

Modelos de lenguaje multimodal: explorando el uso de la IA en oftalmología

Large language model: exploring the use of AI in Ophthalmology

R. Bouchikh-El Jarroudi^{1,2}, FJ. Valentín-Bravo^{1,3}, L. Broc-Iturralde¹, T. Croitoru-Croitoru¹, A. Ross Jenkinson-Macia¹, X. Valldeperas¹

¹Servicio de Oftalmología. Hospital Universitari Germans Trias i Pujol. Badalona. Barcelona. España. ²Unidad de Oftalmología Basada en Evidencias (Oftalmoevidencia). Scientia Clinical and Epidemiological Research Institute. Trujillo. Perú. ³Institut Català de la Retina (ICR). Barcelona. España.

Correspondencia:

Rachid Bouchikh El Jarroudi

E-mail: rachid.bascara@gmail.com

Resumen

Introducción: La inteligencia artificial (IA) ha emergido como una herramienta crucial en diversas disciplinas médicas, incluyendo la oftalmología. Este avance se sustenta en la capacidad de las computadoras para realizar tareas tradicionalmente exclusivas del ser humano mediante sistemas de aprendizaje automático. Los modelos de lenguaje multimodal (LLM por su nombre en inglés *Large Language Model*), como ChatGPT-4, destacan por su habilidad para responder a consultas de texto libre y, más recientemente, imágenes.

Caso clínico: Este estudio trata de evaluar potencial diagnóstico de un LLM desarrollado por Open AI, ChatGPT-4, para proporcionar un análisis interpretativo de imágenes oftalmológicas realizadas mediante biomicroscopía de polo anterior.

Conclusión: La IA y los LLM son herramientas prometedoras para mejorar la eficiencia y precisión en el diagnóstico y tratamiento en nuestra disciplina. Sin embargo, a pesar del conocimiento demostrado por esta herramienta, y la gran capacidad para definir diferentes estructuras oculares, la precisión y fiabilidad de los informes aportados actualmente son limitados.

Palabras clave: Inteligencia artificial. IA. ChatGPT. Grandes modelos de lenguaje. Imágenes. Oftalmología.

Resum

Introducció: La intel·ligència artificial (IA) ha emergit com una eina crucial en diverses especialitats mèdiques, inclosa l'oftalmologia. Aquest avanç es sustenta en la capacitat dels ordinadors per executar tasques tradicionalment exclusives del ésser humà mitjançant sistemes d'aprenentatge automàtic. Els models de llenguatge multimodal (LLM pel seu nom en anglès *Large Language Model*), com ChatGPT-4, destaquen per la seva habilitat per respondre a consultes de text lliure i, més recentment, imatges.

Cas clínic: Aquest estudi tracta d'avaluar el potencial diagnòstic d'un LLM desenvolupat per OPEN AI, ChatGPT-4, per proporcionar una anàlisi interpretativa de les imatges oftalmològiques realitzades mitjançant biomicroscòpia de pol anterior.

Conclusió: La IA i els LLM són eines prometedores per millorar l'eficiència i precisió en el diagnòstic i tractament en la nostra disciplina. No obstant, tot i el coneixement demostrat per aquesta eina, i la gran capacitat per definir diferents estructures oculars, la precisió i fiabilitat dels informes aportats actualment són limitats.

Paraules clau: Intel·ligència artificial. IA. ChatGPT. Models de llenguatge multimodal. Imatges. Oftalmologia.

Abstract

Introduction: Artificial intelligence (AI) has emerged as a crucial tool in several medical disciplines, including ophthalmology. This advance is based on the ability of computers to perform tasks traditionally exclusive to humans using machine learning systems. Multimodal language models (LLMs), such as ChatGPT-4, are notable for their ability to respond to free-text queries and, more recently, images.

Clinical case: This study aims to evaluate the diagnostic potential of an Open AI-developed LLM, ChatGPT-4, to provide interpretive analysis of ophthalmological images performed by anterior pole biomicroscopy.

Conclusion: AI and LLMs are promising tools to improve the efficiency and accuracy of diagnosis and treatment in our discipline. However, despite the knowledge demonstrated by this tool, and the great capacity to define different ocular structures, the accuracy and reliability of the reports currently provided are limited.

Key words: Artificial intelligence. AI. ChatGPT. Large language models. Imaging. Ophthalmology.

