INVITED REVIEW

Towards a Paradigm Shift with Generative Artificial Intelligence in Ophthalmology: Opportunities, Challenges, and Future Directions

Hakan Özdemir[1], *Furkan Kırık*[1]

ABSTRACT

Generative artificial intelligence (GenAI) is driving a major transformation in ophthalmology by employing models such as generative adversarial networks, diffusion models and large language models (LLMs) to create novel yet realistic synthetic data. These systems, including emerging architectures capable of modality conversion (such as text-to-image generation), provide a foundation for diverse applications. Applications based on images encompass generation of synthetic ophthalmic images to augment data sets for rare conditions, enhancement of image quality to improve clinical assessment, conversion between imaging modalities to reduce equipment costs, and simulation of disease progression or prediction of post-treatment appearance to support surgical planning and patient counselling. Concurrently, LLMs significantly influence clinical practice by supporting diagnostic workflows and differential diagnoses within clinical decision support systems, assisting with patient triage, automating clinical documentation and reporting, and enhancing patient communication and education through personalized, multilingual content. GenAI also shows promise in medical education and research by facilitating the creation of diverse teaching materials and streamlining literature review, data analysis, and manuscript preparation. However, successful deployment of GenAI requires careful attention to ethical, safety, and regulatory challenges, including model reliability, data bias, patient privacy, and establishing clear legal frameworks and human oversight. Future developments are likely to include truly multimodal systems that integrate the use of synthetic data sources, personlized medicine approaches, and expanded use in tele-ophthalmology, together with the widespread adoption of purpose-specific custom-GPT models and exploration of agentic AI's potential in ophthalmic practice, underscoring the crucial role of AI literate ophthalmologists in these emerging fields.

Key Words: diffusion models, generative adversarial networks, generative artificial intelligence, large language models, synthetic data

1- INTRODUCTION

Recent advances in artificial intelligence (AI) have triggered revolutionary changes in ophthalmology, as in many other medical fields. Central to these developments are generative AI (GenAI) models that learn from existing data to produce realistic content such as images, text, and video.[1, 2] We believe that GenAI should be treated as a distinct category from other AI paradigms because it can generate entirely new, synthetic, and realistic data instead of classifying or detecting. This capability enables GenAI to support various ophthalmic applications, from disease diagnosis and clinical decision-making to patient information and medical education.[1-4]

Although numerous GenAI architectures have been introduced, ophthalmology most commonly employs three groups of generative models: Generative adversarial networks (GANs), diffusion models, and large language models (LLMs).[4] GANs and diffusion models are typically

1 Department of Ophthalmology, Bezmialem Vakif University, Istanbul, Türkiye

Received: 06.05.2025
Accepted: 23.05.2025

J Ret-Vit 2025; 34: 69-83

DOI:10.37845/ret.vit.2025.34.13

Correspondence author:
Hakan Özdemir
Email: hozdemir72@hotmail.com

chosen for image-based applications, whereas LLMs are used effectively in text-based tasks such as clinical documentation, development of educational materials, and patient communication.[4] We also expect that new generative models capable of changing data modality, for example, text-to-image or image-to-video, will soon attract attention in ophthalmology through many innovative applications.[1, 2, 5, 6]

Ophthalmology provides an exceptionally suitable setting for deploying generative models, as it contains large amounts of digital images and clinical data. The potential uses of these technologies are wide-ranging, and current research is shaped by the needs of both ophthalmologists and patients. However, despite this exciting potential, the clinical use of GenAI models is still in its early stages and faces several fundamental limitations. In particular, the clinical accuracy of data produced by AI, ethical and legal regulations, patient privacy, and safety are viewed as issues that must be resolved before these technologies can be adopted on a wide scale.[7]

In this review, we aim to provide a perspective on the integration of GenAI into ophthalmology practice by examining its current applications, basic principles, potential opportunities, and the ethical and legal debates it raises.

2- FUNDAMENTALS OF GENERATIVE AI MODELS

GenAI models are algorithms that learn patterns from existing datasets and generate entirely new, original, but synthetic content. This content can either maintain the same modality as the input (for example, text-to-text, image-to-image) or transition across different modalities (such as text-to-image or text-to-video), producing outputs that align with the structure or semantics of the input.[1, 2, 6] Currently, the most common generative model types in ophthalmology are GANs, diffusion-based models, and LLMs.

2a- Generative Adversarial Networks

GANs are a deep learning architecture proposed by Goodfellow et al. in 2014 that operates through a competitive interaction between two neural networks, a generator and a discriminator.[8] The generator network first maps a random noise vector into a small stack of feature maps with fully connected layers and then progressively up-samples them through transposed-convolution blocks, each followed by normalization and activation, to refine detail, whereas the discriminator network attempts to distinguish real data from the synthetic data produced by the generator. Throughout training, the generator learns to produce images that mirror the true data distribution, and the discriminator becomes skilled at distinguishing real from fake images (Figure 1). This adversarial training process eventually allows the generator to create synthetic examples that closely resemble real images, enabling the production of data that faithfully approximates the underlying distribution of genuine samples.

Until recently, GANs had been the most widely used GenAI approach for generating synthetic medical images.[5, 9, 10] Their main technical strengths are rapid data synthesis and broad adaptability. Architectural variants such as StyleGAN, CycleGAN, and Conditional GAN make it possible to tailor outputs for diverse problems.[4, 9, 11] Nevertheless, GANs also present important technical limitations.[4, 12, 13] Training is often delicate and unstable. If the competitive balance between the two networks breaks down, a situation known as mode collapse can occur, and the generator begins producing limited and monotonous data. When the balance of power between the discriminator and the generator is not adjusted correctly, the training can suffer from vanishing gradient or oscillation problems. Furthermore, because GANs rely on unsupervised training, they sometimes generate images that contain anatomic or structural inconsistencies. Indeed, studies have documented ocular images produced by GANs that appear realistic at first glance yet reveal anatomical impossibilities upon close inspection.[14, 15]

2b- Diffusion Models

Diffusion models are probabilistic generative models that have emerged as an alternative to GANs and have recently achieved notable success in data synthesis. In a diffusion technique, noise is added to the original data in a stepwise manner, creating a corruption process, the model then learns to reverse that process and reconstruct the original data distribution. During training, the model learns to denoise corrupted inputs, whereas during generation, it starts from pure random noise, progressively structures

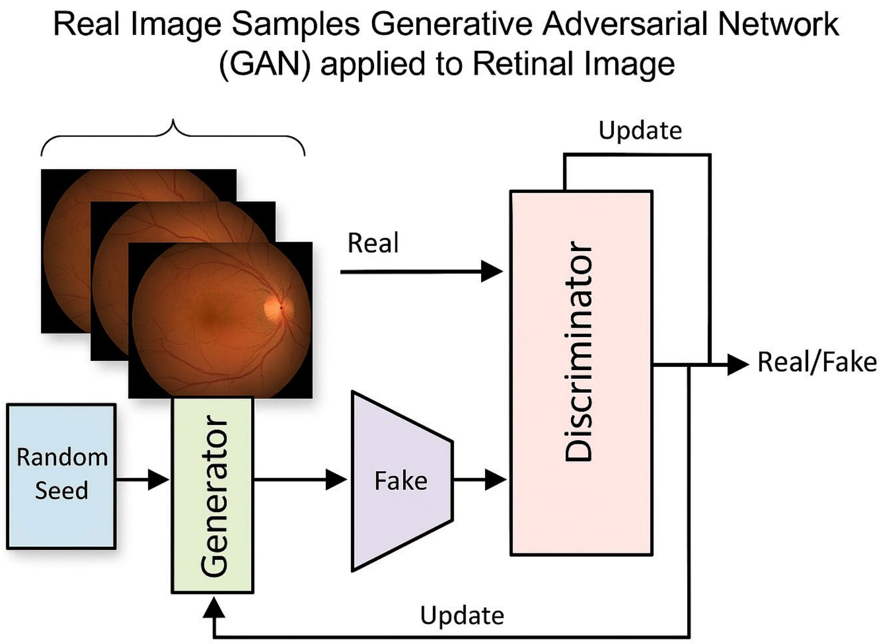

Figure 1: *A schematic representation of the overall architecture of generative adversarial network models. This image was created entirely using the ChatGPT image generation model (gpt-image-1).*

the data, and produces realistic synthetic outputs. In short, diffusion-based generators operate on a theoretical framework grounded in denoising score matching and probabilistic modelling. Their stepwise strategy enables them to capture fine details of high-dimensional and complex data distributions incrementally.[16-18]

