

<https://doi.org/10.1038/s41746-025-01754-4>

Embodied artificial intelligence in ophthalmology



Yao Qiu^{1,2,7}, Xiaolan Chen^{1,7}, Xinyuan Wu^{1,7}, Yunqian Li², Pusheng Xu¹, Kai Jin³, Xianwen Shang¹,
Peranut Chotcomwongse⁴, Mingguang He^{1,5,6} ✉ & Danli Shi^{1,5} ✉

Embodied artificial intelligence (EAI) integrates perception, memory, reasoning and action through physical interaction, enabling multimodal dynamic learning and real-time feedback. In ophthalmology, EAI supports data-driven decision-making, improving the precision and personalization of diagnosis, surgery, and treatment. It also holds transformative potential in medical education and scientific research by simulating clinical scenarios and accelerating discovery. This perspective highlights EAI's unique potential while addressing current challenges in data, interpretation, and ethics, and outlines future directions for its clinical integration.

Artificial intelligence (AI) has advanced rapidly and is increasingly applied in healthcare. Traditional AI systems typically rely on data-driven algorithms that operate within fixed computational environments. Although they have achieved notable success in tasks such as image recognition and risk prediction, their inability to physically interact with the environment limits their effectiveness in dynamic clinical settings.

Embodied AI (EAI) represents a new paradigm by integrating AI into physical entities, such as robots, enabling perception, learning and dynamically interaction with their environments¹. Its theoretical foundation originated in late-1980s and coalesced into the *Embodiment Hypothesis*, which emphasizes the core role of physical interaction in cognition². Unlike conventional AI models that rely on static datasets, EAI uses real-time multimodal inputs for context-aware adaptation, enabling it to navigate changing environments and solve complex real-world problems³. Recent breakthroughs in large language models (LLMs), vision-language models (VLMs), and reinforcement learning (RL) further advanced the development of next-generation EAI systems, equipping them with “intelligent brain” particularly valuable for clinical applications requiring personalization and responsiveness^{4,5}.

In healthcare, EAI offers promising applications. It can support disease diagnosis, monitor recovery, and personalize patient care, especially improving accessibility and efficiency in remote or underserved regions⁶. In ophthalmology, EAI holds unique potential due to the field's demand for micron-level surgical precision, extensive multimodal imaging data, and the widespread use of portable imaging devices. With real-time perception and adaptive capabilities, EAI can assist in tasks such as multi-modal image-

based diagnosis, surgical planning, intraoperative guidance, and individualized treatment optimization. Summarizing the core technologies, current development, and existing challenges in ophthalmology is therefore essential to inform future research and clinical implementation.

This paper provides an introductory guide to EAI in ophthalmology. Section I outlines the concept and fundamental components of EAI, including perception, reasoning, action and memory. Section II explores its potential applications across clinical care, education, and research. Section III discusses existing challenges and future directions in this field. By emphasizing the “intelligent brain” aspect of EAI, we aim to provide valuable insights for clinicians and vision researchers engaging with this emerging technology.

Introduction to EAI

Characteristics and advantages of EAI

EAI addresses traditional AI challenges by endowing virtual agents with the ability to perceive, move, speak, and interact in simulated environments and ultimately in the real world. These agents, initially trained in virtual settings, can seamlessly transfer their learned behaviors to physical robotic systems.

Unlike traditional ophthalmic AI, which depends on large-scale, static datasets such as images, videos, or text curated from the internet, EAI emphasizes real-time, embodied interaction with the environment, closely mirroring human experiential learning. Traditional internet-based AI systems often struggle to generalize across dynamic clinical contexts due to their heavy reliance on the quality and diversity of pre-existing training data. While EAI systems may still utilize internet connectivity for updates or

¹School of Optometry, The Hong Kong Polytechnic University, Kowloon, Hong Kong. ²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangdong Provincial Key Laboratory of Ophthalmology and Visual Science, Guangdong Provincial Clinical Research Center for Ocular Diseases, Guangzhou, China. ³Eye Center, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China. ⁴Vitreoretina Unit, Department of Ophthalmology, Rajavithi Hospital, Rungsit University, Bangkok, Thailand. ⁵Research Centre for SHARP Vision (RCSV), The Hong Kong Polytechnic University, Kowloon, Hong Kong. ⁶Centre for Eye and Vision Research (CEVR), 17W Hong Kong Science Park, Hong Kong. ⁷These authors contributed equally: Yao Qiu, Xiaolan Chen, Xinyuan Wu. ✉ e-mail: mingguang.he@polyu.edu.hk; danli.shi@polyu.edu.hk

remote computations, their distinguishing strength lies in continuous interaction with the environment.

In an EAI framework, multimodal sensors integrated within intelligent agents gather real-time sensory data, including visual, tactile, and auditory signals^{7–9}. Through a closed-loop perception-action cycle, EAI systems integrate real-time observations with historical knowledge to determine appropriate action strategies, such as providing diagnostic assistance, recommending further examinations, or adjusting treatment plans. Following each action, the system evaluates the outcome, compares it with expected results, and refines future behavior. This iterative loop of sensing, acting, and learning allows EAI to progressively capture individual patient characteristics and disease progression, supporting more precise and personalized medical interventions.

The transformative potential of EAI lies in its capacity to enable real-time, adaptive ophthalmic care assistants. By tailoring responses to individual examination results, EAI can offer immediate clinical support, personalized treatment guidance, and dynamic patient management. A comparative summary of traditional AI and EAI-based technologies is presented in Table 1.

Advancements in EAI

EAI builds on advances across multiple disciplines, including intelligent algorithms, sensory perception, decision-making, and robotic control¹⁰. The development of EAI systems is inherently interdisciplinary, integrating insights from neuroscience, psychology, robotics, machine learning, LLMs, and related technologies. These systems continually evolve to enhance key capabilities such as perception, language understanding, reasoning, planning, navigation, and motor execution. At the center of EAI lie four fundamental processes: perception, memory, reasoning, and action (Fig. 1).

Perception: enhancing multimodal sensing in medicine

Advanced perception technologies are crucial for EAI systems to recognize and integrate diverse sensory inputs, such as visual, auditory, tactile, and even olfactory, from complex environments¹¹.

Convolutional neural networks (CNNs) and transformer-based architectures demonstrated superior performance in detecting subtle disease features in medical images, outperforming traditional image processing techniques^{12,13}. In ophthalmology, the integration of multimodal data further improves diagnostic accuracy^{14–16}. Additionally, generative AI facilitates data augmentation, enhancing model performance in underrepresented or rare disease categories. Emerging foundation models and representation learning offer data-efficient improvements across a range of diagnostic

tasks^{17–19}. Meanwhile, advancements in 3D perception technologies enhance spatial awareness and enable accurate interpretation of volumetric data from first-person perspectives²⁰.

