



**Implementación y evaluación de un sistema de detección mediante la  
captura de imágenes para la clasificación de estrabismo, utilizando redes  
neuronales convolucionales**

**Jaime Alejandro Córdoba Tamayo**  
20441428158

**Juan José Jáuregui Yustes**  
20441613634

**Universidad Antonio Nariño**

Programa Ingeniería Electrónica

Facultad de Ingeniería Mecánica, Electrónica y Biomédica

Neiva, Colombia

2021



**Implementación y evaluación de un sistema de detección mediante la  
captura de imágenes para la clasificación de estrabismo, utilizando redes  
neuronales convolucionales**

**Jaime Alejandro Córdoba Tamayo**

**Juan José Jáuregui Yustes**

Proyecto de grado presentado como requisito parcial para optar al título de:

**Ingeniero Electrónico**

Director:

PhD. Ing. José Fernando Barrera Campo

Línea de Investigación:

Inteligencia Computacional

**Universidad Antonio Nariño**

Programa Ingeniería Electrónica

Facultad de Ingeniería Mecánica, Electrónica y Biomédica

Neiva, Colombia

2021

## NOTA DE ACEPTACIÓN

El trabajo de grado titulado  
Implementación y evaluación de un sistema de detección mediante la captura  
de imágenes para la clasificación de estrabismo, utilizando redes neuronales  
convolucionales  
Cumple con los requisitos para optar  
Al título de Ingeniero Electrónico.



---

Firma del Tutor



---

Firma Jurado



---

Firma Jurado

# Contenido

	<u>Pág.</u>
<b>Resumen</b> .....	<b>7</b>
<b>Abstract</b> .....	<b>8</b>
<b>Introducción</b> .....	<b>9</b>
Planteamiento del problema .....	13
Objetivos .....	17
Objetivo general .....	17
Objetivos específicos .....	17
Justificación .....	18
<b>Capítulo 1</b> .....	<b>19</b>
<b>Marco referencial</b> .....	<b>19</b>
1.1 Marco conceptual .....	19
1.1.1 Fisiología del ojo humano .....	19
1.1.2 Estrabismo .....	23
1.1.3 Ambliopía .....	25
1.1.4 Redes Neuronales Artificiales .....	26
1.1.5 Redes Neuronales Convolucionales .....	27
1.1.6 Modelo De Capas y Clasificación .....	29
1.2 Estado de arte .....	30
<b>Capítulo 2</b> .....	<b>36</b>
2.1 Materiales y métodos .....	36
2.1.1 Pacientes .....	36
2.1.2 Protocolo de adquisición .....	36
2.1.3 Adquisición de las imágenes .....	37
2.1.4 Metodología propuesta .....	39
2.1.5 Segmentación facial .....	40
2.1.6 Detección de la región ocular .....	41

2.1.7 Entrenamiento de la red neuronal convolucional .....	42
2.1.8 Configuración de entrenamiento .....	49
2.1.9 Desarrollo de la aplicación de diagnóstico del estrabismo.....	50
<b>Capítulo 3.....</b>	<b>53</b>
3.1 Métricas de evaluación .....	53
3.2 Resultados.....	54
<b>Capítulo 4.....</b>	<b>58</b>
4.1 Conclusiones.....	58
<b>Anexos .....</b>	<b>60</b>
4.2 Anexo 1.....	60
4.3 Anexo 2.....	61
<b>Referencias bibliograficas .....</b>	<b>62</b>



## Lista de Figuras

	<u>Pág.</u>
Figura 1 - 1 Exterior del ojo.....	20
Figura 1 - 2 Anatomía interna del ojo.....	21
Figura 1 - 3 Niño con esotropía congénita (Estrabismo).....	23
Figura 1 - 4 Estructura de modelo de capas de clasificación.....	30
Figura 2 - 1 Imágenes de las tres posiciones de miradas analizadas a) frontal (PF), b) posición izquierda (PLI), y c) posición derecha (PLD)......	38
Figura 2 - 2 Diagrama de bloques del método propuesto.....	39
Figura 2 - 3 Estructura generalizada de la red neuronal convolucional implementada.....	43
Figura 2 - 4 Algunos ejemplos de regiones oculares segmentadas con etiquetas normales.....	44
Figura 2 - 5 Algunos ejemplos de regiones oculares segmentadas con etiquetas estrabismo.....	44
Figura 2 - 6 Estructura de modelo Densenet.....	46
Figura 2 - 7 Mapa Secuencial del algoritmo de entrenamiento.....	48
Figura 2 - 8 Módulo de detección de estrabismo desarrollado en Tkinter de Python.....	51
Figura 2 - 9 Mapa de secuencia del algoritmo de implementación.....	51

Figura 3 - 1 Grafica de resultado de exactitud. ....	54
Figura 3 - 2 Graficas de perdida.....	54
Figura 3 - 3 Ventana principal de interfaz grafica .....	56
Figura 3 - 4 Selección de tipo de imagen en ventana principal de interfaz grafica	56
Figura 3 - 5 Diagnostico positivo de estrabismo.....	57
Figura 3 - 6 Diagnostico negativo de estrabismo.....	57
Figura 3 - 7 Código para configuración y entrenamiento de la red neuronal. ....	60
Figura 3 - 8 Código de procesamiento de la red neuronal en Tkinter. ....	61

## **Lista de tablas**

	<b><u>Pág.</u></b>
Tabla 2 - 1 Características del modelo Densenet 201 .....	46
Tabla 3 - 1 Resultados de métricas de evaluación del sistema. ....	55

## **Lista de ecuaciones**

Ecuación 2- 1 Función objetivo .....	45
--------------------------------------	----

*(Dedicatoria)*

*Como primera instancia quiero darle las gracias a Dios todo poderoso por haberme permitido llegar al lugar que me encuentro actualmente, a mis padres José Jair Jauregui y Olga lucia Yustes, que me brindaron todo su apoyo, amor y comprensión para enfrentar las dificultades que se presentaron, a mis hermanos por su ejemplo y apoyo incondicional, gracias a ellos le puse todo mi empeño en este trabajo.*

*Juan Jose Jauregui Yustes*

*En primer lugar, agradecer a Dios y a la Santísima Virgen María, por darme la fortaleza necesaria durante mi ciclo estudiantil, total agradecimiento a mis padres Jaime Humberto Córdoba y Mariela Tamayo por brindarme su comprensión y apoyo en los momentos más difíciles, igualmente a los docentes que hicieron posible el aprendizaje a cada uno de ellos los llevo en el corazón.*

*Jaime Alejandro Córdoba Tamayo*

## **Agradecimientos**

Como primero dar gracias a Dios y a la Santísima virgen María por su intersección durante el desarrollo del trabajo, a mi padre José Jair Jáuregui por sus fuerzas, amor y dedicación por hacerme una persona correcta formada desde los principios, a mi madre Olga Lucia Yustes por su apoyo incondicional, paciencia, amor, a mi compañero Jaime Alejandro Córdoba Tamayo por su entrega, dedicación, amor y dedicación al proyecto.

Juan José Jáuregui Yustes

Agradezco eternamente a mi Dios todo poderoso y a la Santísima Virgen María por permitir avanzar en mi proyecto de vida, a mi padre Jaime Humberto Córdoba, por servir de ejemplo y dedicación, a mi madre Mariela Tamayo, por apoyo y comprensión, los amo mucho; a mi compañero Juan José Jáuregui Yustes, por su compromiso, dedicación y trabajo en equipo, igualmente agradecer a mis docentes por impartir su conocimiento, entrega y compromiso, a cada uno de ellos mi más sincero respeto y admiración.

Jaime Alejandro Córdoba Tamayo



## **Resumen**

El diagnóstico de estrabismo, es muy importante realizarlo a tiempo durante la infancia, el estrabismo afecta entre el 2% y 4% de la población mundial infantil debido a que esta condición produce ambliopía, que consiste en la pérdida de la visión del ojo desviado, la ambliopía una vez desarrollada no se puede tratar, debido a que el cerebro inhibe la señal proveniente del ojo desviado, dando como resultado la pérdida gradual y permanente de la agudeza visual en el ojo afectado, para este estudio se utilizaron redes neuronales convolucionales, con el fin de detectar el estrabismo en imágenes de pacientes, el modelo que se utilizó es el DenseNet 201, una arquitectura diseñada para labores de clasificación de imágenes, entrenada por un conjunto de imágenes propio, adquirida por los autores que consiste en 332 imágenes. Las imágenes provienen de personas voluntarias entre hombres y mujeres con estrabismo y sin esta condición, cuyas edades oscilan entre los 8 y 42 años, la red una vez entrenada es implementada utilizando el paquete Tkinter, donde se realiza una aplicación para el diagnóstico de estrabismo, obteniéndose una sensibilidad del 0.9619, una especificidad del 0.9329 y una precisión del 0.9388, con 1500 ciclos de entrenamiento, siendo esto resultados muy favorables.

***Palabras clave: estrabismo, red neuronal convolucional, detección.***

## **Abstract**

The diagnosis of strabismus, is very important to do in time during childhood, strabismus affects between 2% and 4% of the world population in children because this condition produces amblyopia, which consists of the loss of vision in the deviated eye, once developed the amblyopia, it cannot be treated, because the brain inhibits the signal from the deviated eye, resulting in the gradual and permanent loss of visual acuity in the affected eye. Convolutional neural networks were used for this study, In order to detect strabismus in patient images, the model used is DenseNet 201, an architecture designed for image classification tasks, trained by a set of own images, acquired by the authors, consisting of 332 images. The images come from volunteers between men and women with strabismus and without this condition, whose ages range is between 8 and 42 years, the network once trained is implemented using the Tkinter package, where an application is made for the diagnosis of strabismus, obtaining a sensitivity of 0.9619, a specificity of 0.9329 and a precision of 0.9388, with 1500 training cycles, these being very favorable results.

***Key words: strabismus, convolutional neural network, detection.***

## Introducción

El cerebro humano y los diferentes sistemas que él regula, dentro de ellos el sistema ocular, son cada día motivo de diversas investigaciones para aumentar la calidad de vida del ser humano. Se ha demostrado que, a través del análisis de las diferentes características fisiológicas del sistema visual, se pueden diagnosticar enfermedades que afectan al sistema nervioso central. En la última década, el desarrollo de tecnologías y herramientas para el rastreo ocular ha sido un área en constante investigación (Leggett, 2010).

Una de las patologías más comunes y que comprometen la desviación de los ojos es el estrabismo, que consiste en la deficiencia ocular, para mantener la alineación del punto de visión correcto, junto con la alteración de los movimientos oculares y en algunos casos posición anómala de la cabeza. Esta condición afecta el desarrollo psicosocial de las personas en especial, los infantes. Debido que pueden ocasionar baja autoestima y trastorno hiperactivo con déficit atencional, siendo esto una cifra de entre el 2% y 4% de la población infantil a nivel mundial (Mohney, Greenberg, & Diehl, 2007), los especialistas, buscan corregir esta patología a temprana edad para evitar la disminución de la agudeza visual. condición característica de la ambliopía que es la pérdida gradual de la vista producida por la inhibición de la señal proveniente del ojo afectado (ojo desviado), esta patología es degenerativa al transcurrir el tiempo, dicha condición es muy importante ser detectada en los primeros años de la infancia entre el nacimiento, hasta los siete años de edad. La importancia de detectar esta condición a temprana edad, es que hace posible su corrección sin que ocurran daños permanentes en la visión del paciente, por consiguiente se ve la necesidad de

diagnosticar y elegir el tratamiento adecuado para cada caso que se presente, con el fin de facilitar la detección y la causa de estrabismo en el paciente, y con ello realizar un tratamiento que logre tener éxito en la corrección del estrabismo (Qanat et al., 2020). Una de las herramientas computacionales más adecuadas para la detección de esta condición son las redes neuronales convolucionales.

Una red neuronal convolucional (CNN) es un tipo de red neuronal artificial que, mediante capas, imita la corteza visual para identificar características en las entradas, haciendo identificación de objetos, mediante un entrenamiento que puede ser tanto supervisado como no supervisado, puede aprender características de interés en las imágenes. En otras palabras, es un modelo de aprendizaje profundo automático que hace referencia a las redes neuronales artificiales (ANN) con múltiples capas, estos tipos de red son consideradas como una de las herramientas más adecuadas a la hora de manejar un gran volumen de datos. Las redes neuronales convolucionales se constituyen por múltiples capas de filtros, con el fin de que por cada capa se estipule una función específica para realizar un mapeo no lineal.

Al momento de emplear una red de clasificación, se debe de contener una fase de extracción de características, conformadas por neuronas convolucionales, luego, se procede a realizar una reducción, con el fin de obtener neuronas tipo perceptrón para poder realizar la clasificación final. Es decir que al momento en que la CNN realiza la convolución, consiste en tomar grupos de píxeles de la imagen de entrada para posteriormente ser operadas mediante producto escalar contra una matriz denominada kernel, la cual ira recorriendo la imagen de izquierda a derecha y de arriba abajo, para generar una nueva

matriz de salida, que dará como resultado una nueva capa de neuronas ocultas o feature mapping para poder aplicar una función de activación ReLu (rectifier linear unit) esta función de activación convierte los datos negativos provenientes de las imágenes en cero, debido a que estos no son importantes en este proceso. Todo esto permite generar los datos provenientes de las imágenes de forma linealmente rectificadas, ya que con esto lo que se busca es mantener una pila de datos sin valores negativos.

Las redes neuronales convolucionales, poseen múltiples aplicaciones enfocadas en la clasificación de imágenes como, por ejemplo, el diagnóstico automatizado de retinopatía en bebés prematuros mediante redes neuronales convolucionales profundas, se obtuvo como resultado, que el modelo propuesto diagnosticó correctamente 91 de 100 imágenes en su validación, lo que equivale a un 91% de exactitud al realizar la clasificación, mientras que los ocho expertos obtuvieron una exactitud media del 82% clasificando correctamente un rango de 77 a 94 imágenes de forma correcta (Brown et al., 2018).