Diffusion models have a more stable training process than GANs and largely eliminate the mode collapse problem that is often observed with GANs. Because of their architecture, these models are inherently stable and can generate high-resolution and much more diverse synthetic images.[2, 4, 16] A study by Dhariwal and Nichol showed that diffusion models outperform GANs in both image synthesis quality and diversity.[16] Diffusion models are also particularly strong at text-to-image conversion. Systems such as Stable Diffusion (Runway, Stability AI, and LMU Munich) and DALL-E (OpenAI) apply the diffusion principle to transform a written description (text) into a matching synthetic image. For instance, given the prompt "an optical coherence tomography scan demonstrating dry age-related macular degeneration," the model can produce an artificial optical coherence tomography (OCT) image that fits the description. These vision language models are trained on very large visual and textual datasets and can support many creative uses in ophthalmology.[19-21] Another advantage of the diffusion approach is its controllability and capacity for incremental refinement. Because an intermediate output is produced at each generation step, the user can intervene when needed, giving diffusion models flexibility in creative applications and scenarios requiring progressive editing.[22]

Alongside these innovative strengths, diffusion models also present several technical limitations. They are generally more computationally expensive, and their generation process takes longer than that of GANs, because the image is built in gradual steps. Achieving high-resolution outputs, therefore, demands powerful hardware and large datasets. Another limitation is the greater mathematical complexity of diffusion models. To understand and, when necessary, adapt the training procedure, one must be comfortable with sequential probabilistic processes and continuous random variables. This suggests that developing applications may require more theoretical expertise and experience than working with GANs. Moreover, although diffusion models operate more stably and produce more consistent outputs than GANs, their multi-step reconstruction can permit the accumulation of errors, and these accumulated errors may degrade the final results.[23-25]

2c- Large Language Models

LLMs are deep-learning systems that have achieved a breakthrough in processing and producing human-like language in recent years. Built on the transformer architecture, LLMs are trained on massive text corpora containing trillions of words and learn the statistical patterns of language, enabling them to generate fluent and coherent text in response to a given prompt, their most striking feature is the ability to produce human-like language output. Through exposure to vast data, they acquire a distributed representation that encodes grammar, semantic consistency, and knowledge.[3, 19, 26, 27] The technical strength of LLMs rests on extensive pre-training with large datasets followed by task-specific fine-tuning. In the first stage, the model internalizes broad linguistic patterns from very large text collections, and in the second stage, it is retrained and customized on small datasets tailored to particular tasks. This two-phase training strategy allows the models to achieve successful results across various applications.[28] When you examine the technical definition above, you can see how the abbreviation GPT, which we often use, originated: Generative Pre-trained Transformer.

Well-known applications in this category include ChatGPT (OpenAI), Llama (Meta), Gemini (Google DeepMind), DeepSeek (Hangzhou DeepSeek Artificial Intelligence Basic Technology Research Co., Ltd.), and similar systems. Unlike earlier rule-based or single-task-focused health-care AI solutions, LLM-based chatbots offer an interactive and versatile experience for users. Containing billions of parameters, these models attract attention for their ability to extend a given prompt with sentences and paragraphs that approach human writing quality.[28] Thanks to this scale and capability, LLMs are considered foundation models that address many language processing problems under a single umbrella model.[1, 29] Because their pre-training is general enough to underpin many downstream tasks, they are considered the generative subset within the broader family of foundation models. However, not every foundation model is generative. A foundation model refers to a versatile deep learning core that is trained in a self-supervised manner on large and diverse data and can later be adapted to numerous tasks with only minimal fine-tuning.[29] By learning general representations from text, images, audio, or multimodal inputs, these models can create synthetic outputs and transfer readily to new applications. In clinical settings, for example, an LLM that can understand free-text questions from a clinician or patient and deliver logical, detailed answers can serve as a general-purpose assistant for information access and communication. This capability has brought the potential uses of LLMs to the forefront across every branch of medicine, including ophthalmology.[3,19,26]

The first major issue with LLMs is their dependence on the data used for training. These models can absorb and reproduce the biases present in those datasets, so the accuracy and impartiality of their outputs must be monitored continuously. In fact, an LLM is considered only as reliable as the data on which it is trained.[19] Achieving the desired result also hinges on writing effective text prompts, and the growing use of LLMs has even given rise to a new profession called prompt engineering.[30] A second main concern is the tendency of LLMs to produce erroneous outputs, a phenomenon often referred to as hallucination.[7] When a model lacks sufficient correct information or cannot locate a direct answer, it may confidently fabricate content that does not exist. LLMs further suffer from interpretability challenges because their enormous parameter counts make it difficult to understand why a particular output was generated.[27] Finally, training a multibillion-parameter LLM demands vast amounts of data and extended access to powerful compute infrastructure such as GPU or TPU farms, resulting in high costs for researchers.[19] Although the recent DeepSeek-R1 model is said to offer a more cost-effective approach, there is still speculation about how the model was developed.[31, 32]

2d- Other Generative AI Models

In GAN and diffusion frameworks, the input is an image, and the output is again an image, whereas in LLMs, the input is text, and the output is likewise text. Over the last few years, GenAI systems have become available to handle text-to-video, text-to-image, and image-to-video tasks. As their names imply, these models accept one data modality and return a different one. Beyond ChatGPT-4o (OpenAI), DALL-E 3 (OpenAI), gpt-image-1 (OpenAI) and Midjourney (Midjourney, Inc.), companies such as Adobe (https://helpx.adobe.com/acrobat/using/generate-image-from-text.html) and Google-DeepMind offer comparable text-to-image or text-to-video tools. After receiving a natural-language prompt, the model returns an image, for instance, if asked to create a depiction of diabetic retinopathy, the system can produce a synthetic fundus photograph that shows this disease (Figure 2). Conversely, specific image-to-text models (traditional

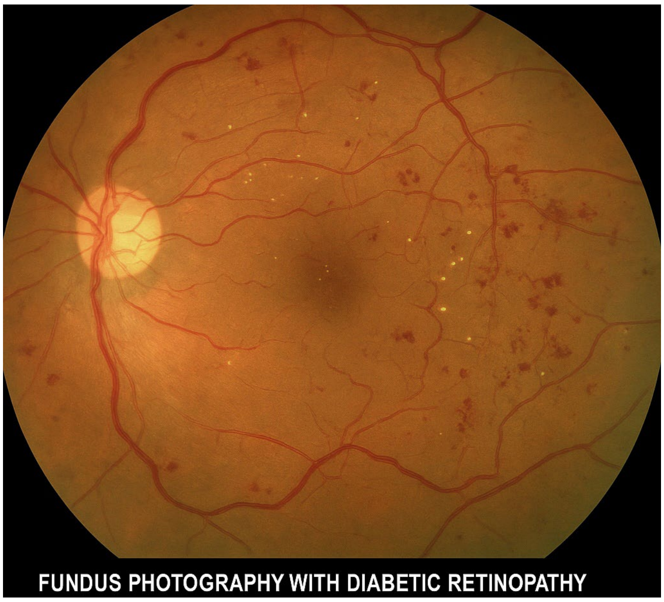

Figure 2: *A synthetic color fundus photograph of diabetic retinopathy was generated using the gpt-image-1 vision-language model with text-prompt–supported image-to-image generation. This image was created entirely using the ChatGPT image generation model (gpt-image-1).*

optical character recognition [OCR] systems are omitted here) also belong to the GenAI family. One can envision, for example, a LLM system built to interpret an OCT scan. Because such architectures process both visual and linguistic representations and ultimately provide synthetic outputs, they are often grouped under the term vision language models.[1, 21]