Beyond visual input, haptic and auditory perception are increasingly vital, particularly in surgical settings. EAI systems equipped with these capabilities can deliver real-time tactile or verbal feedback to surgeons, enabling robotic systems to adapt dynamically during intricate ocular procedures and minimize the risk of tissue damage²¹. Visual navigation, critical for autonomous systems, combines visual, tactile, and linguistic information to map surroundings, recognize obstacles, and respond to voice commands²². These capabilities are essential for tasks such as indoor navigation and medical procedures.

Memory: facilitating continuous learning and context retention

Memory is essential for EAI systems to operate effectively in dynamic clinical environments. Unlike static AI models, EAI systems maintain both short-term and long-term memory to support continual learning and context-aware decision-making²³.

Short-term memory captures recent interactions and transient environmental changes, enabling respond rapidly to evolving scenarios. In ophthalmology, this is critical during procedures such as intraoperative monitoring, where real-time adaptation is essential²⁴.

Long-term memory stores accumulated knowledge from multiple clinical encounters and human feedback, allowing EAI systems to generalize from past experiences, refine clinical reasoning, and improve treatment strategies over time²⁵. Integrating long-term memory supports the continual adaptation of diagnostic models to evolving patient populations or rare disease presentations in the target scenario.

The combination of short- and long-term memory provides a foundation for lifelong learning, enabling continuous updates to the model’s knowledge base without the need for complete retraining.

Reasoning: empowering clinical decision-making

Reasoning is fundamental to the ability of EAI systems to operate in complex and uncertain medical environments. To ensure safe and explainable decision-making, EAI systems should use advanced techniques (Table 2) to generate transparent, context-aware responses with real-time multimodal data process, including imaging, clinical text, and patient feedback.

LLMs and VLMs play a pivotal role in multimodal reasoning, natural language and visual inputs translation into clinically meaningful actions²⁶. Techniques such as chain-of-thought (CoT) prompting enhance logical reasoning by guiding models through step-by-step inferences²⁷, while self-

Table 1 | Comparison between traditional ophthalmic techniques and Embodied Artificial Intelligence (EAI) ophthalmic techniques

	Traditional ophthalmic techniques	EAI-enhanced ophthalmic techniques
Data Source	Passive recording (manual examination, input by ophthalmologists)	Active multimodal sensing (intraoperative OCT, eye tracking, etc.)
Information Integration	Unimodal or limited multimodal (e.g., OCT and CFP)	Multimodal fusion (eye tracking, biomechanics, etc)
Feedback Mechanism	Doctor explains results, patient passively receives	Device dynamically adjusts exams, interactive with users
Learning Mechanism	One-way learning, requires large scale datasets for retraining	Closed-loop perception-action reinforcement learning
Decision Timeliness	Retrospective analysis (minutes)	Real-time feedback (milliseconds)
Decision Capability	Relies on subjective judgment (affected by fatigue/ experience)	Instant data acquisition and feedback
Adaptability	Limited adaptability to novel or rare diseases.	Requires extensive training and sensor integration for practical deployment.
Cost	Relatively low for standard techniques (e.g., slit-lamp examination).	Relatively high, but it can reduce human workload in long-term use through automation.
Clinical application	Standardized screening and diagnosis (e.g., DR grading, OCT analysis)	Personalized treatment assistance (e.g., retinal injections, robotic obstacle avoidance)
Pros and Cons	Well-established, cost-effective, but limited real-time adaptability.	Enables real-time, data-driven decision-making, but requires high computational resources and advanced integration.

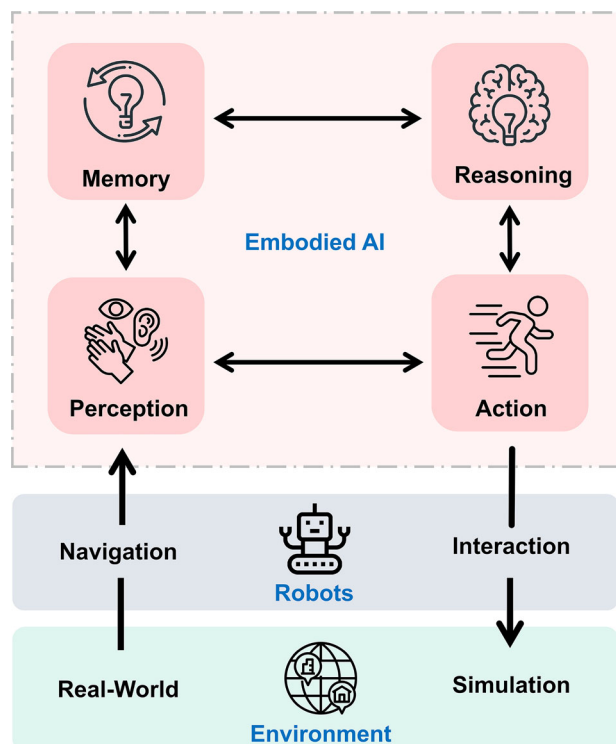


Fig. 1 | The core process of embodied artificial intelligence (EAI). EAI interact with environments through physical embodiment, integrating perception, memory, reasoning, and action to achieve autonomous learning and adaptive decision-making. The graphic was made with assets from Freepik.com.

consistency methods further improve decision reliability by selecting the most coherent output from multiple reasoning paths²⁸.

Graph neural networks support causal modeling between clinical variables, facilitating personalized diagnostics and treatment planning²⁹. RL, particularly when integrated with deep learning, enables EAI systems to optimize clinical strategies through trial-and-error interactions³⁰. To reduce dependence on hand-crafted rewards and improve stability, human-in-the-loop feedback aligns system behavior with expert knowledge and patient-specific contexts³¹. These advancements enable continuous refinement of EAI systems' reasoning capabilities^{32–34}.

Spatial reasoning is essential for surgical planning and autonomous navigation. Recent developments in LLM-based spatial reasoning have reduced navigation errors and supported dynamic, sequential decision-making^{35,36}. High-level task planning tools, such as ProgPrompt and Socratic Planner, enable systems to decompose clinical objectives into executable subtasks and dynamically adjust plans based on environmental feedback^{37,38}. Frameworks like ISR-LLM further enhance performance in uncertain settings by employing iterative self-evaluation and real-time adaptation³⁹.

Action: closing the perception-decision loop

Action modules convert the outputs of perception and reasoning into precise, context-aware physical operations, enabling EAI systems to interact with physical or virtual environment.

Low-level actions are typically direct responses to sensory inputs or external instructions and are governed by control policy representations, which encode robotic behavior. An effective control policy balances expressive capacity required for executing complex tasks with computational efficiency and adaptability⁴⁰. By precisely regulating joint positions, robotic systems can follow predefined motion trajectories with high accuracy. Force-feedback mechanisms further enhance safety and precision by enabling real-time adjustments during physical interactions, maintaining stable contact with delicate tissues in surgical settings⁴¹. These low-level actions often execute deterministic tasks, such as instrument manipulation

or gaze tracking in ophthalmic devices. Vision-language-action models extend this capability by integrating sensory perception with motor control, facilitating real-time intraoperative adjustments in response to dynamic changes⁴².