Introducción

La inteligencia artificial (IA) hace referencia a la capacidad de una computadora de realizar tareas que eran exclusivas o propias del ser humano¹. Para tal fin, los sistemas de IA se basan en un sistema de aprendizaje automático mediante algoritmos derivados de una base de datos previa y de la experiencia con esta base de datos². Estos métodos de autoaprendizaje son los llamados "gran modelo de lenguaje" o "*large language model*" (LLM), que permiten responder a consultas sin un entrenamiento específico^{3,4}, con una alta capacidad de acierto. De hecho, ChatGPT-4, ejemplo de LLM, ha demostrado ser una herramienta efectiva para responder a preguntas clínicas oftalmológicas⁵, demostrando una gran capacidad de razonamiento en este campo⁶.

Actualmente, la medicina se enfrenta al reto de analizar gran cantidad de información para resolver problemas clínicos complejos. El desarrollo de la IA médica puede permitir el desarrollo de programas destinados a ayudar al profesional en la toma de decisiones clínicas¹.

En oftalmología, las imágenes son un componente fundamental de la atención oftalmológica, ya que permiten el examen detallado de las estructuras oculares⁷. Debido al creciente volumen de imágenes diagnósticas y a la mayor demanda asistencial, cada vez es más necesario agilizar el proceso de interpretación para ayudar a realizar diagnósticos oportunos y precisos. Por ello, la IA puede ser útil para la toma de decisiones de casos complejos y así agilizar muchos aspectos de nuestra práctica diaria^{2,4,8}.

El presente estudio pretende valorar las potenciales ventajas y limitaciones del uso de un LLM como ChatGPT-4 en la práctica diaria.

Material y método

Se recopilan tres imágenes oftalmológicas de casos consecutivos entre 31 de diciembre de 2023 y 5 de enero de 2024, que se presentaron en la plataforma virtual de ChatGPT-4 (<https://chat.openai.com/chat>). Se obtuvo el consentimiento informado de los pacientes antes de la toma de las imágenes. El estudio se adhiere a los principios de Helsinki⁹.

Se incluyeron imágenes de alteraciones del segmento anterior. Se le solicitó describir a ChatGPT-4 las características observables en las imágenes, sugerir posibles diagnósticos e identificar cualquier anomalía aparente.

Casos clínicos

Caso 1

Varón de 40 años acude por sensación de cuerpo extraño en el ojo izquierdo (OI), tras haber estado realizando obras. La biomicroscopía revela un cuerpo extraño corneal entre las 2 y 3 horas. ChatGPT-4 identifica las diferentes estructuras del segmento anterior y la presencia de una anomalía corneal. Sin embargo, no precisa el diagnóstico y solicita información adicional. Además, al preguntarle por la localización del cuerpo extraño, no lo localiza adecuadamente dibujando un círculo en un lugar completamente erróneo (Figura 1).

Caso 2

Paciente intervenido de queratoplastia penetrante en OI hace 3 años acude a urgencias por molestias oculares en el mismo. En la exploración oftalmológica se detecta la dehiscencia de un punto de sutura corneal a las 10 horas. Al introducir la imagen en el LLM, sorprendentemente identifica el trasplante de córnea, su falta de transparencia y la presencia de las suturas corneales. No obstante, no identifica el punto suelto (Figura 2).

Caso 3

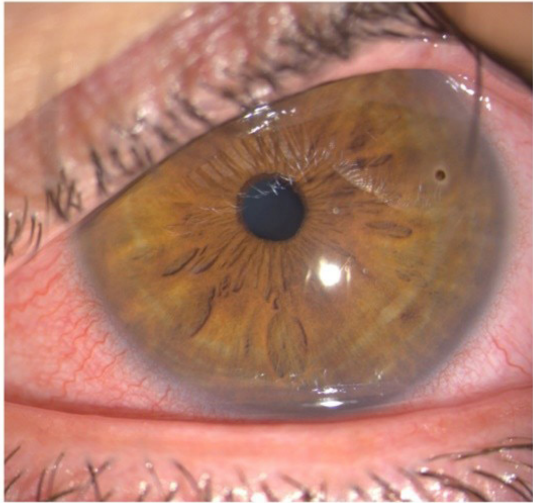
Paciente intervenido hace 13 años de cataratas con implante de lente intraocular (LIO) acude a urgencias por visión borrosa. En la exploración se aprecia una luxación de la LIO. ChatGPT-4 es capaz de detectar de nuevo las diferentes estructuras oculares, pero su respuesta es muy genérica y no objetiva el verdadero problema (Figura 3).

