




OPEN ACCESS

Generative artificial intelligence in ophthalmology: current innovations, future applications and challenges

Sadi Can Sonmez ¹, Mertcan Sevgi ^{2,3}, Fares Antaki ^{2,3,4}, Josef Huemer ^{3,5},
Pearse A Keane ^{2,3}

¹Department of Public Health, Ege University, Izmir, Turkey

²Institute of Ophthalmology, University College London, London, UK

³Moorfields Eye Hospital, NIHR Moorfields Biomedical Research Centre, London, UK

⁴The CHUM School of Artificial Intelligence in Healthcare, Montreal, Quebec, Canada

⁵Department of Ophthalmology and Optometry, Kepler University Hospital, Linz, Austria

Correspondence to

Pearse A Keane, Institute of Ophthalmology, University College London, London, UK; p.keane@ucl.ac.uk

SCS and MS are joint first authors.

Received 2 March 2024

Accepted 3 June 2024

ABSTRACT

The rapid advancements in generative artificial intelligence are set to significantly influence the medical sector, particularly ophthalmology. Generative adversarial networks and diffusion models enable the creation of synthetic images, aiding the development of deep learning models tailored for specific imaging tasks. Additionally, the advent of multimodal foundational models, capable of generating images, text and videos, presents a broad spectrum of applications within ophthalmology. These range from enhancing diagnostic accuracy to improving patient education and training healthcare professionals. Despite the promising potential, this area of technology is still in its infancy, and there are several challenges to be addressed, including data bias, safety concerns and the practical implementation of these technologies in clinical settings.

INTRODUCTION

Generative models have revolutionised the landscape of artificial intelligence (AI), offering groundbreaking capabilities in image generation that hold transformative potential.^{1,2} This innovation opened new possibilities in medical imaging, including ophthalmology, where generative adversarial networks (GANs) have been employed for tasks such as image synthesis, including the generation of ocular fundus photographs.³ Despite their success, GANs faced limitations like artefact generation and limited variety of image outputs (referred to as mode collapse), leading to the advent of diffusion models.⁴ Diffusion models have achieved high fidelity and diversity in image generation of colour fundus photographs (CFP) and optical coherence tomography (OCT) images that are indistinguishable from actual clinical photographs.⁵

The utility of synthetic images in healthcare, particularly for augmenting real training data, addresses critical challenges like sample bias and the under-representation of specific patient groups.⁶ In ophthalmology, leveraging synthetic images not only improves the training of deep learning (DL) models but also facilitates the development of innovative solutions for rare diseases such as inherited retinal disease (IRD).⁷ Concurrently, the advent of vision-language models (VLMs) has propelled significant advancements in natural language processing.⁸ These models, trained on extensive datasets of images paired with textual captions, show promising capabilities in generating visual content from

text descriptions. This emerging technology along with future text-to-video enhancements holds potential for ophthalmology, offering novel ways to enhance surgical education and patient care.^{9–11}

In this work, we explore the advancements of generative AI, focusing on the pivotal role of GANs, diffusion models and the burgeoning area of text-to-image models. We also discuss challenges of generative AI including bias, data privacy and patient safety.

HISTORY AND OVERVIEW OF GENERATIVE MODELS

A foundational model is an AI system that is developed through extensive self-supervised training on large amounts of unlabelled data that can be later adapted for a variety of downstream tasks.² Its capability to be trained on multimodal data—such as text and images—enables it to perform tasks like generating text from images and creating images from textual descriptions. Broadly, there are four categories of foundation models, as summarised in [table 1](#): large language models (LLMs), large vision models (LVMs), VLMs and large multimodal models (LMMs).

LLMs are text-focused models, trained on extensive textual corpora amassing trillions of words, enabling them to generate human-like text.^{2 12–16} LVMs specialise in image processing and are trained on vast image datasets for tasks like image recognition and generation. VLMs can generate unique images based on textual prompts after getting trained on datasets containing image-text pairs.¹⁷ These datasets comprise millions of paired images corresponding textual descriptions to enable models like DALL-E and Stable Diffusion to understand and create complex visual content in response to textual prompts. Another architecture, LMMs, represents the most advanced category. They are designed to process and generate content across multiple formats, including text, images, videos and music.¹⁸

The literature on foundation models in ophthalmology is expanding. A significant focus has been on LLM applications for medical question answering and clinical reasoning, as well as agentic AI that assists clinicians with workflow tasks, such as summarising notes and medical documentation.^{19–21} While VLMs using Gemini have been tested for interpreting OCT pathology, their reliability remains unproven.²² In contrast, LVMs like



