

Revolucionando el Diagnóstico: el Potencial de la Inteligencia Artificial y la Visión por Computador en Medicina

Xavier Sevillano

Resumen—A raíz del descubrimiento accidental de los rayos X a finales del siglo XIX, la práctica médica se dotó de una herramienta que permitía lo que hasta aquel momento era impensable: ver el interior del cuerpo de los pacientes para poder diagnosticar sin necesidad de hacer una sola incisión. Al tiempo que los avances en imagen médica han permitido obtener cada vez imágenes más detalladas en dos y tres dimensiones, la evolución de las tecnologías relacionadas con la visión por computador y la inteligencia artificial han impulsado el desarrollo de técnicas de apoyo al diagnóstico. La convergencia entre estas dos disciplinas da lugar a un campo científico a caballo de la ingeniería y la medicina, donde algoritmos entrenados con grandes cantidades de datos pueden contribuir a un diagnóstico más preciso, incrementando la eficiencia del personal sanitario y revirtiendo en una mejor atención sanitaria y un mayor bienestar de los pacientes.

Palabras clave: imagen médica, diagnosis, inteligencia artificial, visión por computador

I. INTRODUCCIÓN

GRACIAS a las tecnologías de imagen médica es posible estudiar las estructuras y órganos internos del cuerpo humano sin causar daño alguno al paciente. Por este motivo, la imagen médica es hoy en día una herramienta diagnóstica y terapéutica de uso común en la práctica clínica: radiografía, ecografía, resonancia magnética, tomografía por emisión de positrones ... Todas ellas son modalidades de imagen médica que se diferencian en los procesos físicos en los que se basan para obtener imágenes del interior del cuerpo. En este sentido, se pueden clasificar dichas modalidades según i) el proceso de propagación de dichas ondas respecto al tejido que se pretende analizar, y ii) la frecuencia de las ondas físicas empleadas en el proceso.

Respecto al proceso de propagación, podemos distinguir entre técnicas basadas en:

- **Atenuación:** las ondas atraviesan el tejido, y dependiendo de la densidad de este se ven atenuadas en mayor o menor medida. Un conjunto de sensores situados tras el tejido capta la cantidad de radiación que lo ha atravesado, generando la imagen resultante. Este es el proceso de funcionamiento de las radiografías, que se basan en rayos X (Fig. 1a).
- **Reflexión:** las ondas penetran en el tejido, y rebotan en las estructuras internas de este, siendo recogidas por un sensor. Así es como funcionan las ecografías, basadas en ondas de ultrasonidos (Fig. 1b).
- **Emisión:** en este caso, el origen de las ondas se sitúa en el interior del cuerpo, previa inyección de una sustancia ligeramente radiactiva denominada contraste. La radiación de este material es captada por sensores que generan la imagen correspondiente. Este es el principio de funcionamiento de la tomografía por emisión de positrones (Fig. 1c).

En cuanto a la frecuencia de las ondas físicas empleadas en el proceso de captura de la imagen (Fig. 2), ésta va desde el rango de las ondas de radio (alrededor de 10.000 Hercios o ciclos por segundo) empleadas en las resonancias magnéticas hasta el rango de los rayos gamma (frecuencias 16 órdenes de magnitud mayores) que se usan en la tomografía por emisión de positrones, pasando por el rango

de los infrarrojos (como en la imagen térmica que se emplea para detectar personas con fiebre) o el de la imagen visible (usada, por ejemplo, en endoscopias).

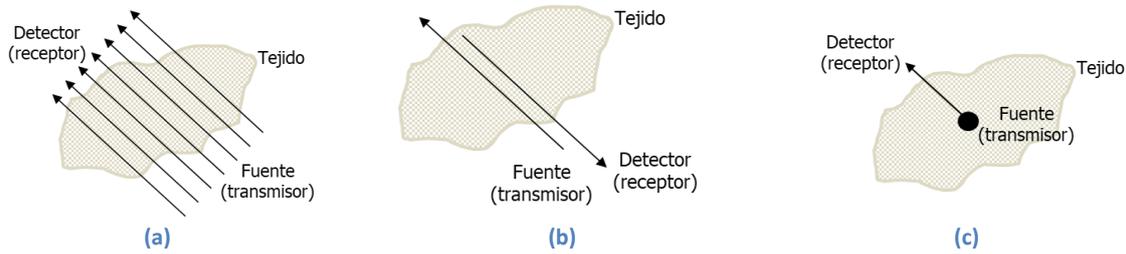


Fig. 1 Principios físicos de la imagen médica (basado en [1])

