

Programas informáticos RETIPROGRAM y MIRA

RETIPROGRAM and MIRA software

A. Valls, A. Moreno, J. Pascual-Fontanilles, J. Cristiano, D. Puig, P. Romero Aroca

Resumen

Los datos de pacientes y de su diagnóstico disponibles en los sistemas informáticos de los centros de salud son una fuente de información y conocimiento para desarrollar programas informáticos de soporte a los oftalmólogos. Los métodos de aprendizaje automático (inteligencia artificial [IA]) permiten definir modelos de clasificación de pacientes a partir del análisis de estos datos. La tipología de la información determina las técnicas a utilizar para elaborar estos modelos. En este capítulo, se explicarán dos tipos de sistemas. Primero, el uso de datos clínicos y analíticos de pacientes diabéticos para construir el sistema RETIPROGRAM, que calcula el riesgo de padecer retinopatía diabética. En segundo lugar, las técnicas usadas para construir el sistema MIRA, basadas en el análisis de imágenes del fondo de ojo proporcionadas por un retinógrafo, para determinar el grado de retinopatía diabética (RD).

Palabras clave: Retinopatía diabética. Inteligencia artificial. Sistemas de soporte a la decisión. Aprendizaje profundo. Visión por computadora.

Resum

Les dades de pacients i del seu diagnòstic disponibles als sistemes informàtics dels centres de salut són una font d'informació i coneixement per desenvolupar programes informàtics de suport als oftalmòlegs. Els mètodes d'aprenentatge automàtic (intel·ligència artificial [IA]), permeten definir models de classificació de pacients a partir de l'anàlisi d'aquestes dades. La tipologia de la informació determina les tècniques a utilitzar per elaborar aquests models. En aquest capítol s'explicaran dos tipus de sistemes. Primer, l'ús de dades clíniques i analítiques de pacients diabètics per construir el sistema RETIPROGRAM que calcula el risc de desenvolupar retinopatia diabètica. En segon lloc, les tècniques utilitzades per construir MIRA, basades en l'anàlisi d'imatges del fons d'ull proporcionades per un retinògraf, per determinar el grau de retinopatia diabètica (RD).

Paraules clau: Retinopatia diabètica. Intel·ligència artificial. Sistemes de suport a la decisió. Aprenentatge profund. Visió per ordinador.

Abstract

Data about patients and diagnosis available in the electronic health records is a source of information and knowledge for developing computer programs to support ophthalmologists. Machine learning methods (from artificial intelligence [AI]) can create patient classification models based on the analysis of this data. The typology of the information determines the techniques to be used to elaborate these models. In this chapter, two types of clinical decision support systems will be explained. First, the use of clinical and analytical data from diabetic patients to build the RETIPROGRAM system, which calculates the risk of developing diabetic retinopathy. Secondly, the techniques used to construct MIRA, based on the analysis of fundus images provided by a non-mydratic camera, to determine the degree of diabetic retinopathy (DR).

Key words: Diabetic retinopathy. Artificial intelligence. Decision support systems. Deep learning. Computer vision.

3.2. Programas informáticos RETIPROGRAM y MIRA

RETIPROGRAM and MIRA software

A. Valls^{1,2}, A. Moreno^{1,2}, J. Pascual-Fontanilles¹, J. Cristiano^{1,2}, D. Puig^{1,2}, P. Romero Aroca^{1,2,3}

¹Universitat Rovira i Virgili, Tarragona. ²Institut d'Investigació Sanitària Pere Virgili (IISPV), Reus. Tarragona. ³Servicio de Oftalmologia, Hospital Universitario Sant Joan de Reus, Reus. Tarragona.

Correspondencia:

Pere Romero-Aroca

E-mail: romeropere@gmail.com

Introducción

La IA es la disciplina informática que proporciona métodos para la construcción automatizada de modelos de clasificación de pacientes. Estos modelos se entrenan a partir de un conjunto de datos que contenga información de los pacientes y de su diagnóstico. Así pues, recogiendo y almacenando de forma sistemática las evidencias sobre los factores de riesgo o indicadores de una enfermedad, se puede crear un modelo de clasificación que distinga los pacientes sanos y los que no lo son.

Dentro de la IA, el área que elabora los métodos para construir automáticamente modelos de clasificación se denomina aprendizaje automático. Estos métodos tienen como objetivo imitar a los expertos en la asignación de una etiqueta (sano, enfermedad leve, moderada, etc.) y pertenecen al modelo de aprendizaje supervisado¹. Existen diferentes metodologías de aprendizaje supervisado, como son: redes neuronales, árboles de decisión y máquinas de soporte vectorial, entre otros. No existe una técnica mejor que otra para todos los problemas, sino que para cada caso se debe escoger la más adecuada o comparar los resultados de varias de ellas.