High-level actions integrate outputs from memory and reasoning modules, often leveraging LLMs, RL, or hybrid strategies. Advanced planners such as EmbodiedGPT and LLM-Planner synthesize environmental and clinician inputs to generate adaptive, context-aware action sequences⁴³. These approaches reduce dependence on exhaustive pretraining and improve flexibility, especially in high-stakes applications such as retinal repair and intraocular laser surgery^{42,44}.

Application of EAI in Ophthalmology

Emerging AI technologies have already been applied in the diagnosis and treatment of ophthalmic diseases^{45–47}. With the rise of EAI, its potential applications in ophthalmology are drawing attention, including disease screening and diagnosis, surgical assistance, support for the visually impaired, medical education, and clinical research (Fig. 2).

EAI in clinical applications

Diagnosis and screening of eye diseases. The global aging population has led to a rising prevalence of vision-threatening conditions such as diabetic retinopathy, glaucoma, age-related macular degeneration, and cataracts⁴⁸. Early detection through population-based screening is critical to preserving vision and improving long-term outcomes.

Recent applications of deep learning algorithms in eye disease screening have achieved high accuracy⁴⁹. However, large-scale implementation remains limited by the need for skilled personnel and the associated costs^{50,51}. EAI holds promise for autonomous screening, minimizing human intervention. For example, EAI systems can automatically acquire OCT images and interpret them using integrated AI algorithms⁵², facilitating scalable, rapid, and cost-effective screening⁵⁰. Low-cost, portable devices are also emerging, such as the lightweight fundus camera for self-examination⁵³, and the TRDS system, which combines a deep learning model with a handheld infrared eccentric photorefraction device to deliver refraction measurements comparable to tabletop autorefractors under varied lighting condition⁵⁴. These tools are especially valuable in underserved regions, including rural, mountainous, high-altitude, and conflict-affected areas, where routine screening is logistically challenging.

In addition, EAI can generate automated referral recommendations in primary care and non-ophthalmic settings^{55,56}. For instance, the RobOCT-Net system integrates AI and robotics to detect referable posterior segment pathologies in emergency departments with 95% sensitivity and 76% specificity⁵⁶. Similarly, the IOMIDS platform combines an AI chatbot with multimodal learning to support self-diagnosis and triage based on user history and imaging data⁵⁷. These systems show promise for at-home monitoring and broader community-level screening.

Future EAI systems are expected to incorporate LLMs to facilitate natural, human-like interactions, improving user trust and accessibility^{58–64}. LLM-powered chatbots could allow patients to query symptoms before or after appointments, serving as cognitive engines to improve EAI's capabilities in environmental understanding, decision-making, and multimodal information integration^{60–63}.

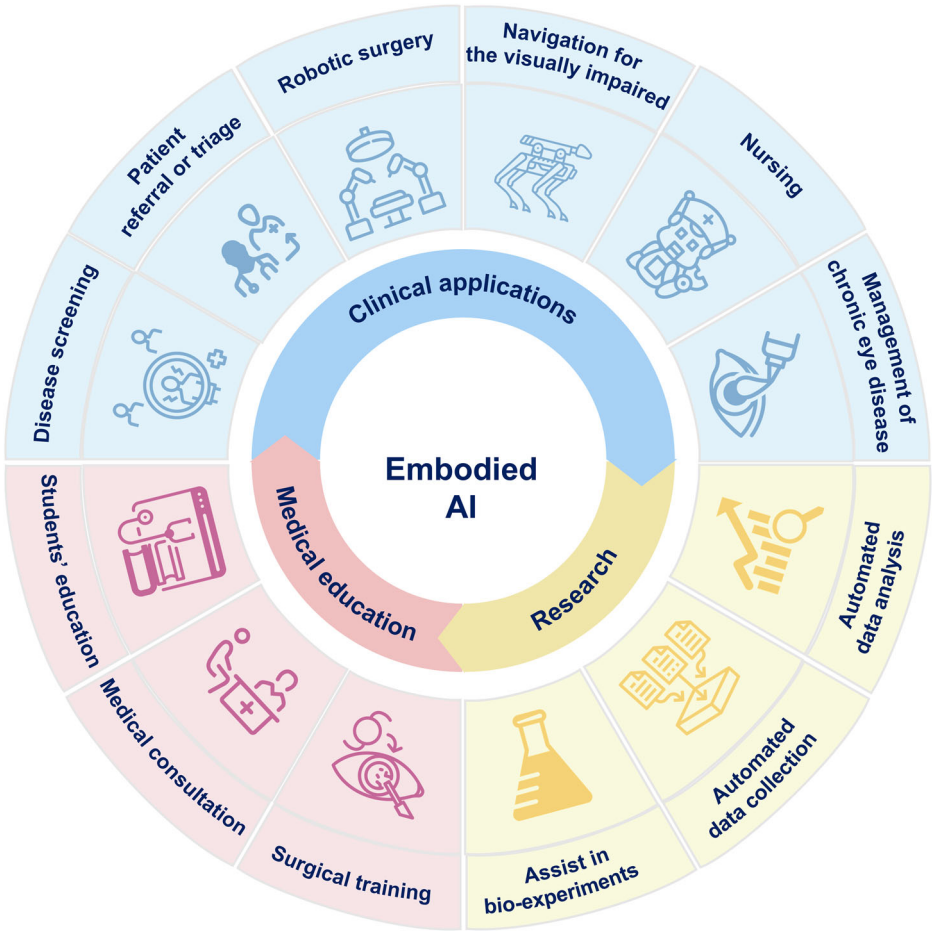
Assistant in ophthalmic surgery. Various surgical robots have been utilized to assist in ophthalmic procedures, including vitreoretinal surgery, cataract surgery, corneal transplantation, and strabismus surgery⁴⁵. Robotic assistance can improve surgical precision by mitigating the effects of physiological hand tremor⁶⁵. However, current surgical robots usually require operator control or supervision, and lack the ability to make autonomous decisions⁶⁶.

By integrating surgical robots with real-time intraoperative monitoring techniques, such as OCT, fundus photography, and surgical video, alongside AI technologies like multimodal large models, EAI can enhance the ability to perceive and interpret the surgical environment in real time. For instance,

Table 2 | Summary definitions of technical terms

Abbreviation	Full term	Definition
CNN	Convolutional neural network	A deep learning architecture specialized in processing grid-like data (e.g., images) via convolutional layers for feature extraction.
LLM	large language model	A type of AI model trained on vast text corpora to generate, comprehend, and manipulate human-like language.
VLM	Vision-language model	A multimodal AI system that jointly processes visual and textual data for tasks like image captioning or visual questions and answers.
RL	Reinforcement learning	A machine learning paradigm where agents learn optimal actions through trial-and-error interactions with an environment, guided by rewards.
CoT	Chain-of-thought	A reasoning framework that prompts LLMs to decompose problems into intermediate steps, enhancing logical inference.
GNN	Graph neural network	A neural network designed to operate on graph-structured data, capturing dependencies via message passing between nodes.
ISR-LLM	Iterative Self-Refined LLM	An LLM variant that iteratively refines its outputs through self-evaluation or external feedback loops.