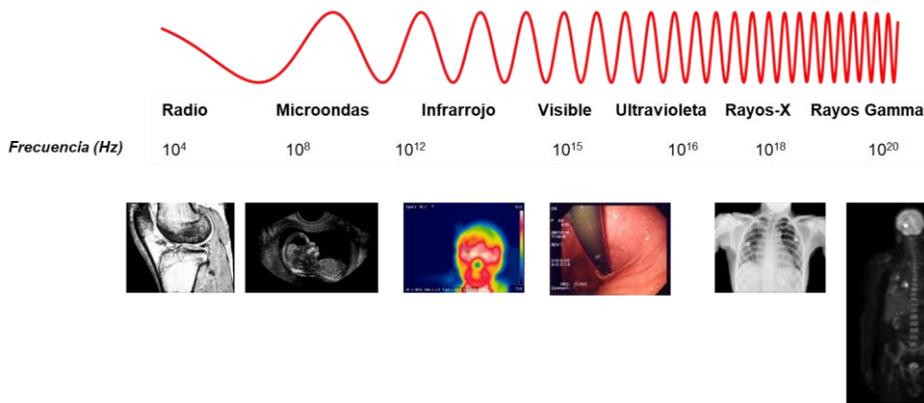


Fig. 2 Rangos de frecuencias empleados en imagen médica y ejemplos

Tradicionalmente, las imágenes captadas a través de estos medios han sido analizadas por parte de personal médico y sanitario, quienes a partir de su experiencia interpretan las imágenes.

No obstante, los recientes avances en inteligencia artificial y visión por computador han dado lugar a algoritmos que, puestos al servicio de la interpretación de estas imágenes, ofrecen al profesional de la salud una valiosa herramienta adicional en los procesos diagnósticos y terapéuticos basados en imagen médica.

II. DOS INSTANTES CLAVE EN LA HISTORIA DEL DIAGNÓSTICO MÉDICO POR IMAGEN

El primer paso hacia el diagnóstico médico por imagen fue el descubrimiento accidental de los rayos X por parte del físico alemán Wilhelm Conrad Roentgen en 1895, hallazgo que le valió el Premio Nobel de Física en 1901. Los rayos X son una radiación electromagnética ionizante (por lo que es preciso limitar la exposición a ellos), invisible para el ojo humano, capaz de penetrar cuerpos opacos en mayor medida cuanto menor es la densidad de éstos e imprimir películas fotográficas. Es por ello que los rayos X supusieron una revolución en el campo médico, ya que permitieron por primera vez generar imágenes de órganos y huesos.

Otro salto de gigante en este ámbito fue el descubrimiento de la tomografía axial computerizada (TAC) en los años 70 del siglo XX. Esta invención, atribuida a los científicos Godfrey Hounsfield y Allan M. Cormack (y por la que lograron el Premio Nobel de Medicina en 1979), permite obtener imágenes tridimensionales del interior del cuerpo del paciente, ofreciendo al personal sanitario una información mucho más rica y detallada acerca de la estructura, posición y forma de tumores, órganos, etc. A grandes rasgos, la TAC se basa en una fuente de rayos X motorizada que dispara haces estrechos de rayos X a medida que gira alrededor del paciente, obteniendo distintos cortes de imagen que pueden mostrarse individualmente en forma bidimensional, o apilados juntos para generar una imagen tridimensional que puede revelar estructuras anormales, o ayudar al médico a planear y monitorear los tratamientos.

III. ¿QUÉ ES LA INTELIGENCIA ARTIFICIAL? ¿Y LA VISIÓN POR COMPUTADOR?

Aunque pueda parecer un concepto nuevo, el término “inteligencia artificial” (IA) se acuñó en la década de los 50 del siglo XX. En concreto, en 1955, un grupo de investigadores norteamericanos punteros en el ámbito de la informática, la electrónica y las comunicaciones organizaron un encuentro para discutir qué se entendía por IA [1].

De hecho, la IA comprende muchos ámbitos, entre los que destaca el aprendizaje automático (o *machine learning*, en inglés), que es la disciplina en la que los investigadores en IA centraron sus principales esfuerzos desde la década de 1980 hasta la de 2010.