En este capítulo, explicaremos dos sistemas informáticos construidos mediante algoritmos de aprendizaje automático supervisado,

para la ayuda al diagnóstico de la RD, que es una enfermedad ocular derivada de la diabetes *mellitus*². Estos sistemas son:

- RETIPROGRAM: calcula el riesgo de RD a partir de nueve variables clínicas de pacientes diabéticos tipo 2. Clasifica en dos grupos: RD = 0 y RD = 1.
- MIRA: determina si una imagen de fondo de ojo es graduable o no, y en caso positivo, indica el nivel de RD en cuatro posibles niveles: no RD, RD leve, RD moderada o RD severa.

RETIPROGRAM es un sistema de ayuda al cribado para los médicos de familia que visitan regularmente a los diabéticos, y que no son necesariamente expertos en detectar la retinopatía en etapas muy tempranas de su desarrollo. Usando RETIPROGRAM, se puede determinar qué pacientes necesitan hacerse una retinografía, evitando la situación actual en que todos los diabéticos se hacen retinografías. En una segunda fase, el programa MIRA permite ayudar a analizar las retinografías automáticamente. Si el riesgo identificado a partir de la imagen del fondo del ojo es moderado o alto, el paciente se remite al oftalmólogo. En caso de que el riesgo sea nulo o bajo, el sistema programa una visita dentro de uno o dos años. El esquema del funcionamiento de ambos sistemas se puede ver en la Figura 1.

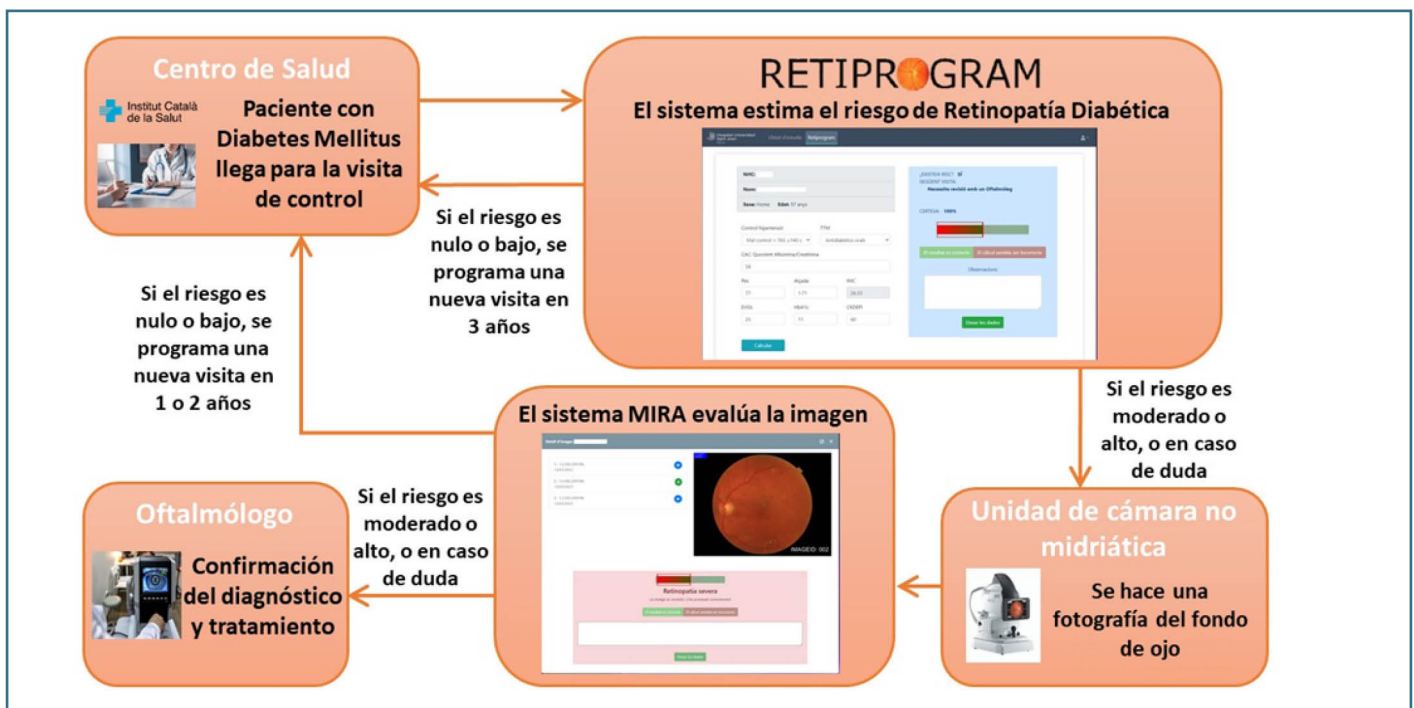


Figura 1. Diagrama de flujo de los sistemas de soporte al diagnóstico de RD.

