




## COMMENT OPEN



# OpenAI's Sora in ophthalmology: revolutionary generative AI in eye health

Ethan Waisberg <sup>1</sup>✉, Joshua Ong <sup>2</sup>, Mouayad Masalkhi <sup>3</sup> and Andrew G. Lee<sup>4,5,6,7,8,9,10</sup>

© The Author(s) 2024

Eye (2024) 38:2502–2503; <https://doi.org/10.1038/s41433-024-03098-x>

## INTRODUCTION

Sora is a state-of-the-art AI model developed by Open AI that has been engineered to generate realistic and imaginative scenes purely based on textual instructions [1]. This innovative application is a remarkable leap forward and highlights the advanced capabilities of modern AI in interpreting and visualizing complex narratives [1]. The technological foundation of Sora rests on Large Language Models (LLMs) and artificial video generation techniques. LLMs are advanced neural network architectures designed to understand, generate, and interpret human language in a highly sophisticated manner [2–5]. When combined with diffusion models for video generation, these AI systems can create detailed and dynamic visual content from text descriptions [1]. This involves processing the text to understand its meaning and context, and then translating them into a series of images that form a coherent video sequence.

The implications of such technology extend across various fields, including ophthalmology (Fig. 1). Sora and similar AI models could revolutionize patient education, surgical training, and the visualization of complex eye conditions and visual phenomena. By generating detailed visual simulations based on textual case descriptions or surgical procedures, practitioners can enhance their understanding and teaching of intricate ophthalmic concepts, thereby improving patient care and outcomes. Similarly, it could enable practitioners get an accurate first-person perspective into what their patients are seeing/experiencing, which could provide guidance for improved care and empathy. Despite its ground-breaking potential, at the time of writing, Sora is currently inaccessible for public use and is only available to select individuals.

## Surgical training

Ophthalmic surgery is a highly technical and learning new surgical techniques can be time intensive. Sora can be used to generate step-by-step surgical technique videos from text descriptions. While there are written step-by-step text explanations, with a few photos to describe ophthalmic surgical techniques, an AI-generated surgical video can provide an invaluable visual aid to ophthalmology trainees. A study by Reck-Burneo et al. [6] found that surgical trainees reported having

significantly higher levels of confidence following watching operative videos, rather than reading a peer-reviewed manuscript on a surgical technique.

## Patient Education

Strong ophthalmologist-patient communication is essential in the management of eye disease. Helping to educate and empower patients with conditions such as glaucoma, has been shown to improve both clinical outcomes [7] and treatment adherence [8]. A systematic review by Farwana et al. [9] showed that video-based media can be a useful ophthalmic patient education tool, with 71% of studies showing a significant improvement in comprehension following a video intervention. The current standard of providing additional written information, which is usually written in small text on leaflets, is also not particularly well-suited for individuals with visual impairments, non-native English speakers, or individuals with low literacy levels [9].

## Public awareness campaigns

To reduce preventable blindness and vision impairments, the general public must be aware of the importance of regular eye examinations. This was also a key recommendation made by the World Health Organization's World Report on Vision [10], to empower people and improve eye health literacy worldwide as early detection and timely management can help reduce preventable visual impairments. Ophthalmologists can use Sora to rapidly generate high-quality public awareness campaigns to educate the general public about various ophthalmic disorders and preventative measures that can be taken.

## Clinician education

A video generated by Sora can potentially illustrate symptoms and signs of rare/uncommon ophthalmic diseases, to help improve the ability of ophthalmologists-in-training to recognize them. Ophthalmology residents could then observe and diagnose these conditions in a supportive and controlled environment.

It is also important to also consider possible limitations of Sora. Like all LLMs, minor misunderstandings in written text can lead to the production of inaccurate videos [11–14]. Future research will

<sup>1</sup>Department of Ophthalmology, University of Cambridge, Cambridge, UK. <sup>2</sup>Department of Ophthalmology and Visual Sciences, University of Michigan Kellogg Eye Center, Ann Arbor, MI, USA. <sup>3</sup>School of Medicine, University College Dublin, Dublin, Ireland. <sup>4</sup>Department of Ophthalmology, Blanton Eye Institute, Houston Methodist Hospital, Houston, TX, USA. <sup>5</sup>The Houston Methodist Research Institute, Houston Methodist Hospital, Houston, TX, USA. <sup>6</sup>Departments of Ophthalmology, Neurology, and Neurosurgery, Weill Cornell Medicine, New York, NY, USA. <sup>7</sup>Department of Ophthalmology, University of Texas Medical Branch, Galveston, TX, USA. <sup>8</sup>University of Texas MD Anderson Cancer Center, Houston, TX, USA. <sup>9</sup>Texas A&M College of Medicine, Bryan, TX, USA. <sup>10</sup>Department of Ophthalmology, The University of Iowa Hospitals and Clinics, Iowa City, IA, USA.

