

Cribado de la retinopatía diabética mediante la inteligencia artificial

Screening for diabetic retinopathy by applying the artificial intelligence

M. Baget Bernaldiz, P. Romero Aroca

Resumen

La aplicación de la inteligencia artificial (IA) en el cribado de la retinopatía diabética (RD) va a suponer un cambio profundo del modelo actual. Permitirá su descentralización con el objeto de estudiar a una mayor proporción de enfermos diabéticos de manera precoz y con un coste asociado menor.

Palabras clave: Inteligencia artificial. Retinopatía diabética. Cribado ocular.

Resum

L'aplicació de la intel·ligència artificial (IA) en el cribratge de la retinopatia diabètica (RD) suposarà un canvi profund del model actual. Permetrà la seva descentralització amb l'objectiu d'estudiar una major proporció de malalts diabètics de manera precoç i amb un cost associat menor.

Paraules clau: Intel·ligència artificial. Retinopatia diabètica. Cribratge ocular.

Abstract

The application of artificial intelligence to the ocular screening of diabetic retinopathy will greatly change the current eye screening system. It will allow its decentralization with the goal of controlling a more significant proportion of diabetic patients on time and at a lower cost.

Key words: Artificial intelligence. Diabetic retinopathy. Eye screening.

3.5. Cribado de la retinopatía diabética mediante la inteligencia artificial

Screening for diabetic retinopathy by applying the artificial intelligence

M. Baget Bernaldiz¹, P. Romero Aroca²

¹Médico adjunto. Servicio de Oftalmología. Hospital Universitari Sant Joan. Reus. Universitat Rovira i Virgili. Reus. Institut d'Investigació Sanitària Pere Virgili (IISPV). Reus. Tarragona. ²Jefe del Servicio de Oftalmología. Hospital Universitari Sant Joan. Reus. Universitat Rovira i Virgili. Reus. Institut d'Investigació Sanitària Pere Virgili (IISPV). Reus. Tarragona.

Correspondencia:

Marc Baget Bernaldiz

E-mail: mbaget@gmail.com

Introducción

La diabetes *mellitus* afecta a un 10% de la población mundial y se prevé que su prevalencia aumente de manera significativa en las próximas décadas¹. En la actualidad, la detección de la RD en fases tempranas no es posible con los recursos humanos y tecnológicos actuales². Además, hay que tener en cuenta que una proporción importante de pacientes diabéticos viven en áreas rurales lejos de los centros de cribado. En España y en el resto de Europa, esta cifra se estima en un 30%³. Un tercio de ellos no acude al chequeo ocular. Por lo tanto, es necesario un cambio de paradigma en el cribado de la RD, con el objeto de prevenir la pérdida de visión que esta ocasiona a escala mundial.

Para lograr un cribado masivo y precoz de la RD, es necesario aproximar los chequeos a los pacientes, automatizar el proceso de cribado y reducir el coste asociado. Ello requiere el desarrollo de dispositivos portátiles que puedan tomar retinografías de alta calidad en el ámbito de la medicina primaria para su posterior lectura. Con esta finalidad, las diferentes casas comerciales de referencia mundial en la toma y procesamiento de imágenes retinianas ya están desarrollando retinógrafos portátiles y lentes

adaptadas a *smartphones* (teléfonos inteligentes) con un coste menor respecto a las cámaras no midriáticas convencionales.

Para lograr el cribado masivo de la RD, es imprescindible la implementación de sistemas automáticos de lectura e interpretación de las imágenes de fondo de ojo que previamente hayan demostrado su eficacia en poblaciones de pacientes diabéticos del mundo real. Desde la aceptación en el año 2018 del primer algoritmo basado en la IA para la lectura automatizada de las retinografías de los pacientes diabéticos⁴, existe un gran esfuerzo por parte de diversos grupos, también en nuestro país, para desarrollar modelos algorítmicos para la detección y clasificación de la RD.

La mayoría de los algoritmos actuales que interpretan las imágenes de fondo de ojo están basados en las redes neuronales. Algunos de ellos han demostrado ser tan efectivos como los oftalmólogos especialistas en retina a la hora de detectar y clasificar las retinografías de los pacientes diabéticos en los diferentes tipos de RD⁵.