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Sonmez SC, Sevgi M, Antaki F, *et al.* *Br J Ophthalmol* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bjo-2024-325458

Table 1 A concise summary of foundation models and their features

Model type	Focus	Examples	Data used for training
Large language models (LLMs)	Text	PaLM, LLaMA, GPT-3.5	Text corpora (books, articles, etc)
Large vision models (LVMs)	Images	RETFound ²³ , SAM ⁷⁰ , ViT ⁷¹ , Vision ⁷²	Image datasets
Vision-language models (VLMs)	Text and images	CLIP ⁴¹ , Vision, DALL-E ⁷³ , Stable Diffusion ⁷⁴	Image-text pairs
Large multimodal models (LMMs)	Multiple formats	GPT-4, Gemini	Various (text, images, audio, etc)

CLIP, Contrastive Language-Image Pre-Training; GPT, Generative Pre-trained Transformer; LLaMA, Large Language Model Meta AI; PaLM, Pathways Language Model; SAM, Segment Anything Model; ViT, Vision Transformer;

RETFound represent a highly promising development. This pioneering model for CFP and OCT has demonstrated superior performance compared with traditional deep learning models. This achievement is largely due to the use of self-supervised learning, which allows the model to learn effectively from unlabelled images, thus bypassing the need for extensively annotated datasets. Additionally, the incorporation of Vision Transformers into the model architecture leverages their ability to handle varied image resolutions and complexities, further enhancing the model's performance in analysing medical images.²³ The subsequent sections will concentrate on the image-generation capabilities of AI in ophthalmology, focusing on image-to-image and text-to-image models.

IMAGE-TO-IMAGE MODELS

A pivotal moment in AI-driven image generation was the introduction of GANs in 2014 by Goodfellow *et al.*²⁴ GANs consist of two neural networks: the generator and the discriminator. The generator creates images, while the discriminator evaluates them. The generator's goal is to produce images so realistic that the discriminator cannot differentiate them from real images. This adversarial process results in the generator producing increasingly refined images. While GANs marked a significant advancement, they had limitations, including the potential for generating artefacts and the challenge of mode collapse, where the generator produces limited varieties of outputs. Diffusion models effectively mitigate common GAN issues such as artefact generation and mode collapse. For artefact generation, diffusion models enhance image quality by gradually removing noise added to an image in a controlled way. This method involves a series of small, precise adjustments, resulting in superior performance metrics in tests of image diversity and realism.²⁵ Regarding mode collapse, the iterative noise reduction process helps maintain a wide variety of image types, preventing the model from limiting its output variety—a frequent limitation with GANs. The training method of diffusion models is specifically designed to enhance both the detail and diversity of images, thereby effectively overcoming these major drawbacks observed in GAN-based methods.^{4 26}

We provide a comparative summary of GANs and diffusion models in [table 2](#).

Use cases

GANs and diffusion models are becoming increasingly pivotal in ophthalmology, as they offer a broad range of applications and capabilities. GANs excel in numerous areas: they can predict post-intervention outcomes such as reduced macular oedema and the results of oculoplastic surgeries, remove artefacts from CFPs, generate angiography images from CFPs through domain transfer, create synthetic images of rare diseases to augment data and perform segmentation tasks.^{3 27–33} These diverse applications underscore their versatility and value in advancing medical imaging and treatment prediction.

Diffusion models in ophthalmology are particularly noted for their ability to produce colour fundus photographs and OCT images that are virtually indistinguishable from actual clinical images.⁵ This capability not only enhances the realism of synthetic images but also supports more robust training of DL models.

The generation of synthetic images addresses several critical challenges in ophthalmology. Real training datasets often suffer from limitations such as sample bias, including under-representation of certain patient demographics and rare diseases.^{34 35} Synthetic images can augment these datasets, thereby reducing the number of real images needed and enhancing model performance across more diverse scenarios.^{36 37} This approach helps overcome the generalisability issues in AI, where models trained on limited datasets may not perform well on unrepresented image types. This lack of generalisability is a significant barrier to the deployment of AI in real-world settings.³⁸

A particularly pressing issue is the under-representation of populations from low and middle-income countries in publicly available datasets, which could exacerbate health inequalities.³⁹ Diffusion models show promise in mitigating this through the generation of CFPs from limited datasets, thus improving equal representation in training data.⁴⁰

