

Trabajo Fin de Máster

Máster En Subespecialidades Oftalmológicas

Glaucoma

Curso 2021-2022

Universidad de Valladolid

Instituto Universitario de Oftalmobiología Aplicada



Universidad de Valladolid



Análisis crítico de los sistemas de inteligencia artificial entrenados para el diagnóstico de glaucoma.

Autora: Deborah N. Caamaño Diaz

Tutor: Dr. José Carlos Pastor Jimeno



ÍNDICE

Resumen.....	3
Abstract.....	4
Introducción.....	5
<i>Glaucoma.....</i>	<i>5</i>
Inteligencia artificial.....	8
<i>Machine Learning (ML).....</i>	<i>9</i>
<i>Inteligencia artificial aplicada para el diagnóstico y la progresión del glaucoma.....</i>	<i>10</i>
Justificación.....	11
Hipótesis.....	12
<i>Objetivo General:.....</i>	<i>12</i>
<i>Objetivos Específicos:.....</i>	<i>12</i>
Materiales y métodos.....	13
<i>Diseño del Estudio.....</i>	<i>13</i>
<i>Recopilación de datos.....</i>	<i>13</i>
Resultados.....	14
<i>Resultados bibliométricos.....</i>	<i>14</i>
<i>Resultados de contenido.....</i>	<i>14</i>
Discusión.....	16
<i>Bases de Datos.....</i>	<i>16</i>
<i>Extracción de características.....</i>	<i>17</i>
<i>Desempeño frente a los humanos.....</i>	<i>18</i>
Limitaciones.....	18
Conclusión.....	19
Bibliografía.....	20



Agradecimientos.....	25
Material Complementario.....	26
<i>Figura 1. Diagrama de flujo de la selección de los estudios con la consulta de investigación y los criterios de inclusión.</i>	<i>26</i>
<i>Tabla 1. Resumen de los estudios revisados.</i>	<i>27</i>
<i>Tabla 2. Sistemas comercializados.....</i>	<i>32</i>
Anexos.....	33
.....	33



Resumen.

Objetivo: Esta revisión bibliográfica tiene como objetivo comparar la eficacia diagnóstica de los algoritmos de inteligencia artificial entrenados con imágenes de tomografía de coherencia óptica o con más de una prueba diagnóstica para la detección de glaucoma frente aquellos entrenados solo con imágenes de la papila óptica. **Materiales y métodos:** Se realizó una búsqueda en las bases de datos electrónicas *PubMed*, *Web of Science*, *Scopus*, combinando las palabras claves: *Artificial Intelligence*, *Glaucoma*, *Detection*, *Diagnosis*, *Machine Learning*, *Deep Learning*. Fueron seleccionados aquellos artículos sobre entrenamiento de algoritmos para diagnóstico de glaucoma, en inglés, disponibles en texto completo, que expusieron valores de sensibilidad y especificidad. Se excluyeron revisiones sistemáticas y/o metaanálisis, cartas al editor, documentos de conferencias o congresos, estudios sobre predicción y progresión, diagnóstico de otras enfermedades oculares en conjunto con glaucoma o diferentes a glaucoma. **Resultados:** La búsqueda electrónica arrojó 136 registros, después de excluir 24 duplicados, se examinaron los títulos y resúmenes de 112 artículos donde 62 publicaciones no cumplieron los criterios de inclusión. Luego se realizó una revisión de texto completo de 50 artículos donde se excluyeron otros 15 artículos y finalmente se realizó la recopilación de datos de 35 artículos. **Conclusiones:** Se concluyó que no hay diferencias en los resultados del entrenamiento de los algoritmos sea utilizando imágenes de la papila; tomografía de coherencia óptica; o pruebas combinadas. Estos sistemas tienen la capacidad de diferenciar una papila normal de una glaucomatosa. Sin embargo, aún no son efectivos para estadios iniciales. Se consideran accesibles, de bajo costo. La capacidad de los algoritmos de examinar una imagen píxel a píxel supera la percepción visual del ojo humano y también la capacidad de atención.

Palabras claves: inteligencia artificial, glaucoma, diagnóstico, *deep learning*, *machine learning*.



Abstract.

Objective: This literature review aims to compare the diagnostic efficacy of artificial intelligence algorithms trained with optical coherence tomography images or with more than one diagnostic test for the detection of glaucoma against those trained only with images of the optic papilla. **Materials and methods:** A search was carried out in the electronic databases *PubMed*, *Web of Science*, and *Scopus* combining the keywords: *Artificial Intelligence*, *Glaucoma*, *Detection*, *Diagnosis*, *Machine Learning*, and *Deep Learning*. Articles on training algorithms for glaucoma diagnosis, in English, available in full text, were selected, which presented values of sensitivity and specificity. We excluded systematic reviews and/or meta-analyses, letters to the editor, conference or conference papers, studies on prediction and progression, and diagnosis of other eye diseases in conjunction with glaucoma or other than glaucoma. **Results:** The electronic search yielded 136 records, after excluding 24 duplicates, we examined the titles and abstracts of 112 articles where 62 publications did not meet the inclusion criteria. A full-text review of 50 articles was then conducted where another 15 articles were excluded and finally, data collection of 35 articles was performed. **Conclusions:** It was concluded that there are no differences in the results of the training of the algorithms either using images of the papilla; optical coherence tomography; or combined tests. These systems can differentiate a normal papilla from a glaucomatous one. However, they are not yet effective in the early stages. They are considered accessible, low-cost. The ability of algorithms to examine an image pixel by pixel exceeds the human eye's visual perception and attention span.

Keywords: artificial intelligence, glaucoma, diagnosis, deep learning, machine learning.



Introducción.

Glaucoma.

Hoy se considera como glaucoma a un conjunto de enfermedades que tienen en común el presentar una neuropatía óptica, crónica progresiva, cuyo origen aún no está bien establecido, asociado a múltiples factores entre ellos vasculares y genéticos y sobre todo al aumento de la presión intraocular y que conduce a la pérdida progresiva e irreversible de las fibras del nervio óptico y por lo tanto a la ceguera. El 11% de todos los casos de ceguera total están producidos por el glaucoma, posicionándose como la segunda causa de ceguera y la cuarta causa de discapacidad visual moderada y severa en adultos de 50 años y mayores en todo el mundo. Se estima que habrá un aproximado de 111.5 millones de afectados para el año 2040 (1), convirtiéndose en un problema significativo de salud pública. El glaucoma se caracteriza por un inicio generalmente asintomático e insidioso, y con frecuencia es diagnosticado en estadios evolucionados con la existencia de cambios ya irreversibles en la visión. Se estima que la mitad de los individuos que padecen de glaucoma aún no recibe tratamiento.

El diagnóstico del glaucoma se basa en identificar la degeneración progresiva en la cabeza del nervio óptico que implica la pérdida de fibras de las células ganglionares con cambios característicos correspondientes a las pérdidas del campo visual (2), pudiendo estar alterados o no los niveles de tensión intraocular. En la actualidad la visualización de la papila óptica, la campimetría y la tomografía de coherencia óptica (OCT) constituyen las pruebas básicas para el diagnóstico y seguimiento del glaucoma.

En el examen de la papila se debe tomar en cuenta el tamaño de la papila y la relación vertical y horizontal de la excavación, la coloración, la inclinación, la salida y el calibre vascular, la presencia de adelgazamiento de los vasos, el grosor del anillo neuroretiniano, la existencia de escotaduras, el grado y el tipo de atrofia peripapilar, y las características de la lámina cribosa (3).



La perimetría automatizada es una prueba subjetiva fundamental para el diagnóstico y manejo del glaucoma y se considera el estándar para diagnóstico del glaucoma por su alta sensibilidad para detectar pérdidas de visión (2). Los cambios en el campo visual representan el resultado del daño funcional más importante en el glaucoma, pero pueden ser alterados fácilmente al ser una medida dependiente, casi en su totalidad, de la atención del paciente, así como de la agudeza visual, que a menudo se encuentra disminuida en la población en riesgo de glaucoma debido a comorbilidades como deterioro cognitivo y catarata, sin olvidar que las primeras fases del glaucoma no presentan síntomas ni cambios en el campo visual. Y aun teniendo buenos resultados necesita, además, una cuidadosa interpretación por parte del oftalmólogo.