Fig. 2 | Applications of embodied artificial intelligence (EAI) in ophthalmology. EAI-driven systems in ophthalmology encompass three domains: clinical application, medical education, research. The graphic was made with assets from Iconfinder.com.



Nespolo et al. developed a surgical navigation platform based on YOLACT ++, capable of detecting, classifying, and segmenting instruments and tissues, achieving high-speed processing (38.77–87.44 FPS) and high precision (AUPR up to 0.972)⁶⁷. Another study used a region-based CNN to analyze video frames in real time, locating the pupil, identifying the surgical phase, and providing real-time feedback during cataract surgery⁶⁸. This system achieved area under the receiver operating characteristic curve values of 0.996 for capsulorhexis, 0.972 for phacoemulsification, 0.997 for cortex removal, and 0.880 for idle phase recognition. Although these studies do not fully represent EAI systems, they provide crucial references for future EAI development. EAI can leverage similar CNN structures to enhance the environmental understanding and use real-time instance segmentation models for support surgical decision-making.

EAI has the potential to further improve the accuracy and safety of ophthalmic robotic surgery. Intelligent robotic systems can assist in performing delicate ophthalmic procedures like subretinal injections^{69–71}. Furthermore, AI-based video analysis systems can monitor the movement, position and depth of surgical instruments, minimizing accidental collisions with ocular tissues and improving surgical safety⁷². For example, Wu et al. introduced an autonomous robotic system for subretinal injection that incorporates intraocular OCT with deep learning-based motion prediction to enable precise needle control even under dynamic retinal conditions⁷³. In terms of autonomy, Kim et al. proposed an imitation learning-based navigation system that learned expert tool trajectories on the retinal surface, achieving sub-millimeter accuracy and maintained robustness under varying lighting conditions and instrument interference⁷⁴. More recently,

RL and imitation learning agents trained on intraoperative imaging data have enabled robots to autonomously perform tasks during the incision phase of cataract surgery. By incorporating surgeon's actions and preferences into the training process, these models enable the robotic system to adapt to individual surgical techniques and personalize procedural execution⁷⁵. These findings collectively represent important steps toward greater autonomy and deeper integration of EAI in ophthalmic surgery.

Embodied navigation for the visually impaired. It is predicted that 61 million people will be blind and 474 million will experience moderate or severe visual impairment by 2050⁴⁸. Navigation for the visually impaired has been a longstanding challenge. Assistive tools such as guide dogs, canes, smartphones, and wearable devices can help blind individuals avoid obstacles, recognize objects, and navigate both indoor and outdoor environments⁷⁶. However, the limited availability and lifespan of guide dogs are insufficient to meet global mobility needs. Traditional assistive devices also fall short, as they lack adaptability and the ability to learn from users or changing environments.

Various intelligent technologies can be integrated into EAI to enhance current navigation tools for the visually impaired. For instance, intelligent robots and special canes are anticipated to provide more effective navigation assistance^{77,78}. LLMs can support EAI in complex environmental perception and provide decision, improving environmental understanding, path planning, and dynamic adjustment²². EAI can also interact with the users through tactile feedback, pressure sensing, or voice recognition systems. For example, cane-based force perception systems enable bidirectional interaction, allowing robotic guides to lead users by moving forward or turning, while users can adjust the robot's walking speed with push or pull gestures. Compared to conventional assistive technologies, EAI-based solutions provide greater intelligence, personalization, and adaptability by continuously learning from user feedback and behavior. These systems can respond in real time to dynamic and unpredictable environments, enhancing navigation safety and improving overall user experience by minimizing the need for manual adjustments or external technical support.

Beyond navigation, EAI can assist visually impaired individuals with daily activities such as cooking, eating, and dressing, thereby enhancing their independence⁷⁹. It can also monitor the environment and track user movements to prevent falls⁸⁰. Additionally, EAI can support the management of chronic eye diseases by enabling home-based visual acuity monitoring and assisting with reminders and administration of eye drops. By integrating intelligent health chatbots⁸¹, EAI is expected to engage in more natural, human-like communication with visually impaired users, helping to alleviate their negative emotions and promote mental well-being.

EAI in medical education

For patients, intelligent agents have been employed in medical consultations and nursing education⁸². However, traditional internet-based AI interactions rely solely on text or voice, lacking the nuances of body language and emotional expression. EAI enables the development of intelligent agents within immersive 3D simulators, enhancing environmental perception and conveying emotions through facial expressions, voice modulations, and other channels. These embodied agents can provide more engaging and human-like interactive, potentially improving patients' participation in disease management. In addition to interactive support, EAI can assist patients in self-monitoring disease activity, providing valuable data for follow-up and clinical decision-making. For instance, Notal Vision Home OCT (NVHO) enables patients with neovascular age-related macular degeneration to perform daily retinal imaging at home. The acquired scans are analyzed by AI-based software to detect and quantify macular edema, achieving over 94% concordance with expert human grading⁸³. Additionally, a robot-mounted OCT scanner has been developed to image free-standing individuals from a safe distance without the need for operator intervention or head stabilization. Although current resolution remains limited, this approach shows potential for future at-home monitoring of eye diseases⁵².

Medical students can interact with EAI that simulates ophthalmic patients, improving their skills in history-taking, physical examination, and clinical communication while increasing their learning engagement⁸⁴. Besides, EAI can also be applied to ophthalmic surgical training. DeepSurgery, an AI-based video analysis platform⁸⁵, provides real-time supervision during cataract surgery and demonstrated expert-level evaluation performance (κ 0.58–0.77). When integrated with robotic systems, it can further assist novice ophthalmologists by guiding surgical steps and alerting them to incorrect operations, facilitating the acquisition of standardized and efficient cataract surgical techniques.

EAI in medical research

EAI holds potential to transform research in ophthalmology by enabling automation and large-scale data analysis. Traditional research processes are often slow and resource-intensive, requiring repetitive tasks such as dosage testing, patient monitoring, and manually data collection. EAI can improve efficiently by automating these tasks, delivering precise outcomes, and accelerating the development of new treatments and drugs. High-throughput screening powered by EAI allows simultaneous testing of multiple variables, further enhancing research productivity. Additionally, some biological experiments that require skilled operators can be performed remotely through robotic arms, expanding the capabilities of research laboratories⁸⁶.