Moreover, synthetic images have shown potential in training models for rare conditions, such as IRD. For instance, the

Table 2 A comparative summary of GANs vs diffusion models

Feature	GANs	Diffusion models
Aim	Generate data mimicking training data via a generator-discriminator competition	Generate data by reversing a process that adds noise to data
Architecture	Two networks: generator (creates data) and discriminator (evaluates data)	Single network learns to remove noise over many steps
Mode collapse	Prone, leading to less diverse outputs	Less prone, ensures diverse sample generation
Data efficiency	Requires large datasets to train effectively	More efficient, works well with smaller datasets
Input noise	Begins with a noise vector, transformed into data by the generator	Begins with noise, gradually denoised to form data
Applications	Image generation, style transfer, super-resolution	Image/audio generation, text-to-image synthesis
Other features	Unstable training, single-step generation	Stable training, multistep generation process

GANs, generative adversarial networks.

Eye2Gene DL model uses images to predict the genetic profiles of diseases but struggles with under-represented genetic subtypes. To address this, GANs like SynthEye are being developed to generate high-quality synthetic images that diversify the training datasets of diagnostic models, enhancing their accuracy and reliability.^{6,7}

TEXT-TO-IMAGE MODELS

Simultaneously, the evolution of VLMs has marked a significant leap in the field of natural language processing and image generation.⁴¹ These models, incorporating both the intricate understanding of natural language and visual elements, are trained on vast datasets of images paired with textual captions.⁴² By integrating the capabilities of diffusion models for high-quality image generation with nuanced language understanding, VLMs such as DALL-E, Stable Diffusion and Midjourney have the ability to produce detailed and contextually relevant images from textual descriptions. Available for public use, these models have garnered considerable attention for their innovative approach to bridging the gap between textual input and visual output, showcasing the potential of combining language and vision in AI.⁴³

Use cases

Currently, text-to-image models such as DALL-E, Midjourney and StabilityAI have generalist knowledge and are able to create meaningful images from text prompts. However, for medical use, which is a niche and specific field, text-to-image generation of medical imaging is not yet fully developed. Using DALL-E 2 to generate X-ray images of anatomical areas such as the skull, hands, chest and ankle to name a few, shows realistic and

well-proportioned images; however, detailed structure of bones is not correct.⁴⁴ When tasked with more complex modalities such as CT and MRI, the model only shows images with features that reveal the modality; however, overall, the image is non-sensical. These widely available large text-to-image models seem to have some medical domain knowledge in the training data but not sufficient to create realistic medical images.⁴⁵ We likely need to see models trained on a greater number of medical images for accurate text-to-image generation.

However, there are some current promising features of text-to-image models that could benefit the field of ophthalmology. As most ophthalmic disease is associated with a change in vision and visual defects, it will be useful to create images to demonstrate visual defects to accompany textual descriptions. This could be useful for both patients and for healthcare professionals. As an example, ophthalmologists could generate images for clinicians to aid in understanding of neuro-ophthalmological conditions such as oscillopsia in multiple sclerosis, Charles Bonnet syndrome, Pulfrich phenomena and rare conditions with no examination findings such as visual snow syndrome.^{10,46} Simulations of eye disorders do exist;⁴⁷ however, ophthalmologists could create personalised prognostic markers in the form of images for patients to understand how their vision may change if their condition worsens or what they could expect if treatment is started. Another area is improving informed consent for operations. For procedures that will change a person's appearance, such as ptosis surgery or orbital decompression surgery, ophthalmologists can generate personalised images for the patient to show how they could look following their surgery.⁴⁸ We provide examples of generating images with ChatGPT4 and built-in DALL-E in figure 1.