A partir de ese momento, en el ámbito de la IA se impuso un cambio de paradigma a causa de la evolución del aprendizaje profundo (o *deep learning*, en inglés), un subtipo de aprendizaje automático basado en unos modelos llamados *redes neuronales profundas*.

Las redes neuronales no eran nada nuevo en 2010. De hecho, eran uno de los modelos fundacionales del aprendizaje automático, y se habían usado con frecuencia durante décadas, aunque sus prestaciones como algoritmos de aprendizaje y clasificación automática eran consideradas inferiores a los de otros algoritmos más avanzados, como las máquinas de soporte vectorial [2].

No obstante, los avances tecnológicos de inicios del siglo XXI dieron lugar a arquitecturas de ordenadores con mucha más potencia de cálculo, y esto a su vez permitió evolucionar e implementar modelos de redes neuronales *profundas*, mucho más complejas que sus predecesoras. A su vez, se beneficiaron de la creciente disponibilidad de datos necesarios para el entrenamiento de dichos modelos. Esta doble conjunción (ordenadores más potentes y gran disponibilidad de datos) alumbró la era del *deep learning*, responsable en gran medida del *boom* de la IA producido en la segunda década del siglo XXI.

El primer campo de aplicación en que las redes neuronales profundas demostraron su potencial y dieron pie al cambio de paradigma desde el *machine learning* hacia el *deep learning* fue la visión por computador (VC), que es la disciplina que trata de que los ordenadores interpreten el contenido de una imagen de la misma forma que un humano. En 2012, los investigadores Alex Krizhevsky, Ilya Sutskever y Geoffrey Hinton presentaron AlexNet [3], una red neuronal profunda capaz de resolver el mayor reto de clasificación automática de imágenes -el llamado “ImageNet Large Scale Visual Recognition Challenge”- con una precisión 10 puntos superior al del resto de participantes en dicho reto. A partir de ese momento, la VC y todas las disciplinas en las que se aplicaba hasta entonces el aprendizaje automático (el procesamiento de señales, de lenguaje natural, etc.) iniciaron su transición hacia el aprendizaje profundo.

IV. VISIÓN POR COMPUTADOR E IMAGEN MÉDICA: UN ENCAJE PERFECTO

La imagen es una de las principales modalidades de adquisición de información en el ámbito de la biología y la medicina. Bien sea usando un microscopio en un laboratorio o un equipo de resonancia magnética en un centro hospitalario, cada día se adquieren ingentes cantidades de imágenes vinculadas al estudio, diagnóstico o tratamiento de múltiples patologías. En la mayoría de las ocasiones, el análisis de estas imágenes lo realiza personal experto en la interpretación de las mismas, lo que supone un cuello de botella en el proceso, pues la cantidad de imágenes disponibles supera la capacidad de análisis humano.

Por este motivo, existe lo que consideramos un encaje perfecto entre los ámbitos de la imagen médica y la visión por computador: una disciplina que genera grandes cantidades de imágenes que necesitan ser analizadas, y otra que ofrece herramientas para el automatizar el análisis de estas. En este contexto, la literatura científica reciente ofrece ejemplos de cómo la visión por computador permite resolver diversos problemas relacionados con la imagen médica.

A. Ejemplo#1: IA y VC para reducir la carga de trabajo en radiología en “screenings” de cancer de mama.

En un trabajo conjunto entre investigadores de la Universidad de Haifa (Israel), IBM Research y el hospital Johns Hopkins de Estados Unidos, se creó un sistema de IA para reducir la carga de trabajo de los departamentos de radiología en el proceso de “screening” de cáncer de mama [4]. Cuando una paciente se somete a una mamografía, el primer paso realizado por los radiólogos es el “screening”, tarea consistente en detectar y descartar los casos negativos. El reto en este proceso es doble: en primer lugar, no dar como negativo ningún caso positivo. Y, en segundo lugar, minimizar la tasa de rellamada, es decir, evitar en la medida de lo posible tomar un caso negativo como positivo. Lógicamente, esta es una tarea que requiere de profesionales muy experimentados, y supone una carga de trabajo notable que se beneficiaría en gran medida de la automatización. Gracias a la herramienta presentada en este artículo se redujo un 40% la carga de trabajo de los radiólogos, ya que el sistema automático de análisis de las tomografías de mama permitió detectar con gran precisión los casos negativos. Además, la colaboración entre el sistema de IA y los radiólogos permitió mantener cerca del 90% la identificación de casos positivos, aumentar del 91% al 93% la identificación de casos negativos y reducir la tasa de rellamada en un 25%.