El tercer pilar necesario para lograr un chequeo masivo de los pacientes diabéticos es tener un sistema de cribado que sea coste-efectivo. Los grupos de Liu, Raumviboonsuk y Tufail, demostraron

que la implementación de sus algoritmos basados en la IA en zonas de la China rural, India y Singapur, fueron coste-efectivos⁶⁻⁸.

Los dos mayores inconvenientes actuales para la implementación de los modelos de IA son la falta de robustez y explicabilidad de los algoritmos. La mayoría de ellos han demostrado ser eficaces en la fase de validación y testado, pero baja su nivel predictivo cuando se les somete a poblaciones de pacientes diabéticos del mundo real^{9,10}. Además, el no poder explicar cómo decide el algoritmo el resultado final para cada paciente dificulta su aceptación por parte del médico y del propio paciente¹¹. Esta falta de explicabilidad ha sido en parte mitigada por la creación de mapas de resaltación o *saliency heatmaps*. En estos, se resaltan las áreas de las retinografías en las que se fija un algoritmo a la hora de clasificar la RD¹² (Figura 1).

En este capítulo, describimos las fases requeridas en la construcción de los algoritmos de lectura para el cribado de la RD junto a los principales algoritmos validados a escala mundial y también en España.

Procedencia de las imágenes oculares analizadas por los algoritmos

Retinografías adquiridas con cámaras no midriáticas

La mayoría de los algoritmos actuales que clasifican la RD lo hacen a partir de retinografías obtenidas con cámaras no midriáticas. Esto es normal, ya que, para entrenarlos, se precisan imágenes de gran calidad que tradicionalmente han proporcionado este tipo de cámaras. No obstante, en algunas ocasiones, se precisa la dilatación pupilar para mejorar la calidad de la imagen y, así, su gradabilidad¹³. Entre los algoritmos que han sido entrenados mediante retinografías tomadas con este tipo de cámaras, cabe destacar IDX-DR¹⁴, Retmarker¹⁵ y EyeArt¹⁶. Todos ellos han obtenido la certificación para funcionar de manera autónoma en diferentes territorios. En España, los grupos de Romero y Zapata han validado sus algoritmos para la detección de la RD a partir de retinografías de pacientes diabéticos de bases de datos propias y públicas^{17,18}.

Retinografías adquiridas con dispositivos móviles

Hay grupos que, de manera complementaria, han entrenado sus algoritmos a partir de retinografías obtenidas con dispositivos

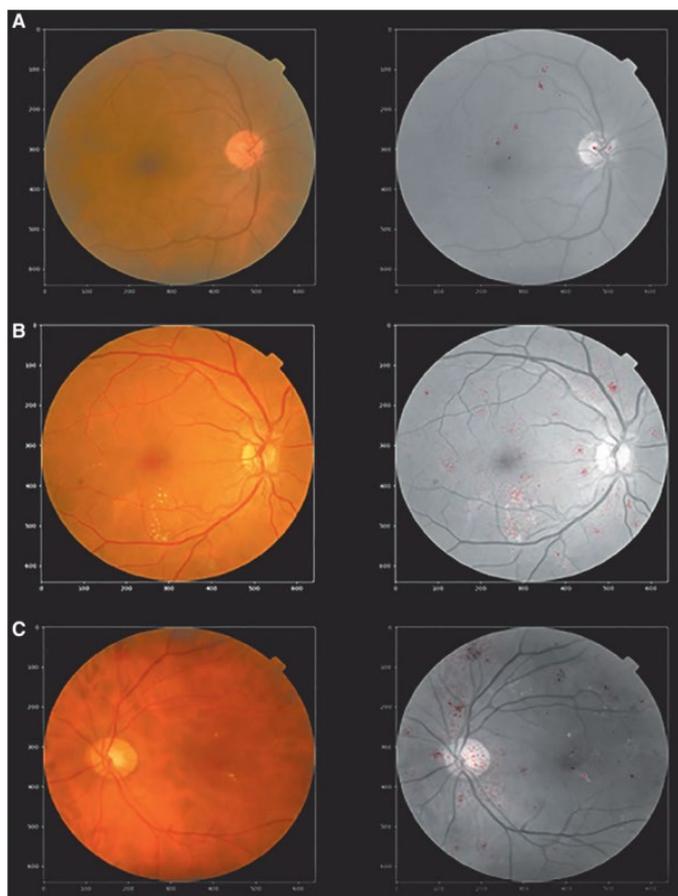


Figura 1. Se muestran tres retinografías con diferentes niveles de retinopatía diabética (RD), con su correspondiente imagen anérita, en las cuales se resaltan las lesiones identificadas por el algoritmo como patológicas (imagen cedida por los autores del capítulo). **(A)** Es una imagen de RD leve, donde se resaltan únicamente microaneurismas. **(B)** La imagen corresponde a una RD moderada, donde se exhiben microaneurismas con alguna hemorragia asociada. **(C)** La imagen es una RD severa, con la presencia abundante de microaneurismas y hemorragias junto a exudación lipídica foveal.