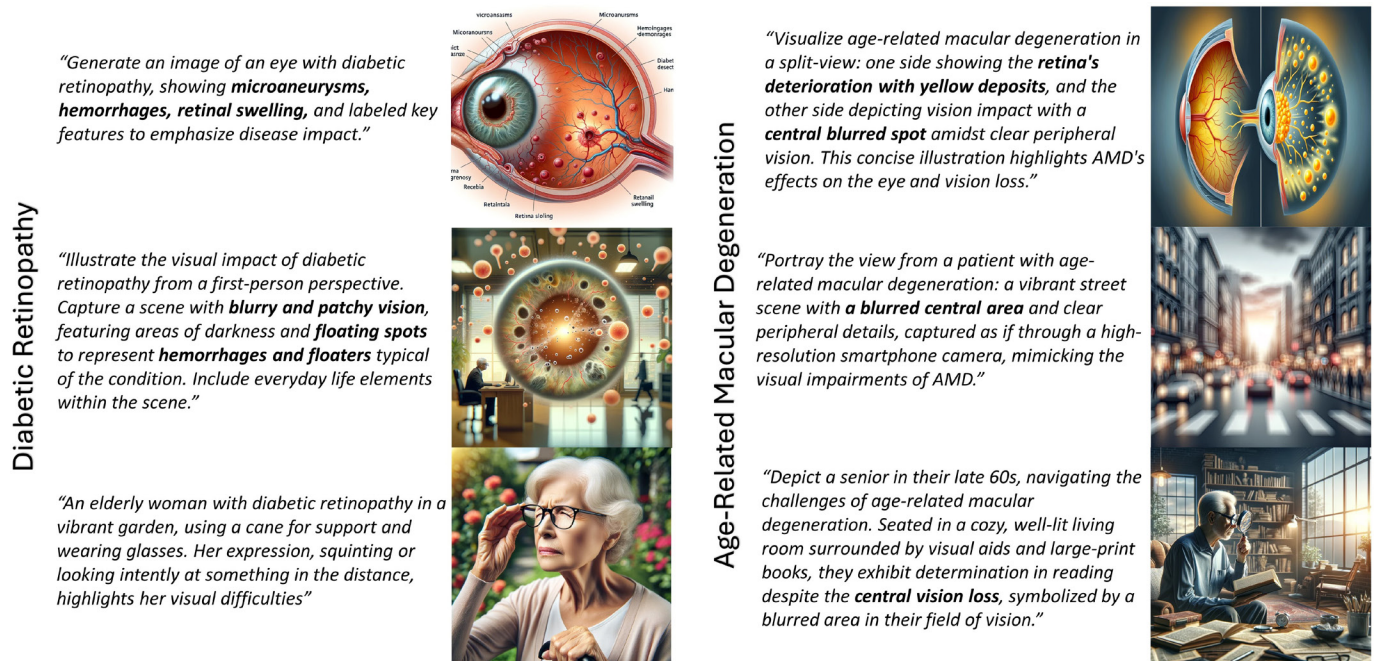


Figure 1 Sample prompts and images related to two ophthalmic conditions generated using DALL-E within ChatGPT-4. We consecutively prompted ChatGPT-4 to generate a written prompt to later use as a prompt for an image, separately for each condition. Our prompts read as: 1. "Could you give me a prompt for an image to describe Diabetic Retinopathy/Age-Related Macular Degeneration under 50 words and generate an image for it?" 2. "Could you give me a prompt under 50 words for an image to describe how a person with Diabetic Retinopathy/Age-Related Macular Degeneration sees and generate an image for it?" 3. "Could you give me a prompt under 50 words for an image of a person with Diabetic Retinopathy/Age-Related Macular Degeneration and generate an image for it?" Note that certain pathognomonic features (in bold) have been successfully included by ChatGPT-4 despite its non-medical training, highlighting potential for public and professional education. The notable limitations in this figure include the inability to generate accurate anatomical details of the eye and incorrectly spelled words.

Further developments include text-to-video generation.^{49 50} In addition to providing a more immersive experience to the examples above, video generation can be a useful learning tool for surgery simulation and learning. Within ophthalmology, practising through surgical simulation is well recognised.^{9 51} AI is also being developed to provide real-time analysis during operations.¹¹ Generative AI could enable generation of operations and procedures, which could be a tool for learning and simulation.

ADDRESSING THE CHALLENGES AND FUTURE DIRECTIONS

While generative AI holds immense potential in healthcare, certain challenges demand attention.^{19 52–54} The emerging issues can be grouped under data bias, safety and implementation. These should not be overlooked by any stakeholder including academics and private entities which have increased in prevalence in AI research.^{55 56} As a whole, open communication, transparency and regulation are needed to safeguard public health.