La tomografía de coherencia óptica (OCT) es una prueba de imagen no invasiva que permite, entre otras cosas, cuantificar de manera objetiva el espesor de las capas retinianas en tiempo real. En el glaucoma, el grosor de la capa de fibras nerviosas de la retina (RNFL) y el grosor de la capa de células ganglionares combinada con la capa plexiforme interna (GCIPL) se consideran biomarcadores útiles para el diagnóstico, ya que en estas capas se evidencia un adelgazamiento significativo a medida que avanza la enfermedad. Las medidas de la RNFL en la región inferior tienen la mejor capacidad para identificar ojos con glaucoma en un estadio moderado, con una sensibilidad del 64-79% y una especificidad del 90% o superior (4). Además, la estimación del volumen macular y del espesor de la coroides, se consideran también datos de valor para la identificación de los cambios inducidos por el glaucoma (5). Aunque esta afirmación todavía genera controversias (6). Independientemente de los valores reportados, es evidente que en estos momentos es una prueba complementaria esencial para el diagnóstico y el seguimiento del glaucoma.



El glaucoma es una enfermedad progresiva, generalmente de evolución lenta, y en el momento del diagnóstico puede encontrarse en diferentes etapas, siendo las primeras fases las más difíciles de reconocer. En las últimas décadas se han descrito y estandarizado diferentes métodos de estadificación para categorizar la gravedad del daño funcional y estructural glaucomatoso. Dentro de los sistemas más utilizados para la clasificación del daño funcional se describen: el método de Hodapp-Parrish-Anderson, que se basa en el valor de la desviación media, la cantidad de puntos con pérdidas de sensibilidad en el mapa de desviación del patrón, teniendo en cuenta los defectos al punto de fijación para clasificar en inicial, moderado o avanzado (7); el modelo de Mills et al. (8) que es una evolución del método derivado del Estudio de Intervención Avanzada para el Glaucoma (*Advanced Glaucoma Intervention Study*) en el cual se utiliza el número de puntos con pérdidas de sensibilidad en varias áreas del campo visual central, y proporciona una puntuación que oscila entre 0 y 20 (9) y el sistema de estadificación del glaucoma 2 (*Glaucoma Staging System*), basado en los valores de desviación media y desviación estándar del patrón trazados en un gráfico x-y (10).

En cuanto a la categorización del daño estructural, se dispone actualmente de los siguientes métodos: El método propuesto por Elbendary y Elal, Sistema global de estadificación del glaucoma (*Global Glaucoma Staging System*), que se basa en los valores promedio de RNFL en el OCT (11), y el sistema de estadificación del glaucoma por OCT (*OCT Glaucoma Staging System*) que se basa en los valores del grosor de la capa de fibras nerviosas del cuadrante inferior (12).



Inteligencia artificial.

La inteligencia artificial (IA) se define como la capacidad de un sistema de interpretar datos externos, aprender de estos y lograr tareas específicas a partir de lo aprendido, lo que se considera una adaptación flexible (13). Según la propuesta de varios autores se diferencian cuatro tipos de IA, donde se encuentran los sistemas que piensan como humanos (por ejemplo, las Redes Neuronales Artificiales), los sistemas que actúan como humanos (por ejemplo, la visión por computadora y robótica), los sistemas que piensan racionalmente (por ejemplo, los sistemas expertos) y los sistemas que actúan racionalmente (por ejemplo, los agentes inteligentes) (14).

El término IA aparece por primera vez en 1956 durante una conferencia en Dartmouth, aunque, anteriormente en 1943 M. Warren y P. Walter habían dado a conocer su modelo de neuronas artificiales considerado el primer trabajo en el campo de la IA. En 1959 Rosenblatt introduce el “perceptron”, un sistema capaz de identificar patrones geométricos, la unidad de donde nacen las redes neuronales. También en 1959 se describe la rama de la inteligencia artificial conocida como *Machine Learning* (ML), creada por Arthur Samuel.

En la década de los 1960-1970 se crea el sistema MYCIN, considerado uno de los primeros contactos de la IA con la medicina, éste basado en la estructura del primer sistema experto, llamado Dendral. Años más tarde, para 1982 John J. Hopfield describe el algoritmo *Backpropagation*, un sistema de red neuronal recurrente formado por procesadores interconectados. Más adelante en 1989 aparecen las redes de neuronas convolucionales (15).

La introducción de la IA en el campo de la oftalmología surge en la década de los 90, debido al desarrollo de nuevas técnicas informáticas y al gran acumulo de datos, rápidos y complejos, difíciles de procesar de la manera tradicional, Estos conjuntos de datos son conocidos como *Big Data* (16). Inicialmente la IA en oftalmología se emplea para el diagnóstico del glaucoma, la retinopatía diabética y el queratocono (17).



Machine Learning (ML)

Machine learning (ML) es una rama de la IA enfocada en crear sistemas que aprenden de forma automática mediante el uso de datos y algoritmos, se utiliza ampliamente en el campo del diagnóstico médico. Dentro de las ML se describen subconjuntos como el *Deep Learning* (DL) que se basa en asimilar representaciones de datos. Un ejemplo es la representación de imágenes en píxeles, a su vez del DL se considera como un subcampo las Redes Neuronales Artificiales (RNA) que consisten en sistemas que imitan la estructura neuronal cerebral a través de unidades conectadas entre sí para transmitir señales. Precisamente se debe al surgimiento de estas estructuras el reciente impacto que está teniendo la IA en las ciencias médicas. Dentro de los sistemas de RNA sobresalen los modelos de Redes Neuronales Convolucionales (CNN) cuya estructura de neuronas artificiales simula los campos receptivos de la corteza visual primaria del cerebro biológico (18).

La fuente de interés en estos sistemas computacionales consiste en su capacidad de aprendizaje, la posibilidad de resolver tareas de forma óptima y sus múltiples aplicaciones. Para que estos sistemas puedan realizar dichas tareas deben ser entrenados con un gran número de datos a partir de los cuales podrán reconocer y extraer características de interés. Este proceso de validación puede ejecutarse de tres maneras; aprendizaje Supervisado, es el más utilizado, consiste en aprender de los datos de entrada de objetivos conocidos o un conjunto de anotaciones, a menudo son dadas por humanos; aprendizaje no Supervisado, consiste en encontrar, sin ayuda externa, transformaciones interesantes de los datos de entrada, con la finalidad de visualización, compresión o eliminación de ruido de datos; y aprendizaje parcialmente supervisado o por refuerzo donde existe una pequeña cantidad de los datos clasificados) (19).



Inteligencia artificial aplicada para el diagnóstico y la progresión del glaucoma.

El desarrollo de la tecnología en el procesamiento de imágenes ha favorecido la detección de enfermedades incluso sin la intervención de expertos. Actualmente, los algoritmos basados en aprendizaje profundo son un campo de investigación activo por su notable capacidad de generar y aprender de forma automática de las características más complejas desde los datos de entrada. Particularmente, esta facultad les permite demostrar un alto rendimiento en la detección de objetos o en la detección e identificación de los rasgos relevantes de los datos sin el requisito de extracción manual de características (16).

La aplicación de la inteligencia artificial para la detección y progresión del glaucoma abarca diferentes modelos. Con el desarrollo de las Redes Neuronales Convolucionales el diagnóstico del glaucoma se ha convertido en el área más explorada para utilizar estos sistemas. En años recientes se han entrenado algoritmos con fotografías de fondo de ojo para diferenciar papilas ópticas sanas de las glaucomatosas utilizando las medidas de relación excavación papila (20), combinando los indicadores proporción de excavación papila, verificando la regla Inferior-Superior-Nasal-Temporal (ISNT) y la escala de probabilidad de daño de la papila (DDLS, *Disc Damage Likelihood Scale*), y la extracción de la región de interés (21). Otros algoritmos se han entrenado para identificar daño glaucomatoso en exploraciones de OCT (22) y cambios funcionales en el campo visual (23) y algunos con combinaciones de estos datos (24).



Justificación.