Although no EAI systems specifically designed for ophthalmic drug development have been reported to date, recent advances in other fields have demonstrated transferable frameworks and inspiring potential. A-Lab, an autonomous platform for novel material synthesis, automates sample transfer with robotic arms, generates initial synthesis recipes using a natural language processing model, and employs active learning for a closed-loop optimization mechanism⁸⁷. Within 17 days, it achieved a 71% success rate in synthesizing 41 novel materials, significantly accelerating the materials discovery process. Similarly, in protein engineering and organic synthesis, the SAMPLE platform integrates intelligent agents, protein design, and fully automated robotic experimentation to enable autonomous protein engineering⁸⁸. It successfully accelerated the discovery of thermostable enzymes by iteratively refining its understanding of sequence-function relationships. These interdisciplinary explorations not only validate the effectiveness of EAI in automating scientific research but also offer valuable insights for potential applications in ophthalmology, particularly in drug screening, disease modeling, and therapeutic optimization.

Challenges and future directions

The development of EAI in ophthalmology faces several challenges. At the data and algorithm level, there are difficulties in acquiring large-scale, real-world interactive datasets and ensuring data representativeness. In terms of system capabilities, EAI struggles with understanding complex environments and integrating specialized ophthalmic medical knowledge. Furthermore, existing evaluation metrics are inadequate for assessing real-time, interactive AI systems. Lastly, ethical concerns, including privacy, informed consent, accountability, AI bias, and health equity, remain significant.

Training EAI algorithms typically requires the collection and processing of large amounts of interaction data from real-world environments, which is often costly and resource-intensive. Moreover, it remains challenging for collected datasets to comprehensively capture the complexity and variability of real-world scenarios. While database sharing may help alleviate the data shortage, it raises ethical concerns, particularly regarding privacy protection. Interaction with virtual simulation environments offers a promising approach to accelerate EAI development⁸⁹. However, it is essential to recognize that real-world environments are often more complex and unpredictable than their simulated counterparts.

Currently, EAI faces several critical challenges which limit its clinical utility. One major challenge is its insufficient ability to interpret complex and dynamic environments, which may lead to delays in decision-making and task execution⁹⁰. Furthermore, addressing these challenges in ophthalmology requires a profound understanding of specialized medical

knowledge⁶¹. Future research should prioritize the development of adaptive and scalable embodied agent architectures to improve generalization capabilities. In addition, expanding the medical knowledge base within EAI is essential to ensure its accuracy, safety, and reliability in ophthalmic practice.

Research on EAI in ophthalmology remains in its early stages. Most studies primarily evaluate the performance of AI-based medical devices in clinical applications with traditional performance metrics, such as accuracy and sensitivity. However, given that EAI systems involve real-time interaction with the environment, their assessment requires more comprehensive and multidimensional protocols. Future research should consider adapting existing evaluation benchmarks for ophthalmic LLMs and robotic systems while incorporating the unique characteristics of EAI^{91–93}. Expanding assessments beyond static data to include interactive scenarios will significantly facilitate the establishment of a standardized and holistic evaluation framework for ophthalmic EAI⁹⁴.

AI bias and fairness are critical concerns in the application of EAI in healthcare. A lack of diversity in training data, such as in gender, age, ethnicity, or geographic distribution, may lead to degraded performance in certain populations, potentially exacerbating existing health disparities^{95,96}. To promote fair and inclusive healthcare, it is imperative to adopt bias mitigation strategies and develop data governance frameworks that emphasize diversity and transparency. Furthermore, healthcare professionals must remain vigilant against automation bias and be prepared to critically evaluate and override AI-generated outputs when necessary.

Ethical and regulatory challenges in EAI are becoming increasingly prominent. Compared to traditional AI systems, EAI require more extensive evaluation and stricter regulation oversight. For example, the use of surgical robots in ophthalmic procedures introduces additional risks related to robot malfunctions⁹⁷. When EAI is involved in disease diagnosis, treatment decision-making, or medical education, it is crucial that its outputs remain objective, clinically relevant, and demonstrate humanistic care. To solve these issues, autonomous smart devices intended for clinical use should obtain appropriate regulatory approval, such as FDA or CE certification, with regulatory standards tailored to their classification^{98,99}. FDA employs the “Levels of Autonomous Surgical Robots” (LASR) classification system, which categorizes robotic autonomy from Level 1 (robot-assisted) to Level 5 (fully autonomous)¹⁰⁰. Computer scientists and healthcare professionals need to collaborate to refine classification criteria and ensure the safe and effective integration of EAI into clinical practice. In addition to technical oversight, robust governance models must be established to define accountability for clinical outcomes. In cases of errors or adverse events occur involving EAI systems, it remains unclear whether responsibility lies with developers, healthcare providers, or deploying institutions. Therefore, clear legal and ethical accountability frameworks are urgently needed. Furthermore, developers should incorporate explainability and human-in-the-loop mechanisms into EAI systems to enhance transparency, reliability and trust in clinical environments.

Lastly, privacy protection and informed consent present additional complexities in EAI systems^{101,102}. These systems often require real-time collection and processing of continuous multimodal data, including facial images, physiological parameters, and environmental information. However, the traditional model of “specific and explicit” consent is insufficient for the highly dynamic data environments associated with EAI systems. Therefore, a tiered consent framework should be established, distinguishing between physiologic data that requires strict authorization and environmental interaction metadata that can be broadly authorized. Moreover, adopting a dynamic consent process which allows patients to modify or withdraw their consent at any time would better respect patient autonomy¹⁰³. Ensuring that patients make informed decisions based on a clear understanding of the EAI system’s functions, related risks, and their rights is an important step toward achieving ethical and responsible AI development.

Conclusion

Unlike traditional Internet-based AI, EAI is characterized by real-time, context-specific interactions. EAI integrates multidisciplinary approaches

and leverage its physical embodiment to interact with the environment through real-time perception, reasoning, and action, enabling self-learning and autonomous decision-making. Although still in its early stages, EAI holds significant promise in ophthalmology, with potential applications spanning rapid multimodal screening, precise surgical navigation, assistance for visually impaired individuals, personalized medical education, and efficient medical research. However, the implementation of EAI still faces challenges related to data acquisition, system architecture, and clinical integration. Future developments should focus on enhancing adaptability and decision-making, while also ensuring safe and responsible deployment through regulatory and ethical frameworks.

Data Availability

No datasets were generated or analysed during the current study.