RETIPROGRAM

Este programa informático se basa en nueve atributos que son indicadores de factores de riesgo de RD, según los estudios médicos². Estos atributos son: sexo, edad, duración de la diabetes, tipo de tratamiento, control de la hipertensión, índice de masa corporal, HbA_{1c}, CKD-EPI y microalbuminuria.

Para construir el sistema, se recopilaron los valores de estos atributos para 2.323 pacientes controlados en los centros de salud de sector sanitario Baix Camp-Priorat, 579 de ellos con RD (clase 1) y 1.744 sin RD (clase 0). Este conjunto de datos se dividió en dos: un conjunto de entrenamiento con 1.212 pacientes y otro conjunto de prueba con 1.111 pacientes, manteniendo en ambos una distribución similar entre clase 0 y 1.

El método de aprendizaje automático utilizado se denomina *fuzzy random forest* (FRF)³, y combina dos técnicas:

- *Fuzzy logic* (lógica difusa): es una técnica de modelado de datos que permite trabajar con etiquetas lingüísticas asociadas a intervalos numéricos. La principal característica es que los límites entre los intervalos no son estrictos, sino

C:\Users\Lluís\Downloads\correccions autors CDS2023\capítulo 3-2\Figura 1.png

difusos, es decir, permiten un cierto grado de solapamiento, que representa la transición no binaria entre una etiqueta y la siguiente. La Figura 2 muestra un ejemplo de la definición de una variable lingüística difusa para el índice de masa corporal. Este tipo de representación permite que el sistema informático sea más flexible a la hora de evaluar los valores de un paciente durante el proceso de decisión.

- *Decision tree* (árbol de decisión): es una estructura como la de la Figura 3, en la que cada nodo corresponde a un atributo, y la rama que desciende del nodo corresponde a un valor posible del atributo. Finalmente, cada hoja del árbol corresponde a una de las dos clases posibles. Cuando un paciente es clasificado por un árbol de decisión, se debe seguir la rama del árbol que corresponde a los valores de los atributos del paciente, hasta que se encuentre la hoja que nos indica cuál es la clase a la que pertenece.

En un FRF, se tienen múltiples (centenares) de árboles de decisión difusos. Para un nuevo caso, cada árbol de decisión del bosque comprueba el grado de satisfacción (difuso) de sus ramas y hace una predicción individual con un cierto grado de certeza.

