# **The Convergence of AI and Ophthalmic Imaging: An Expert Report on Deep Learning and Large Language Models in Optical Coherence Tomography**

## **Part I: Foundational Technologies in Ophthalmic AI**

### **Section 1.1: Artificial Intelligence in Medical Imaging: A Paradigm Shift in Diagnostics**

The integration of Artificial Intelligence (AI) into medicine represents one of the most significant technological shifts in modern healthcare, fundamentally altering the processes of diagnosis, treatment planning, and patient monitoring. At its core, AI in healthcare refers to the application of complex algorithms and software designed to emulate human cognition in the analysis and interpretation of complex medical data.1 This technology has demonstrated a profound capacity to process vast datasets with a speed and precision that can exceed human capabilities, identifying subtle patterns and making predictions that enhance clinical decision-making across numerous specialties.1

The application of AI is not a monolithic concept but rather a spectrum of technologies. It encompasses the broad field of Machine Learning (ML), where algorithms learn from data without being explicitly programmed, and more specifically, Deep Learning (DL), a sophisticated subset of ML that utilizes multi-layered artificial neural networks. It is DL that has been the primary catalyst for the recent revolution in medical image analysis.2 These advanced models are being applied across a wide array of imaging modalities, providing a new layer of insight for clinicians. For instance, AI algorithms are used to analyze X-rays for bone fracture detection, computed tomography (CT) scans for the identification of pulmonary nodules, and magnetic resonance imaging (MRI) for the segmentation of brain tumors.1 The success of AI in these areas is indicative of a broader transformation, of which ophthalmology is a key beneficiary.

The engine driving this transformation in image analysis is a class of DL models known as Convolutional Neural Networks (CNNs). CNNs are uniquely suited for visual data because they automatically and adaptively learn hierarchical features directly from images. The initial layers of a CNN might learn to recognize simple features like edges and textures, while deeper layers combine these to identify more complex structures and, ultimately, pathological findings.7 This process of automatic feature extraction obviates the need for manual feature engineering, a laborious and often subjective process that characterized earlier computer-aided diagnostic systems. Many advanced CNNs employ an encoder-decoder architecture, where an encoder network progressively reduces the spatial dimensions of an image to capture abstract, semantic information, and a decoder network gradually restores the spatial dimensions to produce a detailed, pixel-level output, such as a segmentation map.10

The clinical and operational benefits derived from the integration of AI into medical imaging are multifaceted and substantial. A primary advantage is the enhancement of diagnostic accuracy. By analyzing imaging data with high precision, AI algorithms can identify minute abnormalities that may be overlooked by the human eye, thereby reducing the risk of false negatives and facilitating the early detection of diseases like cancer and cardiovascular conditions.1 This leads to more timely interventions and improved patient outcomes. Concurrently, AI improves operational efficiency and productivity within clinical workflows. The automation of routine and time-consuming tasks, such as the segmentation of anatomical structures or the measurement of lesion volumes, liberates clinicians from tedious work, allowing them to dedicate more time to complex cases, interdisciplinary collaboration, and direct patient care.11

Furthermore, AI introduces a level of standardization and consistency that is difficult to achieve with human interpretation alone. By providing objective and reproducible analysis, AI reduces the inter-reader and intra-reader variability that is inherent among radiologists and ophthalmologists, ensuring more reliable and uniform interpretations across different cases and time points.1 Finally, AI serves as a powerful integration tool, capable of synthesizing imaging data with other sources of patient information, such as electronic health records (EHRs) and genetic data. This creates a holistic, comprehensive patient profile that supports a more personalized approach to medicine.1

The adoption of AI in medical imaging constitutes more than just an incremental improvement in existing processes; it represents a fundamental paradigm shift. Traditional image interpretation has long been a qualitative practice, heavily reliant on the subjective experience and trained eye of a clinician. This approach, while highly skilled, is susceptible to variability and has inherent limits in detecting and quantifying subtle changes. AI, and particularly DL, introduces the ability to perform precise, automated quantification of pathological and anatomical features. This transforms a descriptive clinical finding, such as "intraretinal fluid is present," into a quantitative data point, such as "0.898 mm3 of intraretinal fluid is present".13 This move from a qualitative to a quantitative framework enables more objective disease monitoring, the establishment of standardized endpoints for clinical trials, and the discovery of novel quantitative biomarkers that were previously inaccessible. In this context, AI is not merely a "second reader" or a diagnostic aid; it is an enabling technology that fundamentally changes the nature of the data extracted from medical images, propelling the field toward a new era of precision and objectivity.