El auge de las CNN inicio en el año 1989, con reconocimiento de caracteres numéricos manuscritos, como el reconocimiento de dígitos árabes utilizando una red neuronal de convolución profunda (Mahdi Haref, Srayyih Al-Maliki, Tuaama, & Albehadili, 2020) desde entonces, se ha ido incrementando sus aplicaciones en diferentes áreas como: la ingeniería informática (Su et al., 2020), ingeniería biomédica (Hu et al., 2021), biotecnología (Savojardo, Bruciaferri, Tartari, Martelli, & Casadio, 2020), multimedia (Sokolova, Kharchevnikova, & Savchenko, 2018), etc.

En particular, desde el 2012 su uso se ha extendido al área de la salud donde se han empleado CNN para: análisis de sangre, para la detección de enfermedades, clasificación de Alzheimer mediante datos de resonancia magnética funcional (Sarraf & Tofighi, 2016), clasificación automatizada de tuberculosis pulmonar (Lakhani & Sundaram, 2017), detección de carcinoma ductal invasivo (Cruz-Roa et al., 2014), detección de retinopatía diabética (Tamkin, Usiri, & Fufa, 2017) y diagnóstico rápido de enfermedades respiratorias pediátricas (El-Bakry & Hamada, 2013).

Los trabajos mencionados con la generación de registros médicos sintéticos por medio de CNN son pocos, sin embargo, esta aplicación es muy prometedora para este tipo de red, pues representa una alternativa para la creación de un sistema de clasificación ocular (Smaida & Yaroshchak, 2020).

En el área de la salud, en la optometría, con el apoyo de la ingeniería, se han desarrollado múltiples proyectos utilizando inteligencia artificial, para apoyar al personal médico en el diagnóstico de enfermedades, por medio de algoritmos de aprendizaje automático (Lavric & Valentin, 2019), (Grewal, Oloumi, Rubin, & Tennant, 2018), (Thakoor, Li, Tsamis, Sajda, & Hood, 2019), (Huang, He, Fang, Rabbani, & Chen, 2019). Sin embargo, tales trabajos reportan entre sus limitantes, como primero escasez de imágenes anotadas, junto con la exactitud del algoritmo, y por ende no se subestima el porcentaje de subjetividad humana al momento de realizar los exámenes.

Los registros generados por algoritmos de aprendizaje automático, pueden emplearse con fines investigativos, especialmente en proyectos que involucran el desarrollo de

algoritmos de inteligencia artificial, como apoyo al diagnóstico médico, sin embargo, la utilidad que le puedan dar a esta, es dependiente de los parámetros requeridos por el investigador o la persona que haga uso de estos, debido a las necesidades particulares que se presente, así mismo será su porcentaje de eficiencia.

## **Planteamiento del problema**

En la actualidad han sido propuestos una gran cantidad de métodos y algoritmos con el objetivo de detectar la posición de la pupila (Omran & AlShemmary, 2020), (Lee, Chen, Yu, & Wu, 2020), (Petrov & Minakova, 2020), (Gokul Rajan & Vijayalakshmi, 2020). Una de las dificultades en el proceso de detección del centro de la pupila tiene relación con el proceso de adquisición de la imagen, por consecuencia de: cambio de la posición ocular respecto de la cámara, condiciones desfavorables de iluminación, ruido en los datos adquiridos, obstrucción parcial de la pupila por el párpado, pestañas o cabellos, así como destellos de luz, entre otros, por ende, estos han sido algunos de los parámetros que influyen en la detección de estrabismo, ya que debido a la falta de robustez del sistema, estos factores podrían verse afectados al momento de la ejecución del software.

Estas causas se han evidenciado en proyectos que se han llevado a cabo, durante diferentes investigaciones relacionadas con el seguimiento ocular, para pacientes con esta condición Jung y Co (Jung, Umirzakova, & Whangbo, 2019), propusieron un método para clasificar el estrabismo mediante características faciales, analizando el ángulo de inclinación de la cara, sabiendo que los rostros de las personas con estrabismo normal mente no son

simétricos se podría determinar si padecen de esta condición, los resultados arrojaron un 95% de precisión lo que demuestra la efectividad de este método.

Una investigación llevada a cabo por un grupo de universidades Chinas como la universidad de la Academia de Ciencias de China, realizaron un reconocimiento de estrabismo utilizando datos de seguimiento ocular y redes neuronales convolucionales. Conociendo la dificultad de caracterizar los datos, este método demostró una especificidad del 84% y una sensibilidad del 47,1%, sensibilidad comprometida por movimientos imperceptibles y sombras presentes en los videos de muestra (Chen, Fu, Lo, & Chi, 2018).

Zheng y Co (Zheng et al., 2019), realizaron una evaluación inteligente del estrabismo en videos basada en una prueba de cobertura automatizada, que consiste en realizar un video con una cámara infrarroja mientras uno de sus ojos se encuentra ocluido, con el objetivo de diagnosticar el estrabismo, medir la magnitud de movimientos oculares inconscientes. Para ello se implementaron seis etapas: la primera etapa se encarga de la extracción de la imagen de la región del ojo, en la segunda etapa se realiza la detección del límite del iris, la tercera etapa realiza la selección de los fotogramas de interés, la cuarta etapa realiza la localización de la pupila en los diferentes fotogramas, en la quinta etapa se calcula la desviación ocular, para posteriormente en la sexta etapa se pueda diagnosticar el estrabismo. Este método demostró una precisión superior al 91% en dirección horizontal y 86% en dirección vertical, demostrando ser muy prometedor para cumplir su objetivo.

Zolkifli y Nazari (Zolkifli & Nazari, 2020), en su estudio para la detección del estrabismo mediante la transformada de Hough. Diseñaron un sistema compuesto por dos

etapas aplicadas a un conjunto de imágenes, estas etapas corresponden el pre procesamiento, aplicado a imágenes provenientes del conjunto de imágenes CAVE. El sistema obtuvo un Error Cuadrático Medio (MSE) de 0.0003 y una Relación Señal- Ruido Máxima (PSNR) de 84,35% y la extracción de características, para ello se aplicó el detector de bordes de Sobel, luego la transformada de Hough para detectar los centros y el radio de los límites. Este sistema demostró realizar satisfactoriamente la función de preprocesamiento y extracción de características, pero no fue diseñado para clasificar el estrabismo.

Teniendo en consideración la escasez de imágenes de dominio público para realizar estudios con imágenes de estrabismo, Santos y Frago (Santos & Frango, 2020), propusieron un método para la generación de imágenes fotorrealistas de los ojos de personas con estrabismo utilizando redes neuronales adversas generativas convolucionales profundas, obtuvieron una precisión de 50% luego de realizar 2500 épocas de las 3000 programadas, antes de que el discriminador pierda precisión en reconocer las imágenes generadas.

Los trabajos mencionados anteriormente, se realizaron para el control y seguimiento ocular, con el fin de detectar y medir el estrabismo, debido a la importancia de detectarlo antes de los 7 años, debido a que la agudeza visual, la capacidad de acomodación, la fijación y la correspondencia sensorial se desarrolla a estas edades, entonces ante un estrabismo, todas estas condiciones se ven alteradas. Para que haya visión binocular ambos ojos deben trabajar uniformemente, de lo contrario no habrá estereopsis, y por tal motivo la persona estaría en desventaja a comparación con una persona normal influyendo en lo emocional y laboral.

Teniendo en cuenta como problema principal el avance de una enfermedad degenerativa, que a futuro puede afectar la calidad de la visión en el ojo afectado, desarrollando una ambliopía que a futuro puede causar la disminución de la agudeza visual y el desgastamiento de el mismo, si no se detecta o se realiza el tratamiento adecuado para el estrabismo los daños producidos son irreparables.

De acuerdo a lo previamente mencionado, un sistema experimental debe ofrecer una exactitud igual o superior al 70% para ser considerado útil, para clasificar imágenes, por esta razón, ¿Qué beneficios tiene un sistema de bajo costo para la detección de estrabismo en niños de 3 a 7 años de edad, a través de CNN y con un porcentaje de exactitud esperado del 70 %?

## Objetivos

### *Objetivo general*

Implementar y evaluar un sistema de detección mediante la captura de imágenes para clasificar estrabismo en infantes a través de redes neuronales convolucionales.

### *Objetivos específicos*

- Construir un conjunto de datos de entrada a través de imágenes anotadas con los diferentes tipos de estrabismo y sujetos sanos.
- Desarrollar una metodología de procesamiento de las imágenes mediante redes neuronales convolucionales (CNN) para la caracterización y análisis, utilizando herramientas de desarrollo de código abierto basadas en Python, específicas para aprendizaje de máquina y redes neuronales.
- Ajustar los parámetros necesarios del sistema en aras de obtener mejor desempeño de las técnicas escogidas al igual que resultados confiables.
- Validar el sistema con el fin de verificar la exactitud mayor o igual al 70% planteada en la hipótesis, haciendo uso de herramientas estadísticas normalmente empleadas para este fin

## **Justificación**

Como pasos preliminares en el diseño de un sistema que asista al personal médico en el diagnóstico de enfermedades oftalmológicas en infantes mediante la utilización de redes neuronales convolucionales, debido a la capacidad de analizar una gran cantidad de datos, facilita descartar los pacientes sanos de los que poseen estrabismo, permitiéndole a los médicos enfocarse en los pacientes de interés en una brigada de salud.

El método más utilizado por los especialistas para detectar estrabismo es el test de Hirschberg, que consiste en captar el reflejo luminoso en el ojo proveniente de un haz de luz. Debido a la forma en la que se realiza este examen requiere la participación del personal médico (Luo, Pundlik, Tomasi, & Houston, 2020), lo cual conlleva a una medición sujeta a errores, debido a que depende de las habilidades del médico tratante.

Por ende, lo que se implementó en este proyecto, es un sistema que ofrece la facilidad a la hora de realizar la detección, ya que con este sistema lo que se busca es eliminar la subjetividad humana al momento de tomar las pruebas y reducir el porcentaje de error humano.

Nuestro principal aporte es desplegar un sistema para potenciar las capacidades de diagnóstico por parte de los especialistas. Ya que es de suma importancia contar con un dispositivo que permita realizar de forma rápida y eficiente una revisión preliminar de pacientes, con el fin de dar un prediagnóstico del estado en que se encuentre, para posteriormente ser analizada por expertos de la salud.

# Capítulo 1

## Marco referencial

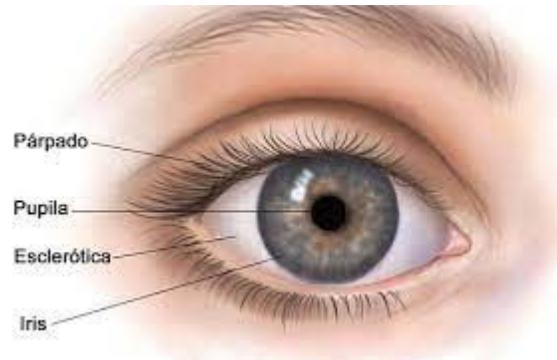
En este capítulo se describen algunas características fisiológicas del sistema de la visión haciéndose énfasis en la estructura del ojo humano, también se menciona información respecto al estrabismo congénito y acomodativo, así mismo se adiciona información con respecto a la ambliopía, condición derivada del estrabismo y para finalizar se agrega contenido sobre las redes neuronales, tanto artificiales como convolucionales, complementados con conceptos de modelado de capas de clasificación, muy importante para la comprensión del modelo de redes neuronales.

### 1.1 Marco conceptual

#### *1.1.1 Fisiología del ojo humano*

El conocimiento completo de la anatomía del ojo es requisito indispensable para la interpretación adecuada de enfermedades con manifestaciones oftálmicas. El ojo humano es un órgano que detecta, filtra y convierte un pequeño espectro de energía electromagnética en información significativa, que puede ser interpretada por el cerebro. Seis músculos, organizados en tres pares, dan al ojo tres grados de libertad. Los diferentes pares de músculos controlan los movimientos horizontales, verticales y de rotación alrededor de la dirección de la vista. Tres partes del ojo son visibles desde el exterior: la esclerótica (blanca y curvada), la pupila (negra), y el iris (de colores) tal como se muestra en la figura 1.1.

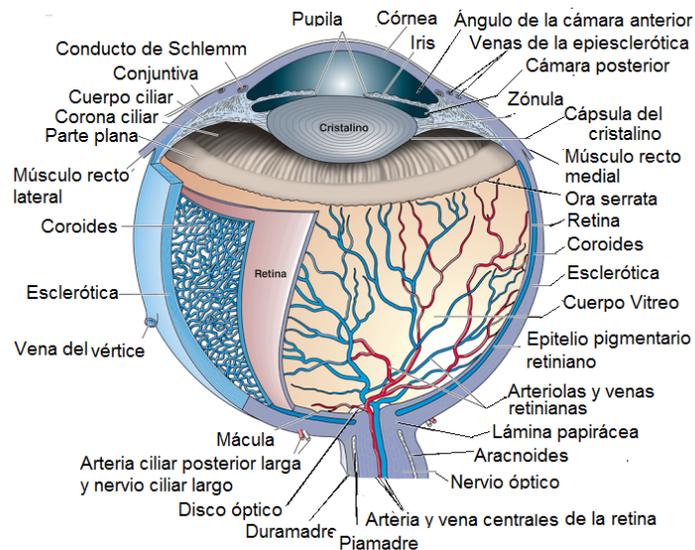
**Figura 1 - 1 Exterior del ojo**



Fuente: fotonostra.com

La forma del ojo, en el adulto normal, es casi esférica y su diámetro anteroposterior mide, en promedio, 24 milímetros. Se encuentra compuesta por cuatro membranas, la córnea, la esclerótica, la coroides y la retina. Esta última es la región con mayor sensibilidad del ojo, ya que cuenta con una gran variedad de células receptoras. La retina, o parte interna del globo ocular está especialmente diseñada, para obtener la mayor resolución del detalle del espacio, de esta manera se pueda obtener un punto o área de clara visión conocida como fovea.