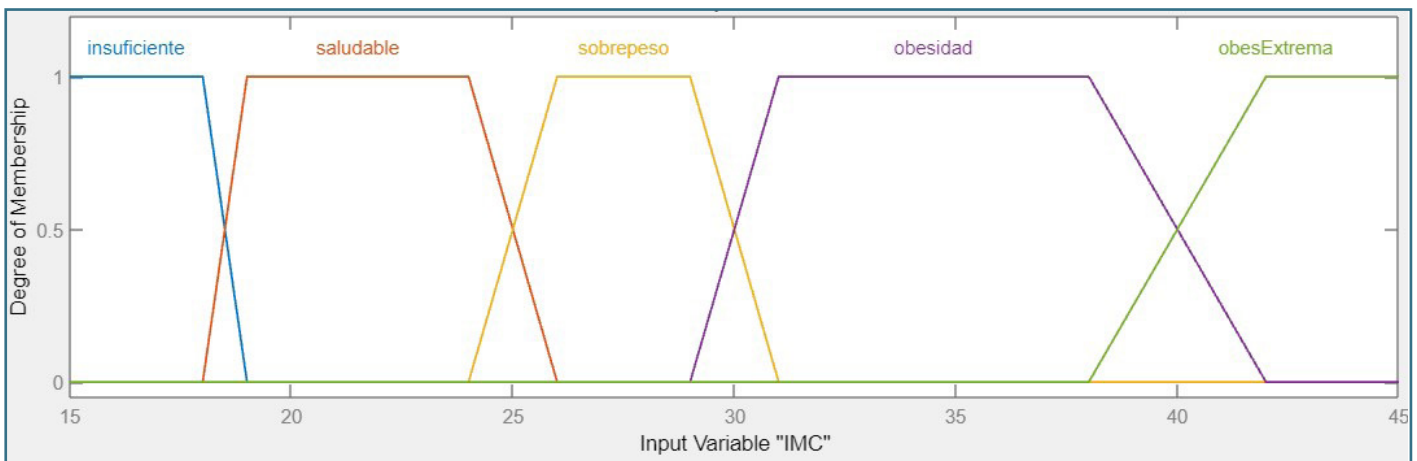


Figura 2. Definición con lógica difusa del índice de masa corporal.

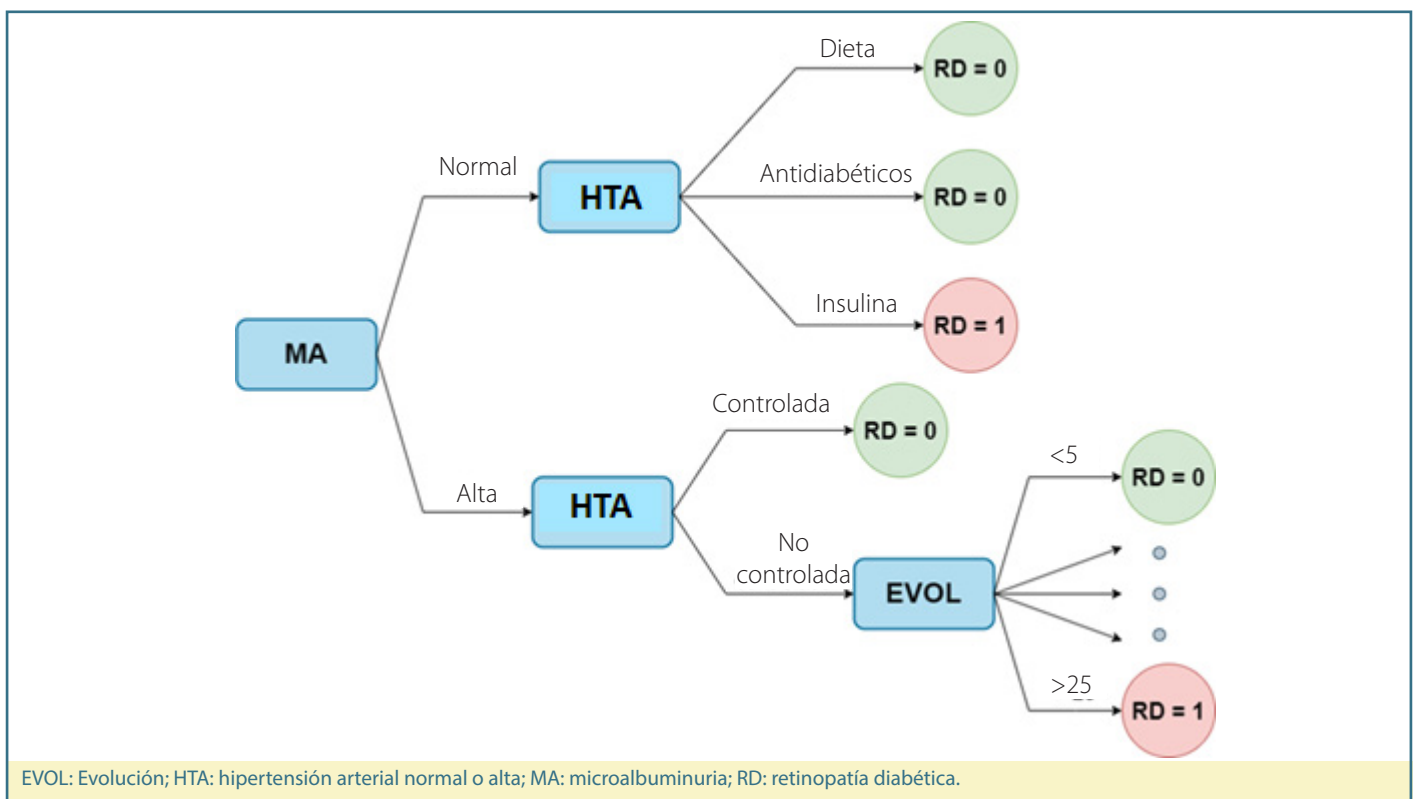


Figura 3. Modelo de árbol de decisión. Se puede observar que a partir de un valor de microalbuminuria (MA) positivo, se añaden dos valores de HTA (hipertensión arterial normal o alta) y a partir de aquí se obtienen tres valores de riesgo para cada combinación de microalbuminuria + HTA.