### **Section 1.2: Optical Coherence Tomography (OCT): Principles and Clinical Utility**

Optical Coherence Tomography (OCT) is a non-invasive medical imaging modality that has revolutionized the field of ophthalmology, and is increasingly finding applications in other medical specialties like cardiology and neurology.14 The technology functions as a form of "optical ultrasound," utilizing near-infrared light waves to capture micrometer-resolution, cross-sectional images of biological tissue in real time.14 By measuring the echo time delay and intensity of light reflected and backscattered from different tissue layers, a process known as low-coherence interferometry, OCT constructs detailed two- and three-dimensional maps of tissue microstructure.16 Its ability to visualize the transparent structures of the eye, particularly the retina, with such high fidelity has established it as an indispensable standard of care for the diagnosis and management of a vast array of ocular diseases.14

The clinical utility of OCT has grown in lockstep with its technological evolution. The technology has progressed through several distinct generations, each offering significant improvements in speed, resolution, and imaging depth, which in turn have direct implications for the quality of data available for AI analysis.

* **Time-Domain OCT (TD-OCT):** As the first-generation technology, TD-OCT utilized a moving reference mirror to acquire data, resulting in relatively slow acquisition speeds of approximately 400 A-scans (axial scans) per second and an axial resolution of 10-15 microns.16 Imaging was typically limited to a series of radial B-scans (cross-sectional images), creating a risk of missing pathology located between the scan lines.17
* **Spectral-Domain OCT (SD-OCT):** This generation represented a major technological leap, replacing the moving mirror with a fixed mirror and a high-speed spectrometer. This innovation increased scan speeds by orders of magnitude to between 27,000 and 70,000 A-scans per second, with axial resolution improving to as fine as 3-7 microns.17 The speed of SD-OCT enabled the acquisition of dense, volumetric datasets, significantly reducing motion artifacts and improving the ability to comprehensively map retinal structures. Further enhancements, such as Enhanced Depth Imaging (EDI), were developed to improve the visualization of deeper structures like the choroid by adjusting the focal point of the scan.16
* **Swept-Source OCT (SS-OCT):** The current state-of-the-art, SS-OCT employs a rapidly tunable, wavelength-sweeping laser source, pushing acquisition speeds to 100,000-400,000 A-scans per second.17 SS-OCT typically uses a longer wavelength of light (e.g., 1050 nm) compared to SD-OCT (~840 nm), which allows for deeper penetration into tissues. This provides superior image quality of deep structures, such as the choroid and sclera, without the need for specialized techniques like EDI.16

This technological progression has profound implications for the development and application of AI models, as summarized in Table 1. The higher speed and resolution of SD-OCT and SS-OCT provide cleaner, more detailed, and artifact-free data, which is essential for training robust AI algorithms and for the sensitive task of tracking disease progression over time.

| **Table 1: Comparison of OCT Modalities and Implications for AI** |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Modality** | **Key Technology** | **Scan Speed (A-scans/sec)** | **Axial Resolution (microns)** | **Key Advantages** | **Limitations** | **Prime Suitability for AI Applications** |
| **Time-Domain (TD-OCT)** | Moving reference mirror, single detector | ~400 | 10–15 | First-generation, established baseline | Slow, prone to motion artifacts, sparse sampling (risk of missing lesions) | Limited; suitable for basic classification tasks on legacy datasets. |
| **Spectral-Domain (SD-OCT)** | Fixed mirror, spectrometer (detector array) | 27,000–70,000 | 3–7 | Fast, high-resolution, volumetric scanning, reduced artifacts | Limited penetration depth without EDI | Excellent for high-resolution retinal layer segmentation, fluid quantification, and training robust classification models. |
| **Swept-Source (SS-OCT)** | Wavelength-sweeping laser, single photodetector | 100,000–400,000 | ~5 | Very fast, deep tissue penetration, wide-field imaging | Higher cost, slightly lower axial resolution than some SD-OCTs | Ideal for AI models analyzing deep structures (choroid, lamina cribrosa), wide-field analysis, and applications requiring minimal artifacts (e.g., OCTA). |

The clinical applications of OCT are extensive. It is the preferred imaging study for diagnosing and monitoring a wide range of posterior segment diseases affecting the macula and optic nerve.14 This includes conditions such as age-related macular degeneration (AMD), diabetic retinopathy and macular edema, macular holes, and epiretinal membranes.14 In glaucoma management, OCT is essential for quantitatively tracking the thickness of the retinal nerve fiber layer (RNFL) and ganglion cell complex, providing objective evidence of structural damage.16 A key innovation built upon OCT technology is