**Figura 1 - 2 Anatomía interna del ojo**



Fuente: [eca-centrodeaprendizaje.blogspot.com](http://eca-centrodeaprendizaje.blogspot.com)

La túnica media del ojo, o coroides constituye una membrana altamente pigmentada. Se trata de un conjunto de vasos capilares, encargados de la irrigación sanguínea del ojo. En su porción anterior la coroides se encuentra dividida en dos estructuras: el cuerpo ciliar y el iris. El cuerpo ciliar está compuesto por múltiples fibras musculares radiales y circulares que, junto a otros ligamentos, se encargan de mantener el cristalino posicionado en su sitio. El iris, que rodea la pupila reacciona abriéndose o cerrándose con el propósito de moderar la intensidad de la luz que ingresa al globo ocular y cuenta entre sus propiedades con el pigmento visible del ojo.

La esclerótica está compuesta de tejido fibroso blanco y duro, exceptuando su porción frontal denominada córnea, cuya propiedad es ser de color transparente, para

permitir el paso de luz y la se localiza sobre el iris. La córnea cubre la pupila y el iris con una capa transparente, que refracta la luz antes de que entre en el ojo. La abertura en el centro del iris es la pupila, a través de ella pasa la luz hacia el cristalino y la retina, esta no es más que un agujero que se encuentra rodeado por el cuerpo muscular del iris.

El diámetro de la pupila aumenta a poca intensidad de luz, y disminuye cuando la luz es brillante e intensa. La pupila puede llegar a comprimirse hasta los 1,5 mm de diámetro, o dilatarse hasta los 8 ó 9 mm y tiene la capacidad de responder a estímulos de 0,2 segundos. La cantidad de luz que ingresa al globo ocular es proporcional al área de la pupila, por lo cual, considerando las dimensiones anteriormente mencionadas, la cantidad de luz que ingresa al ojo puede variar unas 30 veces como resultado de la variación del diámetro de la pupila (Laeng & Endestad, 2012).

La constricción y dilatación de la pupila dependen del ajuste controlado ejercido por el sistema nervioso autónomo aplicado sobre los músculos del iris (“Psychophysiology: Human Behavior & Physiological Response - John L. Andreassi - Google Libros,” 2000). Esta acción va ligada entre el músculo esfínter de la pupila y las fibras del iris, que tienen únicamente inervación autónoma (Leon-Sarmiento, Prada, & Gutiérrez, 2008).

El tamaño de la pupila, cambia de un individuo a otro y en algunos casos entre los ojos de un mismo sujeto. Aproximadamente el diámetro se encuentra entre los 3mm y 4mm y tiende a disminuir con el paso del tiempo. Las respuestas de la pupila más comunes, que constituyen la variación en su diámetro, están relacionadas con los cambios en los niveles

de luz, presentes en el ambiente que estimulan el sistema ocular, variaciones de color y movimientos, aunque hay que tener en cuenta que no son las únicas condiciones relacionadas con la reactividad pupilar (Leon-Sarmiento et al., 2008).

### ***1.1.2 Estrabismo***

El estrabismo es un trastorno ocular muy común, que consiste en la deficiencia ocular para alinear correctamente los ojos, esta condición produce graves problemas durante el desarrollo de los infantes, como puede ser: conductas de baja autoestima y trastorno hiperactivo y déficit atencional. El estrabismo afecta entre el 2% y 4% de la población infantil (Mohney et al., 2007). El estrabismo fuera de producir un impacto psicológico en las personas que la padecen, puede producir alteraciones en los movimientos oculares y posiciones anómalas de la cabeza, por nombrar algunos.

**Figura 1 - 3 Niño con esotropía congénita (Estrabismo)**



Fuente: Autores

La causa del padecimiento de estrabismo, puede ser muy variada, la más frecuente es el estrabismo congénito, presente en infantes y cuya condición se encuentra desde su nacimiento, normalmente causado por problemas motores, que afectan la región muscular encargados de realizar los movimientos del globo ocular, que se puede atribuir a una lesión

en los músculos extraoculares, la unión neuromuscular o los nervios craneales encargados de inervarlos, otro tipo de trastorno es en los mecanismos neuronales, que permite mover los ojos juntos (Ropper, Samuels, Klein, & Prasad, 2019a), así como también el estrabismo acomodativo, que se produce cuando el sujeto debido a problemas en la visión, desvía un ojo momentáneamente para poder enfocar un objeto, por esta razón, el estrabismo acomodativo se puede catalogar como un tipo de estrabismo alternante.

Los movimientos oculares, necesitan una coordinación de los músculos motores, para realizar un movimiento es necesaria la contracción del músculo agonista de esa dirección y la relajación del músculo antagonista de la posición contraria, por ejemplo: cuando el ojo se gira hacia afuera, el elevador es el recto superior y el depresor es el recto inferior. Una de las causas del estrabismo, es la debilidad de un músculo del ojo individual conocido como (estrabismo paralítico), o por un posible desequilibrio del tono muscular causado por defectos en el sistema nervioso que se encarga de controlar los movimientos oculares; existe una forma de endotropía, conocida como endotropía acomodativa, esta condición es producida por la hipermetropía para compensar la respuesta cercana que hacen los ojos, la endotropía se puede tratar con uso de gafas dentro de los seis meses luego del estrabismo restaura la visión y conduce a la realineación de los ejes si el grado de endotropía es muy grande es necesario realizar una realineación quirúrgica. (Ropper et al., 2019a).

En los infantes menores de siete años, que posean estrabismo evidente es necesario ser examinado de forma inmediata, para evitar la pérdida de visión del ojo desviado. Aproximadamente el 3% de los niños nacen con estrabismo o es desarrollado durante su infancia, (Pitha & Tsai, 2020), los ojos se pueden desviar hacia adentro (esotropía), hacia

afuera (exotropía), hacia arriba (hipertropía) y hacia abajo (hipotropía). El estrabismo debido a la desviación ocular es uno de los causantes del desarrollo de la ambliopía, que produce el déficit monocular de la visión. El paciente al perder la estereopsis, confunde al cerebro al recibir dos señales distintas, lo que produce la inhibición de la señal proveniente del ojo desviado o vago, para evitar pérdida del equilibrio o desorientación del paciente (Ropper, Samuels, Klein, & Prasad, 2019b).

El deterioro de la vista, se expresa como defectos de la agudeza y alteraciones de los campos visuales, básicamente es el síntoma más importante en las enfermedades oculares. Para referirse a la pérdida visual, existen diferentes términos, la amaurosis es un término utilizado para referirse a la pérdida total o parcial de la vista. La ambliopía se refiere a un déficit monocular en la visión, una de las causas de la ambliopía es la supresión de la información, proveniente de un ojo durante la primera infancia causada por el estrabismo. (Ropper et al., 2019b).

### ***1.1.3 Ambliopía***

La ambliopía compromete de forma negativa la agudeza visual normal en un ojo sano, esta afectación visual no se es posible la corrección por medio de anteojos o lentes de contacto. Existen múltiples causas que producen la ambliopía, donde se incluyen: la ambliopía estrábica, la ambliopía refractiva y la ambliopía por privación de formas, por nombrar algunas; debido a que el estrabismo es la incapacidad de fijar los dos ojos al mismo objetivo, se desarrolla una preferencia por fijar el objeto con un solo ojo, esta condición produce ambliopía estrábica. El cerebro tiende a ignorar la señal proveniente del ojo no

fijado para evitar la diplopía, que consiste en la mirada doble que produce dos puntos de visión diferentes, que confunden al cerebro. En el caso de la ambliopía refractiva, esta es causada por la diferencia de refracción de ambos ojos, donde uno de ellos suele tener una mejor agudeza visual que el otro; por lo tanto, uno de los ojos se desarrolla normalmente mientras que la señal proveniente del otro ojo suele ser ignorada por el cerebro. Por lo tanto, las vías centrales para el desarrollo de la agudeza visual, no se desarrolla de forma adecuada (Samuels, 2020).

La deficiencia en el desarrollo de la corteza occipital que da como resultado la ambliopía, se ha estudiado a profundidad en animales y humanos. Los neurólogos, saben la importancia de realizar la detección de la ambliopía en niños, incluso si el tratamiento aplicado para tratar la ambliopía no da los resultados esperados. Los expertos en primer lugar, realizan los procedimientos para corregir errores refractivos, cataratas o problemas oculares que son posibles de corregir. Luego se utiliza un método para forzar el uso del ojo desfavorecido, utilizando parches y gotas de atropina, que hacen parte del procedimiento típico para esta labor (Ropper et al., 2019b).

#### ***1.1.4 Redes Neuronales Artificiales***

Una red neuronal biológica, es el sistema de conexión que posee el cerebro para el funcionamiento de todas las partes del cuerpo humano, desde los músculos hasta los procesos más complejos. Ahora bien, una red neuronal artificial también es conocida como un sistema de conexión, pero esta vez de modelos matemáticos conectados entre sí, cuyo fundamento básico, es el proceso de aprendizaje e imitación de datos reales por medio de una computadora. Tal herramienta, hace parte de la actual Inteligencia Artificial (AI), que

imita las respuestas biológicas que se generan dentro de la máquina de mando del cuerpo humano llamado cerebro (Amato et al., 2013).

Así mismo, dentro de estas nuevas tecnologías de la Inteligencia Artificial, se encuentran las diferentes ramas como el machine learning que es una forma de aprendizaje de la inteligencia computacional el cual se da para el aprendizaje automático por medio de un algoritmo, la cual se considera como un proceso complejo de entrenamiento, al que hay que brindarle un modelamiento de entrada y se proporciona de la misma manera, y un modelamiento para predecir una salida. Así mismo, el machine learning tiene unos enfoques específicos que se dividen en:

**Aprendizaje Supervisado:** Su principal función es la búsqueda de patrones dentro de un conjunto de datos de entrada etiquetados, para que estos puedan entrar a un proceso de análisis de información.

**Aprendizaje No Supervisado:** A diferencia del aprendizaje supervisado, se trabaja con un conjunto de datos masivos sin etiquetar, pero tiene las mismas características y funcionalidad.

**Deep Learning:** Enfocado en Redes Neuronales, para la creación e imitación de respuestas biológicas (IBM, 2020).

### ***1.1.5 Redes Neuronales Convolucionales***

Una red neuronal convolucional (CNN o ConvNet) es un tipo de algoritmo el cual consiste en la técnica de aprendizaje profundo con la capacidad de tomar imágenes de entrada, asignar importancia (pesos y sesgos) a varios aspectos u objetos en la imagen y poder diferenciar uno del otro. El procesamiento previo requerido en una CNN computacionalmente es menor comparandolo con otros algoritmos de clasificación. Mientras que, en los métodos anteriormente utilizados, los filtros eran diseñados a mano,

con suficiente entrenamiento, ConvNet tiene la capacidad de aprender estos filtros(características).

Las CNN son las arquitecturas de uso común de redes neuronales profundas. Inicialmente se aplican para resolver problemas desafiantes como el reconocimiento de caracteres escritos a mano. En la actualidad, las CNN se han desarrollado rápidamente y se han utilizado para un amplio espectro de problemas de visión, como la teledetección y las aplicaciones médicas. Generalmente, las CNN consisten en capas convolucionales, capas agrupadas y capas completamente conectadas (como se muestra en la **Figura 1-4**). Las capas convolucionales pueden extraer características significativas y efectivas. Si la entrada a la capa convolucional es una imagen  $h_i \times w_i \times c_i$  y el tamaño del núcleo es  $h_k \times w_k$ , entonces la capa convolucional puede obtener k mapas de características con tamaño  $[h_i - h_k; w_i - w_k]$ . La capa de agrupación agrega la salida de las neuronas dentro de un vecindario rectangular, lo que puede reducir la cantidad de parámetros para las CNN. La agrupación máxima, es la función de agrupación más utilizada. La capa completamente conectada mapea la excitación en neuronas de salida, cada una de las cuales corresponde a una clase de decisión.

La arquitectura de una ConvNet es similar a la fisiología de las neuronas en el cerebro humano y se basa en la organización de la corteza visual. Las neuronas de forma individual, tienen una respuesta a los estímulos solo en una región delimitada del campo visual conocida como campo receptivo. Los diferentes campos de visión se superponen para cubrir el total del área visual.

Una CNN es capaz de tomar eficazmente las dependencias espaciales mediante la aplicación de filtros. Con el fin de que cualquier arquitectura se adapte al conjunto de imágenes, debido a la simplificación de parámetros como la resolución y reutilización de pesos de las imágenes.

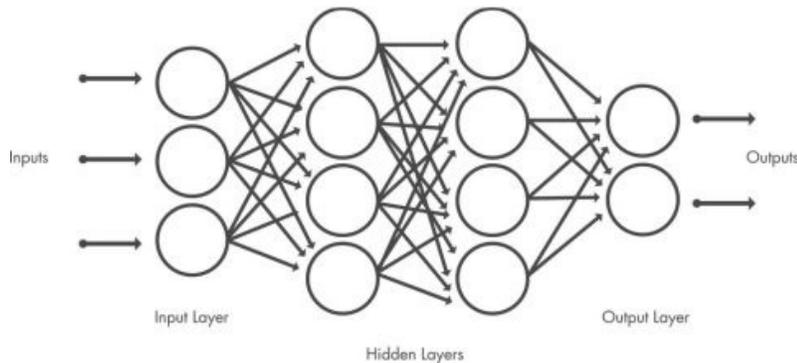
En otras palabras, una red neuronal convolucional CNN es una arquitectura de red para el aprendizaje profundo, la cual aprende directamente de los datos eliminando la extracción de características manualmente.

Las CNN son muy utilizadas para encontrar patrones y con esto poder reconocer objetos, siluetas, formas, rostros, escenas, etc. El uso de CNN para el aprendizaje profundo ha tenido bastante prestigio debido a tres factores, como, ya anteriormente mencionado. No hay necesidad de introducir o extraer funciones, debido a que la CNN las aprende directamente, además, los resultados que puede ofrecer son muy precisos y en una CNN se puede reentrenar para nuevas tareas.