Among the best-known text-to-video models are Sora (OpenAI), Kling AI, and Veo 2 (Google DeepMind).[33] In these systems, the input is a written prompt, and the model generates a video that matches the description. Image-to-video models operate similarly: a single still image serves as input, and the output is a video. Most such systems need an additional prompt that explains the target clip, although a few, such as Ray2 (Luma AI), can create video without any text input. Modern diffusion-based architectures used for text-to-image, image-to-video, and text-to-video generation share a common backbone composed of a variational auto-encoder, a diffusion transformer, and a cascade up-sampler. Depending on the input modality, they integrate extra conditioning layers, such as text embeddings or motion adapters. These layers allow the system to produce high-resolution training videos, patient education animations, synthetic OCT or fundus images, and even generate a textual interpretation of an OCT scan.[1,33]

3- IMAGE-BASED GENERATIVE AI APPLICATIONS IN OPHTHALMOLOGY

As previously mentioned, it is essential to distinguish Generative AI (GenAI) models from machine learning and deep learning models used for image classification and segmentation. In our previous studies, we demonstrated that various machine learning models can be utilized for the detection and classification of vitreomacular interface disorders, diabetic macular edema, and strabismus.[34-36] On the other hand, Generative AI (GenAI) models generate entirely new synthetic images or perform data transformations across different modalities as their output. GenAI algorithms used for synthetic image generation, such as GANs and diffusion models, can be grouped into four main ophthalmic application categories:

a. Disease-specific synthetic image generation

b. Image enhancement

c. Prediction of post-treatment appearance

d. Modality conversion in ocular imaging

3a. Synthetic Image Generation and Data Augmentation

GenAI models, especially GANs, are widely used in ophthalmology to generate synthetic images and augment datasets. Their most significant clinical potential emerges in rare diseases and in situations where training data are limited.[1, 9] When a sufficiently large training set is available (the exact number is never straightforward to specify because it depends on the characteristic imaging features of the disease and on the technical architecture of the model), disease-specific images can be synthesized to expand the dataset. For example, rare retinal pathologies such as North Carolina macular dystrophy or Doyne honeycomb retinal dystrophy are represented by only a few cases in real-world datasets. Starting from the existing limited images, GANs and diffusion models can generate realistic and clinically meaningful synthetic images. Diffusion models can likewise create high-resolution OCT or fundus photographs synthetically, helping clinicians and trainees learn the imaging features of rare diseases.[37]

The same synthetic data can also facilitate the training of diagnosis algorithms that rely on machine learning. For instance, deep learning systems designed to detect diabetic retinopathy can improve generalization and real-world performance when supplemented with additional synthetic fundus images produced by GAN or diffusion models.[4, 38]

3b. Image Enhancement

GANs and diffusion models are not limited to producing synthetic images in ophthalmology; they can also be used to enhance existing images. Reducing speckle noise in OCT scans, for example, can markedly improve clinical assessment, and algorithms that transform low-quality OCT images into high-resolution versions allow more precise automatic segmentation of retinal layers and make disease monitoring easier (Figure 3).[1, 4, 9] Consider a model designed to quantify intra-subretinal fluid volume. Because the fluid appears as dark pixels, a cleaner high-resolution image free of noise artefacts (white pixels within the fluid space caused by speckle noise) should allow the model to measure the volume more accurately. Additionally, various GAN algorithms can be used for ocular image upscaling (to obtain a high-resolution image from a low-resolution image) (Figure 4). In one of our previous studies, we demonstrated that a GAN-based approach, the Noise2Noise algorithm, can effectively diminish noise in OCT images (Figure 5).[39]

3c. Modality conversion

Modality conversion is another important innovation that generative AI models introduce to ophthalmology. These algorithms transform images acquired with one technique (for example, structural OCT scans) into images corresponding to a different modality (for example, OCT-angiography).[4] A wide spectrum of ocular imaging methods is used for diagnosis, and the importance of multimodal imaging is increasing. Because each modality generally requires its own hardware, GAN-based converters can generate the needed images with only one or a few existing devices, eliminating the expense of additional equipment. This capability reduces costs and offers patients alternatives to costly or invasive procedures. For instance, synthesizing OCT-angiography views from structural OCT data may lessen the need for new device purchases.[4] Many cross-modality converters will likely be tailored to specific clinical requirements, so diverse examples are expected to emerge. Nevertheless, these models remain less mature than other GenAI tools intended for clinical deployment and are still under development (Figure 6).

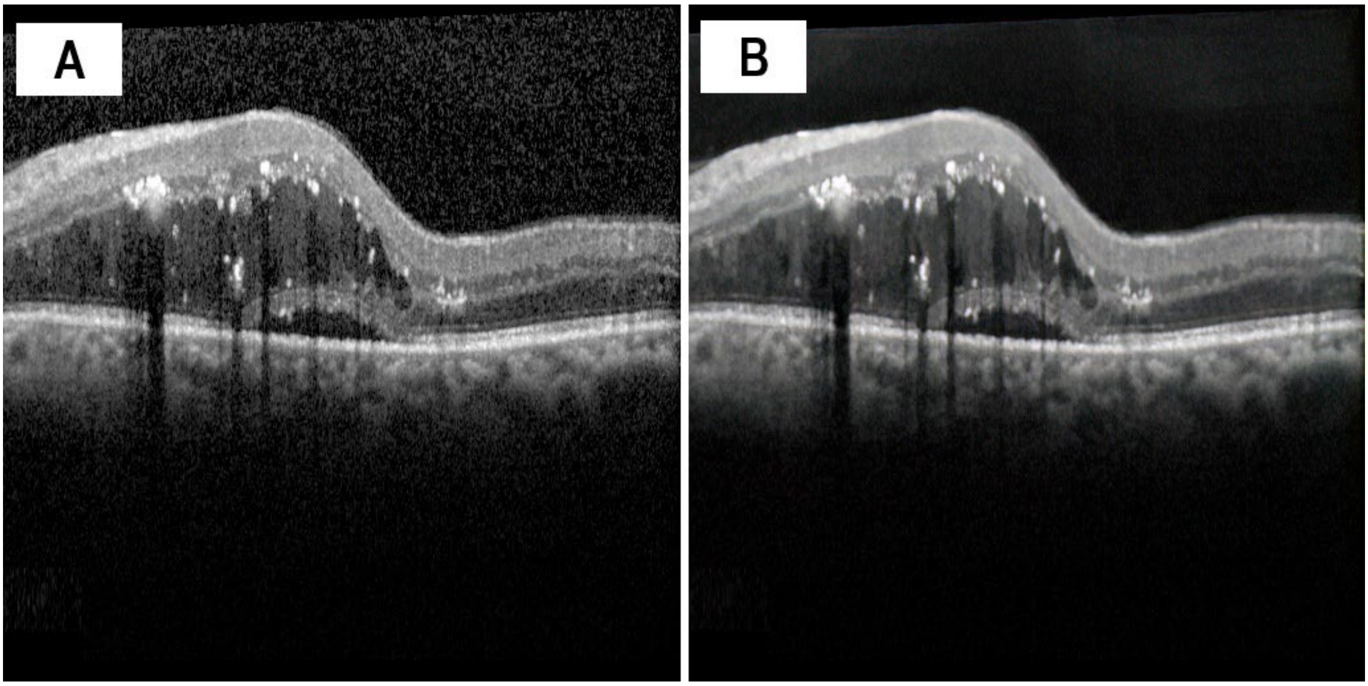

Figure 3: *The Noise2Noise GAN model was used to reduce speckle-noise in cross-sectional spectral-domain optical coherence tomography (SD-OCT) images with diabetic macular edema, enabling clearer visualization of retinal structures and cystic spaces.* ***A)*** *Original SD-OCT image.* ***B)*** *Synthetic image with noise reduction with Noise2Noise algorithm.*