La predicción final se calcula agregando las predicciones de los árboles individuales, teniendo en cuenta su grado de confianza. La predicción de la clase final se acompaña de un valor de certeza. En caso de duda, el sistema puede no dar una respuesta.

El método utilizado para desarrollar el FRF está basado en técnicas de inducción de árboles a partir de ejemplos. Brevemente, cada uno de los árboles se construye de la siguiente manera. En primer lugar, se seleccionan aleatoriamente 2/3 de los pacientes del

conjunto de entrenamiento. La información sobre esos pacientes se usa para construir un árbol de decisión, iterando los siguientes pasos en cada nodo del árbol:

- Se selecciona aleatoriamente un pequeño conjunto de los atributos restantes.
- Se verifica qué atributo de ese conjunto discrimina mejor entre las dos clases.
- Se crea un nodo con ese atributo, con tantos hijos como valores pueda tomar ese atributo.

En cada nodo, se repite el mismo proceso con los atributos restantes y los ejemplos de entrenamiento que tienen ese valor como atributo, a menos que la mayoría de los ejemplos ya pertenezcan a una clase, en cuyo caso, ya se pone una hoja en el árbol con una predicción de esa clase. Se puede encontrar una descripción más detallada del algoritmo de construcción del FRF en el artículo de Saleh *et al.*³.

Para comparar el rendimiento de diferentes FRF y para comparar los FRF con otros modelos computacionales, se emplearon las medidas estándar de sensibilidad, especificidad y precisión de la clasificación. Hay tres parámetros básicos a considerar en la construcción de un FRF: el número de árboles en el bosque, el número de atributos que se eligen al azar en cada nodo del árbol para continuar el análisis, y el umbral que controla la creación de las hojas del árbol (el porcentaje mínimo de muestras correspondientes a esa rama del árbol, que debe pertenecer a una clase para finalizar el análisis con una hoja). Realizamos un análisis empírico, utilizando una validación cruzada de diez veces en el conjunto de entrenamiento, para determinar los valores de los parámetros. Los mejores resultados en esta validación se obtuvieron con 100 árboles, obteniendo una precisión del 81%, una sensibilidad del 80% y una especificidad del 84%^{4,5}.

La Figura 4 muestra la interfaz actual de RETIPROGRAM. En la parte izquierda, el médico solo debe indicar los valores de los nueve atributos. Después de presionar el botón para ejecutar el clasificador, cada uno de los árboles del FRF hace una predicción (RD o no-RD) con un cierto nivel de confianza, y toda esta información se fusiona en una única respuesta sobre el riesgo del paciente (sí/no) y un nivel de certeza global.

Si la mayoría de los árboles predicen una falta de riesgo, el sistema mostrará una barra con un rectángulo verde a la derecha, cuya longitud depende del nivel de certeza. Si la mayoría de los árboles

Figura 4. Interfaz gráfica de RETIPROGRAM en el programa piloto en Reus.

consideran que el paciente está en riesgo, la barra muestra un rectángulo rojo a la izquierda, cuyo tamaño también depende del nivel de certeza (cuanto mayor sea el riesgo, mayor será el rectángulo). Si no hay riesgo, el sistema sugiere el número de meses hasta la próxima visita del paciente. Los dos botones de la parte derecha permiten confirmar o no el resultado dado por el sistema. Esto se almacena y nos sirve para poder estudiar la calidad de los resultados del RETIPROGRAM durante la práctica clínica real.

MIRA

El cribado preventivo mediante imágenes del fondo de ojo en pacientes con diabetes *mellitus* tipo 2 es una técnica habitual para la detección temprana de la RD, la cual permite el tratamiento oportuno de la enfermedad, permitiendo prevenir la pérdida de la visión y, sobre todo, la ceguera permanente. Las herramientas automáticas de procesamiento y clasificación de imágenes facilitan el arduo trabajo de análisis manual de este tipo de imágenes. Dicha clasificación manual requiere una gran experiencia y mucho tiempo por parte de médicos expertos, para poder analizar detalladamente cada imagen de fondo de ojo.