### ***1.1.6 Modelo De Capas y Clasificación***

Una red neuronal convolucional está compuesta por numerosas capas que aprenden a detectar diferentes características de una imagen. Los filtros que se utilizan en cada imagen de entrada en la etapa de entrenamiento serán las salidas de cada imagen convolucionada la cual se utilizara como entrada a la siguiente capa. Los filtros se encargan de caracterizar las imágenes, como brillo y bordes, aumentando la complejidad de características que definen de forma única el objeto, como se muestra en la siguiente imagen.

**Figura 1 - 4 Estructura de modelo de capas de clasificacion**



Fuente: researchgate.net

## 1.2 Estado de arte

A continuación, se presenta una revisión de los trabajos científicos más recientes y destacados de aplicaciones relacionadas con Redes Neuronales Convolucionales en el diagnóstico y detección de enfermedades oculares, donde se exponen los objetivos de cada estudio y su aporte al área de investigación.

En la revisión sistemática titulada: “Evaluación automática de la alineación ocular para la detección del estrabismo, mediante redes U-NET y ResNet (CNN)” (De Oliveira Simoes, Carvalho Souza, Sousa De Almeida, Correa Silva, & Cardoso De Paiva, 2019), se consultaron basándose en la anterior referencia, más de 35 trabajos contextualizados con la aplicación de la red CNN para imágenes de pacientes con estrabismo, investigados en fuentes indexadas como IEEEExplore, ScienceDirect y Google académico, con palabras claves como CNN, Strabismus, Medical Imaging, que fueron clasificados de acuerdo a dos pautas, la primera fue que estuvieran en una fuente verídica como este tipo de fuentes y que

estuvieran presentes en una revista de publicaciones de alto impacto, y la segunda fue que todo artículo que se fuera a analizar debería estar en lengua inglesa, por tal motivo se realizó la exclusión de la mayoría de trabajos. Sin embargo, se encontró 19 trabajos relacionados con el procesamiento de imágenes, 7 de detección de siluetas y contornos, 8 de detección de patologías y 9 de estudios de las arquitecturas de redes neuronales, los trabajos más destacados se mencionan a continuación.

En la investigación denominada como: “Diagnostico automatizado de la enfermedad plus en la retinopatía del prematuro mediante redes neuronales convolucionales profundas” (Brown et al., 2018), el objetivo fue detectar la retinopatía del prematuro (ROP) que es una patología vascular retiniana proliferativa que afecta a un aproximado de dos tercios de los bebés prematuros que pesan menos de 1250 gramos al nacer, aunque la mayoría de los casos son leves, del 5% al 10% de los casos progresan a una retinopatía del prematuro grave, que puede producir desprendimiento de la retina y ceguera permanente si no se trata. El desafío es el diagnóstico de esta condición que se basa en la aparición de vasos retinianos en el examen oftalmoscópico.

Para este proyecto se utilizaron dos arquitecturas de redes neuronales, la primera se utilizó para segmentar los vasos sanguíneos, que se entrenó para generar una imagen donde, cada pixel representa la probabilidad de que pertenezca a un vaso retiniano y puede generar una nueva imagen de pixeles que oscilan entre 0 y 1, el valor de cada pixel simboliza si pertenece o no a un vaso sanguíneo, para esta tarea se seleccionó la arquitectura U- Net.

La segunda red se utilizó para diagnosticar la enfermedad plus a partir de las imágenes preprocesadas, que se encarga de clasificar en tres valores que pueden ser: normal, enfermedad pre- plus y enfermedad plus, para esta labor se utilizó la arquitectura de red Inception versión 1 de Szegedy y co (Szegedy et al., 2015), esta red fue entrenada con antelación utilizando la base de datos ImageNet, demostrando de esta manera, que este proceso de aprendizaje mejora el rendimiento de clasificación, aprendiendo características de imágenes generales de un conjunto de datos grande y diverso, que no está relacionado con el objetivo del proyecto.

El rendimiento del diagnóstico de la enfermedad plus se evaluó a fondo con ocho expertos con una experiencia de más de diez años en la detección de retinopatía del prematuro (ROP), estos expertos habían publicado más de cinco artículos revisados por pares sobre la ROP; de 100 imágenes, con diagnóstico de ocho expertos internacionales en ROP, el algoritmo propuesto para detectar la enfermedad plus obtuvo una sensibilidad de 93% y la especificidad de 94%, para la detección de la enfermedad pre- plus, se obtuvo una sensibilidad del 100% y una especificidad del 94%, el algoritmo diagnóstico correctamente 91 de 100 imágenes (91%), mientras ocho expertos tuvieron una precisión media del 82%, ninguna de las clasificaciones erróneas dio como resultado la enfermedad plus como normal o viceversa.

Un segundo trabajo de interés fue: “Clasificación de imágenes de neumonía infantil utilizando redes neuronales convolucionales” (Saraiva et al., 2019), si entramos en la problemática, encontramos que la neumonía es una enfermedad que puede producir la muerte en los infantes de no ser diagnosticada y tratada de forma adecuada, para realizar un

tratamiento adecuado es necesario diagnosticarlo a una corta edad, durante la revisión de historias clínicas se evidenció que el 22% de los pacientes se presentó algún motivo de incertidumbre en el diagnóstico, la prueba más adecuada para diagnosticar la neumonía, es realizar una tomografía computarizada de tórax, pero esta prueba no siempre se encuentra disponible, y el paciente debe someterse a altas dosis de radiación, que en gran medida pueden ser perjudiciales para la salud, considerando también el alto costo que posee esta prueba, es una de las limitantes a la hora de realizar un diagnóstico correcto de la neumonía.

El método propuesto, utilizó el conjunto de datos que utilizó Kermay y su equipo en sus investigaciones (Kermay et al., 2018), este conjunto posee 5863 imágenes de radiografías y dos categorías: normal y neumonía, estas imágenes son de pacientes pediátricos de uno a cinco años de edad, provenientes del Centro médico de Guangzhou (China), hechas como parte de la atención clínica, estas imágenes se sometieron a un tratamiento de para eliminar posibles escaneos de baja calidad, posteriormente fueron evaluados por dos médicos especialistas y un especialista externo para obtener el material de entrenamiento para la red neuronal. Para realizar el diagnóstico, se observa si los alveolos se encuentran llenos de secreción, si es así esta aparece como una mancha blanca en la radiografía de tórax, indicando el caso como positivo.

El modelo propuesto (Saraiva et al., 2019), utilizó radiografías de tórax, para la detectar y clasificar la neumonía, también utilizó una red neuronal convolucional. Para comprobar los resultados de este modelo se realizó la validación cruzada con otro método ya existente, este modelo obtuvo una precisión promedio de 95,3% superando el 92,8% del modelo realizado por Kermay y su equipo, con el cual se estaba comparando.

Como tercer trabajo destacado se encuentra el documento titulado: “Aplicación de una red neuronal convolucional profunda para la detección automatizada de infarto de miocardio utilizando señales de ECG” (Acharya et al., 2017), El infarto de miocardio es una condición que se produce por la interrupción del flujo sanguíneo hacia un segmento del miocardio, las arterias coronarias se encargan de la irrigación de sangre rica en oxígeno al musculo cardiaco, en algunos caso se produce un bloqueo de la arteria coronaria causado por la acumulación de placas, ese acontecimiento conlleva a la muerte del musculo cardiaco, si el flujo de sangre no se restablece a tiempo.

Para realizar el diagnóstico temprano de esta enfermedad, se utiliza el electrocardiograma (ECG). El ECG es un dispositivo no invasivo que permite detectar y registrar pequeñas señales provenientes del corazón, al tratarse de señales con amplitudes muy pequeñas, en el orden de los milivoltios mV y en algunos casos microvoltios  $\mu\text{V}$ , estas señales están expuestas al ruido del ambiente y el método de recepción de la señal por medio de electrodos, que pueden hacer mal contacto, por esta razón uno de los requisitos para este proyecto es que la red neuronal debe ser capaz de detectar la patología aun con señales ruidosas, la dificultad de inspeccionar las señales de ECG se pueden superar por medio de sistemas de diagnóstico asistido por computadora (CAD), debido a la capacidad de ofrecer resultados rápidos, objetivos y confiables.

Para realizar este trabajo se obtuvo la información de la base de datos de ECG (base de datos de ECG de diagnóstico Physikalisch- Technische Bundesanstalt), esta base de datos proporciona información de 200 sujetos de los cuales 148 poseen infarto de miocardio y 52 de sujetos sanos, cada una de estas señales cuenta con 1000 muestras por segundo y se

utilizaron un total de 10546 latidos normales y 40182 latidos con infarto de miocardio, la arquitectura de red estándar que se utilizó posee cuatro etapas, convolución, función de activación lineal rectificadora, función de agrupación y capa totalmente conectada.

En este estudio se utilizó un ordenador con dos procesadores Intel Xeon 2.40 GHz y 24 GB de memoria RAM, luego de realizar la validación cruzada se obtuvo que de 10546 latidos normales aproximadamente el 7,17%, el sistema clasificó erróneamente como infarto de miocardio, siendo estos latidos normales, del mismo modo, con los latidos de infarto de miocardio, un total de 6,29% fueron clasificados de manera errónea como latidos normales. De forma similar el 94,19% de los latidos se clasifican correctamente como latidos normales y el 4,51% de los latidos de infarto de miocardio, se clasificaron de forma incorrecta como latidos normales.

## Capítulo 2

### 2.1 Materiales y métodos

#### 2.1.1 *Pacientes*

Las imágenes utilizadas en este estudio provienen de pacientes de una botica oftalmológica privada ubicada en la ciudad de Neiva (Huila). Hasta el momento la base de datos está constituida, por imágenes de personas tanto con estrabismo como por personas sin esta condición. El conjunto de datos de estrabismo obtenido contiene 332 imágenes. Cada imagen contiene solo un rostro humano. Este se divide en un conjunto de datos de entrenamiento que contiene 232 imágenes y un conjunto de datos de prueba que contiene 100 imágenes. El conjunto de datos de entrenamiento consta de 162 imágenes de estrabismo y 70 imágenes normales, mientras que el conjunto de datos de prueba se compone de 79 imágenes de estrabismo y 21 imágenes normales. La condición de estrabismo fue sssssdiagnosticada por el personal especialista y se aplicaron los criterios descritos en el siguiente apartado. Los pacientes que se ofrecieron como voluntarios para participar en el estudio firmaron un consentimiento informado.

#### 2.1.2 *Protocolo de adquisición*

El médico especialista estableció los criterios de inclusión y exclusión para los pacientes de la base de datos de forma previa, a la imposición del protocolo de adquisición y de bioseguridad. Los pacientes pertenecientes a la base de datos fueron examinados mediante la observación de agudeza visual, biomicroscopía, fundoscopia (polo posterior),

tonometría de aplanación (cuando fue posible) y examen de motilidad ocular mediante el test de Cover para detectar la presencia o ausencia de desviación.

La exclusión de los pacientes del estudio se realizó según los siguientes criterios de exclusión: i. presentar desviaciones por debajo de  $90\Delta$  horizontal o verticalmente, ii. opacidad o cualquier otra alteración en la córnea, iii. contorno irregular del limbo, iv. microftalmía, v. nistagmo perceptible, vi. incapacidad de lograr 1.0/1.0 sobre la mesa de Snellen, y vii. incapacidad de alcanzar los 40 "de arco en la prueba de Titmus de agudeza visual estereoscópica.

### ***2.1.3 Adquisición de las imágenes***

La adquisición de imágenes se realizó en la clínica oftalmológica Lafam, utilizando una cámara Sony  $\alpha 7R$  full-frame de 35 mm y 61 Mega Píxeles, zoom óptico 3x, configurada en modo de captura de imágenes (para proporcionar una imagen clara y más detallada) con una resolución de  $2048 \times 1536$  píxeles.

Las fotografías se tomaron con el paciente sentado en la silla de reconocimiento, utilizada para realizar la adquisición de imágenes. No se utilizó foco de luz adicional, la luz de la habitación permaneció encendida. Se pidió al paciente que mirara una figura utilizada para alinear el rostro del paciente, esta figura se encuentra unida lateralmente a la lente de la cámara. El flash se encendió para generar brillo, necesario para generar la primera imagen de Purkinje, que es un método utilizado para analizar la forma del ojo a través de reflejos

luminosos. Se utilizó la función macro de la cámara para asegurar un enfoque perfecto del rostro del paciente, incluso cuando este estaba aproximadamente a 40 ó 50 cm de la cámara. En la **Figura 2-1**, se puede observar las imágenes que, para cada paciente se tomaron en total tres imágenes, con la finalidad de evidenciar el desvío ocular sin importar la dirección de la mirada que el paciente pueda tener, para esto se tomaron imágenes en las siguientes posiciones: frente, lateral izquierdo y lateral derecho.

**Figura 2 - 1 Imágenes de las tres posiciones de miradas analizadas a) frontal (PF), b) posición izquierda (PLI), y c) posición derecha (PLD).**



Fuente: Autores

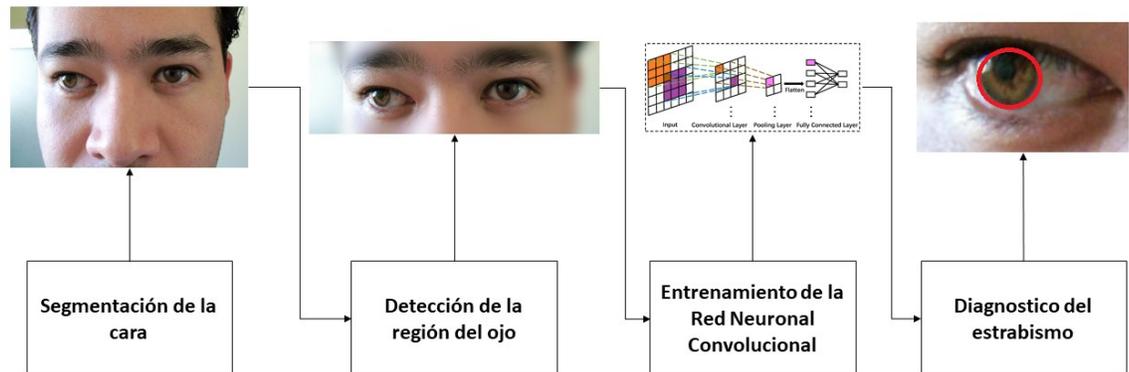
Se utilizó una regla y un transportador para guiar el ángulo de inclinación y la rotación de la cabeza del paciente a los valores estándar de las posiciones deseadas de la mirada de diagnóstico instruidos por el médico especialista. La base de datos de las imágenes estaba formada por 45 pacientes fotografiados solo en la posición de frente (PF). En este

trabajo se utilizaron 200 imágenes nuevas de 40 pacientes de ambos sexos y edades variadas. La **Figura 2-1** muestra las posiciones de la mirada a las que fueron sometidos los pacientes.