B. Ejemplo#2: IA y VC para la prognosis de la enfermedad de Alzheimer mediante resonancias magnéticas cerebrales.

Otro ejemplo es el artículo presentado por investigadores de las universidades chinas de Xi'an Jiaotong y ShanghaiTech, en colaboración con científicos de la Universidad de Carolina del Norte en Chapel Hill y la Korea University, en la que presentaron una herramienta de IA para la prognosis de Alzheimer basada en el análisis automático de resonancias magnéticas (RM) del cerebro [5]. Dicho sistema facilita realizar un diagnóstico y un seguimiento preciso en etapas tempranas para retrasar la progresión de la enfermedad de Alzheimer. Actualmente, el estado de progresión de la enfermedad se analiza mediante tests neuropsicológicos (cuestionarios, realización de tareas, etc.). Aunque la puntuación obtenida por los pacientes en estos test es un indicador muy fiable del estado de la enfermedad, su realización supone un coste temporal elevado, por lo que los autores de esta publicación se propusieron predecir las puntuaciones automáticamente a partir de RM cerebrales. Los resultados presentados en este trabajo no solamente indican que el análisis de las imágenes cerebrales permite predecir de forma precisa las puntuaciones obtenidas por los pacientes en los tests neuropsicológicos (especialmente en casos incipientes de demencia), sino que además las regiones cerebrales identificadas como más vinculadas a la enfermedad son específicas de cada paciente y son biológicamente significativas.

C. Ejemplo#3: IA y VC para la interpretación de radiografías de tórax.

Un último ejemplo de investigación en el uso de IA y VC en imágenes médicas es el trabajo desarrollado entre las universidades de Nueva York, Stanford y Harvard Medical School en el que se desarrolló una herramienta para interpretar radiografías de tórax de forma automática [6]. Uno de los puntos de mayor interés de este trabajo reside en que se centra en la interpretabilidad y explicación de los métodos de IA: es decir, el personal sanitario solamente confiará en las herramientas basadas en IA si estas son capaces de “explicar” sus decisiones, y, lógicamente, si estas coinciden con las del propio personal sanitario. En el caso concreto de esta herramienta, se aplican mapas de saliencia para indicar las zonas de la imagen en las que los algoritmos detectan las lesiones relacionadas con la patología torácica del paciente, y los resultados obtenidos muestran que la herramienta tiene dificultades en los casos de lesiones de pequeño tamaño y forma compleja. Es decir, la explicación de la IA es uno de los mayores retos a futuro.

V. HACIA LA PRIMERA LÍNEA DE ATENCIÓN MÉDICA

El siguiente paso en esta evolución consiste en trasladar los avances de la IA y la VC a la primera línea de atención médica. Nuestro grupo de investigación ha desarrollado proyectos y prototipos orientados en esta dirección. Lo que sigue es una breve descripción de algunos de estos trabajos, todos ellos de carácter multidisciplinar como no puede ser de otro modo.

A. Ejemplo#1: Biomarcadores faciales cerebrales y genéticos para el diagnóstico del trastorno psicótico.

La esquizofrenia y el trastorno bipolar pertenecen al espectro de los trastornos psicóticos. Afectan al 3% de la población, comparten un trasfondo genético común y pueden desencadenarse por factores genéticos y/o ambientales como el estrés. Son enfermedades altamente incapacitantes, que generan mucho sufrimiento al paciente y a su entorno social y familiar. Además, su diagnóstico suele ser tardío y tiene una componente subjetiva, puesto que en la actualidad solamente se diagnostica mediante entrevista clínica una vez el paciente ha sufrido uno o varios episodios psicóticos. Por tanto, todos los esfuerzos dirigidos a la obtención de un diagnóstico precoz y fiable redundarán en beneficio de los pacientes y sus familias.

En el proyecto BeGiN, llevado a cabo en colaboración entre la Universitat de Barcelona, FIDMAG Research Foundation y La Salle-Universitat Ramon Llull se aborda la identificación de biomarcadores diagnósticos de trastornos psicóticos desde tres vertientes: biomarcadores cerebrales, faciales y genéticos, que están interrelacionados a través de la teoría del neurodesarrollo [7].