Un ojo con RD puede presentar algunas o varias de las siguientes lesiones: microaneurismas, exudados, hemorragias y neovascularizaciones. Algunas de ellas son de tamaño muy pequeño, y en varias, su color es rojizo como el fondo de ojo, por lo que resulta complicado distinguirlas manualmente. Por eso es tan importante desarrollar sistemas informáticos que ayuden al análisis de este tipo de imágenes, como es el sistema MIRA.

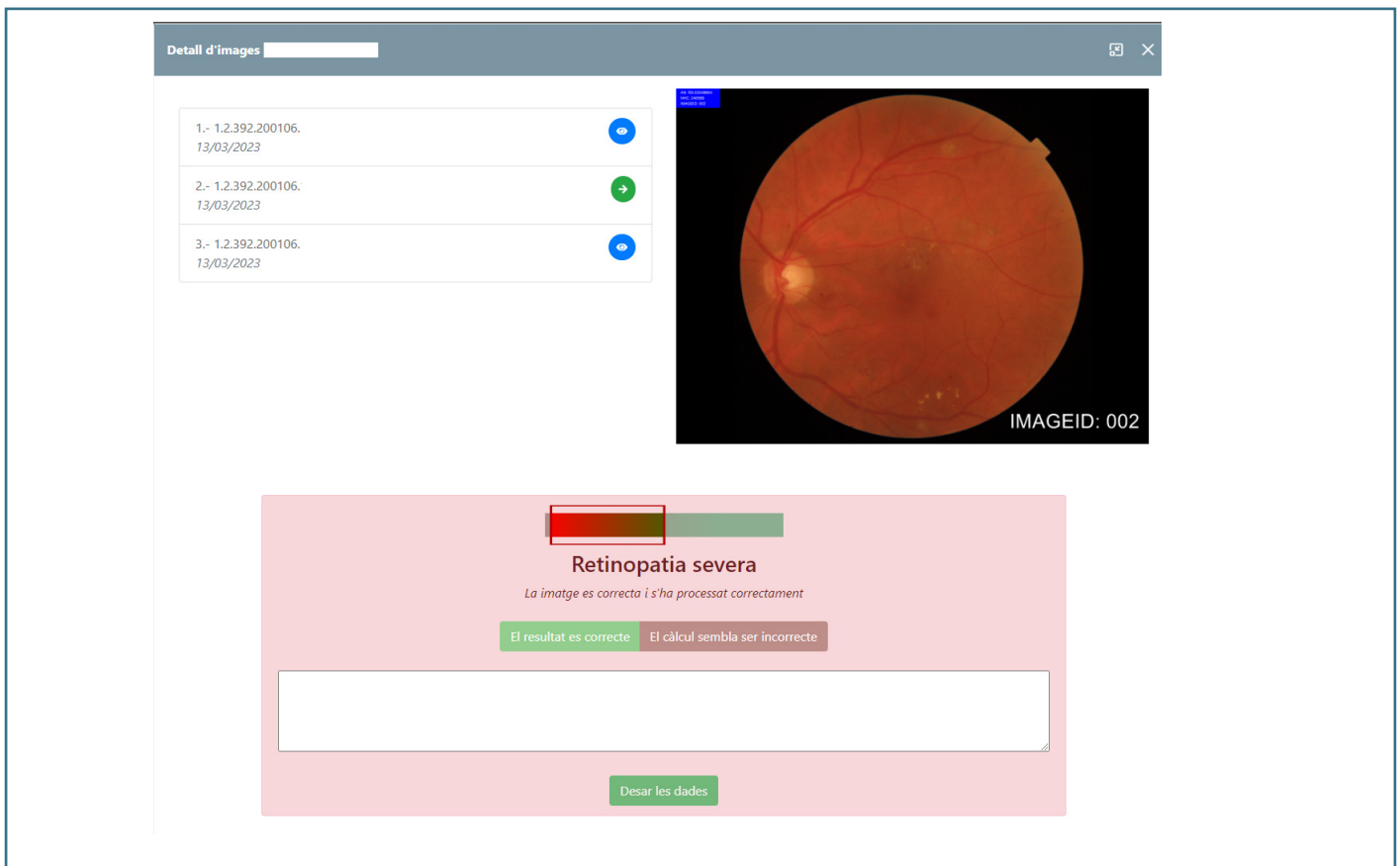


Figura 5. Interfaz gráfica del sistema MIRA en el prototipo del Hospital Universitario Sant Joan de Reus (Cataluña).