### 2.1.4 Metodología propuesta

El diagnóstico de estrabismo por medio de redes neuronales convolucionales de imágenes digitales depende del entrenamiento a partir de las imágenes de pacientes que presentan esta patología, así como de pacientes sanos. Para lograrlo, el método propuesto se organiza en cuatro pasos, como se muestra en la **Figura 2-2**. Las imágenes faciales de los pacientes se difuminaron manualmente en las partes superior e inferior para proteger las identidades de los sujetos mostrados en este trabajo.

**Figura 2 - 2 Diagrama de bloques del método propuesto**



Fuente: Autores

El primer paso es reducir el espacio de búsqueda eliminando el fondo mediante la segmentación del rostro. Una vez hecho esto, se detecta la región del ojo para excluir las

regiones de la cara que no forman parte del método. A continuación, se realiza una ubicación precisa del ojo para restringir aún más el espacio de búsqueda. Se realizará el entrenamiento de redes neuronales convolucionales con las imágenes procesadas, y se realizará el aplicativo para hacer el diagnóstico del estrabismo. Las siguientes secciones están dedicadas a una descripción detallada de los pasos enumerados anteriormente.

### ***2.1.5 Segmentación facial***

El método propuesto inicia con la segmentación del rostro del paciente, esta etapa tiene como objetivo eliminar el fondo existente y delimitar el área de la cara que se utilizará en las etapas posteriores. Para poder minimizar el costo computacional del procesamiento de la imagen, el método comienza con la adquisición de la imagen segmentada seguida de su redimensionamiento de  $2048 \times 1536$  píxeles, a una resolución diez veces menor de  $205 \times 154$  píxeles. Esta imagen reducida solo se usa para el entrenamiento de la red neuronal convolucional, ya que el paso del diagnóstico requiere una imagen con mejor resolución para no perder información de la imagen.

La imagen se convierte del tipo de color RGB a YCbCr para obtener el mapa de color de piel. La detección de piel permite la segmentación de la cara en función de la intensidad de los píxeles que más se asemeja al color que suele presentar la piel. Para la segmentación facial, se utiliza el mismo mapa de color de piel propuesto por Kakumanu y su equipo (Kakumanu, Makrogiannis, & Bourbakis, 2007), que consiste en tomar una imagen a color RGB (Red, Green, Blue) conocidos como colores luz, para seguidamente por medio del

método YCbCr generar una imagen en escala de grises, representando los colores como luminancia, calculada como la suma de valores ponderados de colores luz, y la crominancia que se calcula restando la luminancia de los colores azul y rojo .

### ***2.1.6 Detección de la región ocular***

Este paso tiene como objetivo reducir el espacio de búsqueda en la imagen de la cara segmentada para generar una subimagen que contenga solo la posible región de los ojos, mientras que excluye las regiones sin interés (como la boca, nariz y cabello) para simplificar el siguiente paso del entrenamiento de la red neuronal convolucional.

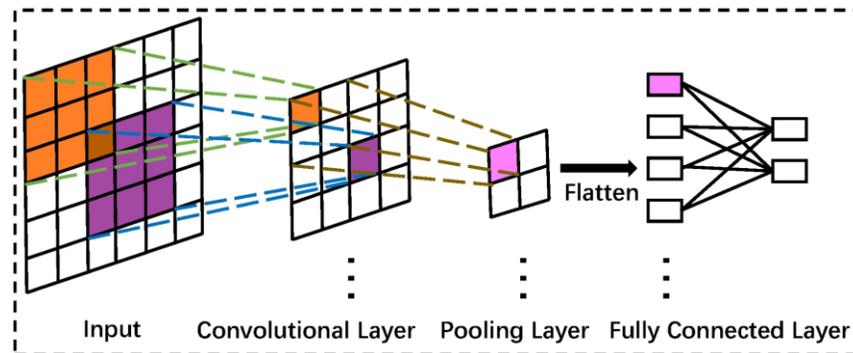
La imagen redimensionada en el paso anterior se convierte de color RGB a escala de grises porque es computacionalmente más eficiente trabajar con un canal monocromático que con tres colores diferentes. Después de la conversión, se aplica un filtrado homomórfico basado en el modelo de luminancia / reflectividad, utilizando un rango de brillo comprimido y un contraste mejorado, para resolver las divergencias de luz. El suavizado de la imagen se realiza utilizando un filtro gaussiano de tamaño  $3 \times 3$ , el filtro gaussiano, es utilizado para funciones de suavizado de la imagen que es afectada por el ruido gaussiano, el resultado del filtrado es igual al promedio ponderado de la imagen y el gradiente de la imagen de entrada, generado en el paso anterior, se calcula utilizando el filtro Sobel que se utiliza para realizar la detección de bordes, calculando la aproximación de los gradientes en la imagen.

Como paso siguiente, es necesario aplicar una proyección horizontal de ese gradiente, esto permite calcular el valor promedio de los tres picos principales de dicha proyección. Es importante saber que los ojos se encuentran en la parte superior del rostro y que, junto con las cejas, corresponden a los dos picos próximos entre sí. Esta información fisiológica, conocida a priori, se puede utilizar para identificar el área de interés, y el pico de la proyección horizontal proporcionará la posición horizontal de los ojos. Para eliminar las interferencias de las regiones superiores del cabello y la boca, la proyección se aplica a partir de  $1/5$  a  $1/2.5$  de la altura de la imagen de la cara. El área de la región de los ojos tiene el mismo ancho y  $1/3$  de la altura de la imagen segmentada en el paso anterior.

### ***2.1.7 Entrenamiento de la red neuronal convolucional***

El marco propuesto para el entrenamiento es el RF-CNN puede considerarse como un método de detección de estrabismo en dos etapas. RF (Random Forest) es un método de aprendizaje automático con la capacidad de realizar clasificaciones, para ello, utiliza múltiples algoritmos de árbol de decisiones para clasificar los objetos, basándose en los atributos que este posea, cada árbol de decisión genera un resultado clasificadorio y la decisión con mayor cantidad de selecciones, es la predicción del algoritmo. Como se muestra en la **Figura 2-3**, las imágenes en el conjunto de datos de estrabismo se alimentan secuencialmente en RF-CNN, que finalmente genera las ubicaciones de las regiones oculares. Luego, las imágenes de la región del ojo se recortan y se introducen en las CNN. RF-CNN se aprende directamente a partir de datos de imágenes y las anotaciones de eventos se utilizan para el entrenamiento en ambas etapas.

**Figura 2 - 3 Estructura generalizada de la red neuronal convolucional implementada**

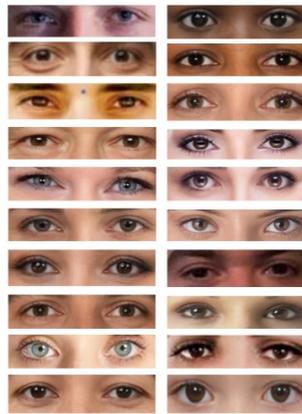


Fuente: Autores

La base de datos ha sido creada con un total de 332 imágenes, provenientes de 234 personas constituida por hombres y mujeres con estrabismo y sanas, las edades de aquellas personas oscilan entre los 8 hasta los 42 años. El conjunto de imágenes de entrenamiento posee 162 imágenes de estrabismo, provenientes de 102 pacientes, que debido a la carencia de voluntarios para la toma de muestras, por motivo de la contingencia sanitaria en el momento de recolectar las imágenes, se le pidió a 30 de esos pacientes, el consentimiento para tomar tres imágenes de ellos dirigiendo su mirada hacia al frente, hacia la izquierda y hacia la derecha, con la finalidad de evidenciar el estrabismo sin importar la dirección de la mirada, mientras se aumenta el número del conjunto de datos. En el conjunto de prueba, se cuenta un total de 79 imágenes de personas con estrabismo, provenientes de 41 pacientes voluntarios, de los cuales se seleccionó a 19 de ellos para realizar la toma de tres imágenes, de la misma forma que se utilizó en el conjunto de entrenamiento. El conjunto de imágenes de personas sin estrabismo, se compone de 70 imágenes destinadas para el entrenamiento y

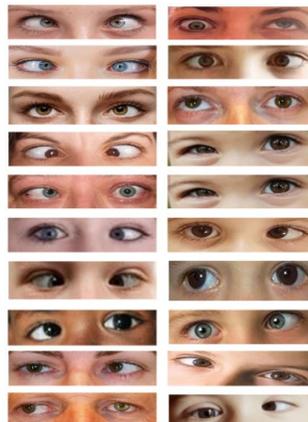
21 imágenes para la validación, cada una de estas imágenes corresponde a un paciente diferente. En la **Figura 2-4** y la **Figura 2-5** se muestran ejemplos del conjunto de datos de imágenes segmentadas de pacientes con estrabismo y sin esta condición.

**Figura 2 - 4 Algunos ejemplos de regiones oculares segmentadas con etiquetas normales.**



Fuente: Autores

**Figura 2 - 5 Algunos ejemplos de regiones oculares segmentadas con etiquetas estrabismo.**



Fuente: Autores

El entrenamiento de estrabismo se realiza aplicando CNN según las regiones oculares segmentadas. Las **Figuras 2-8 y 2-9** muestran algunos ejemplos de regiones oculares segmentadas con estrabismo y marcadores normales.

Después de confirmar la arquitectura CNN, los parámetros  $w$  de las CNN se aprenden optimizando la función objetivo  $J(w)$ , como se muestra en la **Ecuación 2-1**

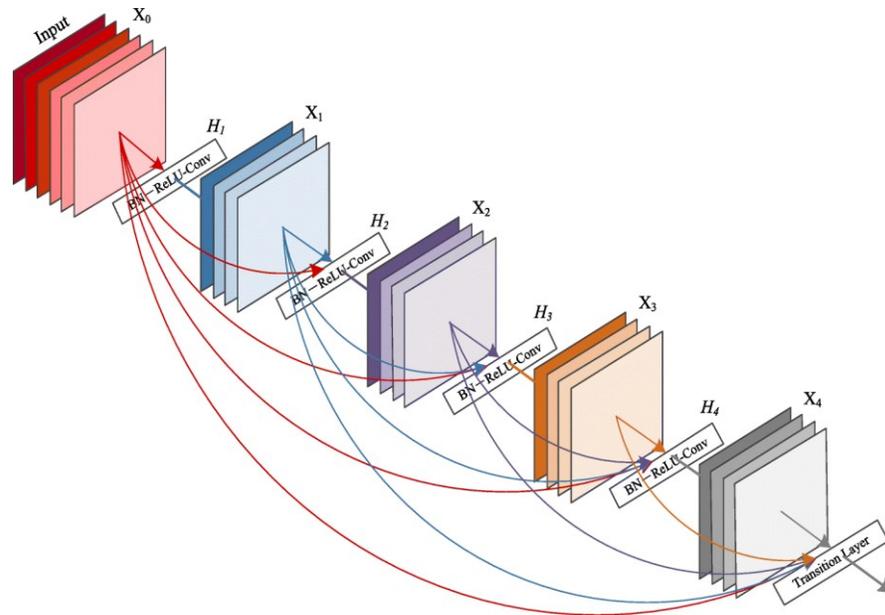
**Ecuación 2- 1 Función objetivo**

$$J(\omega) = \frac{1}{n} \sum_{i=1}^n \|f(\omega, x_i) - y_i\|^2$$

donde  $n$  es el número de ejemplos de entrenamiento.  $(x_i, y_i)$  representa el  $i$ -ésimo ejemplo de entrenamiento.  $f(\cdot)$  es la función de activación.