La aplicación de la IA en oftalmología es actualmente un campo de estudio destacado, y en los últimos años es evidente el aumento de las investigaciones sobre cómo aplicar los algoritmos, ya entrenados para el análisis automatizado de imágenes, para mejorar la detección del glaucoma. Los nuevos modelos basados en estructuras de aprendizaje profundo están reemplazando a los algoritmos tradicionales basados en estructuras de aprendizaje automático, en muchos de los casos han superado a los clasificadores tradicionales. El objetivo principal es desarrollar una herramienta eficaz, rentable y de fácil manejo para el diagnóstico del glaucoma, todo esto a pesar de la evidente dependencia de una cuidadosa evaluación, necesaria para un diagnóstico eficaz de la enfermedad.

La IA es un promete ser un instrumento de gran utilidad, y se espera que los modelos en desarrollo tendrán un impacto positivo en la evaluación del paciente con glaucoma en un futuro no lejano, pudiendo así suponer un paso a la evolución en la práctica de la oftalmología. Muchos estudios han demostrado la eficacia de la IA con resultados favorables, como consecuencia, se ha despertado un especial interés en analizar cómo se han entrenados estos sistemas y cuáles han sido las pautas para regular dicho entrenamiento. Es necesario conocer a fondo la estructuración, el proceso de clasificación y el de distribución de los conjuntos de datos que utilizan estos sistemas. Esta revisión surge de la necesidad de identificar los potenciales desafíos y las posibles limitaciones en el proceso de construcción de las bases de datos y en consecuencia la validación de los algoritmos.

El presente trabajo fin de máster tiene como propósito analizar la información existente para determinar si es confiable la aplicación de estos sistemas para el diagnóstico de glaucoma.



Hipótesis.

Los algoritmos de inteligencia artificial entrenados con más de una prueba diagnóstica para la detección de glaucoma tienen mayor sensibilidad y especificidad que aquellos entrenados solo con imágenes de la papila óptica.

Objetivo General:

- Comparar la eficacia diagnóstica de los algoritmos de inteligencia artificial entrenados con imágenes de tomografía de coherencia óptica o con más de una prueba diagnóstica para la detección de glaucoma frente aquellos entrenados solo con imágenes de la papila óptica.

Objetivos Específicos:

1. Analizar la estructura de las bases de datos utilizadas para el entrenamiento y validación de los algoritmos, cuales se basan exclusivamente en imágenes de papila, de OCT u otro tipo de imagen.
2. Examinar el proceso de clasificación de las imágenes utilizadas para las bases de datos.
3. Comparar los resultados de sensibilidad y especificidad de los algoritmos entrenados con OCT o más de una prueba diagnóstica frente aquellos entrenados con fotografías de la papila.



Materiales y métodos.

Diseño del Estudio.

Esta revisión de la literatura tiene como objetivo el análisis de los artículos sobre la implementación de estructuras de inteligencia artificial para el diagnóstico de glaucoma. La identificación de los artículos se basó en una búsqueda en las bases de datos *PubMed*, *Web of Science*, y *Scopus* combinando las palabras claves: *Artificial Intelligence*, *Glaucoma*, *Detection*, *Diagnosis*, *Machine Learning*, *Deep Learning*, hasta fecha del 31 de diciembre del año 2021, realizada por un único investigador. Fueron seleccionados aquellos artículos resultantes de la búsqueda y en la opción de artículos similares sobre validación de algoritmos para el diagnóstico de glaucoma, en idioma inglés, disponibles en texto completo y que expusieron en sus resultados valores de sensibilidad y especificidad. Se excluyeron las publicaciones de tipo revisiones sistemáticas y/o metaanálisis, cartas al editor y documentos de conferencias o congresos. También se excluyeron los estudios con el objetivo de predicción del glaucoma, detección de progresión del glaucoma, diagnóstico de otras enfermedades oculares en conjunto con glaucoma o diferenciaron el glaucoma de otras enfermedades oculares.

Recopilación de datos.

Los datos fueron recopilados durante la revisión del texto completo, se extrajeron las siguientes variables: autor, año de publicación, país, tipo de imagen utilizada, total de imágenes, total de imágenes con glaucoma, nivel de los profesionales que realizaron la clasificación de los datos, algoritmos utilizados, los valores de sensibilidad y especificidad. Posteriormente a la extracción de datos, los estudios se clasificaron según el tipo de imagen utilizada en foto de papila, OCT, campo visual y combinaciones; luego se clasificaron según el tipo de algoritmo utilizado y la combinación con clasificadores si existiere.



Resultados

Resultados bibliométricos.

La búsqueda de literatura en las bases de datos electrónicas arrojó 100 registros en *PubMed*, 19 registros en *Scopus* y 17 registros en *Web of Science*. Después de excluir 24 duplicados, se analizaron inicialmente los títulos y resúmenes de 112 publicaciones de las que fueron excluidas otras 26 publicaciones por no implicar glaucoma, 13 por no implicar el diagnóstico de glaucoma, 20 por tratarse de revisiones o de metaanálisis y 3 por ser publicaciones diferentes de artículos de revista. Así pues, se sometieron a una revisión de texto completo un total de 50 artículos, de los cuales se excluyeron 9 por tratarse de publicaciones relacionadas con predicción o progresión del glaucoma, 2 por no tener acceso al texto completo y 4 por incluir otras patologías además del glaucoma. Finalmente, se incluyeron 35 publicaciones en total para el análisis y la extracción de los datos, de las cuales una corresponde al año 2017, 6 al año 2018, 9 al año 2019, 8 al año 2020 y 11 al año 2021.

Resultados de contenido.

Se realizó la revisión de texto completo de 35 publicaciones, se encontró que 21 publicaciones utilizaron fotografías de la papila (33-53), 10 usaron imágenes de OCT (29,54-62), uno utilizó imágenes de campo visual combinadas con imágenes de OCT (55), en 2 de los estudios combinaron imágenes de la papila y de OCT (56,57) y uno combinó fotografías de papila con imágenes de campo visual (58). Se revisó la estructura de las bases de datos utilizadas para el entrenamiento y validación de los diferentes modelos encontrando que 22 usaron bases de datos propias, 4 usaron bases de datos públicas, 7 usaron bases de datos propias combinadas con bases de datos públicas y 2 de los estudios revisados no detallan el origen de las bases de datos utilizadas.

Se observó que 8 de los estudios utilizaron bases de datos compuestas por menos de mil imágenes, en 18 sus bases de datos contenían más de mil imágenes, sin alcanzar las diez mil imágenes, y en 9 de los estudios se analizaron con más de diez mil imágenes.



La base de datos con mayor contenido contó con 269.601 fotografías de la papila (32), sin embargo, los resultados son comparables con aquellas de menor cantidad de imágenes. En cuanto a la clasificación de las imágenes, en glaucoma o no glaucoma, 14 publicaciones especifican que dicha tarea fue realizada por expertos en glaucoma, 11 expresan haber tenido un clasificador cualificado, 3 de las publicaciones especifican que fue realizada por oftalmólogos y 7 no ofrecen esta información. En relación con el tipo de algoritmo utilizado, un total de 30 de los estudios revisados utilizaron CNN como el principal clasificador, los 5 restantes utilizaron otro tipo de clasificador.

Referente a los resultados según el tipo de imagen utilizada, los valores obtenidos con imágenes de la papila oscilan desde 76.6% de precisión (ACC) (41) hasta el 100% de ACC; y 96.2% de sensibilidad y 97.7% de especificidad (32), hasta y un 98% de sensibilidad y especificidad (39). Los estudios que tomaron en cuenta la identificación de papilas glaucomatosas según el estadio de la enfermedad obtuvieron mejores resultados diferenciando las papilas con glaucoma avanzado, con valores de hasta 90.9% de sensibilidad y 95.8% de especificidad, por el contrario, para papilas con glaucoma inicial reportaron valores de 83.8% de sensibilidad y 75.3% de especificidad (28). En cuanto a los estudios revisados que utilizaron imágenes de OCT los resultados oscilan de 88% de ACC hasta 97.35% ACC; y 81% de sensibilidad y 95% de especificidad (22), de 94.7% sensibilidad y 100% de especificidad (52). Con referencia a los resultados de los algoritmos entrenados con más de una prueba diagnóstica, la combinación de campo visual más fotografía de la papila alcanzó valores de área bajo la curva (AUC) de 0.801 (95% CI: 0.757-0.845) (58); la combinación de campo visual más OCT obtuvo una AUC de 0.950 (0.931-0.968) (55) y el resultado superior para la combinación de fotografía de la papila y OCT fue de 0.945 (95% CI:0.874–0.980) (56).



Discusión.