**OCT Angiography (OCTA)**, a non-invasive technique that visualizes retinal and choroidal vasculature by detecting the motion of red blood cells. OCTA provides a three-dimensional view of blood flow without the need for dye injections, making it a valuable tool for assessing vascular diseases.16 Beyond the posterior segment, OCT is also used to image anterior structures like the cornea, iris, and anterior chamber angle to aid in surgical planning and the diagnosis of corneal diseases and narrow-angle glaucoma.19 The output of an OCT system provides a rich source of data for AI analysis, including B-scans, which show the cross-sectional retinal layers, and quantitative thickness maps, often displayed on standardized Early Treatment Diabetic Retinopathy Study (ETDRS) grids.17

### **Section 1.3: Large Language Models (LLMs): From Text to Multimodal Understanding**

Large Language Models (LLMs) represent a transformative breakthrough in the field of artificial intelligence, moving beyond narrow, task-specific applications toward more generalist systems with a broad range of capabilities. LLMs are a category of foundation models, which are large-scale models trained on immense quantities of data that can be adapted to a wide variety of downstream tasks.20 The architectural backbone of virtually all modern LLMs is the

**transformer**, introduced in 2017.20 This architecture marked a departure from previous sequential models like Recurrent Neural Networks (RNNs). The transformer's key innovation is the

**self-attention mechanism**, which allows the model to dynamically weigh the importance of different words within an input text sequence, regardless of their position. This enables a far more sophisticated understanding of long-range dependencies and context than was previously possible.20 The transformer architecture, composed of encoder and decoder stacks, also allows for parallel processing of data, which dramatically accelerates the training of these massive models on GPU hardware.23

The training paradigm for LLMs is a multi-stage process. It begins with **self-supervised pre-training** on vast, unlabeled text corpora scraped from the internet, such as the Common Crawl dataset (comprising billions of web pages) and Wikipedia.23 During this phase, the model learns to predict the next word in a sentence or fill in masked-out words. Through this simple objective, performed at an immense scale, the model internalizes the statistical patterns of language, acquiring a deep understanding of grammar, semantics, facts, and reasoning abilities.20 After pre-training, these generalist models are adapted for specific tasks or domains through a process called

**fine-tuning**, where the model's parameters are further adjusted on a smaller, curated dataset.23 This allows for techniques like

**zero-shot learning**, where the model can perform a task it has never been explicitly trained on, and **few-shot learning**, where providing just a few examples dramatically improves its performance on a specific task.23

The capabilities of trained LLMs are remarkably broad and flexible. A single, large-scale model, such as OpenAI's GPT series, Google's Gemini, or Anthropic's Claude, can perform an extensive range of natural language processing (NLP) tasks that once required separate, specialized models.21 These capabilities include:

* **Text Generation:** Creating original, coherent, and contextually appropriate content, from drafting emails and marketing copy to writing poetry or computer code.20
* **Conversational AI and Question Answering:** Engaging in fluent, human-like dialogue and providing answers to complex, knowledge-intensive questions.23
* **Summarization and Information Extraction:** Condensing long documents into concise summaries and extracting structured information from unstructured text.20
* **Text Classification:** Analyzing text to determine attributes like sentiment, topic, or intent.23
* **Translation:** Translating with high fidelity not only between human languages but also between programming languages.20

While these text-based capabilities are impressive, the most critical evolution for the medical field is the development of **Large Multimodal Models (LMMs)**, also known as Vision-Language Models (VLMs). These models extend the transformer architecture to process and integrate information from multiple modalities simultaneously, most notably text and images.26 By representing both images (as a sequence of patches) and words as "tokens" within a shared mathematical space, LMMs can learn the relationships between visual concepts and their textual descriptions.27

This multimodal capability signals a profound shift in the potential of AI in medicine. While earlier AI systems were highly specialized—a CNN for image classification, a separate NLP model for report analysis—they operated in informational silos. The true disruptive power of LMMs lies not just in their mastery of language, but in their capacity to perform **reasoning and synthesis across these different data modalities**. An LMM can concurrently "see" a pathological finding in an OCT scan, "read" the patient's clinical history and prior reports in the EHR, and "write" a diagnostic summary that cohesively integrates and contextualizes all of this information.29 This is a higher-order cognitive task that mirrors the process of a human clinician, who synthesizes diverse data points to form a diagnosis. This evolution transforms medical AI from a specialized

*pattern recognizer* into a generalist *clinical data synthesizer and reasoner*, a development with far more profound implications for augmenting the entirety of the clinical decision-making process.

## **Part II: Deep Learning for Automated Analysis of OCT Images**

### **Section 2.1: Image Segmentation: Delineating Retinal Anatomy and Pathology**

Medical image segmentation is the process of partitioning a digital image into multiple distinct regions or segments. This task is a cornerstone of quantitative medical image analysis, as it enables the precise identification and delineation of anatomical structures, organs, tissues, and pathological lesions.7 In the context of ophthalmology, and specifically for OCT imaging, segmentation is the critical first step for objective measurement and monitoring. It allows for the quantification of retinal layer thickness, a key indicator of glaucomatous damage, and the measurement of pathological biomarkers, such as fluid volumes, which are essential for diagnosing and managing diseases like AMD and DME.