Received: 6 November 2024; Accepted: 26 May 2025;

Published online: 11 June 2025

References

1. Duan, J., Yu, S., Tan, H. L., Zhu, H. & Tan, C. A survey of embodied AI: from simulators to research tasks. *IEEE Trans. Emerg. Top. Comput. Intell.* **6**, 230–244 (2022).
2. Smith, L. & Gasser, M. The development of embodied cognition: six lessons from babies. *Artif. Life* **11**, 13–29 (2005).
3. Strathearn, C. & Ma, M. Modelling user preference for embodied artificial intelligence and appearance in realistic humanoid robots. *Informatics* **7**, 28 (2020).
4. Kumar, K. A., Rajan, J. F., Appala, C., Balurji, S. & Balaiahgari, P. R. Medibot: personal medical assistant. in *Proc. 2nd International Conference on Networking and Communications (ICNWC)* 1–6 (2024).
5. Thirunavukarasu, A. J. et al. Robot-assisted eye surgery: a systematic review of effectiveness, safety, and practicality in clinical settings. *Transl. Vis. Sci. Technol.* **13**, 20 (2024).
6. Vimala, S. et al. Telemedical robot using IoT with live supervision and emergency alert. in *Proc. 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)* 1327–1331 (IEEE, 2023).
7. Wang, W. et al. Neuromorphic sensorimotor loop embodied by monolithically integrated, low-voltage, soft e-skin. *Science* **380**, 735–742 (2023).
8. Liu, T. L. et al. Robot learning to play drums with an open-ended internal model. in *Proc. IEEE International Conference on Robotics and Biomimetics (ROBIO)* 305–311 (IEEE, 2018).
9. Zhuang, Z. Y., Yu, X., Mahony, R. & IEEE. LyRN (Lyapunov Reaching Network): a real-time closed loop approach from monocular vision. in *Proc. IEEE International Conference on Robotics and Automation (ICRA)* 8331–8337 (IEEE, 2020).
10. Zhao, Z. et al. Exploring embodied intelligence in soft robotics: a review. *Biomimetics* **9**, 248 (2024).
11. Liu, Y., Tan, Y. & Lan, H. Self-supervised contrastive learning for audio-visual action recognition. in *30th IEEE International Conference on Image Processing (ICIP)* 1000–1004 (IEEE, 2023).
12. Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N. & Folk, J. C. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *npj Digit. Med.* **1**, 39 (2018).
13. Ting, D. S. W. et al. Deep learning in ophthalmology: the technical and clinical considerations. *Prog. Retin. Eye Res.* **72**, 100759 (2019).
14. Shi, D. et al. Translation of color fundus photography into fluorescein angiography using deep learning for enhanced diabetic retinopathy screening. *Ophthalmol. Sci.* **3**, 100401 (2023).
15. Chen, R. et al. Translating color fundus photography to indocyanine green angiography using deep-learning for age-related macular degeneration screening. *npj Digit. Med.* **7**, 34 (2024).

16. Song, F., Zhang, W., Zheng, Y., Shi, D. & He, M. A deep learning model for generating fundus autofluorescence images from color fundus photography. *Adv. Ophthalmol. Pr. Res.* **3**, 192–198 (2023).
17. Zhou, Y. et al. A foundation model for generalizable disease detection from retinal images. *Nature* **622**, 156–163 (2023).
18. Shi, D. et al. EyeFound: a multimodal generalist foundation model for ophthalmic imaging. *arXiv preprint at* <https://doi.org/10.48550/arXiv.2405.11338> (2024).
19. Shi, D. et al. EyeCLIP: A visual-language foundation model for multi-modal ophthalmic image analysis. *arXiv preprint at* <https://doi.org/10.48550/arXiv.2409.06644> (2024).
20. Wang, T. et al. EmbodiedScan: a holistic multi-modal 3D perception suite towards embodied AI. in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 19757–19767 (IEEE, 2024).
21. Mieling, R. et al. Collaborative robotic biopsy with trajectory guidance and needle tip force feedback. in *Proc. IEEE International Conference on Robotics and Automation (ICRA)* 6893–6900 (IEEE, 2023).
22. Lin, J. et al. Advances in embodied navigation using large language models: a survey. *arXiv preprint at* <https://arxiv.org/abs/2311.00530> (2024).
23. Gao, S. et al. Empowering biomedical discovery with AI agents. *Cell* **187**, 6125–6151 (2024).
24. Liu, S. et al. Long short-term human motion prediction in human-robot co-carrying. in *Proc. International Conference on Advanced Robotics and Mechatronics (ICARM)* 815–820 (IEEE, 2023).
25. Wang, W. et al. Augmenting Language Models with Long-Term Memory. In *Advances in Neural Information Processing Systems* (eds Oh, A. et al.) **36**, 74530–74543 (Curran Associates, Inc., 2023).
26. Wang, J. et al. Large language models for robotics: Opportunities, challenges, and perspectives. *Journal of Automation and Intelligence* **4**, 52–64 (2025).
27. Wei, J. et al. Chain-of-thought prompting elicits reasoning in large language models. in *36th Conference on Neural Information Processing Systems (NeurIPS)* (eds. Koyejo, S. et al.) (Neural Information Processing Systems (NIPS), 2022).
28. Wang, X. et al. Self-Consistency Improves Chain of Thought Reasoning in Language Models. *The Eleventh International Conference on Learning Representations*. <https://openreview.net/forum?id=1PL1NIMMrw> (2023).
29. Wang, D. et al. Hierarchical graph neural networks for causal discovery and root cause localization. *arXiv preprint at* <https://arxiv.org/abs/2302.01987> (2023).
30. Mnih, V. et al. Playing Atari with deep reinforcement learning. *arXiv preprint at* <https://arxiv.org/abs/1312.5602> (2013).
31. Goma, A. & Mahdy, B. Unveiling the role of expert guidance: a comparative analysis of user-centered imitation learning and traditional reinforcement learning. *arXiv preprint at* <https://arxiv.org/abs/2410.21403> (2024).
32. Zhang, R. et al. A graph-based reinforcement learning-enabled approach for adaptive human-robot collaborative assembly operations. *J. Manuf. Syst.* **63**, 491–503 (2022).
33. Zhang, Y. et al. Towards efficient LLM grounding for embodied multi-agent collaboration. *arXiv preprint at* <https://arxiv.org/abs/2405.14314> (2024).
34. Wang, L., Fei, Y., Tang, H. & Yan, R. CLFR-M: Continual learning framework for robots via human feedback and dynamic memory. in *Proc. IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE International Conference on Robotics, Automation and Mechatronics (RAM)* 216–221 (IEEE, 2024).
35. Deng, H., Zhang, H., Ou, J. & Feng, C. Can LLM be a good path planner based on prompt engineering? Mitigating the hallucination for path planning. *arXiv preprint at* <https://arxiv.org/abs/2408.13184> (2024).
36. Chen, L. et al. Towards end-to-end embodied decision making via multi-modal large language model: explorations with GPT4-vision and beyond. *NeurIPS 2023 Foundation Models for Decision Making Workshop*. <https://openreview.net/forum?id=mgOtn5p7t> (2023).
37. Singh, I. et al. ProgPrompt: generating situated robot task plans using large language models. in *Proc. IEEE International Conference on Robotics and Automation (ICRA)* 11523–11530 (IEEE, 2023).
38. Shin, S., Jeon, S., Kim, J., Kang, G.-C. & Zhang, B.-T. Socratic planner: inquiry-based zero-shot planning for embodied instruction following. *arXiv preprint at* <https://arxiv.org/abs/2404.15190> (2024).
39. Zhou, Z., Song, J., Yao, K., Shu, Z. & Ma, L. ISR-LLM: iterative self-refined large language model for long-horizon sequential task planning. in *Proc. IEEE International Conference on Robotics and Automation (ICRA)* 2081–2088 (IEEE, 2024).
40. Yihao, L. et al. From screens to scenes: a survey of embodied AI in healthcare. *Inf. Fusion* **119**, 103033 (2025).
41. Huang, P. I. Y. Enhancement of robot position control for dual-user operation of remote robot system with force. *Feedback* **14**, 9376 (2024).
42. Ding, P. et al. QUAR-VLA: Vision-Language-Action Model for Quadruped Robots. In *Computer Vision – ECCV 2024* (eds Leonardis, A. et al.) Vol. 15063, 352–367 (Springer Nature Switzerland, Cham, 2025).
43. Mu, Y. et al. EmbodiedGPT: Vision-Language Pre-Training via Embodied Chain of Thought. In *Advances in Neural Information Processing Systems* (eds Oh, A. et al.) Vol. 36, 25081–25094 (Curran Associates, Inc., 2023).
44. Song, C. H. et al. LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models. in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* 2986–2997 (IEEE, 2023).
45. Alafaleq, M. Robotics and cybersurgery in ophthalmology: a current perspective. *J. Robot. Surg.* **17**, 1159–1170 (2023).
46. Nielsen, K. B., Lautrup, M. L., Andersen, J. K., Savarimuthu, T. R. & Grauslund, J. Deep learning-based algorithms in screening of diabetic retinopathy: a systematic review of diagnostic performance. *Ophthalmol. Retin.* **3**, 294–304 (2019).
47. Zhu, Y. et al. Advancing glaucoma care: integrating artificial intelligence in diagnosis, management, and progression detection. *Bioengineering* **11**, 122 (2024).
48. GBD 2019 Blindness and Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *Lancet Glob. Health* **9**, e130–e143 (2021).
49. Vujosevic, S., Limoli, C. & Nucci, P. Novel artificial intelligence for diabetic retinopathy and diabetic macular edema: what is new in 2024?. *Curr. Opin. Ophthalmol.* **35**, 472–479 (2024).
50. Liu, H. et al. Economic evaluation of combined population-based screening for multiple blindness-causing eye diseases in China: a cost-effectiveness analysis. *Lancet Glob. Health* **11**, e456–e465 (2023).
51. Kang, E. Y.-C. et al. A multimodal imaging-based deep learning model for detecting treatment-requiring retinal vascular diseases: model development and validation study. *JMIR Med. Inform.* **9**, e28868 (2021).
52. Draelos, M. et al. Contactless optical coherence tomography of the eyes of freestanding individuals with a robotic scanner. *Nat. Biomed. Eng.* **5**, 726–736 (2021).
53. He, S. et al. Bridging the camera domain gap with image-to-image translation improves glaucoma diagnosis. *Transl. Vis. Sci. Technol.* **12**, 20–20 (2023).
54. Zhen, Y., Yan, H., Qilin, S., Hong, C. & Wei, T. Artificial intelligence-enabled low-cost photorefractive for accurate refractive error measurement under complex ambient lighting conditions: a model