Este estudio ha pretendido averiguar cómo se han entrenados los diferentes sistemas para el diagnóstico de glaucoma mediante fotografías de papila y OCT de papila, campo visual y combinaciones de estas pruebas. Actualmente mediante el uso de sistemas de inteligencia artificial se intenta desarrollar una herramienta de bajo costo y de fácil aplicación para la detección precoz del glaucoma (59). En general, los estudios exponen resultados excelentes en cuanto al rendimiento diagnóstico utilizando tanto fotos de la papila óptica como con imágenes de OCT. Algunos estudios presentan, además, resultados de diferentes algoritmos, entrenados con las mismas bases de datos, para distinguir papilas sanas de las papilas glaucomatosas (27,29,38,42,43,48,53,54,57).

Bases de Datos.

A pesar del excelente desempeño, estos modelos se ven influenciados directamente por las bases de datos utilizadas durante el entrenamiento, ya sea por la cantidad, la calidad y la correcta clasificación de los datos que las componen. Entrenar algoritmos con bases de datos pequeñas, menos de mil imágenes, puede suponer un problema si no se encuentra bien codificados, en estos casos se recurre a la técnica de generación de datos artificiales, aumentando así la muestra mediante la variación de los datos originales, como la posición y la orientación, sin embargo, se puede alcanzar resultados razonables aun con la relativa falta de datos (19).

Respecto a las bases de datos compuestas por más de una prueba diagnóstica, se observó con mayor frecuencia la combinación de fotos de la papila y OCT (56,57), y con menor frecuencia la combinación incluyó el campo visual (55,58), sin importar el tipo de combinación utilizada los resultados son semejantes con los algoritmos entrenados con una sola prueba diagnóstica.



Es necesario resaltar que la identificación de las imágenes realizadas por sujetos no especialistas o expertos en otras áreas diferentes a glaucoma (29) motiva a la desconfianza en el proceso de clasificación. Se observó con mayor frecuencia la selección de los sujetos con glaucoma basada en el historial clínico, en algunos de los estudios las normas para la selección de los individuos con glaucoma no fueron claramente descritas, o no se explica en lo absoluto el origen de los datos y algunos son de difícil comprensión. En tres de las publicaciones revisadas, las imágenes fueron clasificadas según los estadios de la enfermedad (35-37), mientras las demás realizaron los análisis con bases de datos que contenían diferentes grados de severidad (30,32,36).

Extracción de características.

En la mayoría de los modelos entrenados la tarea de segmentación de la papila óptica y cálculo del diámetro vertical de la excavación es la más utilizada, a menudo teniendo como parámetro un diámetro de 0,7 o superior, lo que supone una ventaja frente a las cuantiosas imágenes con glaucomas avanzados presentes, principalmente en las bases de datos públicas. Sin embargo, el alto rendimiento en la identificación de casos avanzados no supone una contribución a la práctica diaria, por lo que sería interesante debatir cual puede ser el valor real de estos sistemas en la práctica clínica.

En cuanto a al cálculo de otras medidas útiles para el diagnóstico de glaucoma se toman en cuenta con menor frecuencia (35). Los principales errores en la clasificación realizada por los algoritmos se presentaron como falsos positivos en los casos de miopía elevada y otras causas de palidez de la papila (30), los falsos negativos fueron más frecuentes en imágenes con más de una patología, por ejemplo, con retinopatía diabética y glaucoma, oclusiones vasculares y glaucoma (60).

Todo esto podría mejorarse con la creación de bases de datos exclusivas para glaucoma correctamente clasificadas por estadios, principalmente aquellos que suponen mayor dificultad al momento del diagnóstico, con la finalidad de conseguir un entrenamiento enfocado exclusivamente a la identificación del glaucoma, lo cual pudiera ser beneficioso incluso en imágenes que contengan otras patologías. En adición, sería importante tomar



en cuenta otras características sugerentes de glaucoma, tales como, como la coloración, la presencia de escotaduras y también las características de la lámina cribosa, datos que podrían sumar valor a los resultados.

Desempeño frente a los humanos.

Sin duda alguna, los modelos de aprendizaje profundo son cada vez más precisos en la tarea del análisis y reconocimiento de imágenes, su fortaleza se basa en que, a diferencia de versiones anteriores de clasificadores, cuentan con capas más profundas; especializadas y numerosas (43), dando como resultado un procesamiento óptimo de la información, volviéndose más independientes y eficientes. En ocasiones han demostrado una capacidad comparable con los oftalmólogos (40,55,58), llegando a superar la habilidad diagnóstica de los oftalmólogos en formación, sin embargo, presentan valores inferiores frente a los expertos en glaucoma, siendo más notorio en los estadios moderados e iniciales de la enfermedad. Esto no significa que deban ser considerados como una herramienta del todo confiable, más bien, pueden ser utilizados como una herramienta accesoria en la evaluación oftalmológica habitual o para labores de cribado frente a poblaciones de riesgo.

En resumen, las publicaciones revisadas demuestran el rendimiento de la aplicación de sistemas de inteligencia artificial para la detección de características específicas que pudieran indicar glaucoma en imágenes de fondo de ojo o imágenes de OCT (36,51), los resultados se comparan favorablemente con la experiencia de los oftalmólogos (58), demostrando que pueden diferenciar una papila normal de una papila glaucomatosa.

Limitaciones.

La limitación más importante de este análisis es el número de las fuentes revisadas, sin embargo, la cantidad estudiada puede considerarse representativa y razonable. En segundo lugar, puede considerarse una limitante la exclusión de publicaciones dirigidas al diagnóstico de otras enfermedades oculares en conjunto con glaucoma o que diferenciaran el glaucoma de otras enfermedades oculares.



Conclusión.

En los resultados de esta revisión se comparó la eficacia diagnóstica de los algoritmos de IA entrenados con más de una prueba diagnóstica para la detección de glaucoma frente aquellos entrenados solo con una prueba. Se pudo apreciar que no hay diferencias importantes en los resultados, lo que indica que no existen evidencias suficientes que avalen que los datos deban ser obtenidos con herramientas de alto costo. La bibliografía revisada expresa que estos sistemas tienen la capacidad de diferenciar una papila normal de una glaucomatosa.

Los algoritmos de identificación de imagen logran excelentes resultados al reconocer papilas ópticas con glaucomas avanzados y consiguen resultados aceptables en los estadios moderados, sin embargo, aún no son efectivos para estadios iniciales. Para resolver este problema, sería una ventaja la creación de bases de datos exclusivas para glaucoma correctamente clasificadas por estadios, principalmente aquellos que implican mayor dificultad al momento del diagnóstico y que son más difíciles de estandarizar, esto con la finalidad de obtener un entrenamiento enfocado exclusivamente a la identificación del glaucoma, lo cual pudiera ser beneficioso incluso en imágenes que contengan otras patologías. Es importante resaltar que la cantidad de imágenes utilizadas no presenta diferencias en los resultados gracias a las técnicas de generación de datos artificiales.

El desempeño de estos sistemas puede compararse de manera favorable con el de los oftalmólogos sin llegar a superar a los expertos en glaucoma, aunque la capacidad de los algoritmos de examinar una imagen píxel a píxel demuestra ser superior a la percepción visual del ojo humano y también la capacidad de atención, proporcionándole la habilidad de identificar con mayor facilidad características específicas en los datos provistos.