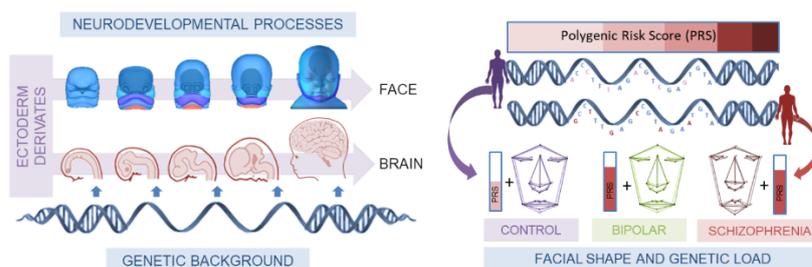


Fig. 3 Biomarcadores faciales, cerebrales y genéticos, vinculados a través de la teoría del neurodesarrollo, para el diagnóstico de las enfermedades del espectro de los trastornos psicóticos

En este contexto, la aplicación de la IA y la VC se centra en las imágenes de RM del cráneo del paciente. En primer lugar, para la obtención de biomarcadores faciales se requiere reconstruir y normalizar [8] de forma automática la anatomía facial 3D del individuo procesando la imagen de RM, y detectar la posición de puntos faciales anatómicos de interés mediante redes neuronales convolucionales. Por otro lado, la obtención de biomarcadores cerebrales requiere de la aplicación de herramientas de *deep learning* para la detección de regiones cerebrales anatómicamente distintas entre individuos sanos y pacientes de trastornos psicóticos, que permitan identificar biomarcadores cerebrales de potencial diagnóstico a partir del análisis automático de las imágenes de RM [9].

B. Ejemplo #2: Diagnóstico del vértigo posicional paroxístico benigno en servicios de urgencia y atención primaria.

El vértigo posicional paroxístico benigno (VPPB) es la causa más común de vértigo, siendo el motivo de cerca del 3% de las visitas a consultas de atención primaria. Se trata de una dolencia muy incapacitante, pero de solución sencilla, consistente en realizar al paciente unas maniobras de diagnóstico y tratamiento. No obstante, un porcentaje bajo del personal sanitario en atención primaria y servicios de urgencia posee formación acerca de la realización de dichas maniobras, lo que conduce a un mal diagnóstico y tratamiento en muchos casos, incluyendo la prescripción de medicamentos innecesarios [10].

En colaboración con el grupo de investigación VERTAP del centro de investigación IDIAP Jordi Gol, se desarrolló un prototipo *software* con el objetivo de ayudar a formar a los profesionales sanitarios en el diagnóstico y tratamiento del VPPB, comenzando por la implementación de un algoritmo de VC para la detección de uno de los síntomas más claros del VPPB: el nistagmo, una serie de movimientos rápidos de la pupila que se producen como respuesta a la maniobra diagnóstica del VPPB. En este contexto, se desarrollaron algoritmos de detección y seguimiento de la pupila, y para la obtención de patrones de dirección y velocidad del movimiento, con el objetivo de posteriormente entrenar modelos de aprendizaje automático para la detección y clasificación del nistagmo.

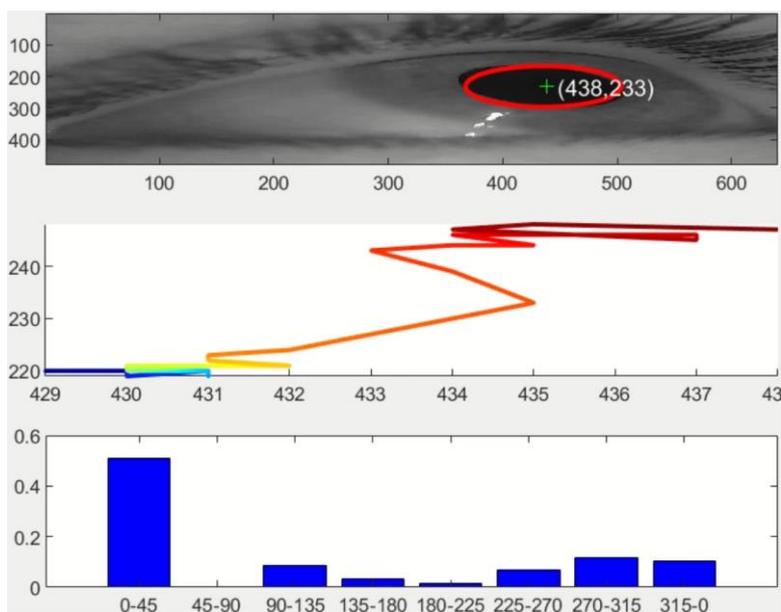


Fig. 4 Detección del nistagmo mediante detección y seguimiento del movimiento de la pupila y la caracterización de este en términos de velocidad y dirección