- development and validation study. Available at SSRN 5064133. (2024).
55. Vought, R., Vought, V., Szirth, B. & Khouri, A. S. Future direction for the deployment of deep learning artificial intelligence: Vision threatening disease detection in underserved communities during COVID-19. *Saudi J. Ophthalmol.* **37**, 193–199 (2023).
 56. Song, A. et al. RobOCTNet: robotics and deep learning for referable posterior segment pathology detection in an emergency department population. *Transl. Vis. Sci. Technol.* **13**, 12 (2024).
 57. Ma, R. et al. Multimodal machine learning enables AI chatbot to diagnose ophthalmic diseases and provide high-quality medical responses. *npj Digit. Med.* **8**, 1–18 (2025).
 58. Yang, Z. et al. Understanding natural language: potential application of large language models to ophthalmology. *Asia Pac. J. Ophthalmol.* **13**, 100085 (2024).
 59. Chotcomwongse, P., Ruamviboonsuk, P. & Grzybowski, A. Utilizing large language models in ophthalmology: the current landscape and challenges. *Ophthalmol. Ther.* **13**, 2543–2558 (2024).
 60. Chen, X. et al. FFA-GPT: an automated pipeline for fundus fluorescein angiography interpretation and question-answer. *npj Digit. Med.* **7**, 111 (2024).
 61. Chen, X. et al. EyeGPT for Patient Inquiries and Medical Education: Development and Validation of an Ophthalmology Large Language Model. *Journal of Medical Internet Research* **26**, e60063 (2024).
 62. Chen, X. et al. ICGA-GPT: report generation and question answering for indocyanine green angiography images. *Br. J. Ophthalmol.* **108**, 1450–1456 (2024).
 63. Chen, X. et al. ChatFFA: an ophthalmic chat system for unified vision-language understanding and question answering for fundus fluorescein angiography. *iScience* **27**, 110021 (2024).
 64. Jin, K., Yuan, L., Wu, H., Grzybowski, A. & Ye, J. Exploring large language model for next generation of artificial intelligence in ophthalmology. *Front. Med.* **10**, 1291404 (2023).
 65. Roizenblatt, M., Grupenmacher, A. T., Belfort Junior, R., Maia, M. & Gehlbach, P. L. Robot-assisted tremor control for performance enhancement of retinal microsurgeons. *Br. J. Ophthalmol.* **103**, 1195–1200 (2019).
 66. Gerber, M. J., Pettenkofer, M. & Hubschman, J. P. Advanced robotic surgical systems in ophthalmology. *Eye* **34**, 1554–1562 (2020).
 67. Nespolo, R. G. et al. Feature Tracking and segmentation in real time via deep learning in vitreoretinal surgery: a platform for artificial intelligence-mediated surgical guidance. *Ophthalmol. Retin.* **7**, 236–242 (2023).
 68. Garcia Nespolo, R. et al. Evaluation of artificial intelligence-based intraoperative guidance tools for phacoemulsification cataract surgery. *JAMA Ophthalmol.* **140**, 170–177 (2022).
 69. Zhou, M. et al. Needle detection and localisation for robot-assisted subretinal injection using deep learning. *CAAI Trans. Intell. Technol.* 1–13 (2023).
 70. Huang, Y., Asaria, R., Stoyanov, D., Sarunic, M. & Bano, S. PseudoSegRT: efficient pseudo-labelling for intraoperative OCT segmentation. *Int. J. Comput. Assist. Radio. Surg.* **18**, 1245–1252 (2023).
 71. Ladha, R., Meenink, T., Smit, J. & de Smet, M. D. Advantages of robotic assistance over a manual approach in simulated subretinal injections and its relevance for gene therapy. *Gene Ther.* **30**, 264–270 (2023).
 72. Baldi, P. F. et al. Vitreoretinal surgical instrument tracking in three dimensions using deep learning. *Transl. Vis. Sci. Technol.* **12**, 20 (2023).
 73. Wu, T. et al. Deep learning-enhanced robotic subretinal injection with real-time retinal motion compensation. *arXiv preprint at: <https://arxiv.org/abs/2504.03939>* (2025).
 74. Kim, J. W. et al. Autonomously navigating a surgical tool inside the eye by learning from demonstration. in *Proc. IEEE International Conference on Robotics and Automation (ICRA)* 7351–7357 (IEEE, 2020).
 75. Gomaa, A., Mahdy, B., Kleer, N. & Krüger, A. Towards a surgeon-in-the-loop ophthalmic robotic apprentice using reinforcement and imitation learning. in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 6939–6946 (IEEE, 2024).
 76. Messaoudi, M. D., Menelas, B. J. & McHeick, H. Review of navigation assistive tools and technologies for the visually impaired. *Sensors* **22**, 7888 (2022).
 77. Tang, T. et al. Special cane with visual odometry for real-time indoor navigation of blind people. in *IEEE International Conference on Visual Communications and Image Processing (VCIP)* 255–255 (IEEE, 2020).
 78. Zhang, Y. et al. Visual Navigation of Mobile Robots in Complex Environments Based on Distributed Deep Reinforcement Learning. in *2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT)* 1–5 (IEEE, 2022).
 79. Guo, C. & Li, H. Application of 5G network combined with AI robots in personalized nursing in China: a literature review. *Front. Public Health* **10**, 948303 (2022).
 80. Juang, L. H. & Wu, M. N. Fall Down Detection Under Smart Home System. *J. Med. Syst.* **39**, 107 (2015).
 81. Chen, X. et al. Visual Question Answering in Ophthalmology: a progressive and practical perspective. *arXiv preprint at: <https://arxiv.org/abs/2410.16662>* (2024).
 82. Tam, W. et al. Nursing education in the age of artificial intelligence powered Chatbots (AI-Chatbots): Are we ready yet?. *Nurse Educ. Today* **129**, 105917 (2023).
 83. Liu, Y., Holekamp, N. M. & Heier, J. S. Prospective, longitudinal study: daily self-imaging with home OCT for neovascular age-related macular degeneration. *Ophthalmol. Retin.* **6**, 575–585 (2022).
 84. Chen, J., Zhan, X., Wang, Y. & Huang, X. Medical robots based on artificial intelligence in the medical education. in *Proc. 2nd International Conference on Artificial Intelligence and Education (ICAIE)* 1–4 (IEEE, 2021).
 85. Wang, T. et al. Intelligent cataract surgery supervision and evaluation via deep learning. *Int. J. Surg.* **104**, 106740 (2022).
 86. Hamm, J. et al. A Modular robotic platform for biological research: cell culture automation and remote experimentation. *Adv. Intell. Syst.* **6**, 2300566 (2024).
 87. Szymanski, N. J. et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **624**, 86–91 (2023).
 88. Rapp, J. T., Bremer, B. J. & Romero, P. A. Self-driving laboratories to autonomously navigate the protein fitness landscape. *Nat. Chem. Eng.* **1**, 97–107 (2024).
 89. Tan, T. F. et al. Metaverse and virtual health care in ophthalmology: opportunities and challenges. *Asia Pac. J. Ophthalmol.* **11**, 237–246 (2022).
 90. Kang, D., Nam, C. & Kwak, S. S. Robot feedback design for response delay. *Int. J. Soc. Robot.* **16**, 341–361 (2023).
 91. Chen, X. et al. Evaluating large language models and agents in healthcare: key challenges in clinical applications. *Intelligent Medicine* **5**, 151–163 (2025).
 92. Xu, P., Chen, X., Zhao, Z. & Shi, D. Unveiling the clinical incapacities: a benchmarking study of GPT-4V(ision) for ophthalmic multimodal image analysis. *Br. J. Ophthalmol.* **108**, 1384–1389 (2024).
 93. Majumdar, A. et al. Openega: Embodied question answering in the era of foundation models. in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition* 16488–16498 (IEEE, 2024).
 94. Cheng, Z. et al. EmbodiedEval: evaluate multimodal LLMs as embodied agents. *arXiv preprint at: <https://arxiv.org/abs/2501.11858>* (2025).