La Figura 5 muestra la interfaz del prototipo del sistema MIRA, actualmente probado por el personal sanitario del Hospital Sant Joan de Reus (Cataluña). En la imagen, se muestra un caso de retinopatía severa identificado a partir de la imagen de fondo de ojo de un paciente. Si la imagen es graduable, el sistema procede a realizar la clasificación de esta como se explica en detalle más adelante.

El sistema MIRA no es para uso por parte de los pacientes, es una herramienta de apoyo para los médicos oftalmólogos, quienes deben corroborar si la clasificación presentada por el sistema realmente corresponde con lo que se observa en la imagen. En ese sentido, MIRA puede ser especialmente útil para evitar falsos negativos, ya que el sistema de clasificación automática es capaz de detectar lesiones muy pequeñas que apenas son visibles para la inspección humana. El sistema MIRA es el resultado de una

extensa investigación multidisciplinar y colaborativa de más de diez años entre oftalmólogos expertos en retina del Hospital Sant Joan de Reus y el equipo técnico experto en inteligencia artificial de la Universitat Rovira i Virgili.

Como se ha dicho, el fondo de ojo puede presentar algunas o varias de las siguientes lesiones: microaneurismas, exudados, hemorragias y neovascularizaciones. El número, la localización y el nivel de las lesiones permite clasificar la RD en diferentes estadios¹.

El sistema MIRA es un clasificador de imágenes de fondo de ojo que permite detectar el grado de RD usando estrategias de aprendizaje profundo (*deep learning*) mediante redes neuronales convolucionales (CNN). Son un tipo de redes especializadas en encontrar las correlaciones locales naturales presentes en las imágenes. Estas redes son capaces de crear abstracciones locales que se combinan capa por capa, formando metaabstracciones más

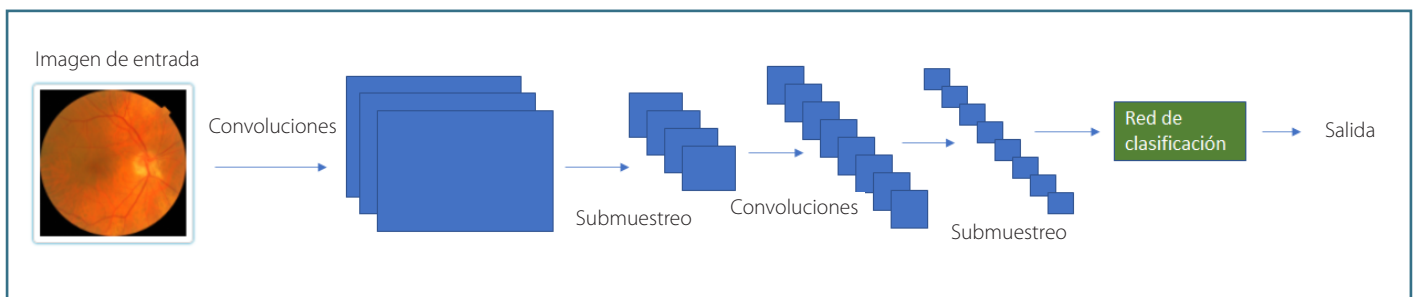


Figura 6. Representación de alto nivel de una red neuronal convolucional.

elaboradas, que representan características que son útiles para resolver problemas de clasificación. En el caso del sistema MIRA, estas redes se usan para clasificar automáticamente una imagen de fondo de ojo en cuatro posibles estadios: no RD, RD leve, RD moderada o RD severa. El sistema también identifica automáticamente si la imagen puede ser clasificada correctamente, o si la misma no puede ser analizada por el sistema debido a problemas en el momento de la captura de esta. Algunos de los problemas que pueden estar presentes en la imagen son: desenfocados, brillos debido a fuentes de iluminación externa, reflejos, artefactos presentes en la imagen, etc.

Para la implementación del sistema MIRA, se han entrenado dos redes CNN, una de ellas es usada para realizar la comprobación automática de la calidad de la imagen suministrada por el oftalmólogo y la otra es usada para estimar el grado de la enfermedad observable en la imagen de fondo de ojo⁶⁷. El sistema MIRA acepta imágenes de fondo de ojo almacenadas en diferentes formatos: jpg, jpeg, png y dicom.

La arquitectura de un sistema de clasificación usando CNN se muestra en la Figura 6. Las capas iniciales de la CNN del sistema MIRA se han diseñado para actuar como extractores de características de las imágenes. Posteriormente, se lleva a cabo la fase de clasificación, la cual es usada para calcular el resultado de clasificación final. Las CNN son un sistema muy robusto y fiable para la clasificación automática de imágenes médicas, el cual es idóneo para la detección de enfermedades oculares.