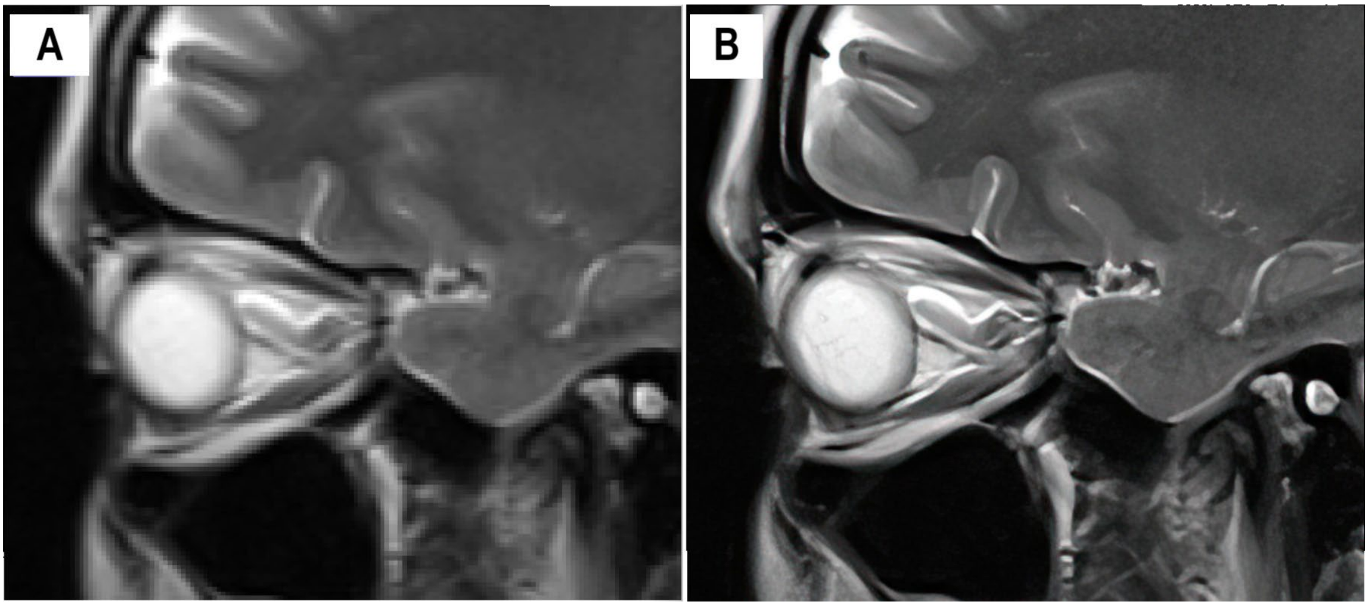

Figure 4: *The low-resolution cranial magnetic resonance image (**A**) of a case of idiopathic intracranial hypertension was upscaled by 600% using the ESRGAN model, resulting in a high-resolution image (**B**).*

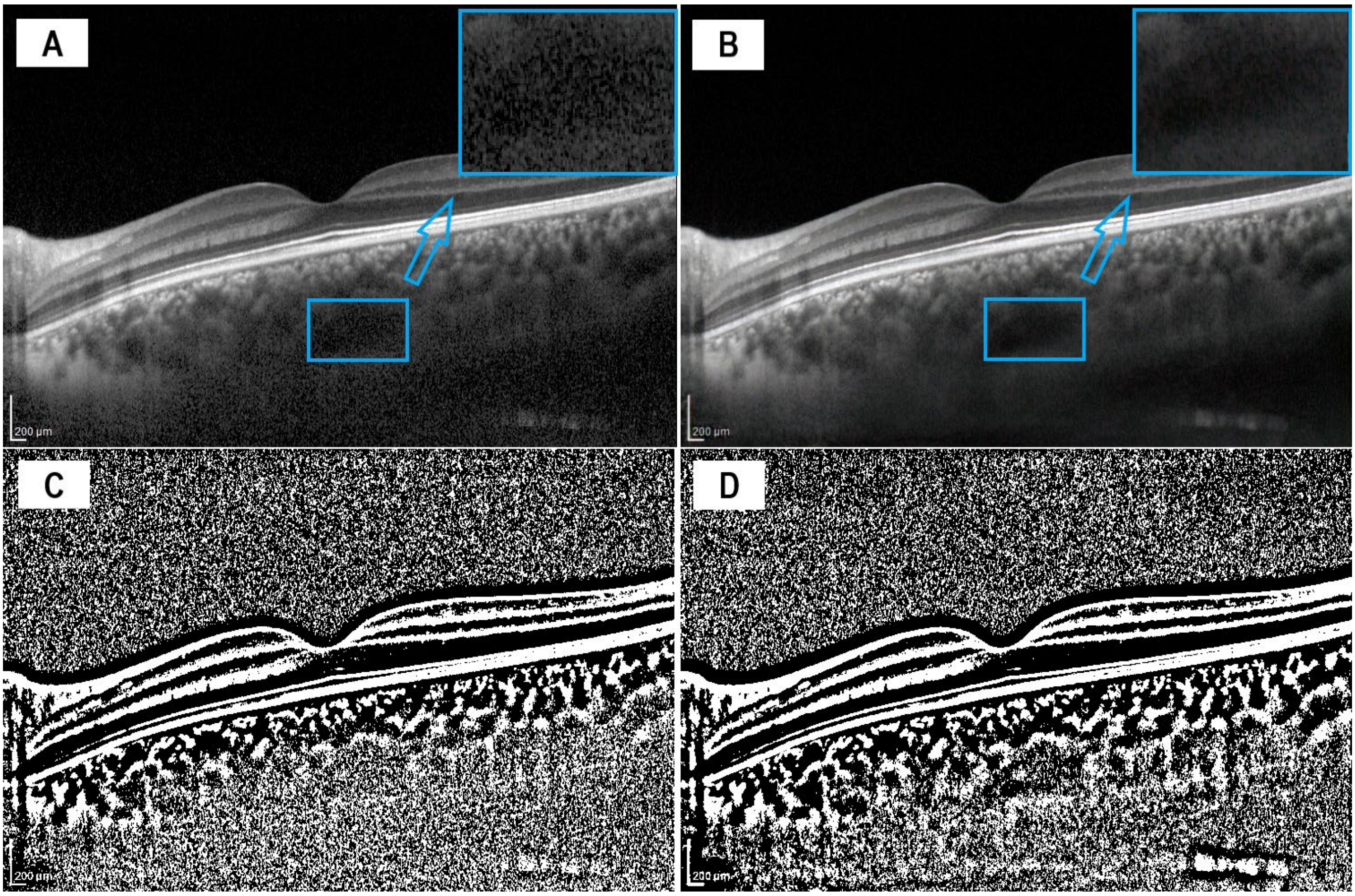

Figure 5: *This figure shows a representative example demonstrating the reduction of speckle noise in enhanced depth imaging–optical coherence tomography (EDI-OCT) images using the Noise2Noise algorithm. In the synthetic image, the speckle noise observed in the vascular structures of the subfoveal choroidal region in the original scan is markedly reduced (blue-framed area). **A.** Original EDI-OCT image, **B.** Speckle-noise-reduced synthetic image, **C.** Binarized original EDI-OCT image with Niblack auto-local thresholding, **D.** Binarized synthetic EDI-OCT image with Niblack auto-local thresholding.*

