopiladas se realiza la etapa de adquisición de imágenes que es la primera en el proceso de desarrollo de un sistema de visión artificial seguido de la etapa de procesamiento en la que se procesan las imágenes de manera

que puedan ser analizadas de mejor manera en las etapas siguientes y facilite el encontrar los objetos de interés para la creación del sistema

La implementación de un sistema de esta índole es de beneficio para instituciones que brindan servicios médicos especializados en la rama de la oftalmología ya que permitiría un diagnóstico oportuno a pacientes con diabetes mellitus que corran el riesgo de perder el sentido de la vista en su totalidad y poder dar acceso a este estudio a la población que tenga inquietud por saber si padece o no retinopatía diabética antes de notar algún síntoma.

## II. Adquisición de imágenes

Para la obtención de imágenes digitales del árbol vascular de la retina se hizo una recopilación de diferentes bases de datos públicas que contenían retinografías con diferentes lesiones de retinopatía diabética y retinografías de retinas sanas.

Las imágenes venían en diferentes formatos dependiendo de la base de datos de origen, en los que se encontraron archivos TIFF, PNG y JPG. De igual manera existían diferentes tamaños de imágenes.

Tabla 1. Características de las BDP.

NOMBRE	CANTIDAD	FORMATO	TAMAÑO
MESSIDOR	1100	TIFF	2240x1488
IDRID	597	JPG	4288x2848
DRIM	126	JPG	760x570
APTOS	5590	PNG	1588x1958
DIARETD1	89	PNG	1500x1152
DIARETB0	130	PNG	1500x1153
<b>total</b>	<b>7632</b>		

En la tabla 1 se muestra la lista de las diversas bases de datos públicas recopiladas para la creación de una sola base de datos que contenga imágenes digitales de la retina con las mismas características para la clasificación de lesiones en la retina. Como resultado se obtuvo una base de datos de imágenes con 7632 retinografías.

## III. Preprocesamiento

El preprocesamiento es la etapa en la que se pretende reparar la imagen obtenida en la adquisición que puede contener desperfectos producidos o no eliminados por el hardware de captura. Toda imagen que se obtiene por medios ópticos, electroópticos o electrónicos lleva en cierta medida los defectos de degradación que se presentan en forma de ruido, pérdida de definición y fidelidad de imagen. La degradación es provocada por el ruido de los sensores de captura, imprecisiones de enfoque de la cámara, movimiento o perturbaciones aleatorias. En un proceso de visión artificial se hace uso de algunos algoritmos con la intención de mejorar ciertas características de la imagen como mejorar el contraste, brillo, niveles de grises, eliminar brillos, resaltar bordes de formas, mejorar texturas, etc. A estas operaciones se les denomina operaciones de mejora de imagen. [4]

Durante esta etapa de procesamiento se realizó la manera de hacer las imágenes lo más limpias posibles para su proceso de clasificación resaltando detalles de la misma imagen y eliminando partes de ellas que no son necesarias para la búsqueda de información por lo que se requiere de una metodología de preprocesamiento

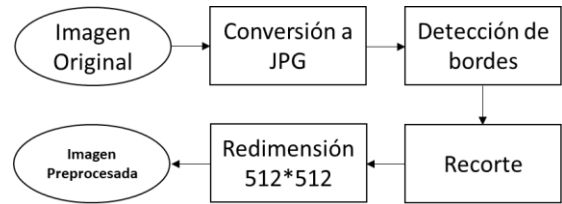


Figura 2. Metodología de preprocesamiento

La metodología sugiere realizar la conversión de formato de imágenes a JPG, debido a que las diferentes bases de datos están desarrolladas en distintas extensiones, definimos el formato final como JPG como estándar para todas las imágenes recolectadas continuando con la eliminación de áreas de no interés para el estudio. Para identificar las áreas a eliminar se optó por detectar primero los bordes de las retinografías ubicando la circunferencia de la escena.

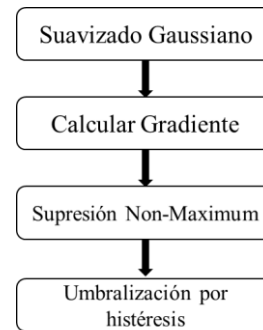


Figura 3. Método de detección de bordes de CANNY

El algoritmo de Canny es usado para la detección y ubicación de bordes que consiste en la conversión a escala de grises de la imagen, aplicación de un suavizado gaussiano para la eliminación de ruido. El filtro Gaussiano, es un filtro con características de filtro pasa-bajo cuyos coeficientes están determinados por los valores de una función de distribución Gaussiana [5].

$$f(x) = ae^{-\frac{(x - \mu)^2}{2\sigma^2}} \quad (1)$$

Este tipo de filtro se utiliza fundamentalmente para el suavizado de imágenes con presencia de ruido, donde  $\alpha$  es igual a las constantes ( $\alpha > 0$ ),  $\mu$  es igual al valor de la media y  $\sigma^2$  equivale a la varianza del filtrado.