Bias, inclusivity and copyright issues in model training

There is the perception that the massive datasets that generative models are trained on are diverse; however, this is difficult to audit.^{2 12} Neglected sampling errors and bias in these datasets can compromise the ground truth and lead to poor generalisability.^{19 52–54} This can affect political, racial, socioeconomic domains and can deepen the inequalities in healthcare in terms of accessibility and quality of care. Minority groups, immigrants, rural populations and individuals with lower socioeconomic status are prone to such biases as they are less represented within online databases and electronic health records.^{57 58} Efforts to diversify the datasets with different real-world resources present another issue centred around copyright. There have been legal claims regarding the use of publicly accessible original work to train generative algorithms without authorisation.^{53 59} This necessitates extended transparency and defined standards in model training, along with heightened sensitivity to sources of bias. There is also an apparent need for a legal foundation governing dataset curation for pre-specified training purposes.^{19 53 54}

Safety concerns regarding output integrity and data privacy

One concern regarding generative AI is the generation of incorrect or misleading results, also referred to as ‘hallucinations.’ Hallucinations can have negative consequences on patient safety and awareness through health-related misinformation.^{19 52–54} More serious complications can arise if these products are used within clinical decision support and reporting or to train other diagnostic algorithms and even physicians.^{19 52–54} Given the current legal framework still being immature, it is also unclear who will be considered liable for such results.^{19 52} This can expectedly lead to reluctance in end-users, physicians and patients, who are unable to discern hallucinations and even at times, AI-generated from real content.^{19 53} To avoid hallucinations, AI models nevertheless require to be trained with comprehensive health data. However, this poses a risk of security breach if the data is shared with multiple centres and especially with third-party developers.⁶⁰ In case of health data violations in which anonymity is lost or the data is used for purposes other than specified, negative outcomes such as mental stress, erosion of trust and even group-based harm and discrimination can arise.⁶¹ These issues therefore need to be discussed under an ethical umbrella to cover consent for data processing and ensuring autonomy of patients as well.^{61 62}

The development of AI models has been rapid; however, the regulatory discussion and response have been slow. It has even

been proposed to cease all development of the technology until more stringent regulation is in place.⁶³ Organisations like the European Union are moving forward with specific frameworks like the ‘AI Act’ where a categorisation of the danger level of AI models and corresponding set of guidelines have been specified.⁶⁴ Other existing legal frameworks for digitalisation in healthcare will likely require AI-specific and particularly generative AI-specific clauses.^{60 64 65} Physician engagement in the layout of health-related AI regulations through specific taskforces can be helpful to uphold patient safety.⁶⁵

Challenges to implementation into Workflow

It is crucial to identify which specific tasks in healthcare can be improved with the integration of generative AI, early on.⁶⁶ These tools need specific positioning within their use cases (ie, care delivery or training) with clear indications and contraindications to delineate appropriate use and accountability.^{61 62 64 65 67} Without such specific guidelines, end-users including physicians can be confused over responsibilities and become discouraged eventually.^{61 62 67} They also need reliable channels for troubleshooting and reporting of adverse operations to assist end-users in times of need.^{67 68} User-friendliness is essential, as a complex system can counterintuitively burden staff already facing demanding workloads.⁶⁸ Physician training on these models and corresponding guidelines should be considered in the medical curricula and as a part of continuous professional training.⁶⁵ Since the capabilities of AI systems are predicted to increase exponentially, definition of such roles and boundaries within the delivery of care should not be delayed.⁶⁹ Initiatives which lack this foresight about workflow as their core standpoint may get impeded when it comes to real-world practice.

CONCLUSION

The fast-emerging field of generative AI has immense potential for progress in ophthalmology including revolutionary advancements in diagnosis, accurate prognostication and professional training. However, there are certain challenges regarding data bias, safety and implementation. Addressing these through open conversations between academia, government and industry is vital for transparent and effective regulation to mitigate risks. Ultimately, as the digital and real world intersect further, we might look back on the generative models as the beginning of a new and brighter chapter in healthcare and ophthalmology.

X Fares Antaki @FaresAntaki and Pearse A Keane @pearsekeane

Contributors Conception and design of the study (SCS, MS, FA, JH and PAK); data collection (SCS and MS); writing of the manuscript (SCS, MS, FA and JH); preparation of figures (SCS and JH); supervision (PAK); review and discussion of the results (all authors); edition and revision of the manuscript (all authors). We have used AI solely in the generation of the Figure 1. We used ChatGPT-4 (OpenAI) to get descriptive written prompts and later to generate images of certain eye conditions based on those prompts. The process is described in detail in the description of the figure summary, submitted alongside the figure.