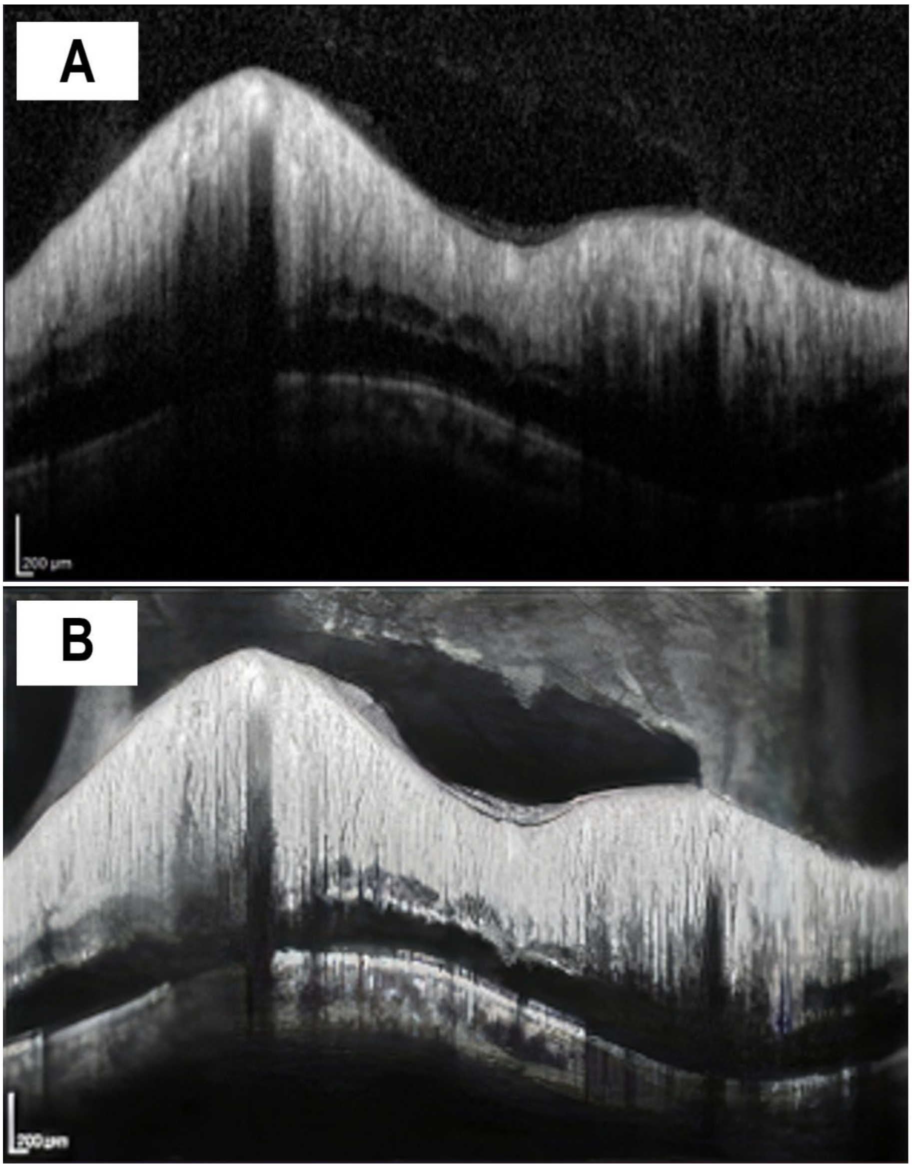

Figure 6: *Representative spectral-domain optical coherence tomography (SD-OCT) images of circumpapillary retinal nerve fiber layer (RNFL) in a case of optic disc edema.* ***A.*** *Original SD-OCT image in which increased RNFL thickness obscures the underlying choroidal structures.* ***B.*** *This image was processed with the BREAD model, a modular deep convolutional neural network (CNN)–based low-light enhancement algorithm, enabling clear visualization of choroidal structures. Thus, it permits assessment of the choroidal region without enhanced-depth imaging–OCT or swept-source OCT.*

3d. Disease Prognosis and Post-treatment Simulations

One of the most promising ophthalmic applications of GenAI is its capacity to forecast disease progression and to simulate ocular or facial appearance after surgical or medical intervention. For example, using GANs, it is possible to generate optical coherence tomography scans that depict the macula of patients with age-related macular degeneration or diabetic retinopathy after anti-VEGF injections. Such simulations allow clinicians to visualize probable therapeutic outcomes in advance and to share them with the patient, fostering realistic expectations and improving adherence.[4] GAN-based systems can predict the cosmetic results that may follow surgery for upper eyelid ptosis, providing both surgeons and patients with valuable preoperative insight.[40] This capability can be crucial in surgical planning and counselling the patient.

4. CLINICAL APPLICATIONS OF LARGE LANGUAGE MODELS IN OPHTHALMOLOGY

The general-purpose capabilities of LLMs are being adapted to various domain-specific tasks in ophthalmology. Current uses can be grouped into six broad areas: (1) diagnostic support and generation of differential diagnoses within clinical decision support systems, (2) triage of incoming cases, (3) clinical documentation and reporting, (4) patient education and communication, (5) medical and specialist training, and (6) academic writing and research assistance. In this section, the clinical applications of LLMs will be discussed, and their roles in academic and medical or ophthalmology education will be addressed in the next section. These categories are not exhaustive, additional use cases will emerge as new clinical needs arise.

Whatever the purpose, the structure and content of the prompt decisively influence the quality of the output.[1, 3, 19] A vague query such as "My right eye does not see well. What could be the cause?" will elicit a markedly different reply from a detailed prompt such as: "My right eye developed sudden painful vision loss one hour ago, the eye is red, I can see partially but the image is blurred, the ocular pain has triggered a headache and vomiting. What are the most likely causes, and should I urgently consult an ophthalmologist?". Well-constructed, informative prompts are therefore essential for obtaining optimal results. The applications discussed below should be evaluated with this principle in mind.

LLM-based clinical decision support systems (CDSS) can analyze patient findings, prior clinical data, and current literature to offer differential-diagnosis suggestions and other forms of diagnostic assistance. When a patient reports symptoms such as blurred central vision and metamorphopsia, the model can propose age-related macular degeneration, diabetic macular oedema, or an epiretinal membrane as possible diagnoses. Acting as a virtual assistant for clinicians, the model can also interpret ocular complaints and return relevant diagnoses and treatment options grounded in up-to-date guidelines. In this way, it supports the differential-diagnosis process and helps ensure that rare eye diseases are not overlooked. The same logic can be applied to triage. By posing structured questions about a patient's symptoms, an LLM can serve as a remote triage and prioritization tool, estimating the urgency of presentation in emergency departments or outpatient clinics.[3, 6, 19, 26, 27]

These models, particularly when integrated with CDSS, can potentially improve the patient experience and accelerate clinical workflows in eye-care settings. LLMs are emerging as powerful tools for optimizing routine documentation in ophthalmology clinics and enhancing decision support efficiency. During note taking, for example, brief dictations or bullet points recorded by the clinician can be automatically transformed by an LLM into a standardized, comprehensive discharge summary. Automating these tasks allows more time to be devoted to patient care, and the administrative workload can be reduced. Similarly, LLMs can automate the assignment of ICD-10 codes.[19, 26, 27]

One key innovation that LLMs can bring to ophthalmic practice is enhancing direct patient communication. LLMs offer clear advantages in producing patient-education materials, translating complex medical conditions into lay language, and answering frequently asked questions.[41-43] By considering each patient's educational background and health literacy, the model can generate personalized content that alleviates anxiety about treatment and supports adherence. Additional GenAI systems, such as text-to-video or image (plus) text-to-video models, can be integrated when creating these resources. Beyond the clinic, LLM-based tools can also assist patients during home treatment and follow-up care. A further benefit of LLMs is their extensive multilingual capability, derived from training on large and diverse datasets, this feature facilitates communication when physician and patient speak different languages and supports the global integration of medical documentation and knowledge.[26]

5. OPHTHALMIC EDUCATION AND ACADEMIC RESEARCH

GenAI technologies can potentially transform medical education, academic research, and scientific communication within ophthalmology. Although most studies examine the use of LLMs, it is clear that other GenAI models will soon be employed for the same purposes. These models perform critical functions by broadening the variety of clinical teaching materials, saving time and resources in scholarly projects, and strengthening academic communication.