Como siguiente paso se realiza la supresión *Non-Maximum* con el fin de eliminar los pixeles que no se consideran parte del borde. El objetivo de la supresión no máxima será encontrar las ubicaciones reales de los objetos y suprimir los demás. Esto

hace que la supresión no máxima sea una etapa muy importante en el reconocimiento de objetos [6]. En este caso se eliminan los píxeles que no son considerados parte del borde para utilizar esta máscara como indicación de la ubicación y límite del borde de la imagen original y por último se realiza la umbralización por histéresis para establecer el valor máximo y mínimo de un píxel que se considera el inicio y fin del borde.



Figura 4. (a)Retinografía original (b)Detección de bordes (c)Imagen recortada

Después de realizado el corte se redimensiona el tamaño de las imágenes debido a las variaciones entre las bases de datos a un valor estándar para todas las imágenes, el tamaño elegido para la resolución de imagen fue 512x512 píxeles de tal manera que las imágenes grandes y pesadas se convirtieran en imágenes ligeras que mantuvieran los detalles de sus características.

#### IV. Clasificación

La clasificación es el proceso de obtener un modelo o función que sea capaz de distinguir los conceptos o tipos de clases de datos. Los modelos son el producto de un análisis de un conjunto de datos para los que se les asigna una etiqueta de clase, el modelo es usado para predecir la etiqueta de clase de los objetos para los que no se tiene conocimiento de la etiqueta de clase que le corresponde. [7]

Una vez preprocesadas las imágenes se hizo una selección manual de cada una dividida en dos grupos “Positivo” en el que se acumulaban las retinografías que presentaban signos de lesiones propias de la retinopatía diabética y en otro llamado “Negativo” se colocaban las retinografías que presentaban retinas sanas.

Para la clasificación y detección se realizó la construcción de una red neuronal convolucional utilizando la biblioteca de software de aprendizaje automático de uso libre Tensorflow, para la detección computarizada basada en imágenes digitales del fondo del ojo.

##### A. Clasificación de etiquetas

Para la clasificación de etiquetas manualmente se seleccionaron imágenes en dos carpetas nombradas como Positivo y Negativo, en la carpeta Positivo se almacenaron las imágenes de la retina que mostraban algún tipo de lesión de retinopatía diabética como lo son microaneurismas, hemorragias y exudados, en la carpeta Negativo se colocaron las imágenes de la retina que no mostraban ningún tipo de lesión, solo retinas sanas para ser analizadas por el algoritmo de entrenamiento.

##### B. Construcción de Red Neuronal Convolucional

Las redes neuronales convolucionales consisten en múltiples capas de filtros convolucionales, uno sobre otro de manera que

va seccionando una imagen en varias capas realizado por un mapeo casual no lineal. Son utilizadas e implementadas en procesos de clasificación en una fase de extracción de características.

Para la construcción de la red neuronal convolucional con las carpetas previamente etiquetadas, se crearon dos carpetas entrenamiento y validación. En la etapa de entrenamiento se seleccionaron 500 imágenes aleatorias de fondo de ojo de la recopilación de bases de datos públicas que fueron etiquetadas en positivo y negativo y al final del entrenamiento almacenar un modelo de aprendizaje los resultados obtenidos por la computadora, consiguiendo en la etapa de validación la red valida por si sola si su entrenamiento fue satisfactorio o no después de entrenar.

```

epocas=20
longitud, altura = 512,512
batch_size = 32
pasos = 1000
pasos_validacion = 300
filtrosConv1 = 32
filtrosConv2 = 64
tamano_filtro1 = (3, 3)
tamano_filtro2 = (2, 2)
tamano_pool = (2, 2)
clases = 2
lr = 0.0005
  
```

En las líneas de código del algoritmo implementado en Python con librerías Tensorflow [8] y keras se muestran los parámetros utilizados para el algoritmo de entrenamiento

##### Epocas

Las épocas son la cantidad de interacciones de entrenamiento con las imágenes clasificadas, en este caso se realizaron 20 épocas realizando entrenamiento y validación.

##### Longitud y Altura

En estas variables se almacena el tamaño del kernel de las imágenes de entrada, el tamaño original de las imágenes es de 512x512.

##### Batch\_size

Es la cantidad de muestras que alimenta la red convolucional para su capa de entrada, para el ingreso de una cantidad de imágenes específica para su procesamiento. En este caso se alimenta de 32 imágenes en cada paso en la capa de entrada de neuronas.

##### Pasos

Los pasos son el número de veces que se va a procesar la información en cada una de las épocas, cada época se le asignaron 1000.

##### Pasos\_validacion

Estos son los pasos que correrán al final el entrenamiento para validar el conocimiento obtenido después de analizar las imágenes ingresadas en la capa de entrada. Se asigna un número menor a los pasos de entrenamiento para su validación rápida.

### filtrosConv1

En esta variable se asigna la cantidad de kernel que se aplican a la primera convolución, en este caso se asignan 32 filtros.

### filtroConv2

En esta línea se asigna la cantidad de filtros del kernel que genera la segunda convolución.

### tamano\_filtro1

En esta variable se asigna el tamaño del filtro de convolución al kernel ocupado en la primera capa de convolución, asignando un filtro de 3 x 3.