✉email: ew690@cam.ac.uk

Received: 10 March 2024 Revised: 13 March 2024 Accepted: 18 April 2024

Published online: 30 April 2024



**Fig. 1** Screenshot from video generated by Sora from the prompt: Extreme close-up of a 24-year-old woman's eye blinking, standing in Marrakech during magic hour, cinematic film shot in 70 mm, depth of field, vivid colors, cinematic. (Open AI).

also need to be conducted on the anatomical accuracy of the ophthalmic AI-generated content.

Other future directions of Sora should include providing audio descriptions of videos to improve accessibility of the content for individuals with vision impairments. All things considered, Sora's artificial video generation has the potential to enhance ophthalmic surgical training, improve patient education, and the visualization of complex eye conditions and visual phenomena.

## REFERENCES

1. Sora: creating video from text. <https://openai.com/sora>.
2. Waisberg E, Ong J, Masalkhi M, Lee AG. Large language model (LLM)-driven chatbots for neuro-ophthalmic medical education. *Eye*. 2023. <https://doi.org/10.1038/s41433-023-02759-7>.
3. Masalkhi M, Ong J, Waisberg E, Lee AG. Google DeepMind's Gemini AI versus ChatGPT: a comparative analysis in ophthalmology. *Eye*. 2024. <https://doi.org/10.1038/s41433-024-02958-w>.
4. Alser M, Waisberg E. Concerns with the usage of ChatGPT in academia and medicine: a viewpoint. *Am J Med Open*. 2023;100036. <https://doi.org/10.1016/j.ajmo.2023.100036>.
5. Waisberg E, Ong E, Masalkhi M, Zaman N, Kamran SA, Sarker P, et al. ChatGPT and medical education: a new frontier for emerging physicians. *Can Med Ed J*. 2023. <https://doi.org/10.36834/cmej.77644>.
6. Reck-Burneo CA, Dingemans AJM, Lane VA, Cooper J, Levitt MA, Wood RJ. The impact of manuscript learning vs. video learning on a surgeon's confidence in performing a difficult procedure. *Front Surg*. 2018;5:67.
7. Friedman DS, Hahn SR, Gelb L, Tan J, Shah SN, Kim EE, et al. Doctor-patient communication, health-related beliefs, and adherence in glaucoma. *Ophthalmology*. 2008;115:1320–27.e3.
8. Sleath B, Blalock SJ, Carpenter DM, Sayner R, Muir KW, Slota C, et al. Ophthalmologist-patient communication, self-efficacy, and glaucoma medication adherence. *Ophthalmology*. 2015;122:748–54.
9. Farwana R, Sheriff A, Manzar H, Farwana M, Yusuf A, Sheriff I. Watch this space: a systematic review of the use of video-based media as a patient education tool in ophthalmology. *Eye*. 2020;34:1563–9.
10. World Health Organization. World report on vision. World Health Organization: Geneva; 2019.
11. Waisberg E, Ong J, Masalkhi M, Zaman N, Sarker P, Lee AG, et al. GPT-4 and medical image analysis: strengths, weaknesses and future directions. *J Med Artif Intell*. 2023;6:29.
12. Masalkhi M, Ong J, Waisberg E, Zaman N, Sarker P, Lee AG, et al. Large language models for post-operative guidance in refractive surgery. *AME Surg J*. 2024;4:2.
13. Paladugu PS, Ong J, Nelson N, Kamran SA, Waisberg E, Zaman N, et al. Generative adversarial networks in medicine: important considerations for this emerging innovation in artificial intelligence. *Ann Biomed Eng*. 2023. <https://doi.org/10.1007/s10439-023-03304-z>.
14. Waisberg E, Ong J, Masalkhi M, Kamran SA, Zaman N, Sarker P, et al. Automated ophthalmic imaging analysis in the era of Generative Pre-Trained Transformer-4. *Pan Am J Ophthalmol*. 2023;5:46.

## AUTHOR CONTRIBUTIONS

EW—Conceptualization, Writing. JO—Conceptualization, Writing. MM—Conceptualization, Writing. AGL—Review, Intellectual Support.

## COMPETING INTERESTS

AGL is a member of the *Eye* editorial board. The authors declare no conflicts of interest.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Ethan Waisberg.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024