Bibliografía

1. Tham Y, Li x, Wong T, Quigley H, Aung T, Cheng C. Global prevalence of glaucoma and projections of glaucoma burden through 2040. *Ophthalmology*. 2014; 121(11): p. 2081-90.
2. Tsai J, Denniston A, Murray P, Huang J, Aldad T. *Oxford American handbook of ophthalmology*. New York: Oxford University Press; 2011.
3. Castañeda R, Jimenez J, Iriarte MJ. [Open angle glaucoma suspect: Definition, diagnosis and treatment.]. *Rev Mex Oftalmol*. 2014; 88(4): p. 153-60.
4. Jeoung JM, Park KH. Comparison of Cirrus OCT and Stratus OCT on the Ability to Detect Localized Retinal Nerve Fiber Layer Defects in Preperimetric Glaucoma. *Invest Ophthalmol Vis Sci*. 2010; 51(2): p. 938-45.
5. Nouri-Mahdavi K, Weiss R. Detection of Glaucoma Deterioration in the Macular Region with Optical Coherence Tomography: Challenges and Solutions. *Am J Ophthalmol*. 2021; 222: p. 277-84.
6. McCann P, Hogg R, Wrihgt D, McGuinness B, Young I, Kee F, et al. Diagnostic Accuracy of Spectral-Domain OCT Circumpapillary, Optic Nerve Head, and Macular Parameters in the Detection of Perimetric Glaucoma. *Ophthalmol Glaucoma*. 2019; 2: p. 336-45.
7. Hodapp E, Parish II R, Anderson D. *Clinical Decisions in Glaucoma*. St. Louis, MO, USA,: The CV Mosby Comp.; 1993.
8. Mills R, Budenz D, Lee P, Noecker R, Walt J, Siegartel L, et al. Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. *Am. J. Ophthalmol*. 2006; 141: p. 24-30.
9. The Advanced Glaucoma Intervention Study Investigators. Advanced glaucoma intervention study. Visual field test scoring and reliability. *Ophthalmology*. 1994; 101(8): p. 1445-55.
10. Brusini P, Filacorda S. Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. *J. Glaucoma*. 2006; 15: p. 40-6.
11. Elbendary A, Mohamed Helal R. Discriminating ability of spectral domain optical coherence tomography in different stages of glaucoma. *J. Ophthalmol. Saudi*. 2013; 27: p. 19-24.
12. Brusini P. OCT Glaucoma Staging System: A new method for retinal nerve fiber layer damage classification using spectral-domain OCT. *Eye*. 2018; 32: p. 113-19.
13. Kaplan A, Haenlein M. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*. 2019; 62(1): p. 15-20.



14. Russell S, Norvig P. Artificial Intelligence. A Modern Approach. 3rd ed. Pearson , editor. New Jersey: Prentice Hall; 2016.
15. Avila-Tomas J, Mayer-Pujadas M, Quesada-Valera V. [Artificial intelligence and its applications in medicine I: introductory background to AI and robotics]. Aten Primaria. 2020; 52(10): p. 778-84.
16. Kaplan A. Artificial Intelligence, Business and Civilization: Our Fate Made in Machines. 1st ed. New York: Routledge; 2022.
17. Issarti I, Rozema J. Basics of Artificial Intelligence for Ophthalmologists. En Grzybowski A, editor. Artificial Intelligence in Ophthalmology.: Springer; 2021. p. 17-30.
18. Sejnowski TJ. The Deep Learning Revolution. Press M, editor. Cambridge, MA; 2018.
19. Chollet F. Deep learning with python. Shelter Island, NY: Manning Publications Company; 2017.
20. Haleem MS, Han L, Hemert J, Li B. Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: A review. Comput Med Imaging Graph. 2013; 37(7-8): p. 581-96.
21. Li L, Xu M, Liu H, Li Y, Wang X, Jiang L, et al. A Large-Scale Database and a CNN Model for Attention-Based Glaucoma Detection. IEEE Trans Med Imaging. 2020; 39(2): p. 413-24.
22. Thompson AC, Jammal AA, Berchuck S, Mariottoni EB, Medeiros FA. Assessment of a Segmentation-Free Deep Learning Algorithm for Diagnosing Glaucoma From Optical Coherence Tomography Scans. JAMA Ophthalmol. 2020; 138(4): p. 333-39.
23. Dixit A, Yohanna J, Boland M. Assessing Glaucoma Progression Using Machine Learning Trained on Longitudinal Visual Field and Clinical Data. Ophthalmology. 2021; 28(7): p. 1016-26.
24. Silva FS, Vidotti VG, Cremasco F, Dias M, Gomi ES, Costo VP. Sensitivity and specificity of machine learning classifiers for glaucoma diagnosis using Spectral Domain OCT and standard automated perimetry. Arq Bras Oftalmol. 2013; 76(3): p. 170-4.
25. Cerentini A, Welfer D, Cordeiro D'Ornellas M, Pereira C, Dotto G. Automatic Identification of Glaucoma Using Deep Learning Methods. Stud Health Technol Inform. 2017; 245: p. 318-21.
26. Liu S, Graham S, Schulz A, Kalloniatis M, Zangerl B, Cai W, et al. A Deep Learning-Based Algorithm Identifies Glaucomatous Discs Using Monoscopic Fundus Photographs. Ophthalmol Glaucoma. 2018; 1(1): p. 15-22.
27. Ahn J, Kim S, Ahn K, Cho S, Lee K, Kim U. A deep learning model for the detection of both advanced and early glaucoma using fundus photography. PLoS One. 2018; 13(11): p. e0207982.



28. Masumoto H, Tabuchi H, Nakakura S, Ishitobi N, Miki M, Enno H. Deep-learning Classifier With an Ultrawide-field Scanning Laser Ophthalmoscope Detects Glaucoma Visual Field Severity. *J. Glaucoma*. 2018; 27(7): p. 647-52.
29. Christopher M, Belghith A, Bowd C, Proudfoot J, Goldbaum M, Weinreb R, et al. Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs. *Sci Rep*. 2018; 8(1): p. 16685.
30. Li Z, He Y, Keel S, Meng W, Chang R, He M. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology*. 2018; 125(8): p. 1199-1206.
31. Fu H, Cheng J, Xu Y, Zhang C, Wong D, Liu J, et al. Disc-Aware Ensemble Network for Glaucoma Screening From Fundus Image. *IEEE Trans Med Imaging*. 2018; 37(11): p. 2493-2501.
32. Liu H, Li L, Wormstone I, Qiao C, Zhang C, Liu P, et al. Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs. *JAMA Ophthalmol*. 2019; 137(12): p. 1353-60.
33. Rogers T, Jaccard N, Carbonaro F, Lemij H, Vermeer K, Reus N, et al. Evaluation of an AI system for the automated detection of glaucoma from stereoscopic optic disc photographs: the European Optic Disc Assessment Study. *Eye (Lond)*. 2019; 33(11): p. 1791-97.
34. Lee J, Kim Y, Kim J, Park K. Screening Glaucoma With Red-free Fundus Photography Using Deep Learning Classifier and Polar Transformation. *J Glaucoma*. 2019; 28(3): p. 258-64.
35. MacCormick I, Williams B, Zheng Y, Li K, Al-Bander B, Czanner S, et al. Accurate, fast, data efficient and interpretable glaucoma diagnosis with automated spatial analysis of the whole cup to disc profile. *PLoS One*. 2019; 14(4): p. e0209409.
36. Li F, Yan L, Shi J, Chen H, Zhang X, Jiang M, et al. Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs. *Graefes Arch Clin Exp Ophthalmol*. 2020; 258(4): p. 851-67.
37. Ko Y, Wey S, Chen W, Chang Y, Chen M, Chiou S, et al. Deep learning assisted detection of glaucomatous optic neuropathy and potential designs for a generalizable model. *PLoS One*. 2020; 15(5): p. e0233079.
38. Christopher M, Nakahara K, Bowd C, Proudfoot J, Belghith A, Goldbaum M, et al. Effects of Study Population, Labeling and Training on Glaucoma Detection Using Deep Learning Algorithms. *Transl Vis Sci Technol*. 2020; 9(2): p. 27.