portátiles. Es el caso de los grupos de Romero y Zapata, han entrenado sus algoritmos en dispositivos portátiles, tales como VistaView de Volk, Aurora de Optomed, Signal de Topcon, Zeiss Visuscout y Optomed Smartscope. De manera análoga, EyeArt entrenó sus algoritmos a partir de retinografías obtenidas con el *smartphone* fabricado por Remidio Fundus on Phone, obteniendo unos valores muy altos de sensibilidad y especificidad para la detección de cualquier tipo de RD y la RD moderada o superior, también denominada RD referible (RDR)¹⁶.

Retinografías de campo ultraamplio

Con estos dispositivos, se permite la visualización, además de la mácula, de la retina periférica, hasta 200°. Se ha documentado que el incremento de lesiones en la periferia retiniana es un marcador de progresión de RD y, por lo tanto, sería de utilidad la monitorización automatizada de estos pacientes¹⁹.

Levenkova desarrolló un algoritmo que detecta con precisión las lesiones brillantes (exudados blandos y duros) y las lesiones oscuras (hemorragias y microaneurismas) en retinografías de campo amplio²⁰. Wang desarrolló un algoritmo para la detección de RDR, obteniendo una buena sensibilidad, pero discreta especificidad²¹.

OCT

La OCT es la herramienta considerada *gold standard* para diagnosticar y monitorizar el edema macular diabético (EMD)²². No obstante, debido a su alto coste y a la necesidad de disponer de expertos para su interpretación, no es una técnica implementada en el cribado del EMD y, mayoritariamente, se continúa confiando en la retinografía para diagnosticarlo²³.

En la actualidad, existen dos grupos de investigadores que trabajan con imágenes multimodales, entre ellas la OCT, con dos objetivos principales: diagnosticar el EMD de manera automática y, sobre todo, encontrar patrones de respuesta en los pacientes diabéticos frente a los diferentes tratamientos del EMD, con el

objetivo último de personalizar el tratamiento. Dentro del primer grupo, encontramos a Zu *et al.*, que validaron la detección automática de los tres tipos de EMD (quístico, esponjiforme y desprendimiento seroso) mediante un modelo de *deep learning* (aprendizaje profundo)²⁴. Los grupos de Hwang y Varadarajan aplicaron un modelo de *deep learning* para la detección del EMD a partir de imágenes de OCT analizadas con un *smartphone*^{25,26}. En el segundo grupo, cabe destacar a Alryalat *et al.*, que han estudiado y clasificado a pacientes *naïves* (no tratados) con EMD en respondedores y no respondedores a la terapia anti-VEGF con un nivel de precisión del 75%²⁷.

Una vez que la IA demuestre mayor robustez en los resultados, será posible establecer protocolos individualizados de tratamiento y seguimiento de estos pacientes.

Construcción de los modelos algorítmicos para el cribado de la retinopatía diabética

Desde que se construye un modelo predictivo de RD hasta que se pone en el mercado listo para aplicarse en poblaciones diabéticas, se requiere de distintas fases, que se detallan a continuación.

Construcción del modelo

La mayoría de los modelos algorítmicos que leen imágenes de fondo de ojo son redes neuronales convolucionales (Figura 2). Este