Discusión

La creciente demanda asistencial precisa de una necesidad de agilizar el proceso diagnóstico⁸. El abordaje de este desafío puede ser facilitado mediante la IA y los LLM como herramientas complementarias en la práctica clínica. Estas tecnologías pueden brindar apoyo en la interpretación de casos por no especialistas y así optimizar el flujo de trabajo de diagnóstico, especialmente en situaciones urgentes y/o emergentes¹.

La IA ha demostrado su utilidad en el cribado y diagnóstico de patologías en diversos ámbitos¹. El desarrollo continuo de las técnicas de imagen y la mayor relevancia de éstas en nuestra

You



ChatGPT

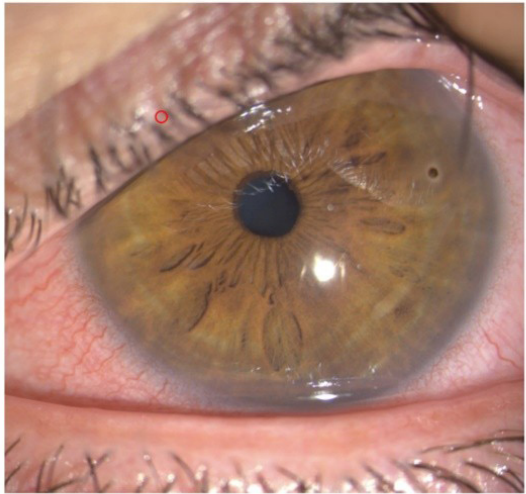
The image you've provided shows a close-up of an eye with several notable features:

- Iris:** The iris has a detailed and intricate pattern, which is normal. There's no visible neovascularization or atrophy.
- Pupil:** It is round and centrally placed within the iris.
- Corneal Reflection:** There are multiple reflections visible on the cornea, indicating a light source used during the examination, likely a slit lamp. This is common in ophthalmic photography.
- Conjunctiva:** The conjunctiva appears clear, without noticeable hyperemia or discharge.
- Cornea:** There is a small, round, darker area on the cornea that could be a corneal scar or pigmentation. Without additional clinical information, it's difficult to determine the exact nature of this finding.
- Tear Film:** There appears to be a healthy tear film on the surface of the eye, indicated by the smooth and even reflection of the light on the cornea.

You

Can you mark this "There is a small, round, darker area on the cornea that could be a corneal scar or pigmentation"?

ChatGPT



The area you inquired about has been marked with a red circle on the cornea. This mark is an approximation based on the visual inspection of the image provided and should not be used as a definitive diagnostic indicator. For precise assessment, a clinical evaluation is necessary. [~]

👍 🗨️ 🔄

Figura 1. Se presenta la imagen a ChatGPT-4 de forma aislada sin especificar detalles de la historia clínica para que sea descrita. En la imagen A define la presencia de una opacidad corneal correspondiente en el cuerpo extraño. Posteriormente se solicita que localice el cuerpo extraño (B) y lo marca con un círculo rojo de forma incorrecta (marcando el párpado y no el cuerpo extraño).

especialidad hace que su conocimiento e interpretación sean fundamentales para mejorar la precisión y el manejo de nuestros pacientes. La IA y los nuevos modelos de LLM, como ChatGPT-4, pueden ayudar al oftalmólogo en la detección y manejo de los pacientes complejos. Además, pueden ser muy útiles en telemedicina, evitando así el desplazamiento de los pacientes y disminuyendo la sobresaturación hospitalaria¹⁰. En situaciones donde los recursos sociosanitarios sean limitados, podría ayudar a priorizar la atención y asignar recursos de manera más eficiente.

En nuestro trabajo se seleccionaron imágenes oftalmológicas de pacientes consecutivos que acudieron a urgencias por diversas alteraciones del segmento anterior, entre las que se incluyeron un cuerpo extraño en la córnea, una dehiscencia de sutura y una subluxación de LIO. ChatGPT-4 fue capaz de describir con precisión las estructuras anatómicas observadas en las imágenes e identificó la presencia de cuerpo extraño, de trasplante corneal y de LIO.