La arquitectura DenseNet (Red convolucional densa) es un desarrollo conjunto entre la universidad de Cornwell, la Universidad de Tsinghua y Facebook AI Research. La conexión densa, logra una alta precisión con menos parámetros comparándola con arquitecturas que se utilizan para el mismo fin como por ejemplo ResNet y ResNet de pre-activación. (Sik - Ho Tsang, 2018)

**Figura 2 - 6 Estructura de modelo Densenet.**



Fuente: researchgate.net

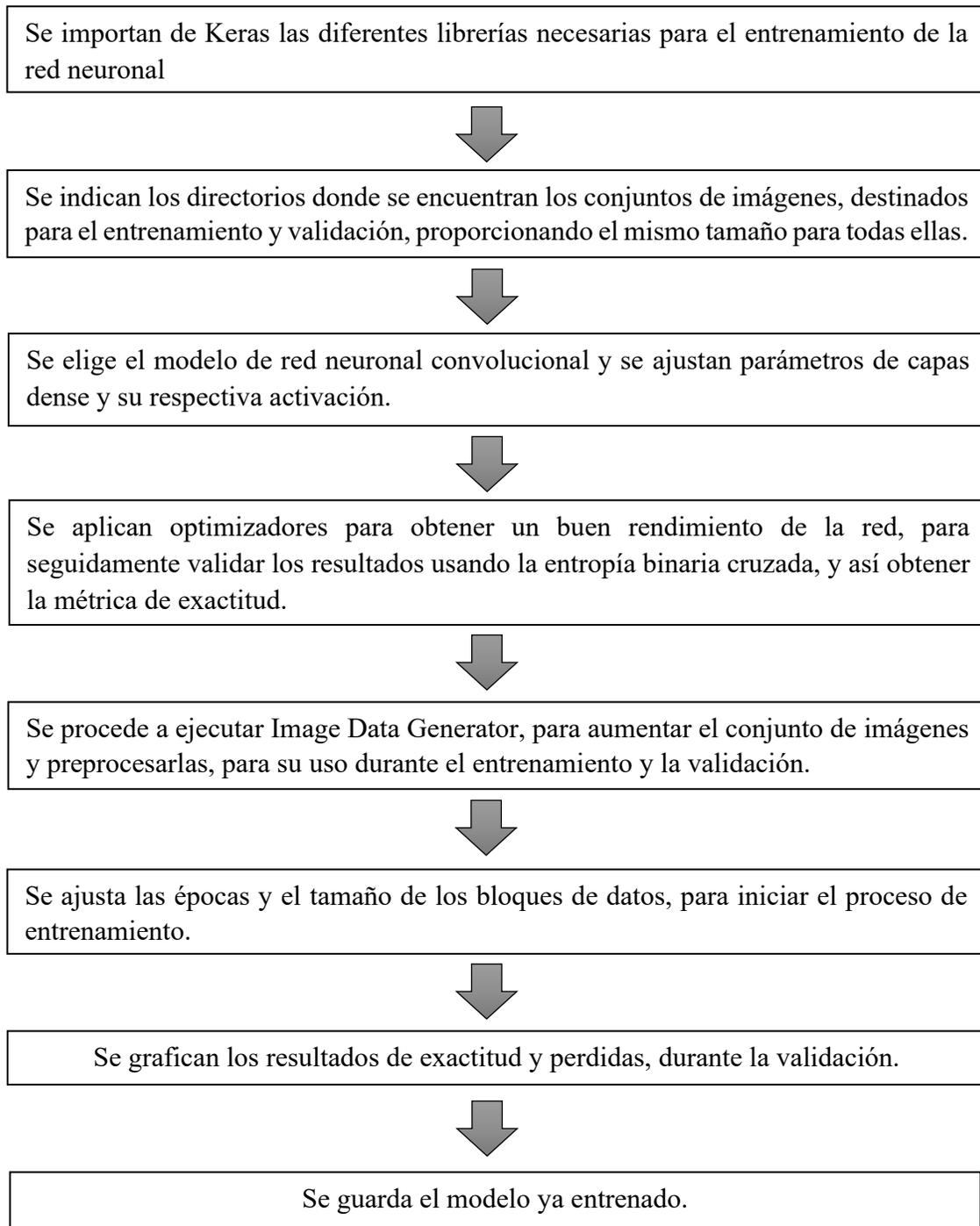
La arquitectura de red consta de cinco capas convolucionales y tres capas agrupadas, seguidas de tres capas completamente conectadas, como se muestra en la **Figura 2-6**. Cada capa convolucional es seguida por una capa Relu (Bressem et al., 2020), una función de activación efectiva para mejorar el rendimiento de las CNN. Además, la estrategia de abandono (Zerouaoui & Idri, 2022) se utiliza en las dos primeras capas completamente conectadas para evitar un ajuste excesivo.

**{Tabla 2 - 1 Características del modelo Densenet 201**

Layers	Output Size	DenseNet- 201 ( $k = 32$ )
Convolution	112 x 112	7 x 7 conv, stride 2

Pooling	56 x 56	3 x 3 max pool, stride 2
Dense Block (1)	56 x 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 x 56	1 x 1 conv
	28 x 28	2 x 2 average pool, stride 2
Dense Block (2)	28 x 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 x 28	1 x 1 conv
	14 x 14	2 x 2 average pool, stride 2
Dense Block (3)	14 x 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Transition Layer (3)	14 x 14	1 x 1 conv
	7 x 7	2 x 2 average pool, stride 2
Dense Block (4)	7 x 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$
Classification Layer	1 x 1	7 x 7 global average pool
		1000D fully – connected, softmax

Fuente: kaggle.com

**Figura 2 - 7 Mapa Secuencial del algoritmo de entrenamiento**

**Fuente: Autores**

### ***2.1.8 Configuración de entrenamiento***

La primera herramienta que se utiliza para detectar la cara de un conductor es el clasificador en cascada de Haar. Es uno de los pocos métodos de detección de objetos con capacidad para detectar rostros. Paul Viola y Michael Jones inventaron este método (Wilson & Fernandez, 2006). Está entrenado con miles de caras en diferentes condiciones de luz (Wilson & Fernandez, 2006) (Selvam & Krishna GVT, 2013). Además, el método detecta de manera eficiente a las personas en una habitación luminosa. En este trabajo, se utilizó el clasificador en cascada de Haar para detectar la cara porque se considera que tiene una tasa de precisión muy alta del 95%.

El clasificador en cascada de Haar, que se utiliza para la detección de rostros en este trabajo, el cual proporciona una gran proporción para el área de los ojos. Para igualar varios tamaños de caras, se redimensionan las imágenes en  $200 \times 200$  para obtener la misma proporción de ojos para todos. El área de los ojos se recortará aún más para borrar las secciones no relacionadas. A pesar de que el clasificador en cascada de Haar proporciona una gran proporción del área de los ojos, todavía no es tan constante como necesitan los autores. Como resultado, los autores utilizan la coincidencia de plantillas para garantizar que el área de los ojos pueda mejorar su eficiencia.

La red neuronal convolucional que se adopta es DensNet-201 (Park, 2020), como la base de CNN, y la minería de ejemplo duro en línea, de sus siglas en inglés: Online Hard Example Mining (OHEM) (Shrivastava, Gupta, & Girshick, 2016), se utiliza para entrenar a CNN. Se utiliza una tasa de aprendizaje de 0,0003, cuya función es ajustar el tamaño de los pesos, ya que, si este valor es muy bajo, provoca una disminución en la velocidad de

convergencia y en caso de ser muy alto, evitará que produzca convergencia, que por consiguiente produce inestabilidad, debido a que los saltos no alcanzaran al mínimo. El impulso (Momentum) de 0,9, es el encargado de acelerar el descenso del gradiente, como una función optimizadora. Varias regiones oculares están segmentadas por la CNN, que luego se redimensionan a 224x224x3.

El entrenamiento en red se realiza mediante el método de descenso de gradiente estocástico (Bressem et al., 2020). La regularización L con la disminución de peso  $5 \times 10^{-4}$  se utiliza en el entrenamiento en red. La tasa de abandono se establece en 0,5. El tamaño del lote se establece en 32. La tasa de aprendizaje se establece inicialmente en 0.01 y el entrenamiento se detiene después de 5000 iteraciones. La implementación de CNN se basó en el paquete Tkinter para crear el Widgets Tk (Zerouaoui & Idri, 2022). Así mismo, se soporta en el paquete de Tensorflow que es una caja de herramientas eficaz para entrenar redes neuronales profundas. El entrenamiento se llevó a cabo en una CPU Asus con procesador Intel i7, de 2.0 GHz de octava generación, con tarjeta GPU NVIDIA GEFORCE y 12 Gb de RAM.

### ***2.1.9 Desarrollo de la aplicación de diagnóstico del estrabismo***

Una vez establecida la base de datos de los pacientes con estrabismo se realiza una separación del 70% para realizar el entrenamiento y 30% para hacer el testeo de la red construida. Mediante el paquete Tkinter se implementa la aplicación para el diagnóstico del estrabismo. Tkinter es el paquete más utilizado para crear interfaces gráficas en Python. Es

una capa orientada a objetos basada en Tcl (sencillo y versátil lenguaje de programación open-source) y Tk (la herramienta GUI estándar para Tcl).

**Figura 2 - 8 Módulo de detección de estrabismo desarrollado en Tkinter de Python.**



Fuente: Autores

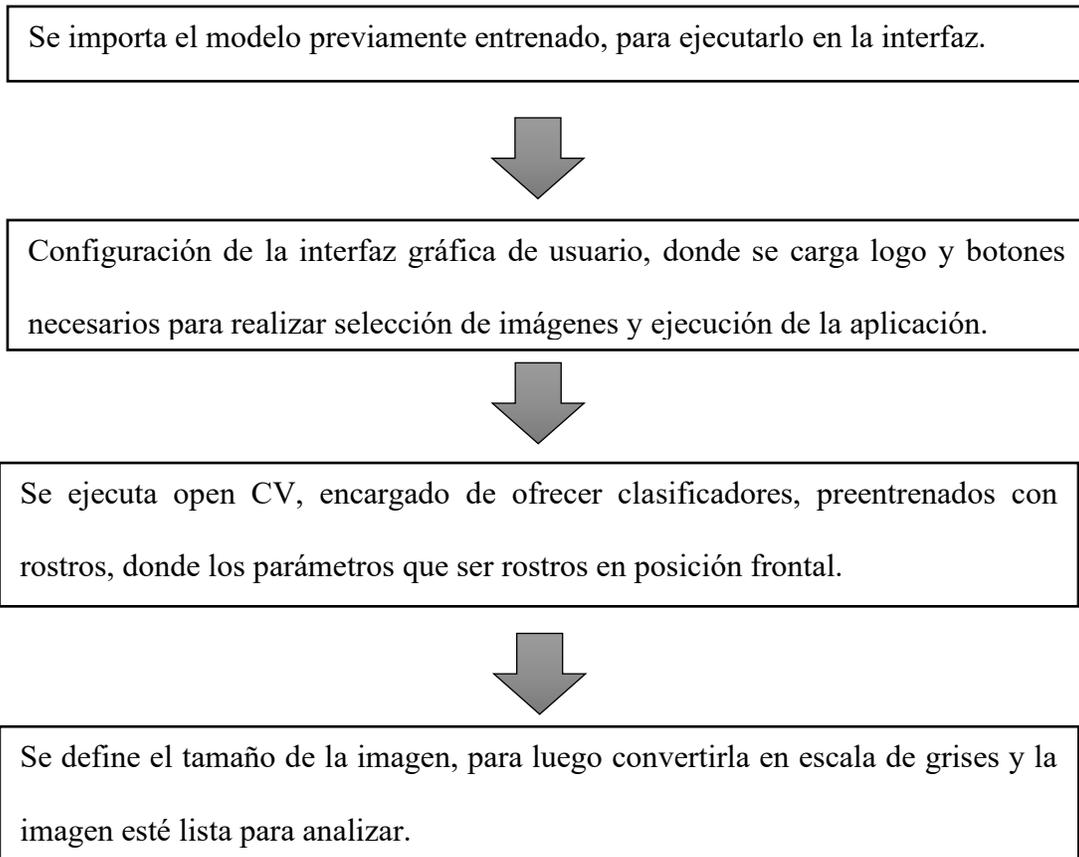
**Figura 2 - 9 Mapa de secuencia del algoritmo de implementación.**

Se importan los módulos necesarios que permiten trabajar de forma portable.



Se importan las librerías necesarias para realizar el preprocesamiento de las imágenes que se ingresan para la prueba.





Fuente: Autores

## Capítulo 3

### 3.1 Métricas de evaluación

En este trabajo se utilizaron cuatro métricas de evaluación de uso común: Sensibilidad, Especificidad, Precisión y AUC para evaluar el rendimiento de RF-CNN:

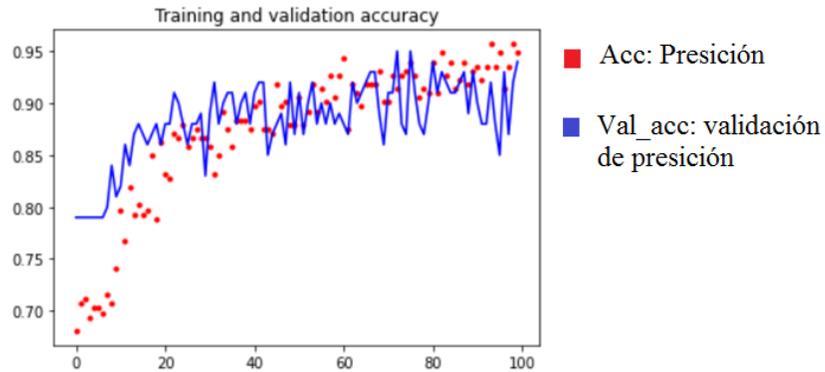
$$\text{Sensibilidad (Se)} = \frac{TP}{TP + FN}$$

$$\text{Especificidad (Sp)} = \frac{TN}{TN + FP}$$

$$\text{Precisión (ACC)} = \frac{TP + TN}{TP + TN + FP + FN}$$

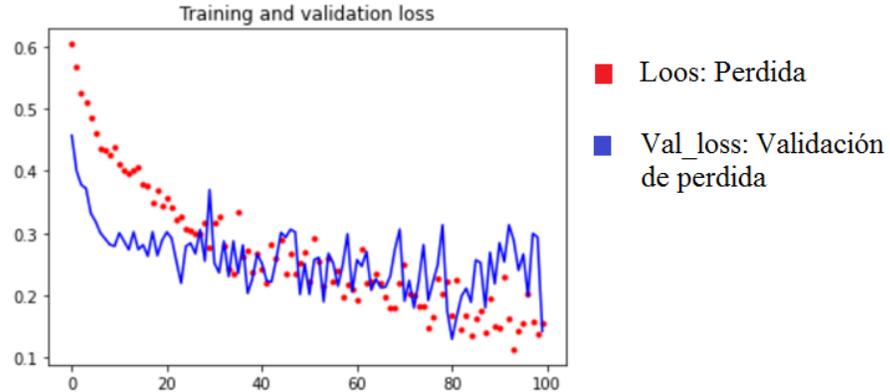
donde TP (verdadero positivo), TN (verdadero negativo), FP (falso positivo) y FN (falso negativo) son los números de imágenes de estrabismo correctamente identificadas, imágenes normales correctamente identificadas, imágenes de estrabismo incorrectamente identificadas e imágenes normales incorrectamente identificadas, respectivamente. La sensibilidad (Se) y la especificidad (Sp) indican la capacidad CNN de identificar imágenes normales e imágenes de estrabismo. La precisión (ACC) se utiliza para evaluar el rendimiento de detección general. Además, el área bajo la curva característica de funcionamiento del receptor (ROC) (AUC) también se aplica para evaluar el rendimiento de detección general de CNN.