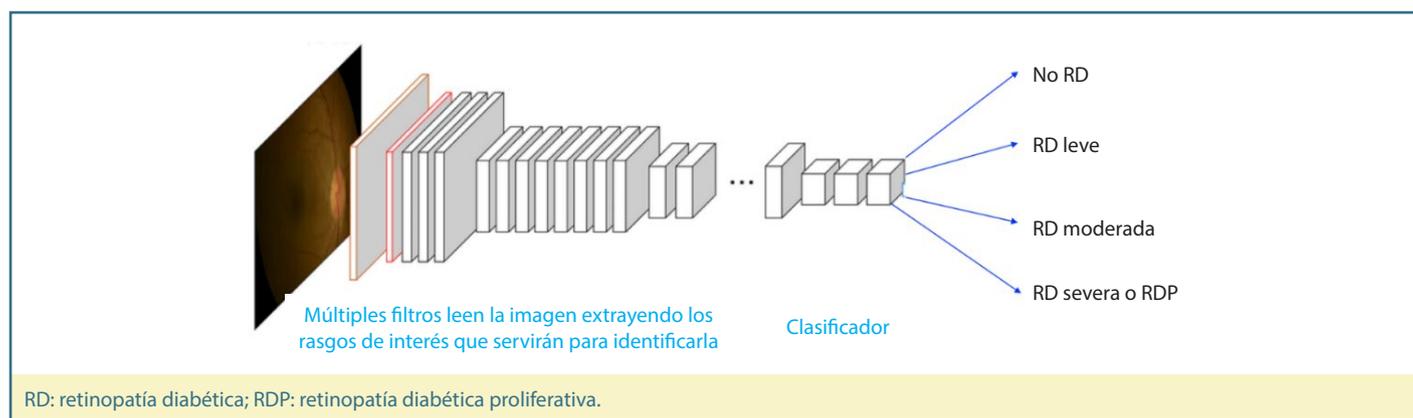


Figura 2. Las redes neuronales convolucionales están diseñadas para la interpretación de imágenes. Constan de dos partes: una primera, donde unos filtros leen la imagen extrayendo los rasgos de interés mediante una operación matemática denominada convolución; y una segunda parte en la que, una vez creado el mapa de rasgos, clasifica la imagen en una de las categorías posibles. (Imagen cedida por los autores del presente capítulo).

tipo de red neuronal consta de dos partes: la primera consiste en un set de filtros que leen las retinografías extrayendo las características de interés, tales como la presencia de microaneurismas y hemorragias, y la segunda clasifica la imagen en el tipo de RD teniendo en cuenta los hallazgos anteriores.

Training (entrenamiento) de los algoritmos

Una vez construida la red neuronal, esta tiene que aprender a identificar las lesiones características de la RD. Para lograrlo, necesita entrenarse leyendo miles de retinografías diferentes. Con el objeto de evitar la introducción de sesgos en esta fase, es necesario disponer de un número similar de retinografías de cada tipo de RD, estando balanceadas por sexo, edad y grupo étnico.

La mayoría de los grupos que se disponen a diseñar redes neuronales para la identificación y clasificación de la RD no dispone de un número tan elevado de retinografías en su base de datos para poder entrenar a los algoritmos. Es por lo que, en la actualidad, se dispone de bases de datos públicas con un número muy elevado de imágenes de fondo de ojo, que pueden ser utilizadas por los investigadores para entrenar a sus algoritmos. Las más utilizadas son la americana EyePACS y la francesa MESSIDOR y MESSIDOR-2. Existen otras desarrolladas en Finlandia (DIARETDB), la India (APTOS 2019) y China (DDR).

Este tipo de aprendizaje es de tipo supervisado, por lo que se requiere el marcado previo de las lesiones en las retinografías y su etiquetado por tipo de RD antes de ser mostradas a la red neuronal para entrenarla. Inicialmente, la red comete muchos errores en la clasificación de las retinografías, los cuales son aprovechados y computados para mejorar la función clasificadora de la red neuronal. Una vez que la red neuronal va leyendo una y otra vez el set de retinografías (iteraciones), las neuronas van ajustando sus parámetros (denominados pesos) hasta convertirse en buena clasificadora de la RD.

La evaluación de la función clasificadora de la red neuronal se hace comparando sus predicciones respecto a las retinografías etiquetadas previamente por retinólogos que sirven de referencia (*ground truth*). Los parámetros más habituales para evaluar a los algoritmos son: la sensibilidad, la especificidad, los valores predictivos positivo y negativo, falsos positivos y negativos, precisión y área bajo la curva (AUC).

Validación de los algoritmos

Una vez que el modelo algorítmico ha estado entrenado y se objetiva que clasifica razonablemente bien las retinografías por tipos de RD, se pasa a la fase de validación con el objeto de afinar un poco más su capacidad predictiva. Para ello, se utilizan retinografías que no han sido previamente mostradas al algoritmo de la propia población de diabéticos o de una base de datos pública.