5a. Education

In medical education, medical students and ophthalmology residents require theoretical knowledge and access to high-quality, diverse visual materials. Obtaining sufficient examples of rare eye diseases or surgical complications is often challenging. GAN and diffusion models can overcome these limitations. Conditions such as inherited retinal dystrophies or uncommon anterior-segment anomalies can be simulated through synthetic images, thereby enriching both the practical and theoretical learning experience.[1, 9, 37]

LLMs can significantly contribute to the preparation of educational resources and to the clear transmission of theoretical knowledge.[27, 30, 44] These models are able to condense extensive scientific texts, clinical guidelines, and review articles in a short time, highlighting essential points and offering didactic support in a question-and-answer format.[30, 44] In case-based teaching sessions, digital educational assistants that provide real-time responses to learners' queries can enhance interaction and promote durable knowledge acquisition. Furthermore, LLMs can support role-play exercises and scenario-based simulations within ophthalmic education programs.

5b. Academic Research and Writing

Academic investigations comprise time-intensive steps such as literature review, data acquisition, analysis, and manuscript preparation. GenAI models can ease each of these stages for researchers. LLMs, for instance, can perform broad literature searches in seconds and return concise, thematically organized summaries of current knowledge. In densely studied domains such as diabetic retinopathy or macular degeneration, investigators can obtain distilled insights from AI-driven platforms (such as Elicit, Consensus, or GPT-4-assisted search tools) instead of reading thousands of individual papers.[19, 45]

Beyond summarization, structured datasets can be assembled, relevant information can be pulled directly from integrated electronic health-record systems, and support can be provided for statistical analysis. AI-assisted writing tools further improve the language and style of academic manuscripts, offering particular value to scholars who publish in languages other than English.[45]

GenAI can also assist academic workflows through synthetic data generation. GAN- or diffusion-based models can create additional synthetic clinical data when clinical studies suffer from limited sample sizes. This approach is especially valuable for expanding small pilot studies and increasing statistical power.

5c. Scientific Presentation and Communication

Effective dissemination of scientific knowledge and research findings is essential in ophthalmic practice. GenAI models can support multiple stages of this process, from preparing presentations to drafting academic abstracts.[27] AI-based tools can convert academic articles directly into presentation slides, turning research results into explicit, engaging visual materials and saving the time of researchers.[19] LLM-based systems can also streamline conference abstract preparation by expressing complex outcomes through concise graphics and explanatory text, thereby accelerating visual and written content production. In recent years, visual abstracts, which are published with research articles and enable readers to obtain information about the article easily, can also be created with GenAI models.

6. ETHICS, SAFETY, AND REGULATORY CONSIDERATIONS

In addition to their technical potential, the use of GenAI models in ophthalmic practice raises substantial debates concerning ethics, safety, and legal regulation. In medical applications, where patient welfare and safety are paramount, these emerging technologies must fully align with ethical obligations. Nevertheless, current legal frameworks remain markedly insufficient to address AI-driven medical applications.

6a. Clinical Ethics and Patient Safety

The clinical deployment of GenAI models directly impacts patient safety and fundamental principles of clinical ethics. One of the most critical issues is the potential of LLMs to produce "hallucinations," that is, statements containing inaccurate or fabricated information. In medical contexts, LLMs may convincingly present nonexistent treatments, medications, or clinical findings, thereby misleading users. Such misinformation can have serious consequences that place patient health at significant risk. Consequently, all diagnoses and treatment recommendations generated by AI must be verified by qualified clinicians before implementation.[3, 7, 30]

Another ethical concern involves data bias during model training. Patient cohorts that are under-represented or misrepresented in training sets may be disadvantaged by the outputs of clinical algorithms.[2, 3] For example, if fundus images from certain ethnic groups are sparse, a model that detects retinal disease may fail to accurately diagnose patients from those backgrounds. Such disparities risk exacerbating existing inequities in health care. Therefore, increasing the diversity of training data and subjecting algorithms to ongoing fairness audits are essential. The integration of AI into clinical workflows also complicates informed-consent procedures. Deploying these systems without explanations that patients can readily understand is ethically problematic. Accordingly, patients must receive clear information about how the AI model operates, what data it uses, and what potential risks it entails, and their explicit consent must be obtained before use.

6b. Regulatory Framework and Approval Pathways

Clinical application of GenAI in ophthalmology requires clear guidance from regulatory authorities. However, the rapid evolution of AI often outpaces legislation, leaving regulatory gaps that create uncertainty in practice and may jeopardize patient safety.

AI-driven clinical decision support systems are usually classified as Software as a Medical Device (SaMD) and are subject to the approval processes applied to conventional medical devices.[3, 7] The capacity of GenAI models for continual learning and adaptation, however, makes evaluation under traditional, static criteria challenging.

Key legal questions remain unresolved, including how extensively AI systems may participate in clinical decision-making and who bears liability when errors occur. This uncertainty can deter clinicians from adopting such systems and slow the integration of AI-enabled tools into routine care.

6c. Data Privacy and Security

The data-driven nature of generative AI creates significant challenges for maintaining the confidentiality and security of patient information. During model training, sensitive materials such as clinical images and electronic health records may be shared with external parties, which introduces serious privacy risks. Current statutes (including the Health Insurance Portability and Accountability Act [HIPAA], the General Data Protection Regulation [GDPR], and Türkiye's Personal Data Protection Law [KVKK]) impose strict requirements for protecting patient data.[7, 30] Nevertheless, common AI data-handling practices, such as cloud-based storage and cross-border transfers, can conflict with these regulations and complicate compliance. Another risk of generative models is the problem of 'memorization', i.e., the risk that the models remember the training data. In this case, models may directly or indirectly disclose the data they were trained on, which may lead to privacy violations.[3]

6d. Proposed Solutions

Regulatory authorities should design tailored assessment pathways that account for the dynamic nature of AI-based systems and create flexible processes capable of evaluating continuously updated algorithms. Developing common international standards and guidelines would allow these technologies to be appraised consistently across jurisdictions. Such frameworks must define the specific role of AI within CDSS, delineate clear limits for its use, and continually emphasize the necessity of human oversight in clinical practice. In the near future, it will be essential to disclose the rationale underlying each recommendation generated by a CDSS. Techniques from explainable AI (XAI) can render decision-making processes transparent, thereby increasing the confidence of clinicians and patients while enhancing system accountability.[4, 7, 30] Establishing algorithmic ethics standards and empowering institutional ethics committees to monitor the deployment of these technologies are additional key measures.

From a data-protection perspective, approaches such as federated learning should be more widely adopted.[46] By keeping data local and training models on aggregated summaries, federated learning prevents the external transfer of sensitive patient information. The use of advanced encryption methods, can further safeguard data by ensuring that it remains encrypted even during processing.[3, 7, 26]

Moreover, patients must receive clear information about how their data will be used in AI systems and must obtain their explicit consent. Institutions should maintain detailed documentation of the datasets used for model training to ensure traceability. In the event of a privacy breach, this documentation enables rapid identification of the source and prompt implementation of corrective measures.

In summary, the ethical, safety, and regulatory challenges surrounding the clinical integration of GenAI require multifaceted solutions, all aimed at applying technological innovations within a patient-centered, safe, and ethically sound framework. Continuous collaboration among clinical, academic, and regulatory stakeholders will be crucial for ensuring the secure and effective use of GenAI systems and achieving successful transformation in the field of ophthalmology.

7. FUTURE DIRECTIONS AND RESEARCH PRIORITIES

Assessing current trends, the future of GenAI in ophthalmology is likely to be shaped by multimodal systems, hybrid methodologies, early-diagnosis tools, personalized treatment strategies, and tele-ophthalmology solutions. Yet forthcoming AI applications will almost certainly extend beyond this shortlist. Emerging clinical needs will spur novel solutions, and advances in model design will gradually minimize today's limitations. Consequently, researchers can expect to encounter new architectures that address existing disadvantages while opening additional avenues for innovation in ophthalmic care.