**Figura 3 - 1 Grafica de resultado de exactitud.**



Fuente: Autores.

**Figura 3 - 2 Graficas de perdida**



Fuente: Autores.

## 3.2 Resultados

Los resultados de detección de CNN en el conjunto de datos establecido se muestran en la **Tabla 3-1** y en la **Figura 3-1**. Se puede observar que CNN puede lograr puntajes altos de Sensibilidad = 0.9619 y Especificidad = 0.9329, lo que indica que CNN puede funcionar

bien en la identificación de imágenes de estrabismo. e imágenes normales. Los puntajes altos de Precisión = 0.9388 y AUC = 0.9864 se obtienen mediante CNN, lo que significa que CNN puede lograr buenos resultados de detección generales en el conjunto de datos de estrabismo establecido.

Además, el análisis de sensibilidad de CNN al número de ejemplos de entrenamiento se muestra en la **Figura 3-2**.

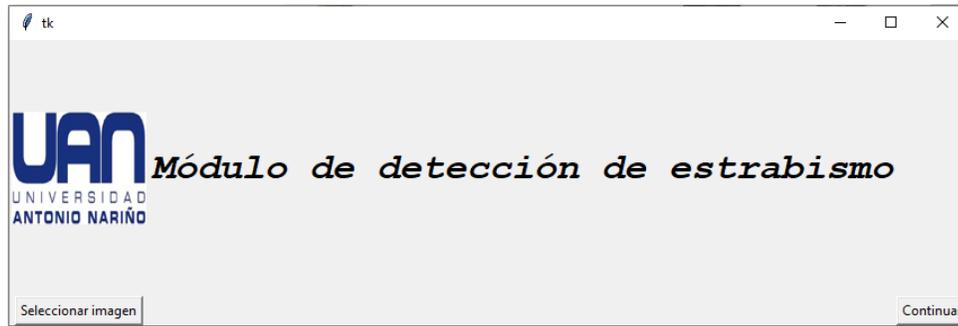
**Tabla 3 - 1 Resultados de métricas de evaluación del sistema.**

Métrica	TP	TN	FP	FN	Se	Sp	Acc	AUC
DensNet	430	1600	115	17	0.9619	0.9329	0.9388	0.9864

En la **Figura 3-2**, se puede observar que las métricas de evaluación Se, Sp, Acc y AUC mejoran con el aumento de ciclos de entrenamiento. Con menos de 1500 ciclos de entrenamiento, los resultados de detección mejoran significativamente con el aumento de ejemplos de entrenamiento. Con más de 1500 ciclos de entrenamiento, los resultados de detección varían ligeramente con el aumento de ciclos de entrenamiento. De la observación anterior, se pueden elegir 1500 ciclos de entrenamiento para entrenar al CNN en el trabajo.

En la **Figura 3-3** se muestra la interfaz gráfica desarrollada en Tkinter donde se observa el nombre de la aplicación, logo de la universidad y botones de funcionamiento.

**Figura 3 - 3 Ventana principal de interfaz grafica**



Fuente: Autores.

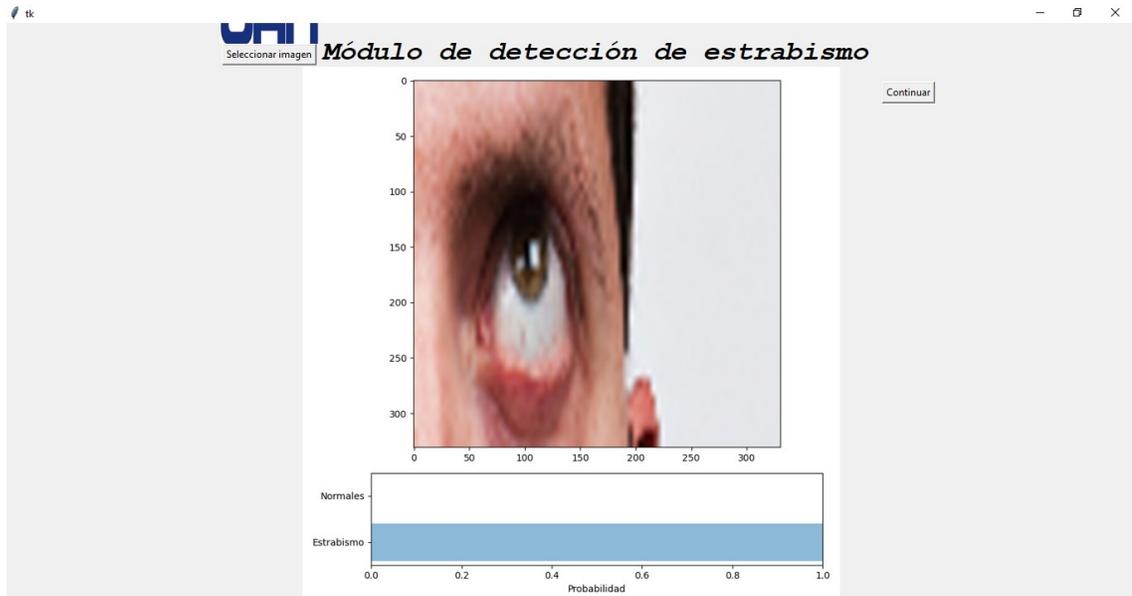
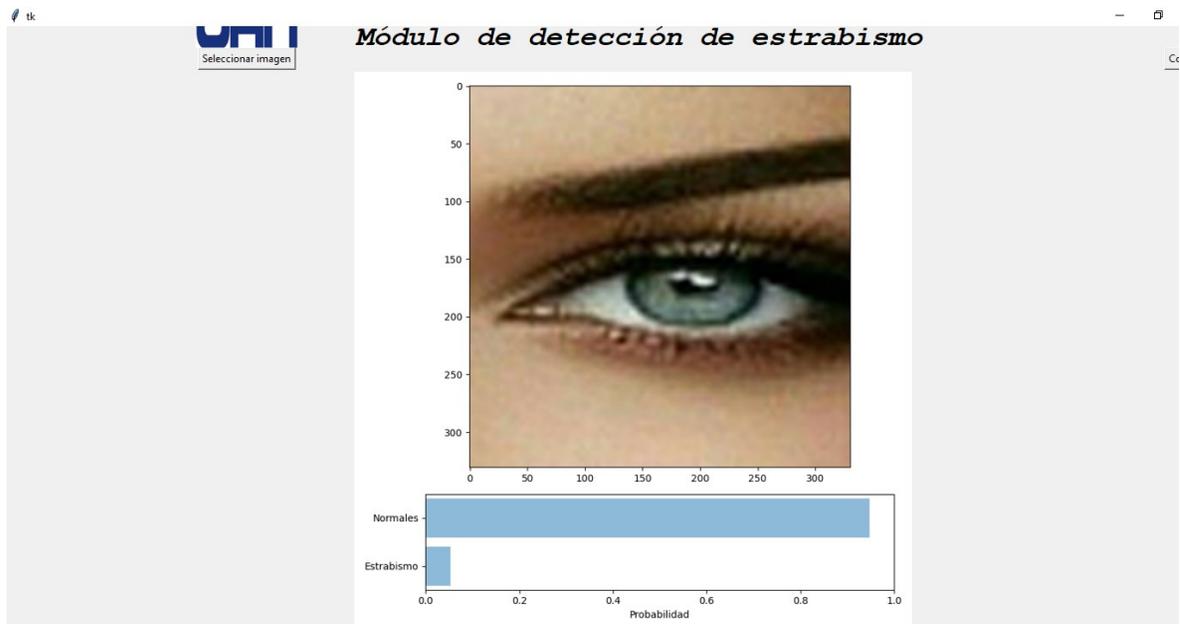
El botón “Seleccionar imagen”, accede a la ruta donde se encuentra las imágenes a procesar como, y una vez seleccionada, se indica que la imagen esta lista para su procesamiento, como se muestra en la **Figura 3-4**.

**Figura 3 - 4 Selección de tipo de imagen en ventana principal de interfaz grafica**



Fuente: Autores.

El modulo una vez cargada la imagen, se inicia con el procesamiento para diagnosticar la probabilidad de que esa persona tenga o no estrabismo como se muestra en las **Figuras 3-5** y **Figura 3-6**.

**Figura 3 - 5 Diagnostico positivo de estrabismo****Figura 3 - 6 Diagnostico negativo de estrabismo**

**Fuente: Autores.**

## Capítulo 4

### 4.1 Conclusiones

Hoy en día el estrabismo se ha convertido en una enfermedad oftalmológica influyente en la vida humana. La detección del estrabismo juega un papel importante en el pronóstico y el tratamiento del estrabismo. Es por esto que el desarrollo de aplicaciones que coadyuven con el profesional médico toma vital importancia. Nuestra aplicación se aplica para lograr la detección oportuna del estrabismo, permite que otros proyectos puedan tener un buen punto de partida para recolectar los datos médicos y luego enviar los datos a los médicos para el diagnóstico y examen físico.

Se construyó una base de datos para el entrenamiento y testeo de la red neuronal convolucional de sujetos con estrabismo y personas sanas constituida por 332 imágenes en total destinándose 162 imágenes de estrabismo y 70 imágenes de personas sin estrabismo, siendo este el 70% del total de imágenes, destinadas para la labor de entrenamiento, mientras que para la validación, se utilizó 79 imágenes de estrabismo y 21 imágenes de personas sin estrabismo, que corresponde al restante 30% del conjunto de imágenes.

Se desarrolló una metodología con redes neuronales convolucionales, utilizando la arquitectura DensNet-201 para lograr la detección automatizada del estrabismo con una sensibilidad del 0.9619, una especificidad del 0.9329 y una precisión del 0.9388 que permitiera entrenarla en la clasificación de personas sanas y sujetos con estrabismo, logrando una exactitud mayor o igual al 70% en el entrenamiento. Para este fin se utilizó programación en el ambiente de Anaconda en el lenguaje Python.

Se realizó una interfaz de programación utilizando Tkinter el cual es un binding de la biblioteca grafica Tcl/Tk para el lenguaje de programación de Python. La aplicación de detección del estrabismo propuesto primero usa el clasificador en cascada de Haar para realizar la segmentación de la región del ojo, y luego clasificar las regiones del ojo segmentadas como estrabismo o normal con una red neuronal convolucional profunda ajustando los parámetros necesarios del sistema para mejorar el desempeño de las técnicas escogidas.

Se realizó validación del sistema con el fin de verificar que la exactitud estuviera por encima del 70%, superando este el 93,8% de lo que fue planteada la hipótesis, estos resultados en la detección del conjunto de datos de estrabismo establecido, demuestran que el CNN propuesto funciona bien en la detección automatizada de estrabismo. Además, hasta donde sabemos, esta es la primera investigación que logra la detección automatizada del estrabismo para cualquier aplicación, como por ejemplo de telemedicina.

## Anexos

### 4.2 Anexo 1

Figura 3 - 7 Código para configuración y entrenamiento de la red neuronal.

```
validation_data_dir = '/content/Datasetestrabismo/val/'
nb_train_samples = 232
nb_validation_samples = 100
#epochs = 50
#batch_size = 16

if K.image_data_format() == 'channels_first':
    input_shape = (3, img_width, img_height)
else:
    input_shape = (img_width, img_height, 3)

CLASSES = 2

# setup model
base_model = DenseNet(weights='imagenet', include_top=False)
x = base_model.output
x = Dense(12, activation='relu')(x)
x = Dense(12, activation='relu')(x)
x = Dense(32, activation='relu')(x)
x = GlobalAveragePooling2D(name='avg_pool')(x)
x = Dropout(0.4)(x)
predictions = Dense(CLASSES, activation='softmax')(x)
model = Model(inputs=base_model.input, outputs=predictions)

# transfer learning
for layer in base_model.layers:
    layer.trainable = False

model.compile(optimizer=keras.optimizers.Nadam(lr=0.001),
              loss='binary_crossentropy',
              metrics=['acc'])
```

## 4.3 Anexo 2

Figura 3 - 8 Código de procesamiento de la red neuronal en Tkinter.

```

def opencv(self,commit=True):
    face_cascade = cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
    eyePair_cascade = cv2.CascadeClassifier('haarcascade_eye.xml')
    img = cv2.imread(self.root.filename)
    res = cv2.resize(img, dsize=(256,256), interpolation=cv2.INTER_CUBIC)
    gray = cv2.cvtColor(res, cv2.COLOR_BGR2GRAY)
    faces = face_cascade.detectMultiScale(gray, 1.3, 5)
    for x,y,w,h in faces:
        roi_gray = gray[y:y+h, x:x+w]
        roi_color = res[y:y+h, x:x+w]
        eyePairs = eyePair_cascade.detectMultiScale(roi_gray)
    for (ex,ey,ew,eh) in eyePairs:
        roiEyes = roi_color[ey :ey + eh, ex: ex+ ew + 65]
    return roiEyes

def prediccion(self,commit=True):
    image = cv2.cvtColor(self.opencv(self), cv2.COLOR_BGR2RGB)
    image = cv2.resize(image, dsize=(331,331), interpolation=cv2.INTER_CUBIC)
    preds = self.predict(new_model1, image)
    self.plot_preds(np.asarray(image), preds)
    self.MostrarResultados(self)

def MostrarResultados(self,commit=True):
    self.root.deiconify()
    time.sleep(10)
    Imagen = Image.open("resultado.png")
    Imagen_Redux = Imagen.resize((650,650))
    imagen1 = ImageTk.PhotoImage(Imagen_Redux)
    fondo = tk.Label(self.root,image=imagen1).grid(row=2,column=0,columnspan=3)
    self.root.mainloop()

def predict(self,model, img):
    #x = image.img_to_array(img)
    x = np.expand_dims(img, axis=0)
    x = preprocess_input(x)
    preds = new_model1.predict(x)