Testado de los algoritmos

Es la última fase requerida antes de poner el modelo algorítmico en producción. Lo ideal es confrontar el algoritmo con imágenes de fondo de ojo de otras poblaciones distintas de las que se han utilizado para entrenarlo y validarlo, demostrando así su robustez. Algunos grupos han utilizado poblaciones propias, bases de datos públicas o una combinación de ambas. La consecuencia de ello es que, en la actualidad, los algoritmos no pueden compararse entre ellos respecto a su capacidad de etiquetar correctamente las retinografías de los pacientes diabéticos. Para poder compararlos, es necesario la creación de una única base de retinografías suficientemente extensa, correctamente etiquetadas y aplicando una única clasificación de la RD para todas ellas. Además, debe estar compensada por grupos de edad, sexo, tipo de RD y grupos étnicos, con el objeto de no introducir sesgos.

Clasificación de la retinopatía diabética por parte de los algoritmos

No existe consenso respecto la clasificación de la RD que se debe emplear para que los algoritmos clasifiquen las retinografías de los pacientes diabéticos. La mayoría de los grupos indican que han utilizado la escala clínica internacional de gravedad de la retinopatía diabética (ICDR, *International Clinical Diabetic Retinopathy Severity Scale*) lo cual es inexacto, ya que, en la actualidad, los algoritmos no detectan correctamente las anomalías microvasculares intrarretinianas ni los arrosamientos venosos. Además, al tratarse de cribado y no de diagnóstico, solo se realizan una o dos retinografías para estudiar el polo posterior y no toda la superficie retiniana. En nuestra opinión, creemos que es más acertado utilizar una clasificación de la RD más sencilla, clara y realista. En esta dirección se encuentra la clasificación utilizada por la base de datos pública MESSIDOR, que tiene en cuenta el

	Microaneurismas	Hemorragias	Neovasos retinianos
No RD	0	0	0
RD leve	1-5	0	0
RD moderada	6-14	1-4	0
RD severa	≥15	≥5	0
RD proliferativa	--	--	1

RD: retinopatía diabética.
Se tiene en cuenta el número y presencia de microaneurismas, hemorragias retinianas y neovasos en la retina. Debido a su sencillez es adecuada para la clasificación de la retinopatía diabética por parte de los algoritmos.

Tabla 1. Clasificación de la retinopatía diabética según la base de datos pública MESSIDOR

número de microaneurismas y hemorragias y la presencia de neovasos retinianos para clasificar la RD (Tabla 1).

La mayoría de los algoritmos existentes hacen una clasificación dicotómica de la RD: presencia o ausencia de RD, o bien RD no referible *versus* referible (presencia de RD moderada o superior). En este caso, se entiende que el paciente necesita ser derivado al oftalmólogo para ser controlado.

Algoritmos validados para el cribado de la retinopatía diabética

Existen una cantidad considerable de algoritmos desarrollados para el cribado de la RD. Presentamos aquellos que han sido validados y que aportan alguna característica diferencial respecto al resto (Tabla 2).

Algoritmos validados en España

MIRA

Es el modelo algorítmico desarrollado por los autores del capítulo juntamente con el Servicio de Oftalmología del Hospital Universitario Sant Joan de Reus, l'Institut d'Investigació Sanitària Pere i Virgili (IISPV) y la Universitat Rovira i Virgili. Esta red neuronal ha sido diseñada para detectar con precisión los casos de RD leve, que es el momento ideal para coordinarse con el médico de familia para minimizar su progresión¹⁷. En la actualidad, se está reentrenando el algoritmo para leer las retinografías procedentes de dispositivos portátiles, tales como VistaView y Aurora.