7a. Multimodal and Hybrid Systems

Multimodal and hybrid AI platforms are promising to improve diagnostic and prognostic accuracy in ophthalmology by integrating heterogeneous data sources. Imaging modalities such as OCT, fundus photography, and ultrasound biomicroscopy can be combined with electronic health records (EHR), structured clinical reports, laboratory values, and demographic variables to build more comprehensive decision-support tools.[1, 4, 6, 21] This integrated strategy can outperform models that rely on a single data stream. Looking forward, one can envision fully integrated "multimodal imaging" systems that automatically generate outputs corresponding to multiple imaging techniques, perhaps even deriving several modality-specific views from a single standard fundus photograph.

Recent studies also underscore the value of incorporating textual data, including clinical notes and discharge summaries, into analytic pipelines, thereby expanding the diagnostic capabilities of AI systems. Such multimodal and hybrid approaches will elevate clinical decision support by furnishing ophthalmologists with multidimensional insights about each patient.[30]

7b. Early Disease Prediction and Personalized Medicine

Early diagnosis, predicting disease progression, and individualized treatment planning are central to contemporary generative AI research. AI-enabled models can identify subclinical alterations in retinal images, such as those seen in diabetic retinopathy (DR), allowing for the detection of the earliest disease manifestations and estimation of the progression rate. Furthermore, by integrating GenAI models with hybrid systems to create "digital twins" of rare or complex cases, clinicians can test therapeutic options in silico before applying them to patients. This approach enhances patient safety and moves ophthalmic practice decisively toward truly personalized care.[3, 4, 30]

GenAI-based prognostic systems can process large datasets to predict patients' future visual acuity and therapeutic responses based on retinal scans. Using GAN-oriented methods, researchers have demonstrated the ability to forecast post-treatment OCT images in conditions such as age-related macular degeneration and diabetic macular edema following anti-VEGF therapy. These personalized approaches will allow therapeutic strategies to be tailored to the genetic profile, lifestyle, and disease dynamics of each patient.[30]

7c. Tele-ophthalmology and Remote Eye-Care

GenAI applications in tele-ophthalmology substantially enhance eye-care delivery in regions where geographical barriers or limited clinic access constrict services. Chatbot systems powered by LLMs can take detailed histories from patients who report blurred vision, ocular pain, or redness and then triage them according to urgency. When a sight-threatening condition is suspected, the system can automatically initiate an emergency tele-consultation and promptly direct the patient to specialist evaluation. Moreover, such platforms can help reduce the burden of non-urgent visits in ophthalmic emergency departments.[26, 30]

Tele-ophthalmology solutions also improve early-detection rates in rural settings by enabling AI-based analysis of remotely acquired ocular images, such as retinal or external ocular photographs.[47] For diseases such as diabetic retinopathy, home monitoring data can be assessed by AI algorithms, and clinicians can be alerted whenever a clinically significant change is detected, thereby facilitating earlier intervention.

Future AI platforms may acquire the capability to perceive emotional state of a patient.[48] By analyzing vocal intonation or textual cues, empathetic virtual assistants could recognize anxiety or distress and provide tailored reassurance, making chronic-disease management more effective and patient-centered.

We are also likely to witness far greater automation of academic workflows. AI-driven analytic engines could extract meaningful scientific findings directly from ophthalmic images and electronic health records, leaving researchers to focus primarily on interpreting results. In parallel, personalized research assistants that adapt to an investigator's interests and working style may take a more active role in project execution. Research teams might soon expand to include virtual AI collaborators that handle routine tasks within the group. It may also be anticipated that purpose-specific custom GPT models, built on the GPT architecture and enriched with a variety of online and offline educational resources, will become increasingly widespread in the near future and be integrated into ophthalmology training.[49]

Beyond the topics discussed above, the potential applicability of GenAI models is also likely to become a focal point of future debate across a wide array of innovative and breakthrough technologies, ranging from robotic surgical instruments and automated imaging devices to augmented or virtual reality-integrated wearables and voice-activated digital or robotic assistants. Additionally, the abovementioned AI systems essentially do not perform autonomous actions, but only provide recommendations that assist the physician's decision-making. Agentic AI, the most intriguing innovation of recent days, are integrated multistage, integrated AI models that can transform their deliberative outputs into autonomous actions.[50] We consider that such systems may have the potential to be primarily integrated into the surgical process in robotic ocular surgery or to provide autonomous treatment decisions. Thus, the scope of human-AI collaboration in ophthalmic practice may be redefined, and the roles of physician and AI may change over time (such as AI actions executed under human supervision). In the future, we are likely to observe the adoption of agentic AI platforms in multiple medical fields, including both robotic ocular surgery and outpatient care.

8. CONCLUSION

GenAI marks the outset of a profound transformation in ophthalmology. Innovative contributions are expected in diagnosis, treatment planning, patient communication, and education. The ethical and practical challenges accompanying these technologies must be taken seriously and addressed through collective effort. With a balanced, evidence-informed strategy supported by close cooperation among academia, clinical practice, and industry, the risks of GenAI can be managed while its benefits are maximized. As evidence and experience accumulate in the years ahead, AI-enabled eye-care services will likely become integral to standard practice. In retrospect, this period may well be remembered as the dawn of a brighter era in ocular health, one in which technology blends harmoniously with the art of medicine to deliver higher-quality, more accessible, and more personalized care to broader populations.

During this transition, physicians bear a critical responsibility: Cultivating an unbiased awareness of AI. Active engagement by physicians is essential for the ongoing development and integration of these technologies. Ophthalmologists who acquire AI literacy and proactively incorporate AI tools into their workflows can become pioneers of AI-assisted practice and play pivotal roles in tomorrow's health-care landscape. In doing so, ophthalmology will stand at the forefront of patient-centered, AI-supported medicine. Realizing this vision depends on ophthalmologists and researchers actively embracing GenAI, remaining continuously updated, and integrating these tools into everyday practice. Education and research will then become more efficient, innovative, and engaging.

In the near future, AI is unlikely to replace ophthalmologists; however, those ophthalmologists who lack AI awareness or resist adopting these systems may find the transition challenging.

Acknowledgements: Original manuscript was drafted entirely by the authors. ChatGPT-4.5 (OpenAI; used April 01–30, 2025), Gemini 2.5 Pro Experimental (Google DeepMind; used April 01–30, 2025), DeepL Pro (DeepL SE; used April 01–30, 2025) and Grammarly Pro (Grammarly Inc.; used April 01–30, 2025) were employed solely for English grammar, stylistic fluency, internal consistency and for simplifying technical expressions to a level understandable by ophthalmologists. The authors conducted the initial literature search; thereafter, Consensus AI (Consensus Systems Inc.; used March 25-28, 2025) was used to identify additional relevant references. All references and content provided via artificial intelligence-assisted technologies were comprehensively reviewed by the authors, and only those fully consistent with the manuscript's content and whose accuracy was verified by the authors were included. Figure 1 and Figure 2 were created entirely using the ChatGPT image generation model (gpt-image-1). The authors take full responsibility for the integrity of all generated and selected material. We also acknowledge the technical support of Eyüb Uzun, MD, and Buse Pulat, MD.