C. Ejemplo #3: Monitorización de úlceras de presión mediante visión por computador.

Las úlceras de presión son lesiones crónicas sufridas por pacientes que pasan mucho tiempo inmovilizados. Los daños causados por la presión causan una isquemia de las estructuras cutáneas subyacentes, haciendo que estas lesiones aparezcan más frecuentemente en prominencias óseas como el área sacra y el talón. En Estados Unidos, se estima que más de 2,5 millones de personas sufren anualmente úlceras de presión, lo que supone un coste sanitario superior a los 25.000 millones de dólares.

En un proyecto llevado a cabo en colaboración entre la Universidad de Deusto y La Salle-Universitat Ramon Llull se aborda el desarrollo de herramientas que permitan monitorizar el estado y la evolución de las úlceras de presión mediante herramientas de VC e IA [11]. Mediante un dispositivo portátil (como una *tablet*) el personal sanitario toma imágenes de las úlceras del paciente, evitando así el contacto, la manipulación dolorosa y minimizando el riesgo de infección de las heridas. Dichas imágenes se procesan automáticamente, detectando y midiendo la superficie de las distintas regiones de la úlcera, identificando los distintos tipos de tejidos que son visibles y ofreciendo un análisis de su evolución, dando soporte al personal sanitario para su tratamiento.

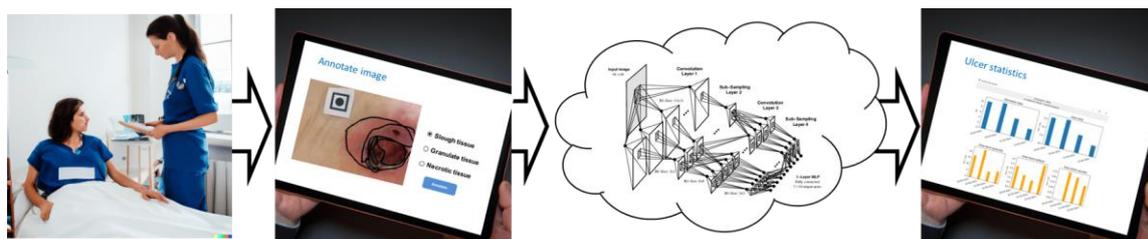


Fig. 5 Monitorización de la evolución de úlceras de presión mediante IA y VC

VI. CONCLUSIONES

- Desde el descubrimiento de los rayos X a finales del siglo XIX y de la tomografía computerizada en los años 70 del siglo XX, la imagen médica se ha convertido en una herramienta diagnóstica y terapéutica de uso corriente en la práctica clínica.
- Existen diversas modalidades de imagen médica, que se diferencian según la naturaleza y frecuencia de las ondas utilizadas: rayos X, resonancia magnética, ultrasonidos, etc.
- En la última década se ha producido un cambio de paradigma clave en la inteligencia artificial y la visión por computador debido a la emergencia del aprendizaje profundo (o *deep learning*). A raíz de esto, cada vez está más clara la convergencia entre estas tecnologías y el diagnóstico por imagen médica.
- Para que esta convergencia sea exitosa, es necesario un diálogo fluido entre ingenieros y biólogos/médicos, lo que requiere nuevos perfiles profesionales a medio camino entre ambos ámbitos como ingenieros de la salud o ingenieros biomédicos.
- Para que estas tecnologías se puedan seguir desarrollando e integrando en los flujos de diagnóstico médica, es necesario seguir trabajando en varios retos, como son la definición de protocolos para su validación, la accesibilidad a datos médicos para entrenar los algoritmos, el desarrollo de mecanismos para la explicación de los mismos o llevar estos avances a la primera línea de atención médica.