Autor	Nombre	Algoritmo	Año	Training set	Validation set	Testing set	Clasificación de RD	Detección	AUC	S (%)	SP (%)
Abràmoff	IDX-DR	CNN	2018	N/A	N/A	Messidor-2	FPRC	RDR	N/A	96,8	59,4
Bhaskaranand	EyeArt 2.0	Image analysis technology	2019	EyePACS	N/A	850.908	ICDR	RDR	0,96	91,3	91,1
Gulshan	Google	Inception v. 4	2019	103.634	40.790	1.983	ICDR	DR/no DR	0,96	88,9	92,2
Oliveira	Retmarker	Recognition of lesions	2011	N/A	N/A	21.514	No RD/RDR	DR/no DR	0,84	95,8	63,2
Bellemo	SELENA+	CNN	2019	ImageNet	N/A	3.556	ICDR	RDR	0,95	91,8	98,7
Li	LabelMe	CNN	N/A	71.043	35.201	N/A	N/A	RDR	0,95	92,5	98,5
Gulshan	ARDA	CNN	N/A	130.000	Messidor-2 EyePACS	N/A	N/A	RDR	0,99	87	98,5
Romero	Mira	CNN	2019	Propia EyePACS	5.000	Propia Messidor-2	Messidor-2	Any-DR RDR	0,96	97,3	94,5
Zapata	Holistic maculopathy screening	Inception v. 3	2021	Propia, ICS, EyePACS	5.918	Optretina	No RD/RDR	DR/no DR	0,95	81,1	94,8

Any-DR: cualquier tipo de retinopatía diabética; AUC: área bajo la curva (*area under the curve*); CNN: red neuronal convolucional (*convolutional neural network*); N/A: no disponible (*not available*); FPRC: *Wisconsin Fundus Photograph Reading Center*; ICDR: escala clínica internacional de gravedad de la retinopatía diabética (*International Clinical Diabetic Retinopathy Severity Scale*); ICS: *Institut Català de la Salut*; RD: retinopatía diabética; RDR: retinopatía diabética referible (*referral diabetic retinopathy*); S: sensibilidad; SP: especificidad.

Tabla 2. Principales algoritmos validados basados en la inteligencia artificial para la detección de la RD.

El algoritmo Mira está diseñado para evaluar al paciente diabético juntamente con el algoritmo clínico Retiprogram²⁸. Si un paciente, supongamos por caso, no presenta RD según el algoritmo Mira, se le somete al algoritmo Retiprogram, que nos va a estimar el riesgo de presentar RD según nueve variables clínicas independientes. De esta manera, podemos personalizar el intervalo de cribado para cada paciente. La aplicación de ambos algoritmos permitiría el cribado de la RD a nivel de la atención primaria y derivar al oftalmólogo solo aquellos pacientes que presenten o tengan riesgo de desarrollar RD. De este modo, se lograría reducir la lista de espera para el oftalmólogo de manera significativa.

Algoritmo del grupo de Zapata

La validación del modelo algorítmico para la detección de la RD del grupo de Zapata se ha llevado a cabo a partir de retinografías obtenidas con los sistemas portátiles Aurora de Optomed, Zeiss Visuscout y Optomed Smartscope. Se trata de un modelo que consta de varias redes neuronales, con el objeto de detectar simultáneamente la RDR, la degeneración macular asociada a la edad (DMAE), el glaucoma y la presencia de nevus. La combinación de algoritmos para la detección de cualquier anomalía, la RDR, la DMAE y los nevus, obtuvieron muy buenos resultados respecto al AUC, la sensibilidad y la especificidad¹⁸.

Principales algoritmos validados fuera de España

IDX-DR

El IDX-DR (Digital Diagnostics, Coralville, IA, Estados Unidos [EE. UU.]) fue el primer algoritmo de cribado de la RD en obtener la aprobación por parte de la *Food and Drug Administration* (FDA)¹⁷, al demostrar buena capacidad predictiva en la fase de validación. Sin embargo, un grupo independiente testó IDX-DR en 900 pacientes para la detección de RDR, obteniendo resultados más discretos¹⁰.

Retmarker

Retmarker es un *software* de detección automatizada implementado en el sistema de cribado de la RD en Portugal. Retmarker establece si existe o no RD y, en caso de detectar patología, un oftalmólogo clasifica el grado de esta.

Una característica interesante de Retmarker es su capacidad de comparar imágenes para detectar la progresión de la RD basado en el *turn over* (la rotación) de los microaneurismas²⁹. Además, Retmarker demostró una disminución del 48% de la carga de trabajo en el cribado de la RD en Portugal.

Eyeart

De manera similar a otros sistemas de detección para la RD, EyeArt ha obtenido el marcaje de conformidad europea (CE), siendo categorizado como dispositivo médico de clase IIa. Está comercializado también en Canadá, siendo su uso en EE. UU. limitado a la investigación. EyeArt ha sido desarrollado por la empresa Eyenuk (Inc., based in Los Angeles, EE. UU.), la cual ofrece el producto Eyemark para la monitorización de la progresión de la RD basado en la cuantificación de los microaneurismas, al igual que Retmarker. EyeArt fue testado en el *National Health Service* (NHS) del Reino Unido, y obtuvo buenos resultados para la detección de cualquier retinopatía y para la RDR.