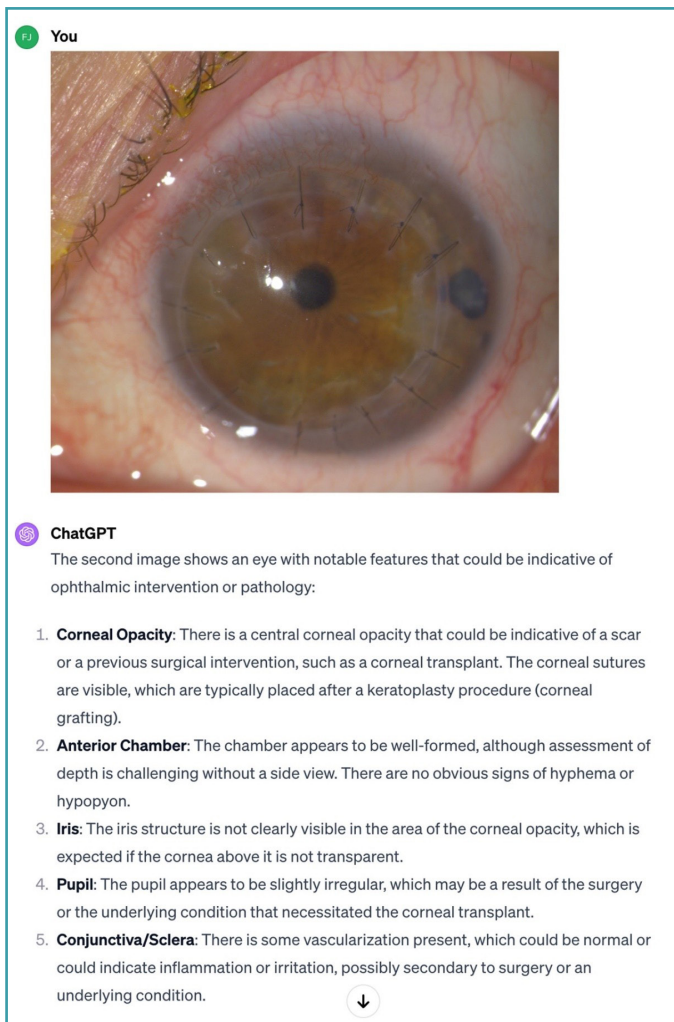


Figura 2. En la imagen ChatGPT-4 define la presencia de una alteración corneal compatible con una posible queratoplastia y detecta la presencia de puntos. También es capaz de detectar alteraciones en la pupila y define los detalles de las diferentes partes del segmento anterior de forma correcta.

Sin embargo, en ningún caso generó un diagnóstico de certeza. Esta deficiencia puso de manifiesto los retos inherentes a la interpretación de imágenes oftálmicas complejas por parte de la IA y subrayó la necesidad de integrar modelos lingüísticos con sistemas de IA médica especializada para mejorar la precisión diagnóstica.

Los resultados indican que, aunque ChatGPT-4 puede proporcionar información descriptiva sobre imágenes oftálmicas, su capacidad de diagnóstico actualmente es limitada.