REFERENCES

1. Sonmez SC, Sevgi M, Antaki F et al. Generative artificial intelligence in ophthalmology: current innovations, future applications and challenges. Br J Ophthalmol 2024; 108:1335-1340.
2. Feng X, Xu K, Luo MJ et al. Latest developments of generative artificial intelligence and applications in ophthalmology. Asia Pac J Ophthalmol (Phila) 2024; 13:100090.
3. Yang Z, Wang D, Zhou F et al. Understanding natural language: Potential application of large language models to ophthalmology. Asia Pac J Ophthalmol (Phila) 2024:100085.
4. Phipps B, Hadoux X, Sheng B et al. AI image generation technology in ophthalmology: Use, misuse and future applications. Prog Retin Eye Res 2025; 106:101353.
5. Waisberg E, Ong J, Kamran SA et al. Generative artificial intelligence in ophthalmology. Surv Ophthalmol 2024.
6. Buess L, Keicher M, Navab N et al. From large language models to multimodal AI: A scoping review on the potential of generative AI in medicine. arXiv preprint arXiv:250209242 2025.
7. Ning Y, Teixayavong S, Shang Y et al. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. Lancet Digit Health 2024; 6:e848-e856.
8. Goodfellow I, Pouget-Abadie J, Mirza M et al. Generative Adversarial Networks. Adv Neural Inf Process Syst 2014; 27.
9. 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; 32:459-467.
10. Paladugu PS, Ong J, Nelson N et al. Generative adversarial networks in medicine: important considerations for this emerging innovation in artificial intelligence. Ann Biomed Eng 2023; 51:2130-2142.
11. Karras T, Aila T, Laine S et al. Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:171010196 2017.
12. Pavan Kumar M, Jayagopal P. Generative adversarial networks: a survey on applications and challenges. Int J Multimed Inf Retr 2021; 10:1-24.
13. Vahdati DS, Stamm MC. Detecting gan-generated synthetic images using semantic inconsistencies. Electronic Imaging 2023; 35:1-6.
14. Hu S, Li Y, Lyu S. Exposing GAN-generated faces using inconsistent corneal specular highlights. ICASSP 2021-2021 Proc IEEE Int Conf Acoust Speech Signal Process (ICASSP). IEEE, 2021; 2500-2504.
15. Guo H, Hu S, Wang X et al. Eyes tell all: Irregular pupil shapes reveal gan-generated faces. ICASSP 2022-2022 Proc IEEE Int Conf Acoust Speech Signal Process (ICASSP). IEEE, 2022; 2904-2908.
16. Dhariwal P, Nichol A. Diffusion models beat gans on image synthesis. Proc Adv Neural Inf Process Syst 2021; 34:8780-8794.
17. Chang Z, Koulieris GA, Shum HP. On the design fundamentals of diffusion models: A survey. arXiv preprint arXiv:230604542 2023.
18. Yang L, Zhang Z, Song Y et al. Diffusion models: A comprehensive survey of methods and applications. ACM Computing Surveys 2023; 56:1-39.
19. Thirunavukarasu AJ, Ting DSJ, Elangovan K et al. Large language models in medicine. Nat Med 2023; 29:1930-1940.
20. Antaki F, Chopra R, Keane PA. Vision-Language Models for Feature Detection of Macular Diseases on Optical Coherence Tomography. JAMA Ophthalmol 2024; 142:573-576.
21. Lim G, Elangovan K, Jin L. Vision language models in ophthalmology. Curr Opin Ophthalmol 2024; 35:487-493.
22. Improved Denoising Diffusion Probabilistic Models (Supplementary). 2021.
23. Xiao H, Wang X, Wang J et al. Single image super-resolution with denoising diffusion GANS. Sci Rep 2024; 14.
24. Lu H, Salah AA, Poppe R. Compensation sampling for improved convergence in diffusion models. Comput Vis ECCV. Springer, 2024; 183-201.
25. Wang W, Eklund A. Beware of diffusion models for synthesizing medical images—a comparison with GANs in terms of memorizing brain MRI and chest x-ray images. Mach Learn Sci Technol 2025; 6.
26. Betzler BK, Chen H, Cheng C-Y et al. Large language models and their impact in ophthalmology. Lancet Digit Health 2023; 5:e917-e924.
27. 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.
28. Chang Y, Wang X, Wang J et al. A survey on evaluation of large language models. ACM T INTEL SYST TEC 2024; 15:1-45.
29. Myers D, Mohawesh R, Chellaboina VI et al. Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. Cluster Computing 2024; 27:1-26.

30. Heinke A, Radgoudarzi N, Huang BB et al. A review of ophthalmology education in the era of generative artificial intelligence. Asia Pac J Ophthalmol (Phila) 2024; 13:100089.

31. Zeng D, Qin Y, Sheng B et al. DeepSeek's "Low-Cost" Adoption Across China's Hospital Systems: Too Fast, Too Soon? JAMA 2025.

32. Young S, Tompros L. DeepSeek, ChatGPT, and the global fight for technological supremacy. Harvard Law Today, https://hls.harvard.edu/today/deepseek-chatgpt-and-the-global-fight-for-technological-supremacy/ Access date: 02.25.2025.

33. Waisberg E, Ong J, Masalkhi M et al. OpenAI's Sora in ophthalmology: revolutionary generative AI in eye health. Eye (Lond) 2024; 38:2502-2503.

34. Kirik F, Ozbas C, Aslanoglu CE et al. Efficiency of web-based code-free artificial intelligence platform in classification of vitreomacular interface diseases. Eur Eye Res 2024; 4:57-62.

35. AKBULUT E, KIRIK F, BAYRAKTAR H et al. Determination of Optimum Hyperparameters in Diagnosis of Strabismus Using Artificial Intelligence Model: Cross-Sectional Study. Türkiye Klinikleri J Ophthalmol 2023; 32:176-183.

36. Kırık F, Demirkıran B, Aslanoğlu CE et al. Detection and Classification of Diabetic Macular Edema with a Desktop-Based Code-Free Machine Learning Tool. Turk J Ophthalmol 2023; 53:301.

37. Pontikos N, Woof W, Krawitz P et al. Eye2Gene: prediction of causal inherited retinal disease gene from multimodal imaging using AI. Invest Ophthalmol Vis Sci 2022; 63:1161-1161.

38. Chen Y, Yang X-H, Wei Z et al. Generative Adversarial Networks in Medical Image augmentation: A review. Comput Biol Med 2022; 144:105382.

39. Kirik F, Iskandarov F, Erturk KM et al. Quantitative analysis of deep learning-based denoising model efficacy on optical coherence tomography images with different noise levels. Photodiagnosis Photodyn Ther 2024; 45:103891.

40. Sun Y, Huang X, Zhang Q et al. A fully automatic postoperative appearance prediction system for blepharoptosis surgery with image-based deep learning. Ophthalmol Sci 2022; 2:100169.

41. Yigit DD, Aykut A, Sevik MO et al. ChatGPT versus strabismus specialist on common questions about strabismus management: a comparative analysis of appropriateness and readability. Marmara Med J 2024; 37:323-326.

42. Aydın FO, Aksoy BK, Ceylan A et al. Readability and Appropriateness of Responses Generated by ChatGPT 3.5, ChatGPT 4.0, Gemini, and Microsoft Copilot for FAQs in Refractive Surgery. Turk J Ophthalmol 2024; 54:313-317.

43. Postacı SA, Dal A. The Ability of Large Language Models to Generate Patient Information Materials for Retinopathy of Prematurity: Evaluation of Readability, Accuracy, and Comprehensiveness. Turk J Ophthalmol 2024; 54:330-336.

44. Sevgi M, Antaki F, Keane PA. Medical education with large language models in ophthalmology: custom instructions and enhanced retrieval capabilities. Br J Ophthalmol 2024; 108:1354-1361.

45. Hu FY, Chen LY, Cheng PJ et al. Utilizing generative AI in ophthalmic medical paper writing: Applications, limitations, and practical tools. Asia Pac J Ophthalmol (Phila) 2025; 14:100174.

46. Teo ZL, Lee AY, Campbell P et al. Developments in Artificial Intelligence for Ophthalmology: Federated Learning. Asia-Pacific Journal of Ophthalmology 2022; 11:500-502.

47. Özdemir H, Kırık F: Tele-Oftalmoloji. In Soysal P, Tükek T: Tele Sağlık. İstanbul Tıp Kitabevleri, İstabul, 2022, 117-174.

48. Mohammad GB, Potluri S, Kumar A et al. An Artificial Intelligence-Based Reactive Health Care System for Emotion Detections. Comput Intell Neurosci 2022; 2022.

49. Taloni A, Alessio G, Coco G et al. New frontiers in generative artificial intelligence - enhancing custom GPTs with ophthalmology guidelines. Eye (Lond) 2025; 39:1235-1237.

50. Acharya DB, Kuppan K, Divya B. Agentic AI: Autonomous Intelligence for Complex Goals–A Comprehensive Survey. IEEE Access 2025; 13:18912-18936.