39. Ganesh S, Kannayeram G, Karthick A, Muhibbullah M. A Novel Context Aware Joint Segmentation and Classification Framework for Glaucoma Detection. *Comput Math Methods Med.* 2021; 2021: p. 2921737.
40. Li Z, Guo C, Lin D, Nie D, Zhu Y, Chen C, et al. Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images. *Br J Ophthalmol.* 2021; 105(11): p. 1548-54.
41. Xu X, Guan Y, Li J, Ma Z, Zhang L, Li L. Automatic glaucoma detection based on transfer induced attention network. *Biomed Eng Online.* 2021; 20(1): p. 39.
42. Ajitha S, Akkara J, Judy M. Identification of glaucoma from fundus images using deep learning techniques. *Indian J Ophthalmol.* 2021; 69(10): p. 2702-09.
43. Gheisari S, Shariflou S, Phu J, Kennedy P, Agar A, Kalloniatis M, et al. A combined convolutional and recurrent neural network for enhanced glaucoma detection. *Sci Rep.* 2021; 11(5): p. 1945.
44. Mariottoni E, Jammal A, Berchuck S, Shigueoka L, Tavares I, Medeiros F. An objective structural and functional reference standard in glaucoma. *Sci Rep.* 2021; 11(1): p. 1752.
45. Hemelings R, Elen B, Barbosa-Breda J, Blaschko M, De Boever P, Stalmans I. Deep learning on fundus images detects glaucoma beyond the optic disc. *Sci Rep.* 2021; 11(1): p. 20313.
46. Ran A, Cheung C, Wang X, Chen H, Lou L, Chan P, et al. Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis. *Lancet Digit Health.* 2019; 1(4): p. e172-82.
47. Wang P, Shen J, Chang R, Moloney M, Torres M, Burkemper B, et al. Machine Learning Models for Diagnosing Glaucoma from Retinal Nerve Fiber Layer Thickness Maps. *Ophthalmol Glaucoma.* 2019; 2(6): p. 422-8.
48. Maetschke S, Antony B, Ishikawa H, Wollstein G, Schuma J, Garnavi R. A feature agnostic approach for glaucoma detection in OCT volumes. *PLoS One.* 2019; 14(7): p. e0219126.
49. Russakoff D, Mannil S, Oakley J, Ran A, Cheung C, Dasari S, et al. A 3D Deep Learning System for Detecting Referable Glaucoma Using Full OCT Macular Cube Scans. *Transl Vis Sci Technol.* 2020; 9(12): p. 12.
50. George Y, Antony B, Ishikawa H, Wollstein G, Schuman J, Garnavi R. Attention-Guided 3D-CNN Framework for Glaucoma Detection and Structural-Functional Association Using Volumetric Images. *IEEE J Biomed Health Inform.* 2020; 24(12): p. 3421-30.
51. Zheng C, Xie X, Huang L, Chen B, Yang J, Lu J, et al. Detecting glaucoma based on spectral domain optical coherence tomography imaging of peripapillary retinal nerve fiber layer: a comparison study



- between hand-crafted features and deep learning model. *Graefes Arch Clin Exp Ophthalmol.* 2020; 258(3): p. 577-85.
52. Lee J, Kim Y, Park K, Jeoung J. Diagnosing Glaucoma With Spectral-Domain Optical Coherence Tomography Using Deep Learning Classifier. *J Glaucoma.* 2020; 29(4): p. 287-94.
 53. Thakoor K, Koorathota S, Hood D, Sajda P. Robust and Interpretable Convolutional Neural Networks to Detect Glaucoma in Optical Coherence Tomography Images. *IEEE Trans Biomed Eng.* 2021; 68(8): p. 2456-66.
 54. Olivas L, Alférez G, Castillo J. Glaucoma detection in Latino population through OCT's RNFL thickness map using transfer learning. *Int Ophthalmol.* 2021; 41(11): p. 3727-41.
 55. Xiong J, Li F, Song D, Tang G, He J, Gao K, et al. Multimodal Machine Learning Using Visual Fields and Peripapillary Circular OCT Scans in Detection of Glaucomatous Optic Neuropathy. *Ophthalmology.* 2022; 129(2): p. 171-80.
 56. Thompson AC, Jammal A, Medeiros F. A Deep Learning Algorithm to Quantify Neuroretinal Rim Loss From Optic Disc Photographs. *Am J Ophthalmol.* 2019; 201: p. 9-18.
 57. Mehta P, Petersen C, Wen J, Banitt M, Chen P, Bojkian K, et al. Automated Detection of Glaucoma With Interpretable Machine Learning Using Clinical Data and Multimodal Retinal Images. *Am J Ophthalmol.* 2021; 231: p. 154-69.
 58. Jammal A, Thompson A, Mariottoni E, Berchuck S, Urat C, Estrela T, et al. Human Versus Machine: Comparing a Deep Learning Algorithm to Human Gratings for Detecting Glaucoma on Fundus Photographs. *Am J Ophthalmol.* 2020; 211: p. 123-31.
 59. Hagiwara Y, Koh J, Tan J, Bhandary S, Laude A, Ciaccio E, et al. Computer-aided diagnosis of glaucoma using fundus images: A review. *Comput Methods Programs Biomed.* 2018; 165: p. 1-12.
 60. Cen L, Ji J, Lin J, Ju S, Lin H, Li T, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nat Commun.* 2021; 12(1): p. 4828.



Agradecimientos.

Agradecer a mi tutor, profesor José Carlos Pastor Jimeno, porque el arte de enseñar pocos lo pueden exteriorizar, tengo que dar las gracias por el tiempo y el esmero que me fue dedicado. Un buen mentor impulsa la confianza en sus alumnos y convierte el aprendizaje en una tarea fascinante.

Agradecer a los Dres. José Ramón Juberías Sánchez y Miguel Maldonado López, que con su conocimiento y experiencia valieron de guía para perfeccionar la calzada hacia el éxito.



Material Complementario.

Figura 1. Diagrama de flujo de la selección de los estudios con la consulta de investigación y los criterios de inclusión.

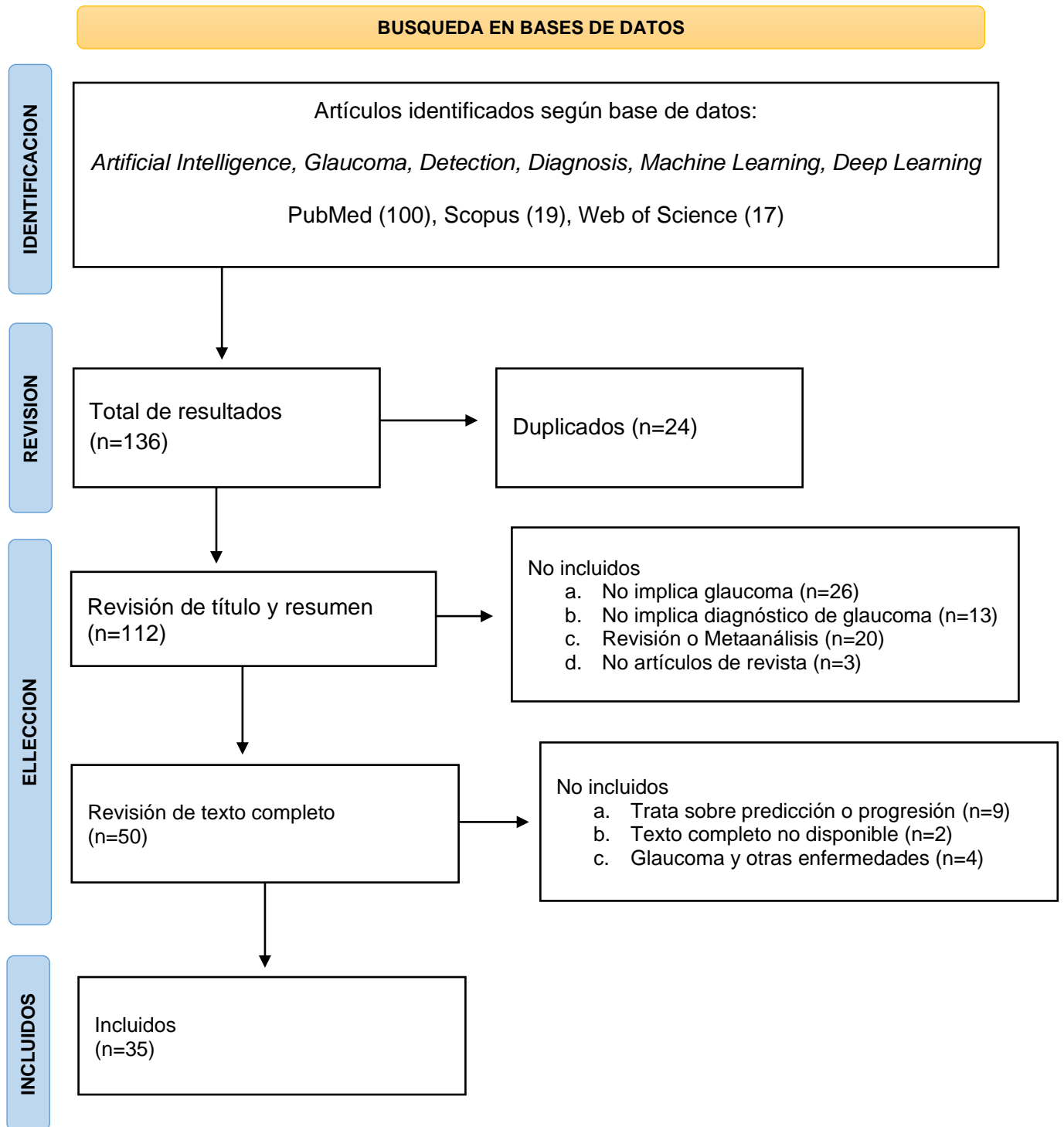




Tabla 1. Resumen de los estudios revisados.