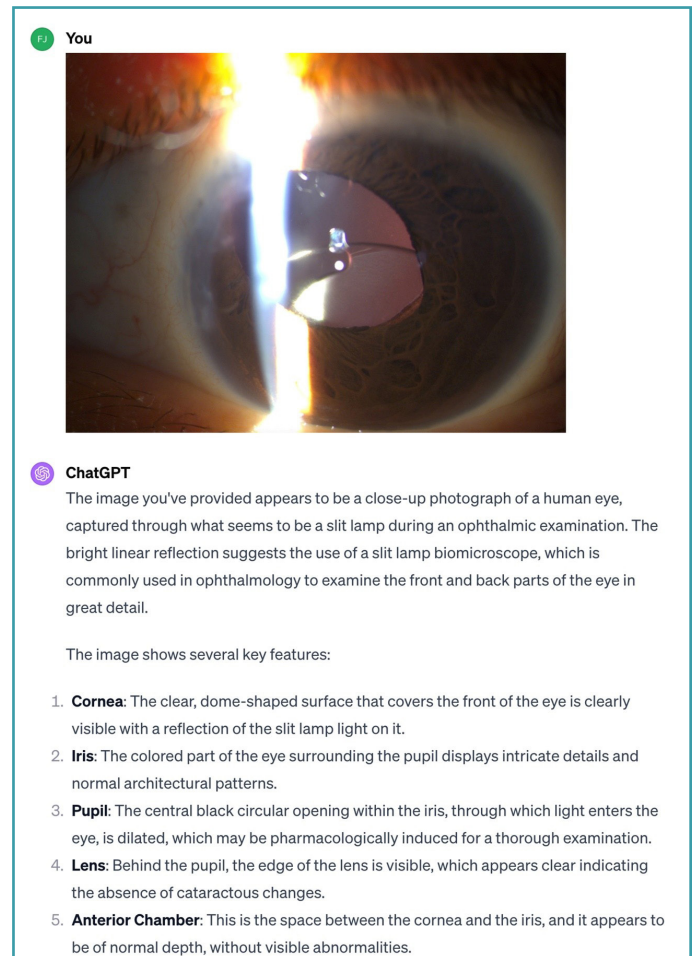


Figura 3. En la imagen ChatGPT-4 define las diferentes estructuras oculares. Es capaz de detectar que el ojo está dilatado y es capaz de detectar los bordes de la lente. No obstante, no detecta la subluxación de lente intraocular.

Su rendimiento como herramienta complementaria para fines educativos o evaluaciones preliminares puede ser suficiente, pero a día de hoy no sustituye la necesidad de una evaluación médica profesional.

Conclusiones

Si bien la utilización de LLM como ChatGPT-4 en la interpretación de imágenes oftálmicas representan un paso prometedor en el diagnóstico de patología ocular, todavía hay importantes desafíos que se deben tener en cuenta antes de que estos sistemas tengan un impacto significativo en la práctica clínica, dada la incapacidad de diagnosticar correctamente los casos presentados.

La investigación futura debe centrarse en la integración de modelos de LLM de reconocimiento de imágenes, así como en la realización de estudios clínicos para validar la precisión y utilidad de estos sistemas en entornos clínicos reales.

Financiación

Los autores declaran no haber recibido ninguna financiación para la realización de este trabajo.

Conflicto de intereses

Los autores no declaran conflicto de interés alguno.

Bibliografía

1. Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl.* septiembre de 2004;86(5):334-8.
2. Bali J, Bali O. Artificial intelligence in ophthalmology and healthcare: An updated review of the techniques in use. *Indian J Ophthalmol.* enero de 2021;69(1):8-13.
3. Valentín-Bravo FJ, Mateos-Álvarez E, Usategui-Martín R, Andrés-Iglesias C, Pastor-Jimeno JC, Pastor-Idoate S. Artificial Intelligence and new language models in Ophthalmology: Complications of the use of silicone oil in vitreoretinal surgery. *Arch Soc Esp Oftalmol.* 1 de mayo de 2023;98(5):298-303.
4. Valentín Bravo FJ, Mateos Álvarez E. Impact of artificial intelligence and language models in medicine. *Arch Soc Esp Oftalmol* (Engl Ed). junio de 2023;98(6):305-6.
5. Cohen SA, Brant A, Fisher AC, Pershing S, Do D, Pan C. Dr. Google vs. Dr. ChatGPT: Exploring the Use of Artificial Intelligence in Ophthalmology by Comparing the Accuracy, Safety, and Readability of Responses to Frequently Asked Patient Questions Regarding Cataracts and Cataract Surgery. *Seminars in Ophthalmology.* 2024;0(0):1-8.
6. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 9 de febrero de 2023;2(2):e0000198.
7. Bennett TJ, Barry CJ. Ophthalmic imaging today: an ophthalmic photographer's viewpoint - a review. *Clin Exp Ophthalmol.* enero de 2009;37(1):2-13.
8. Wawer Matos PA, Reimer RP, Rokohl AC, Caldeira L, Heindl LM, Große Hokamp N. Artificial Intelligence in Ophthalmology - Status Quo and Future Perspectives. *Semin Ophthalmol.* abril de 2023;38(3):226-37.
9. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA.* 27 de noviembre de 2013;310(20):2191-4.
10. Hallak JA, Scanzera A, Azar DT, Chan RVP. Artificial intelligence in ophthalmology during COVID-19 and in the post COVID-19 era. *Curr Opin Ophthalmol.* septiembre de 2020;31(5):447-53.