```

## Referencias bibliograficas

- Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals. *Information Sciences*, 415–416, 190–198. <https://doi.org/10.1016/j.ins.2017.06.027>
- Amato, F., López, A., Peña-Méndez, E. M., Vaňhara, P., Hampl, A., & Havel, J. (2013, January 1). Artificial neural networks in medical diagnosis. *Journal of Applied Biomedicine*. University of South Bohemia. <https://doi.org/10.2478/v10136-012-0031-x>
- Bressem, K. K., Adams, L. C., Erxleben, C., Hamm, B., Niehues, S. M., & Vahldiek, J. L. (2020). Comparing different deep learning architectures for classification of chest radiographs. *Scientific Reports* 2020 10:1, 10(1), 1–16. <https://doi.org/10.1038/s41598-020-70479-z>
- Brown, J. M., Campbell, J. P., Beers, A., Chang, K., Ostmo, S., Chan, R. V. P., ... Chiang, M. F. (2018). Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. In *JAMA Ophthalmology* (Vol. 136, pp. 803–810). American Medical Association. <https://doi.org/10.1001/jamaophthalmol.2018.1934>
- Chen, Z., Fu, H., Lo, W. L., & Chi, Z. (2018). Strabismus Recognition Using Eye-Tracking Data and Convolutional Neural Networks. *Journal of Healthcare Engineering*, 2018, 7692198. <https://doi.org/10.1155/2018/7692198>
- Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., ... Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole

- slide images with convolutional neural networks. In M. N. Gurcan & A. Madabhushi (Eds.), *Medical Imaging 2014: Digital Pathology* (Vol. 9041, p. 904103). SPIE. <https://doi.org/10.1117/12.2043872>
- De Oliveira Simoes, T., Carvalho Souza, J., Sousa De Almeida, J. D., Correa Silva, A., & Cardoso De Paiva, A. (2019). Automatic ocular alignment evaluation for strabismus detection using U-NET and ResNet Networks. In *Proceedings - 2019 Brazilian Conference on Intelligent Systems, BRACIS 2019* (pp. 239–244). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/BRACIS.2019.00050>
- El-Bakry, H. M., & Hamada, M. (2013). Fast diagnosing of pediatric respiratory diseases by using high speed neural networks. In *Proceedings of the International Joint Conference on Neural Networks* (pp. 1–8). IEEE. <https://doi.org/10.1109/IJCNN.2013.6707116>
- Gokul Rajan, V., & Vijayalakshmi, S. (2020). A New Approach for Sclera Segmentation Using Integro Differential Operator. *Journal of Computational and Theoretical Nanoscience*, 17(5), 2330–2335. <https://doi.org/10.1166/jctn.2020.8891>
- Grewal, P. S., Oloumi, F., Rubin, U., & Tennant, M. T. S. (2018, August 1). Deep learning in ophthalmology: a review. *Canadian Journal of Ophthalmology*. Elsevier B.V. <https://doi.org/10.1016/j.jcjo.2018.04.019>
- Hu, H., Zheng, W., Zhang, X., Zhang, X., Liu, J., Hu, W., ... Si, J. (2021). Content-based gastric image retrieval using convolutional neural networks. *International Journal of Imaging Systems and Technology*, 31(1), 439–449. <https://doi.org/10.1002/ima.22470>
- Huang, L., He, X., Fang, L., Rabbani, H., & Chen, X. (2019). Automatic Classification of

- Retinal Optical Coherence Tomography Images with Layer Guided Convolutional Neural Network. *IEEE Signal Processing Letters*, 26(7), 1026–1030. <https://doi.org/10.1109/LSP.2019.2917779>
- IBM. (2020). ¿Qué es Machine Learning? - Argentina | IBM. Retrieved May 14, 2021, from [https://www.ibm.com/ar-es/analytics/machine-learning?p1=Search&p4=43700053692634426&p5=b&cm\\_mmc=Search\\_Google-\\_-1S\\_1S-\\_-LA\\_ISA-\\_-datos%2Bmachine%2Blearning\\_b&cm\\_mmca7=71700000065289299&cm\\_mmca8=kwd-1097762210552&cm\\_mmca9=CjwKCAjwrKr8BRB\\_EiwA7eFapopIBXhO](https://www.ibm.com/ar-es/analytics/machine-learning?p1=Search&p4=43700053692634426&p5=b&cm_mmc=Search_Google-_-1S_1S-_-LA_ISA-_-datos%2Bmachine%2Blearning_b&cm_mmca7=71700000065289299&cm_mmca8=kwd-1097762210552&cm_mmca9=CjwKCAjwrKr8BRB_EiwA7eFapopIBXhO)
- Jung, S. M., Umirzakova, S., & Whangbo, T. K. (2019). Strabismus Classification Using Face Features. In *2019 International Symposium on Multimedia and Communication Technology, ISMAC 2019*. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ISMAC.2019.8836174>
- Kakumanu, P., Makrogiannis, S., & Bourbakis, N. (2007). A survey of skin-color modeling and detection methods. *Pattern Recognition*, 40(3), 1106–1122. <https://doi.org/10.1016/J.PATCOG.2006.06.010>
- Keremany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., ... Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5), 1122–1131.e9. <https://doi.org/10.1016/j.cell.2018.02.010>
- Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks.

- Radiology*, 284(2), 574–582. <https://doi.org/10.1148/radiol.2017162326>
- Lavric, A., & Valentin, P. (2019). KeratoDetect: Keratoconus Detection Algorithm Using Convolutional Neural Networks. *Computational Intelligence and Neuroscience*, 2019. <https://doi.org/10.1155/2019/8162567>
- Lee, K. F., Chen, Y. L., Yu, C. W., & Wu, C. H. (2020). The Eye Tracking and Gaze Estimation System by Low Cost Wearable Devices. In *2020 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-Taiwan 2020* (pp. 1–2). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICCE-Taiwan49838.2020.9258009>
- Leggett, D. (2010). A Brief History of Eye-Tracking | UX Booth. Retrieved October 22, 2020, from <https://www.uxbooth.com/articles/a-brief-history-of-eye-tracking/>
- Luo, G., Pundlik, S., Tomasi, M., & Houston, K. (2020). Using an Automated Hirschberg Test App to Evaluate Ocular Alignment. *Journal of Visualized Experiments*, 2020(157). <https://doi.org/10.3791/60908>
- Mahdi Haref, Q., Srayyih Al-Maliki, M. N., Taaama, M. M., & Albehadili, H. M. (2020). Arabic digit recognition using robust deep convolution neural network. In *Journal of Physics: Conference Series* (Vol. 1530, p. 012085). Institute of Physics Publishing. <https://doi.org/10.1088/1742-6596/1530/1/012085>
- Mohney, B. G., Greenberg, A. E., & Diehl, N. N. (2007). Age at Strabismus Diagnosis in an Incidence Cohort of Children. *American Journal of Ophthalmology*, 144(3), 467–469. <https://doi.org/10.1016/j.ajo.2007.04.022>
- Omran, M., & AlShemmary, E. N. (2020). Towards Accurate Pupil Detection Based on

- Morphology and Hough Transform. *Baghdad Science Journal*, 17(2).
- Park, S.-W. S. S.-C. D.-Y. (2020). Performance Comparison of Commercial and Customized CNN for Detection in Nodular Lung Cancer. *Journal of Korea Multimedia Society*, 23(6), 729–737. <https://doi.org/10.9717/KMMS.2020.23.6.729>
- Petrov, I., & Minakova, N. (2020). Optimization method for non-cooperative iris recognition task using Daugman integro-differential operator. In *Journal of Physics: Conference Series* (Vol. 1615, p. 12007). IOP Publishing Ltd. <https://doi.org/10.1088/1742-6596/1615/1/012007>
- Pitha, I., & Tsai, L. M. (2020). The Eye & Ocular Adnexa. In G. M. Doherty (Ed.), *Current Diagnosis & Treatment: Surgery, 15e*. New York, NY: McGraw Hill LLC. Retrieved from <http://accessmedicine.mhmedical.com/content.aspx?aid=1171283271>
- Qanat, A. S., Alsuheili, A., Alzahrani, A., Faydhi, A. A., Albadri, A., & Alhibshi, N. (2020). Assessment of Different Types of Strabismus Among Pediatric Patients in a Tertiary Hospital in Jeddah. *Cureus*, 12(12). <https://doi.org/10.7759/cureus.11978>
- Ropper, A. H., Samuels, M. A., Klein, J. P., & Prasad, S. (2019a). Disorders of Ocular Movement and Pupillary Function. In *Adams and Victor's Principles of Neurology, 11e*. New York, NY: McGraw-Hill Education. Retrieved from <http://accessmedicine.mhmedical.com/content.aspx?aid=1162591153>
- Ropper, A. H., Samuels, M. A., Klein, J. P., & Prasad, S. (2019b). Disturbances of Vision. In *Adams and Victor's Principles of Neurology, 11e*. New York, NY: McGraw-Hill Education. Retrieved from

- <http://accessmedicine.mhmedical.com/content.aspx?aid=1162590888>
- Samuels, B. (2020). Ocular Disorders. In F. R. Amthor, A. B. Theibert, D. G. Standaert, & E. D. Roberson (Eds.), *Essentials of Modern Neuroscience*. New York, NY: McGraw Hill. Retrieved from <http://accessmedicine.mhmedical.com/content.aspx?aid=1174247652>
- Santos, J., & Frango, I. (2020). Generating photorealistic images of people's eyes with strabismus using Deep Convolutional Generative Adversarial Networks. In *2nd International Conference on Electrical, Communication and Computer Engineering, ICECCE 2020* (pp. 1–4). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICECCE49384.2020.9179425>
- Saraiva, A. A., Fonseca Ferreira, N. M., De Sousa, L. L., Costa, N. C., Sousa, J. V. M., Santos, D. B. S., ... Soares, S. (2019). Classification of images of childhood pneumonia using convolutional neural networks. In *BIOIMAGING 2019 - 6th International Conference on Bioimaging, Proceedings; Part of 12th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2019* (pp. 112–119). <https://doi.org/10.5220/0007404301120119>
- Sarraf, S., & Tofighi, G. (2016). Classification of Alzheimer's Disease using fMRI Data and Deep Learning Convolutional Neural Networks. Retrieved from <http://arxiv.org/abs/1603.08631>
- Savojardo, C., Bruciaferri, N., Tartari, G., Martelli, P. L., & Casadio, R. (2020). DeepMito: accurate prediction of protein sub-mitochondrial localization using convolutional neural networks. *Bioinformatics*, 36(1), 56–64.

- <https://doi.org/10.1093/bioinformatics/btz512>
- Selvam, P. N., & Krishna GVT, G. (2013). *Volume 5 Issue 4, April 2016 www.ijsr.net Licensed Under Creative Commons Attribution CC BY Recycle of E-Waste in Concrete. International Journal of Science and Research (IJSR) ISSN*. Retrieved from [www.ijsr.net](http://www.ijsr.net)
- Shrivastava, A., Gupta, A., & Girshick, R. (2016). Training Region-Based Object Detectors With Online Hard Example Mining.
- Sik - Ho Tsang. (2018). Review: DenseNet — Dense Convolutional Network (Image Classification) | by Sik-Ho Tsang | Towards Data Science. Retrieved October 30, 2021, from <https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803>
- Smaida, M., & Yaroshchak, S. (2020). Bagging of Convolutional Neural Networks for Diagnostic of Eye Diseases. *Undefined*.
- Sokolova, A. D., Kharchevnikova, A. S., & Savchenko, A. V. (2018). Organizing multimedia data in video surveillance systems based on face verification with convolutional neural networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10716 LNCS, pp. 223–230). Springer Verlag. [https://doi.org/10.1007/978-3-319-73013-4\\_20](https://doi.org/10.1007/978-3-319-73013-4_20)
- Su, Z., Yu, S., Chu, J., Zhai, Q., Gong, J., & Fan, H. (2020). A novel architecture: Using convolutional neural networks for Kansei attributes automatic evaluation and labeling. *Advanced Engineering Informatics*, 44, 101055.

- <https://doi.org/10.1016/j.aei.2020.101055>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... Rabinovich, A. (2015). Going Deeper With Convolutions.
- Tamkin, A., Usiri, I., & Fufa, C. (2017). Deep CNNs for Diabetic Retinopathy Detection. Retrieved from <http://cs229.stanford.edu/proj2016/report/tamkinusirifufu-Deep CNNs for diabetic retinopathy-report.pdf>
- Thakoor, K. A., Li, X., Tsamis, E., Sajda, P., & Hood, D. C. (2019). Enhancing the Accuracy of Glaucoma Detection from OCT Probability Maps using Convolutional Neural Networks. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (pp. 2036–2040). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/EMBC.2019.8856899>
- Wilson, P. I., & Fernandez, J. (2006). FACIAL FEATURE DETECTION USING HAAR CLASSIFIERS \*.
- Zerouaoui, H., & Idri, A. (2022). Deep hybrid architectures for binary classification of medical breast cancer images. *Biomedical Signal Processing and Control*, 71, 103226. <https://doi.org/10.1016/J.BSPC.2021.103226>
- Zheng, Y., Fu, H., Li, R., Lo, W.-L., Chi, Z., Feng, D., ... Wen, D. (2019). Intelligent Evaluation of Strabismus in Videos Based on an Automated Cover Test. *Applied Sciences*, 9(4), 731. <https://doi.org/10.3390/app9040731>
- Zolkifli, N. S., & Nazari, A. (2020). Tracing of Strabismus Detection Using Hough Transform. In *2020 IEEE Student Conference on Research and Development, SCOREd*

2020 (pp. 313–318). Institute of Electrical and Electronics Engineers Inc.  
<https://doi.org/10.1109/SCORed50371.2020.9250949>