Para que el algoritmo de clasificación aprenda bien las principales características de las imágenes, este se debe entrenar con una cantidad extensa de datos previamente etiquetados cuidadosamente por los médicos expertos en retina. El entrenamiento de los

algoritmos se realiza usando una o, en lo posible, varias unidades de procesamiento gráfico (GPU, *graphics processing unit*). Estos dispositivos son capaces de realizar millones de operaciones en paralelo, logrando procesar muchos datos simultáneamente, lo cual permite entrenar algoritmos que pueden analizar automáticamente las imágenes de fondo de ojo.

Las imágenes de fondo de ojo usadas para entrenar el sistema han sido capturadas durante varios años por personal especializado en obtención de imágenes de fondo de ojo, y para ello han usado retinógrafos fijos y portátiles.

El sistema MIRA ha sido entrenado utilizando miles de retinografías revisadas y etiquetadas por oftalmólogos expertos en retina que han trabajado en estrecha colaboración con el equipo que trabaja en los algoritmos de IA. Además de las bases de datos públicas, el equipo cuenta con un conjunto de datos privados que se ha utilizado para entrenar y validar el desempeño de los modelos desarrollados. El conjunto de datos privado consta de 40.862 retinografías cuidadosamente etiquetadas. El conjunto de imágenes de fondo de ojo se dividió en un conjunto de entrenamiento conformado por 35.862 retinografías y un conjunto de prueba de 5.000 retinografías. El sistema ha sido validado también con la base de datos Messidor. El sistema MIRA desarrollado tiene una sensibilidad y especificidad superiores al 90% tanto con la base de datos privada como con la base de datos Messidor.

Conclusiones

Los sistemas RETIPROGRAM y MIRA usan técnicas avanzadas de IA para contribuir al cribado masivo de RD en pacientes con diabetes tipo 2. Su uso facilita la detección temprana de la RD, puesto

que permiten detectar el riesgo de padecer la enfermedad para poder prevenir que pacientes diabéticos pierdan visión e inclusive puedan llegar a quedar ciegos.

Por una parte, el sistema RETIPROGRAM se utiliza como un sistema de apoyo a la decisión clínica, utilizando los datos médicos de la historia clínica del paciente, sin necesidad de tomar una imagen del fondo de ojo. Su alta fiabilidad de clasificación, y el hecho de que indique el grado de certidumbre de la respuesta del sistema, permite al profesional determinar qué pacientes no necesitan someterse al cribado por imagen, con el consecuente ahorro en tiempo y dinero.

Por otra parte, el sistema MIRA se utiliza como clasificador de imágenes de fondo de ojo cuando esta prueba se considera necesaria para el diagnóstico. MIRA sirve para detectar el grado de RD mediante el análisis automatizado de imágenes de fondo de ojo capturadas, ya sea con un retinógrafo fijo o portátil. El soporte aportado por MIRA permite al experto hacer una revisión más rápida de la imagen, puesto que solo debe verificar la respuesta dada por el sistema, cosa que permite también una reducción de errores.

Bibliografía

1. Dutt S, Chandramouli S, Das A. *Machine Learning*. Pearson Education India; 2018.
2. Romero-Aroca P, De la Riva-Fernández S, Valls-Mateu A, Sagarra-Alamo R, Moreno-Ribas A, Soler N. Changes observed in diabetic retinopathy: eight-year follow-up of a Spanish population. *Br J Ophthalmol*. 2016;100(10):1366-71.
3. Saleh E, Błaszczyński J, Moreno A, Valls A, Romero-Aroca P, De la Riva-Fernández S, *et al*. Learning ensemble classifiers for diabetic retinopathy assessment. *Artif Intell Med*. 2018;85:50-63.
4. Romero-Aroca P, Valls A, Moreno A, Sagarra-Alamo R, Basora-Gallisa J, Saleh E, *et al*. A Clinical Decision Support System (CDSS) for diabetic retinopathy screening. Creating a clinical support application. *Telemed J E Health*. 2019;25(1):31-40.
5. 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.
6. De La Torre J, Valls A, Puig D. A deep learning interpretable classifier for diabetic retinopathy disease grading. *Neurocomputing*. 2020;396:465-76.
7. Romero-Aroca P, Verges-Puig R, De la Torre J, Valls A, Relaño-Barambio N, Puig D, *et al*. Validation of a Deep Learning Algorithm for Diabetic Retinopathy. *Telemed J E Health*. 2020;26(8):1001-9.