Funding This study was funded by UK Research and Innovation.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Sadi Can Sonmez <http://orcid.org/0000-0002-8891-6691>
Mertcan Sevgi <http://orcid.org/0009-0003-8426-6534>
Fares Antaki <http://orcid.org/0000-0001-6679-7276>
Josef Huemer <http://orcid.org/0000-0001-5238-8295>
Pearse A Keane <http://orcid.org/0000-0002-9239-745X>

REFERENCES

- Acosta JN, Falcone GJ, Rajpurkar P, et al. Multimodal biomedical AI. *Nat Med* 2022;28:1773–84.
- Bommasani R, Hudson DA, Adeli E, et al. On the opportunities and risks of foundation models. *arXiv [CsL.G]* 2021. Available: <http://arxiv.org/abs/2108.07258>
- Wang Z, Lim G, Ng WY, et al. Generative adversarial networks in ophthalmology: what are these and how can they be used? *Curr Opin Ophthalmol* 2021;32:459–67.
- Yang L, Zhang Z, Song Y, et al. Diffusion models: a comprehensive survey of methods and applications. *arXiv [CsL.G]* 2022. Available: <http://arxiv.org/abs/2209.00796v11>
- Ilanchezian I, Boreiko V, Kühlewein L, et al. Generating realistic Counterfactuals for retinal fundus and OCT images using diffusion models. *arXiv [CsCV]* 2023. Available: <http://arxiv.org/abs/2311.11629>
- Veturi YA, Woof W, Lazebnik T, et al. Syntheye: investigating the impact of synthetic data on artificial intelligence-assisted gene diagnosis of inherited retinal disease. *Ophthalmol Sci* 2023;3:100258.
- Pontikos N, Woof W, Veturi A, et al. Eye2Gene: prediction of causal inherited retinal disease gene from multimodal imaging using deep-learning. *In Review [Preprint]* 2022.
- Huggingface. A dive into vision-language models. 2024. Available: https://huggingface.co/blog/vision_language_pretraining [Accessed 12 Feb 2024].
- Radia M, Arunakiranthan M, Sibley D. A guide to eyes: ophthalmic simulators. *Bulletin* 2018;100:169–71.
- Waisberg E, Ong J, Masalkhi M, et al. Text-to-image artificial intelligence to aid clinicians in perceiving unique neuro-ophthalmic visual phenomena. *Ir J Med Sci* 2023;192:3139–42.
- Josef Lindegger D, Alnafise N, Thomsen ASS, et al. The future of virtual reality in cataract surgical training. *Expert Rev Ophthalmol* 2020;15:193–6.
- Brown TB, Mann B, Ryder N, et al. Language models are few-shot learners. *arXiv [CsCL]* 2020. Available: <http://arxiv.org/abs/2005.14165>
- Openai platform. Available: <https://platform.openai.com/docs/models/gpt-3-5-turbo> [Accessed 12 Feb 2024].
- Gemini - Google Deepmind. Available: <https://deepmind.google/technologies/gemini/> [Accessed 12 Feb 2024].
- Barea A. La llama [Bibliotex]. 2021. Available: <https://llama.meta.com/>
- PaLM 2. Google AI, Available: <https://ai.google/discover/palm2/> [Accessed 12 Feb 2024].
- Zhang P, Li X, Hu X, et al. VinVL: revisiting visual representations in vision-language models. *arXiv [CsCV]* 2021. Available: <http://arxiv.org/abs/2101.00529>
- Yang Z, Li L, Lin K, et al. The dawn of Lmms: preliminary explorations with GPT-4V(Ision). 2023. Available: <http://arxiv.org/abs/2309.17421> [Accessed 25 Feb 2024].
- Betzler BK, Chen H, Cheng C-Y, et al. Large language models and their impact in ophthalmology. *Lancet Digit Health* 2023;5:e917–24.
- Mehandru N, Miao BY, Almaraz ER, et al. Evaluating large language models as agents in the clinic. *NPJ Digit Med* 2024;7:84.
- Antaki F, Touma S, Milad D, et al. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. *Ophthalmol Sci* 2023;3:100324.
- Antaki F, Chopra R, Keane PA. Vision-language models for feature detection of macular diseases on optical coherence tomography. *JAMA Ophthalmol* 2024.;e241165.
- Zhou Y, Chia MA, Wagner SK, et al. A foundation model for generalizable disease detection from retinal images. *Nature* 2023;622:156–63.
- Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Adv Neural Inf Process Syst* 2014;27.
- Dhariwal P, Nichol A. Diffusion models beat GANS on image synthesis. *Adv Neural Inf Process Syst* 2021.
- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *arXiv [CsL.G]* 2020. Available: <http://arxiv.org/abs/2006.11239>
- You A, Kim JK, Ryu IH, et al. Application of Generative adversarial networks (GAN) for Ophthalmology image domains: a survey. *Eye and Vis* 2022;9:6.
- Yoo TK, Choi JY, Kim HK. A Generative adversarial network approach to predicting postoperative appearance after orbital decompression surgery for thyroid eye disease. *Comput Biol Med* 2020;118:103628.
- Liu Y, Yang J, Zhou Y, et al. Prediction of OCT images of short-term response to anti-VEGF treatment for neovascular age-related macular degeneration using generative adversarial network. *Br J Ophthalmol* 2020;104:1735–40.
- Yoo TK, Choi JY, Kim HK. CycleGAN-based deep learning technique for artifact reduction in fundus photography. *Graefes Arch Clin Exp Ophthalmol* 2020;258:1631–7.