VII. AGRADECIMIENTOS

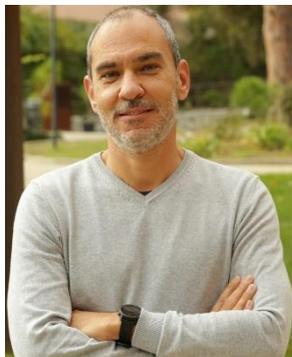
- A la Fondation Jérôme Lejeune (Francia),
- A la Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) de la Generalitat de Catalunya,
- A Aristos Campus Mundus – Campus de Excelencia Internacional,
- Al Programa José Castillejo (Ministerio de Educación),
- Al Ministerio de Ciencia y Universidades,
- Al programa de divulgación de ciencia y tecnología *ciènciaprop*[®]

VIII. REFERENCIAS

- [1] J. McCarthy, M.L. Minsky, N. Rochester & C.E. Shannon (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI magazine*, 27(4), 12-12.
- [2] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt & B. Scholkopf (1998). Support vector machines. *IEEE Intelligent Systems and their Applications*, vol. 13, no. 4, pp. 18-28.
- [3] A. Krizhevsky, I. Sutskever & G.E. Hinton (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, vol. 25.
- [4] Y. Shoshan, R. Bakalo, F. Gilboa-Solomon, V. Ratner, E. Barkan, M. Ozery-Flato, M. Amit, D. Khapun, E.B. Ambinder, E.T. Oluyemi, B. Panigrahi, P.A. DiCarlo, M. Rosen-Zvi & L.A. Mullen (2022). Artificial intelligence for reducing workload in breast cancer screening with digital breast tomosynthesis. *Radiology*, 303(1), 69–77.
- [5] C. Lian, M. Liu, L. Wang & D. Shen (2022). Multi-task weakly-supervised attention network for dementia status estimation with structural MRI. *IEEE Transactions on Neural Networks and Learning Systems*, 33(8), 4056–4068.
- [6] A. Saporta, X. Gui, A. Agrawal *et al.* (2022). Benchmarking saliency methods for chest X-ray interpretation. *Nature Machine Intelligence* 4, 867–878.
- [7] S. Kloiber *et al.* (2020). Neurodevelopmental pathways in bipolar disorder. *Neuroscience & Biobehavioral Reviews* 112, 213–226.
- [8] Á. Heredia-Lidón, A. González, C. Guerrero-Mosquera, R. González-Colom, L.M. Echeverry, N. Hostalet, R. Salvador, E. Pomarol-Clotet, J. Fortea, N. Martínez-Abadías, M. Fatjó-Vilas & X. Sevillano (2023). Automated orientation detection of 3D head reconstructions from sMRI using multiview orthographic projections: an image classification-based approach. *Iberian Conference on Pattern Recognition and Image Analysis 2023*, 603-614.
- [9] J. Malé, Y. Heuzé, J. Fortea, N. Martínez-Abadías & X. Sevillano (2023). Single-landmark vs. multi-landmark deep learning approaches to brain MRI landmarking: a case study with healthy controls and Down syndrome individuals. *British Machine Vision Conference 2023*.
- [10] H. Benecke, S. Agus, D. Kuessner, G. Goodall & M. Strupp (2013). The burden and impact of vertigo: findings from the REVERT patient registry. *Frontiers in Neurology*, 4:136.

- [11] S. Zahia, M.B. Garcia-Zapirain, X. Sevillano, A. González, P.J. Kim & A. Elmaghraby (2020). Pressure injury image analysis with machine learning techniques: a systematic review on previous and possible future methods. *Artificial Intelligence in Medicine*, vol. 102:101742.

IX. AUTOR



XAVIER SEVILLANO nació en Girona en 1975. Cursó sus estudios preuniversitarios en el colegio La Salle de Girona. Se tituló como Ingeniero Técnico de Telecomunicación e Ingeniero Superior en Electrónica en La Salle - Universitat Ramon Llull en los años 1997 y 2000, respectivamente. Después de cursar un Máster en Gestión de Proyectos en 2002, consiguió su doctorado por la Universitat Ramon Llull en 2009. En 2011 fue galardonado por el Ministerio de Educación con una ayuda José Castillejo que le permitió realizar una estancia como investigador postdoctoral en el Multimedia and Vision Research Group, en Queen Mary – University of London. Desde 2015 es profesor titular de la Universitat Ramon Llull, donde actualmente coordina la línea de investigación en Visión por Computador dentro del grupo de investigación HER (Human-Environment Research). Su investigación se centra en la aplicación del procesamiento de la imagen y la visión por

computador en los ámbitos biológico y médico. Ha publicado numerosos artículos en congresos y revistas internacionales, liderando múltiples proyectos de investigación y de transferencia de tecnología.