AUTOR PRINCIPAL	BASE DE DATOS	IMAGEN	NGO	PROFESIONAL	ALGORITMO	ACC / AUC (95%CI)	SENSIBILIDAD / ESPECIFICIDAD (95% CI)
Cerentini, A, 2017 BRASIL	HRF RIM-ONE r1-r2-r3	FOTOS CNO 2.326 imágenes	??	Especialistas cualificados	GoogleNet (CNN)	ACC HRF 90,0% RIM-ONE r1 94,2 % RIM-ONE r2 86,2 % RIM-ONE r3 86,4 % Todos 87,6%	
Liu, S, 2018 AUSTRALIA	Eye Associates, Macquarie University Ophthalmology Clinic, University of New South Wales Optometry Clinic, HRD, RIM-ONE	FOTOS CNO 4.364 imágenes	2.369 imágenes	1 glaucomatólogo	ResNet 50 (CNN)	ACC 0.96 AUC 0.97 (0.96-0.98)	87.9% / 96.50%
Ahn, JM, 2018 KOREA	Kim's Eye Hospital	FOTOS CNO 1.542 imágenes	756 imágenes	2 especialistas en glaucoma	Inception V3 (CNN) Modelo propio (CNN)	ACC Inception V3 0.97 Modelo propuesto 0.922 AUC Inception V3 0.99 Modelo propuesto 0.98	
Christopher, M, 2018 USA	African Descent and Glaucoma Evaluation Study (ADAGES), Diagnostic Innovations in Glaucoma Study (DIGS)	FOTOS CNO 14.822 imágenes	5.633 imágenes	??	ResNet 50 (CNN) VGG16 (CNN) Inception V3 (CNN)	AUC ResNet50 Todos 0.91 (0.90–0.91) Inicial 0.89 (0.88–0.90) Moderado-Severo 0.97 (0.96–0.98)	ResNet 50 Todos 0.84 (0.83–0.85) / 0.83 (0.81–0.84) Inicial 0.82 (0.80–0.84) / 0.82 (0.80–0.83) Moderado-Severo 0.92 (0.89–0.94) / 0.93 (0.92–0.95)
Fu, H, 2018 CHINA	Singapore Chinese Eye Study (SCES) Singapore Indian Eye Study (SINDI) ORIGA	FOTOS CNO 8.109 imágenes	327 imágenes	Especialistas cualificados	ResNet 50 (CNN)	AUC SCES 0.9183 SINDI 0.8173	
Li, Z, 2018 CHINA	Labelme (online)	FOTOS CNO 39.745 imágenes	7.659 imágenes	21 oftalmólogos entrenados	Inception-V3 (CNN)	AUC 0.986	95.6% / 92%
Masumoto, H, 2018 JAPON	Tsukazaki Hospital	FOTOS CNO 1.399 ojos	982 ojos	2 glaucomatólogos	CNN	AUC Normal vs Glaucoma 0.872 Normal vs inicial 0.830 Normal vs moderado 0.864 Normal vs avanzado 0.934	Normal vs Glaucoma 81.3% / 80.2% Normal vs inicial 83.8% / 75.3% Normal vs moderado 77.5% / 90.2% Normal vs avanzado 90.9% / 95.8%



Tabla 1. Continuación.

AUTOR PRINCIPAL	BASE DE DATOS	IMAGEN	NGO	PROFESIONAL	ALGORITMO	ACC / AUC (95%CI)	SENSIBILIDAD / ESPECIFICIDAD (95%CI)
Ran, AR, 2019 CHINA	Chinese University of Hong Kong Eye Centre, Hong Kong Eye Hospital, Prince of Wales Hospital, Tuen Mun Eye Centre, Byers Eye Institute of Stanford University	OCT 6.921 imágenes	4.360 imágenes	3 glaucomatólogos	ResNet (CNN)	ACC 0.969 AUC 0.969	89% / 96%
Liu, H, 2019 CHINA	Chinese Glaucoma Study Alliance, Beijing Tongren Hospital, Peking University Third Hospital, Harbin Medical University First Hospital, Handan Eye Study, Hamilton Glaucoma Center, Website	FOTOS CNO 269.601 imágenes	29.865 imágenes	22 oftalmólogos 2 glaucomatólogos	ResNet (CNN)	ACC 0.996 (0.995-0.998)	96.2% / 97.7%
Rogers, TW, 2019 UK	European Optic Disc Assessment Trial	FOTOS CNO 94 imágenes	48 imágenes	Panel de expertos	ResNet-50 (Pegasus) (CNN)	ACC 83.7% (77.3-89.1)	80.9% / 86.2 %
Wang, P, 2019 USA	Los Ángeles Latino Eye Study	OCT 221 imágenes	93 imágenes	Panel de expertos	Glaucomanet=DenseNet-201+ResNet-152 (CNN)	ACC SVM 0.869 KNN 0.860 ResNet-18 0.905 GlaucomaNet 0.878 AUC SVM 0.915 KNN 0.911 ResNet-18 0.906 GlaucomaNet 0.921	SVM 0.817 / 0.906 KNN 0.828 / 0.883 ResNet-18 0.860 / 0.938 GlaucomaNet 0.828 / 0.914
Jammal, AA, 2019 INDIA	Duke Glaucoma Repository	Campo visual+ Foto CNO 32.820 imágenes	??	Panel de expertos	ResNet 34 (CNN)	AUC 0.801 (0.757-0.845)	
Jinho Lee, 2019 KOREA	Clínica de glaucoma del Seoul National University Hospital	FOTOS CNO 200 ojos	100 ojos	??	NasNet (CNN)	AUC 0.939 (0.881-0.996)	92.9% (82.1-100%) / 84.4% (71.9-96.9%)
MacCormick, IJC, 2019 UK	ORIGA+RIM-ONE	FOTOS CNO 809 imágenes	242 imágenes	Oftalmólogos	Spatial Algorithm+ SVM	AUC 99.6%	96.6% / 99%



Tabla 1. Continuación.

AUTOR PRINCIPAL	BASE DE DATOS	IMAGEN	NGO	PROFESIONAL	ALGORITMO	ACC / AUC (95%CI)	SENSIBILIDAD / ESPECIFICIDAD (95%CI)
Thompson, AC, 2019 USA	Duke Glaucoma Repository Imaging and Performance Laboratory	FOTOS CNO +OCT 9.282 imágenes	5540 imágenes	??	ResNet 34 (CNN)	AUC 0.945 (0.874-0.980)	
Maetschke, S, 2019 AUSTRALIA	??	OCT 1.110 imágenes	847 imágenes	??	Logistic Regression Support Vector Machine Random Forest Extra Trees Naïve Bayes Gradient Boosting	ACC Logistic Regression 0.88 SVM 0.89 Random Forest 0.91 Extra Trees 0.90 Naive Bayes 0.87 Gradient Boosting 0.87	
Russakoff, DB, 2020 USA HONG KONG INDIA	Byers Eye Institute, Stanford School of Medicine	OCT 1.095 imágenes	667 imágenes	1 fellow de glaucoma 1 glaucomatólogo	gNet 3D (CNN)	ACC 0.878 AUC 0.921	0.83 / 0.91
Li, F, 2020 CHINA	Shanghai Zhongshan Hospital, Shanghai First People's Hospital	FOTOS CNO 26.585 imágenes	12.618 imágenes	8 expertos en glaucoma	ResNet 101 (CNN)	ACC 0.954	0.965 / 0.941
Christopher, M, 2020 USA JAPON	The Diagnostic Innovations in Glaucoma Study African Descent and Glaucoma Evaluation Study Matsue Red Cross Hospital, Inan Hospital, ACRYMA	FOTOS CNO 18.340 imágenes	3.391 imágenes	Oftalmólogos	ResNet 50 (CNN) ResNet 34 (CNN)	AUC RESNET 50 0.90 (0.87-0.93) RESNET 34 0.90 (0.870.93)	
Ko, Y, 2020 TAIWAN	Taipei Veterans General Hospital DRISHTI-GS	FOTOS CNO 944 ojos	479 ojos	2 glaucomatólogos	VGGNet (CNN)	AUC 0.984 (0.972-0.998)	92.6% / 95.4%
George, Y, 2020 USA	??	OCT 3.782 imágenes	3.355 imágenes	??	AG-OCT (DL)	AUC 93.8%	
Lee, J, 2020 KOREA	Clinica de glaucoma del Seoul National University Hospital	OCT 657 imágenes	350 imágenes	??	NasNet (CNN)	AUC 0.990 (0.975-1.000)	94.7% / 100%



Análisis crítico de los sistemas de inteligencia artificial entrenados para el diagnóstico de glaucoma.