### tamano\_filtro2

En el segundo kernel se asigna un filtro de 2 x 2, de menor tamaño que en primer filtro.

### tamano\_pool

En esta línea se almacena el tamaño del filtro utilizado en Max Pooling. El Max Pooling tiene como objetivo reducir la muestra de una representación de entrada reduciendo su dimensionalidad y permitiendo suposiciones sobre las características ubicadas en las subregiones agrupadas. Además, disminuye el costo computacional al reducir el número de parámetros a aprender [9]. Se establece un Max Pooling de 2x2.

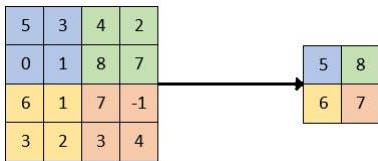


Figura 5. Ejemplo de Max Pooling

### Clases

Se asigna el número de clases establecidas en las carpetas de entrenamiento y validación, en este caso es una clasificación binaria, quiere decir que solo arroja dos resultados posibles.

### Lr

El enfoque LR es capaz de predecir, con solo unas pocas iteraciones de entrenamiento, un rango de tasas de aprendizaje que sería óptimo para una combinación de modelo o conjunto de datos. Lo hace variando la tasa de aprendizaje en las iteraciones de entrenamiento y observando la pérdida para cada tasa de aprendizaje. La forma de la gráfica de pérdida frente a tasas de aprendizaje proporciona pistas sobre el rango óptimo de tasas de aprendizaje, al igual que los gráficos de acciones proporcionan pistas sobre los precios futuros de las acciones.

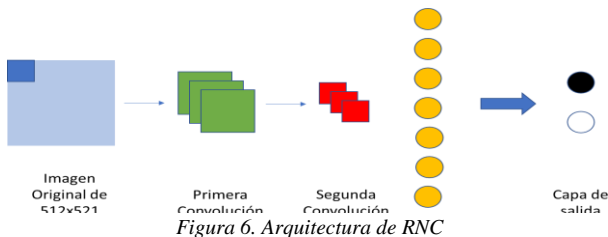


Figura 6. Arquitectura de RNC

### C. Pruebas

Para comprobar la certeza del entrenamiento se realizaron pruebas con 400 imágenes aleatorias que no formaron parte del entrenamiento ni validación, y registrar los casos de éxito de la predicción, así como el grado de certeza de cada una de las pruebas.

	Aciertos	Errores	Promedio de certeza
<b>Con RD</b>	174	26	<b>83.23%</b>
<b>Sin RD</b>	187	13	<b>78.86%</b>

Tabla 2. Pruebas y resultados

### V. Conclusiones

El preprocesamiento fue realizado con el fin de crear un modelo de características estándar para las imágenes ocupadas durante el entrenamiento, de manera que fuera más preciso y rápido analizando solo partes de la imagen que pudieran contener información. El resultado del entrenamiento arroja un grado de certeza de 0.91, se plantea continuar con el desarrollo utilizando una cantidad mayor de imágenes para el aprendizaje del algoritmo de Deep Learning, igual se considera el uso de alguna otra técnica de aprendizaje automático para comparar y validar resultados.

Esta clasificación se realizó de manera binaria, manteniendo solo dos resultados posibles sobre la presencia de la retinopatía diabética y se propone aplicar el entrenamiento sobre imágenes procesadas y segmentadas que logren identificar cada una de las diferentes lesiones que pueda contener la retinografía y hacer énfasis en las características de las imágenes para un mejor entrenamiento y resultados.

### Referencias

- [1] P. M. T. Y. W. Ning Cheung, «Diabetic retinopathy,» *The Lancet*, 10 Julio 2010.
- [2] S. M. d. O. A. P. d. O. Asociación Mexicana de Retina, «Resultados del Día Panamericano de detección de,» *Revista Mexicana de Oftalmología*, vol. 79, n° 2, pp. 88-92, 2005.
- [3] V. L. G. Yoloxochilth Ávila Alcaraz, «Retinopatía que amenaza la visión en pacientes diabéticos no referidos al oftalmólogo,» *Gaceta Médica de México*, pp. 624-629, 2013.
- [4] E. B. Catalán, «Propuesta de un Sistema de Apoyo al Diagnóstico Oportuno de Retinopatía Diabética,» *Congreso Internacional Academia Journals Puebla 2019*, vol. 2, pp. 203-207, 2019.
- [5] A. K. R. Tinku Acharya, *Image Processing: Principles and Applications*, Wiley, 2005.
- [6] M. D. Buil, *NON-MAXIMA SUPPRESSION*, Graz, 2011.
- [7] M. K. J. P. Jiawei Han, *Data Mining: Concepts and Techniques*, Elsevier, 2012, p. 18.
- [8] «Tensorflow,» [En línea]. Available: <https://www.tensorflow.org/>.
- [9] J. Ricco, «Quora,» 15 Junio 2017. [En línea]. Available: <https://www.quora.com/What-is-max-pooling-in-convolutional-neural-networks>.