EyeArt fue testado *offline* en la India, utilizando el dispositivo Remidio Fundus on Phone para la captura de imágenes de fondo de ojo previa dilatación pupilar. Este estudio logró niveles remarcables de sensibilidad y especificidad para la detección de cualquier grado de RD.

SELENA+

El Eye Research Institute y la National University de Singapore desarrollaron SELENA+ para detectar la RDR. Concibieron el sistema para funcionar de manera autónoma o semiautónoma, en el cual hubo asistencia humana. En este último caso, la especificidad fue significativamente superior.

ARDA

ARDA (*Automated Retinal Disease Assessment*) es un algoritmo desarrollado por Verily Life Sciences LLC (South San Francisco, CA, EE. UU.). Para su validación, se usaron retinografías de Messidor-2 y EyePACS, obteniendo valores de sensibilidad y especificidad similares a técnicos entrenados para clasificar la RD.

LabelMe

LabelMe es un algoritmo desarrollado en Guangzhou, China. Para la validación externa, se utilizaron retinografías de pacientes

precedentes de Australia y del sudeste asiático. Los valores de sensibilidad y especificidad alcanzaron niveles estándar.

Retinalyze

Un estudio reciente comparó la precisión de Retinalyze respecto a IDx-DR para la detección de RDR en el mismo grupo de pacientes diabéticos. Retinalyze obtuvo resultados discretos, por lo que los investigadores concluyeron que no puede utilizarse sin supervisión clínica¹⁰.

Conclusiones

La IA es una herramienta necesaria para lograr el cribado masivo de la población diabética con el objeto de detectar la RD en fase precoz. Existen en la actualidad algoritmos capaces de clasificar la RD con un grado de precisión similar al de oftalmólogos expertos en retina. No obstante, será necesario demostrar su robustez en diferentes poblaciones de pacientes diabéticos y mejorar la explicabilidad de los modelos algorítmicos antes de poder ser implementados de manera autónoma y generalizada.