95. Mahamadou, A. J. D. & Trotsyuk, A. A. Revisiting technical bias mitigation strategies. *Annu. Rev. Biomed. Data Sci.* **8**, (2025).
96. Hofmann, V., Kalluri, P. R., Jurafsky, D. & King, S. AI generates covertly racist decisions about people based on their dialect. *Nature* **633**, 147–154 (2024).
97. Di Paolo, M., Boggi, U. & Turillazzi, E. Bioethical approach to robot-assisted surgery. *Br. J. Surg.* **106**, 1271–1272 (2019).
98. O'Sullivan, S. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int. J. Med. Robot. Comput. Assist. Surg* **15**, e1968 (2019).
99. Biswas, P., Sikander, S. & Kulkarni, P. Recent advances in robot-assisted surgical systems. *Biomed. Eng. Adv.* **6**, 100109 (2023).
100. Lee, A., Baker, T. S., Bederson, J. B. & Rapoport, B. I. Levels of autonomy in FDA-cleared surgical robots: a systematic review. *npj Digit. Med.* **7**, 103 (2024).
101. Fiske, A., Henningsen, P. & Buyx, A. Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *J. Med. Internet Res.* **21**, e13216 (2019).
102. Vats, T. et al. Navigating the landscape: Safeguarding privacy and security in the era of ambient intelligence within healthcare settings. *Cyber Security Appl.* **2**, 100046 (2024).
103. Tamuhla, T., Tiffin, N. & Allie, T. An e-consent framework for tiered informed consent for human genomic research in the global south, implemented as a REDCap template. *BMC Med. Ethics* **23**, 119 (2022).

Acknowledgements

The study was supported by the Start-up Fund for RAPs under the Strategic Hiring Scheme (P0048623) from the Hong Kong Polytechnic University, the Global STEM Professorship Scheme (P0046113) from HKSAR and Henry G. Leong Endowed Professorship in Elderly Vision Health. We thank the InnoHK HKSAR Government for providing valuable supports. The research work described in this paper was conducted in the JC STEM Lab of Innovative Light Therapy for Eye Diseases funded by The Hong Kong Jockey Club Charities Trust.

Author contributions

Y.Q., X.C. and X.W. contributed to the study design, data collection and analysis, and the initial and final drafts of the manuscript. Y.L., P.X., K.J., X.S. and P.C. contributed to the management and coordination of the project and helped review the manuscript. D.S. and M.H. contributed to study design, supervision, manuscript review & editing. All the authors performed a critical review and approved the final manuscript and important intellectual input.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Mingguang He or Danli Shi.

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

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

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

© The Author(s) 2025