Universidad de Valladolid

Tabla 1. Continuación.

AUTOR PRINCIPAL	BASE DE DATOS	IMAGEN	NGO	PROFESIONAL	ALGORITMO	ACC / AUC (95%CI)	SENSIBILIDAD / ESPECIFICIDAD (95%CI)
Thompson, AC, 2020 USA	Duke Glaucoma Repository	OCT 20.086 imágenes	612 ojos	3 calificadores experimentados	ResNet 34 (CNN)	AUC 0.96	0.81 / 0.95
Zheng, C, 2020 CHINA	Joint Shantou International Eye Center of Shantou University and the Chinese University of Hong Kong	OCT 1.603 imágenes	740 imágenes	??	Inception-V3 (CNN)	AUC 0.990 (0.974 - 1.000)	
Ganesh, SS, 2021 INDIA	ACRIMA, DRISHTI GS, REFUGE, RIGA, RIM-ONE	FOTOS CNO 1.887 imágenes	764 imágenes	Oftalmólogos	ResNeX (CNN)	ACC 0.998	0.9812 / 0.9801
Li, Z, 2021 CHINA	Chinese Medical Alliance for AI, ZOC clinic in Guangzhou, Xudong Ophthalmic Hospital, Tsukazaki Optos Public Project	FOTOS CNO 22.972 UWF imágenes	4.269 imágenes	2 glaucomatólogos	InceptionResNet V2 (CNN)	ACC 0.983-0.999	97.5-98.2% / 94.3-98.4%
Xu, X, 2021 CHINA	Beijing Tongren Eye Center ORIGA	FOTOS CNO 1.882 imágenes	877 imágenes	1 oftalmólogo	Transfer Induced Attention Network (NN)	ACC Base de datos propia 0.857 ORIGA 0.76 AUC Base de datos propia 0.929 ORIGA 0.835	
Ajitha, S, 2021 INDIA	HRF, ORIGA, DRISHTI Little Flower Hospital and Research Centre	FOTOS CNO 1.113 imágenes	453 imágenes	Oftalmólogos	Google Colab (CNN)	ACC CNN+SoftMax 93.80% CNN+SVM ACC 95.60%	CNN+SoftMax 85.40% / 100% CNN+SVM 89.50% / 100%
Thakoor, KA, 2021 USA	Columbia University Medical Center	OCT 737 imágenes	??	Panel de expertos	InceptionV3 (CNN), ResNet-18 (CNN), VGG-16 (CNN), DenseNet-121 (CNN)	ACC DenseNet-121 85.2 - 95.9 VGG-16 85.2 -97.0 ResNet-18 83.0 -94.9 InceptionV3 91.1 -90.4	
Olivas, LG, 2021 MEXICO	Instituto de la Visión	OCT 260 imágenes	130 imágenes	1 oftalmólogo	MobileNet (CNN) Inception V3 (CNN)	ACC MobileNet OI 86%- OD 90% Inception V3 OI 90%- OD 90%	



Tabla 1. Continuación.

AUTOR PRINCIPAL	BASE DE DATOS	IMAGEN	NGO	PROFESIONAL	ALGORITMO	ACC / AUC (95%CI)	SENSIBILIDAD / ESPECIFICIDAD (95%CI)
Xiong, J, 2021 CHINA	Zhongshan Ophthalmic Center, First Hospital of Shijiazhuang City	CAMPO VISUAL + OCT 2.463 pares	1.423 imágenes	4 glaucomatólogos	ResNet 50 (CNN)	AUC CV+OCT 0.950 (0.931-0.968) Solo CV 0.868 (0.834-0.902)	
Gheisari, S, 2021 AUSTRALIA	Centre for Eye Health of University of New South Wales Marsden Eye Clinic, Research Clinic of University of Technology Sydney	FOTOS CNO 1,810 imágenes+ 295 videos	379 pacientes	1 especialista en glaucoma	VGG16 (CNN) ResNet 50 (CNN) VGG16+RNN ResNet50+RNN		VGG16 0.59 / 0.55 Resnet50 0.72 / 0.70 VGG16+RNN 0.94 / 0.86 Resnet50+RNN 0.88 / 0.79
Hemelings, R, 2021 BELGICA	University Hospitals Laueen	FOTOS CNO 23.930 imágenes	??	1 glaucomatólogo	ResNet (CNN)	AUC 0.94 (0.92-0.96)	
Mariottoni, EB, 2021 USA	Duke Glaucoma Registry	FOTOS CNO 9.830 imágenes	5. 041 imágenes	1 fellow glaucoma+ 1 glaucomatólogo	ResNet 50 (CNN)	AUC 0.92	77% / 95%
Mehta, P, 2021 USA	UK Biobank	FOTOS CNO + OCT 2,476 ojos	1,283 ojos	5 expertos en glaucoma	Densenet Inception + Resnet V4 (CNN)	AUC FOTO 0.74 (0.64-0.84) OCT 0.95 (0.90- 1.0) OCT+FOTO 0.963 (0.91-1.0)	

ACC= precisión (accuracy). AUC= área bajo la curva (area under curve). CI= intervalo de confianza (confidence Interval). CNN= red neuronal convolucional (convolutional neural network). CNO= cabeza del nervio óptico. CV= campo visual. DL= aprendizaje profundo (deep learning). KNN= k-vecinos más cercanos (k-nearest neighbors). NN= red neuronal (neural network). OCT= tomografía de coherencia óptica (Optical coherence tomography). SMV= (máquina vectorial de soporte (support vector machine). UWF= campo de ultra ancho (ultra width field).



Tabla 2. Sistemas comercializados.

NOMBRE	IMAGEN	METODO	EFICACIA	FUENTE	ESTADO
Kalpah's RIA-G	Foto CNO	CDR/ ISNT / DDLS	Sensibilidad 82.3% Especificidad 91.8%	Singh D, Gunasekaran S, Hada M, Gogia V. Clinical validation of RIA-G, an automated optic nerve head analysis software. Indian J Ophthalmol. 2019 Jul;67(7):10891094. doi: 10.4103/ijo.IJO_1509_18. PMID: 31238418; PMCID: PMC6611301.	COMERCIALIZADO
MONA	Foto CNO	??	??	SITIO WEB: https://mona.health/home-light/	EN DESARROLLO
Ophthalmitycs for glaucoma	Foto CNO	??	??	SITIO WEB: https://www.ophtalytics.com/glaucoma/	EN DESARROLLO
PEGASUS	Foto CNO	CDR / ANOMALIAS DEL DISCO	Sensibilidad 80.9% Especificidad 86.0%	Al-Aswad LA, Kapoor R, Chu CK, Walters S, et al. Evaluation of a Deep Learning System For Identifying Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. J Glaucoma. 2019 Dec;28(12): 1029-1034. doi: 10.1097/IJG.0000000000001319. PMID: 31233461.	COMERCIALIZADO
QUANTUS GL	Foto CNO	CDR / ISNT / DDLS / ASIMETRIA DEL DISCO	Sensibilidad 84.1% Especificidad 95.8%	Abad, Pablo & Coronado-Gutiérrez, David & López, Carlos & Burgos-Artizzu, Xavier. (2021). Glaucoma patient screening from online retinal fundus images via Artificial Intelligence. 10.1101/2021.02.11.21251193.	COMERCIALIZADO
TRL 8	OCT	??	Fiabilidad del 98%	SITIO WEB: https://www.cientech.org/inteligencia-artificialde-oct-de-glaucoma/	COMERCIALIZADO