Bibliografía

- Roglic G. WHO Global report on diabetes: A summary. *Int J Noncomm Dis.* 2016;1(1):3-8.
- Wong TY, Bressler NM. Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA.* 2016;316(22):2366-7.
- López MI, Díez A, Velilla S, Rueda A, Álvarez A, Pastor JC. Prevalence of diabetic retinopathy and eye care in a rural area of Spain. *Ophthalmic Epidemiol.* 2002;9(3):205-14.
- Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *Yearbook of Paediatric Endocrinology.* 2019;16.12.1.
- Krause J, Gulshan V, Rahimy E, Karth P, Widner K, Corrado GS, et al. Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. *Ophthalmology.* 2018;125(8):1264-72.
- Liu H, Li R, Zhang Y, Zhang K, Yusufu M, Liu Y, et al. Economic evaluation of combined population-based screening for multiple blindness-causing eye diseases in China: a cost-effectiveness analysis. *Lancet Glob Health.* 2023;11(3):e456-65.
- Raumviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit Med.* 2019;2:25.
- Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated Diabetic Retinopathy Image Assessment Software: Diagnostic Accuracy and Cost-Effectiveness Compared with Human Graders. *Ophthalmology.* 2017;124(3):343-51.
- Lee AY, Yanagihara RT, Lee CS, Blazes M, Jung HC, Chee YE, et al. Multi-center, head-to-head, real-world validation study of seven automated artificial intelligence diabetic retinopathy screening systems. *Diabetes Care.* 2021;44(5):1168-75.
- Soliz P. Comment on Lee et al. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. *Diabetes Care.* 2021;44:1168-1175. *Diabetes Care.* 2021 May;44(5):e107. doi: 10.2337/dc21-0151.
- Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol.* 2019;103(2):167-75.
- Romero-Aroca P, Verges-Puig R, De La Torre J, Valls A, Relano-Barambio N, Puig D, et al. Validation of a Deep Learning Algorithm for Diabetic Retinopathy. *Telemedicine J E-Health.* 2020;26(8):1001-9.
- Murgatroyd H, Ellingford A, Cox A, Binnie M, Ellis JD, MacEwen CJ, et al. Effect of mydriasis and different field strategies on digital image screening of diabetic eye disease. *Br J Ophthalmol.* 2004;88(7):920-4.
- Van der Heijden AA, Abramoff MD, Verbraak F, Van Hecke MV, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmol.* 2018;96(1):63-8.
- Ribeiro L, Oliveira CM, Neves C, Ramos JD, Ferreira H, Cunha-Vaz J. Screening for Diabetic Retinopathy in the Central Region of Portugal. Added Value of Automated 'Disease/No Disease' Grading. *Ophthalmologica.* 2014 Nov 26. doi: 10.1159/000368426. Online ahead of print.
- Lim G, Bellemo V, Xie Y, Lee XQ, Yip MYT, Ting DSW. Different fundus imaging modalities and technical factors in AI screening for diabetic retinopathy: a review. *Eye Vis.* 2020;7:21.
- Baget-Bernaldiz M, Pedro RA, Santos-Blanco E, Navarro-Gil R, Valls A, Moreno A, et al. Testing a deep learning algorithm for detection of diabetic retinopathy in a spanish diabetic population and with messidor database. *Diagnostics.* 2021;11(8):1385.
- Font O, Torrents-Barrena J, Royo D, García SB, Zarranz-Ventura J, Bures A, et al. Validation of an autonomous artificial intelligence-based diagnostic system for holistic maculopathy screening in a routine occupational health checkup context. *Graefes Arch Clin Exp Ophthalmol.* 2022;260(10):3255-65.
- Silva PS, Cavallerano JD, Haddad NMN, Kwak H, Dyer KH, Omar AF, et al. Peripheral lesions identified on ultrawide field imaging predict increased risk of diabetic retinopathy progression over 4 years. *Ophthalmology.* 2015;122(5):949-56.
- Levenkova A, Sowmya A, Kalloniatis M, Ly A, Ho A. Automatic detection of diabetic retinopathy features in ultra-wide field retinal images. En: Armato SG, Patrick NA. *Medical Imaging 2017: Computer-Aided Diagnosis.* SPIE; 2017.
- Wang K, Jayadev C, Nittala MG, Velaga SB, Ramachandra CA, Bhasrakanand M, et al. Automated detection of diabetic retinopathy lesions on ultrawidefield pseudocolour images. *Acta Ophthalmol.* 2018;96(2):e168-73.

22. Goebel W, Kretzchmar-Gross T. Retinal thickness in diabetic retinopathy: A study using optical coherence tomography (OCT). *Retina*. 2002;22(6):759-67.
23. World Health Organisation. *World report on vision*. Geneva: World Health Organisation; 2019. <https://www.who.int/publications/i/item/9789241516570>
24. Wu Q, Zhang B, Hu Y, Liu B, Cao D, Yang D, *et al*. Detection Of Morphologic Patterns Of Diabetic Macular Edema Using A Deep Learning Approach Based On Optical Coherence Tomography Images. *Retina*. 2021;41(5):1110-17.
25. Varadarajan AV, Bavishi P, Ruamviboonsuk P, Chotcomwongse P, Venugopalan S, Narayanaswamy A, *et al*. Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning. *Nat Commun*. 2020;11(1):130.
26. Hwang DK, Yu WK, Lin TC, Chou SJ, Yarmishyn A, Kao ZK, *et al*. Smartphone-based diabetic macula edema screening with an offline artificial intelligence. *J Chin Med Assoc*. 2020;83(12):1102-6.
27. Alryalat SA, Al-Antary M, Arafa Y, Azad B, Boldyreff C, Ghnaimat T, *et al*. Deep Learning Prediction of Response to Anti-VEGF among Diabetic Macular Edema Patients: Treatment Response Analyzer System (TRAS). *Diagnostics*. 2022;12(2):312.
28. Romero-Aroca P, Verges-Pujol R, Santos-Blanco E, Maarof N, Valls A, Mundet X, *et al*. Validation of a diagnostic support system for diabetic retinopathy based on clinical parameters. *Transl Vis Sci Technol*. 2021;10(3):17.
29. Ribeiro ML, Nunes SG, Cunha-Vaz JG. Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in persons with mild nonproliferative diabetic retinopathy. *Diabetes Care*. 2013;36(5):1254-9.