- Tavakkoli A, Kamran SA, Hossain KF, et al. A novel deep learning conditional Generative adversarial network for producing angiography images from retinal fundus photographs. *Sci Rep* 2020;10:21580.
- Yoo TK, Choi JY, Kim HK, et al. Adopting low-shot deep learning for the detection of conjunctival melanoma using ocular surface images. *Comput Methods Programs Biomed* 2021;205:106086.
- Yildiz E, Arslan AT, Yildiz Tas A, et al. Generative adversarial network based automatic segmentation of corneal subbasal nerves on in vivo confocal microscopy images. *Transl Vis Sci Technol* 2021;10:33.
- Norori N, Hu Q, Aellen FM, et al. Addressing bias in big data and AI for health care: a call for open science. *Patterns (N Y)* 2021;2:100347.
- Chen RJ, Lu MY, Chen TY, et al. Synthetic data in machine learning for medicine and healthcare. *Nat Biomed Eng* 2021;5:493–7.
- Anderson JW, Ziolkowski M, Kennedy K, et al. Synthetic image data for deep learning. *arXiv [CsCV]* 2022. Available: <http://arxiv.org/abs/2212.06232>
- Nderitu P, do Rio JMN, Webster L, et al. Conditional diffusion models and retinal image synthesis in diabetic retinopathy. *Invest Ophthalmol Vis Sci* 2023;64:2389.
- Li Z, Wang L, Wu X, et al. Artificial intelligence in ophthalmology: the path to the real-world clinic. *Cell Rep Med* 2023;4:101095.
- Khan SM, Liu X, Nath S, et al. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. *Lancet Digit Health* 2021;3:e51–66.
- Kim HK, Ryu IH, Choi JY, et al. A feasibility study on the adoption of a Generative Denoising diffusion model for the synthesis of fundus photographs using a small Dataset. *Discov Appl Sci* 2024;6:188.
- Radford A, Kim JW, Hallacy C, et al. Learning transferable visual models from natural language supervision. *arXiv [CsCV]* 2021. Available: <http://arxiv.org/abs/2103.00020>
- Ramesh A, Pavlov M, Goh G, et al. Zero-shot text-to-image generation. *arXiv [CsCV]* 2021. Available: <http://arxiv.org/abs/2102.12092>
- Roose K. A coming-out party for Generative A.I., Silicon valley's new craze [The New York Times]. 2022. Available: <https://www.nytimes.com/2022/10/21/technology/generative-ai.html> [Accessed 21 Dec 2023].
- Adams LC, Busch F, Truhn D, et al. What does DALL-E 2 know about radiology? *J Med Internet Res* 2023;25:e43110.
- Kather JN, Ghaffari Laleh N, Foersch S, et al. Medical domain knowledge in domain-agnostic generative AI. *NPJ Digit Med* 2022;5:90.
- Balas M, Micieli JA. Visual snow syndrome: use of text-to-image artificial intelligence models to improve the patient perspective. *Can J Neurol Sci* 2023;50:946–7.
- RichmondEye. Ophthalmologist Richmond VA. 2013. Available: <https://www.richmondeye.com/simulations-of-eye-disorders> [Accessed 24 Dec 2023].
- Balas M, Micieli JA, Wulc A, et al. Text-to-image artificial intelligence models for preoperative counselling in oculoplastics. *Can J Ophthalmol* 2024;59:e75–6.
- StabilityAI. Stable Video. 2024. Available: <https://stability.ai/stable-video> [Accessed 06 Jan 2024].
- Video generation models as world simulators. Available: <https://openai.com/research/video-generation-models-as-world-simulators> [Accessed 17 Feb 2024].
- Antaki F, Doucet C, Milad D, et al. Democratizing vitreoretinal surgery training with a portable and affordable virtual reality simulator in the metaverse. *Trans Vis Sci Tech* 2024;13:5.
- Jin K, Yuan L, Wu H, et al. Exploring large language model for next generation of artificial intelligence in ophthalmology. *Front Med* 2023;10:1291404.
- Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. *Future Internet* 2023;15:286.
- Tan TF, Thirunavukarasu AJ, Campbell JP, et al. Generative artificial intelligence through ChatGPT and other large language models in ophthalmology: clinical applications and challenges. *Ophthalmol Sci* 2023;3:100394.
- Ahmed N, Wahed M, Thompson NC. The growing influence of industry in AI research. *Science* 2023;379:884–6.
- Färber M, Tampakis L. Analyzing the impact of companies on AI research based on publications. *Scientometrics* 2024;129:31–63.
- Boyd AD, Gonzalez-Guarda R, Lawrence K, et al. Potential bias and lack of generalizability in electronic health record data: reflections on health equity from the National Institutes of health pragmatic trials Collaboratory. *J Am Med Inform Assoc* 2023;30:1561–6.
- Yao R, Zhang W, Evans R, et al. Inequities in health care services caused by the adoption of Digital health technologies. *J Med Internet Res* 2022;24:e34144.
- Samuelson P. Generative AI meets copyright. *Science* 2023;381:158–61.
- Tom E, Keane PA, Blazes M, et al. Protecting data privacy in the age of AI-enabled ophthalmology. *Transl Vis Sci Technol* 2020;9:36.
- Abdullah YI, Schuman JS, Shabsigh R, et al. Ethics of artificial intelligence in medicine and ophthalmology. *Asia Pac J Ophthalmol (Phila)* 2021;10:289–98.

- 62 Morley J, Machado CCV, Burr C, *et al.* The ethics of AI in health care: a mapping review. *Soc Sci Med* 2020;260:113172.
- 63 Henshall W. AI experts call for policy action to avoid extreme risks [Time]. 2023. Available: <https://time.com/6328111/open-letter-ai-policy-action-avoid-extreme-risks/> [Accessed 05 Feb 2024].
- 64 Quaranta M, Amantea IA, Grosso M. Obligation for AI systems in healthcare: prepare for trouble and make it double? *Rev Socionetwork Strat* 2023;17:275–95.
- 65 He J, Baxter SL, Xu J, *et al.* The practical implementation of artificial intelligence Technologies in medicine. *Nat Med* 2019;25:30–6.
- 66 Dow ER, Keenan TDL, Lad EM, *et al.* From data to deployment: the collaborative community on ophthalmic imaging roadmap for artificial intelligence in age-related macular degeneration. *Ophthalmology* 2022;129:e43–59.
- 67 Smith H, Downer J, Ives J. Clinicians and AI use: where is the professional guidance. *J Med Ethics* 2023.
- 68 Gandhi TK, Classen D, Sinsky CA, *et al.* How can artificial intelligence decrease cognitive and work burden for front line practitioners. *JAMIA Open* 2023;6:ooad079.
- 69 Bengio Y, Hinton G, Yao A, *et al.* Managing AI risks in an era of rapid progress. *arXiv [CsCY]* 2023. Available: <http://arxiv.org/abs/2310.17688>
- 70 Kirillov A, Mintun E, Ravi N, *et al.* Segment anything. *arXiv [CsCV]* 2023. Available: <http://arxiv.org/abs/2304.02643>
- 71 Dosovitskiy A, Beyer L, Kolesnikov A, *et al.* An image is worth 16X16 words: transformers for image recognition at scale. *arXiv [CsCV]* 2020. Available: <http://arxiv.org/abs/2010.11929>
- 72 Openai platform. Available: <https://platform.openai.com/docs/guides/vision> [Accessed 12 Feb 2024].
- 73 OpenAI. Dall-E 3 [OpenAI]. 2023. Available: <https://openai.com/dall-e-3> [Accessed 21 Dec 2023].
- 74 StabilityAI. Stable diffusion. 2023. Available: <https://stability.ai/stable-diffusion> [Accessed 21 Dec 2023].