Deep learning has revolutionized medical image segmentation, largely supplanting traditional methods with automated, highly accurate algorithms. The development of **Fully Convolutional Networks (FCNs)** was a seminal moment, as these models were the first to perform dense, pixel-wise classification, producing a segmentation map for an entire input image at once rather than classifying individual pixels or patches.32 This laid the groundwork for more advanced architectures.

Among these, the **U-Net** architecture has emerged as the de facto standard and most widely adopted model for medical image segmentation.7 The U-Net is characterized by its elegant U-shaped encoder-decoder structure. The encoder, or contracting path, consists of a series of convolutional and pooling layers that progressively downsample the image to capture high-level, semantic feature information (the "what"). The decoder, or expansive path, uses upsampling convolutions to gradually increase the resolution back to the original image size, creating a detailed segmentation map. The defining feature of the U-Net is its use of "skip connections," which concatenate feature maps from the encoder path directly to the corresponding layers in the decoder path.32 These connections allow the network to combine deep, contextual information from the encoder with precise, high-resolution localization information, which is crucial for accurately delineating the fine boundaries of medical structures.32 More recent developments have seen the integration of transformer-based models, such as

**TransUNet**, which leverage the self-attention mechanism to better capture global context within the image, potentially improving segmentation performance further.31

In OCT image analysis, these segmentation models are applied to two primary tasks:

1. **Retinal Layer Segmentation:** The automatic and precise delineation of the boundaries between the retina's distinct layers is critical for diagnosing and monitoring numerous ophthalmic conditions. For glaucoma, the accurate measurement of the thickness of the Retinal Nerve Fiber Layer (RNFL) and the Ganglion Cell-Inner Plexiform Layer (GC-IPL) is the primary method for quantifying structural damage and tracking disease progression.33 AI models have demonstrated performance on par with human experts in this task. For instance, a study using FCNs to segment retinal layers in patients with diabetic retinopathy achieved a mean unsigned error of just 1.06 pixels, which was comparable to the inter-grader variability of 1.10 pixels, indicating human-level precision.35
2. **Pathological Biomarker Segmentation:** In exudative retinal diseases, the quantification of pathological fluid is a key biomarker for disease activity and treatment response. AI segmentation models are used to automatically identify and measure the volume of **intraretinal fluid (IRF)**, **subretinal fluid (SRF)**, and **pigment epithelial detachments (PED)**.13 This automated quantification provides objective, reproducible endpoints for clinical trials and allows clinicians to precisely monitor a patient's response to therapies like anti-VEGF injections. Beyond fluid, AI can also segment and count other important biomarkers, such as  
   **hyperreflective foci (HRF)**, which are thought to be markers of inflammation.13 This level of detailed, automated analysis supports a more data-driven approach to managing retinal diseases.

### **Section 2.2: Disease Classification and Diagnosis**

Beyond segmentation, the most common application of AI in OCT analysis is for disease classification and diagnosis. This process typically involves training a CNN on a large dataset of OCT scans that have been expertly labeled with their corresponding diagnoses (e.g., Normal, AMD, DME, Drusen, CNV).33 During training, the network learns to identify the subtle textural and morphological patterns characteristic of each disease. Once trained, the model can be presented with a new, unseen OCT scan and will output a probability score for each of the diagnostic classes it has learned, effectively acting as a powerful decision support tool for clinicians.1 AI models have demonstrated high levels of diagnostic accuracy across the most prevalent retinal diseases diagnosed with OCT.

#### **2.2.1 Age-Related Macular Degeneration (AMD)**

AMD is a leading cause of irreversible blindness in the elderly, and OCT is the gold standard for its diagnosis and management.39 AI models are trained to detect and classify the key pathological biomarkers of AMD visible on OCT scans. These include

**drusen** (extracellular deposits under the retina), **geographic atrophy (GA)** (loss of retinal cells), **subretinal hyperreflective material (SHRM)**, and signs of neovascular AMD (nAMD) such as **choroidal neovascularization (CNV)** and associated fluid.38 AI-powered algorithms have shown they can reliably quantify these biomarkers, which is critical for predicting disease progression and guiding treatment decisions.39

The application of AI in AMD clinical trials is particularly impactful. By providing automated, objective, and reproducible quantification of endpoints like GA lesion size or fluid volume in nAMD, AI reduces the variability associated with human grading and can enhance the statistical power of trials.40 A systematic review and meta-analysis of AI algorithms for detecting exudative AMD found that models using OCT images achieved a very high summary specificity of 0.97, indicating a low rate of false positives, which is crucial for a screening tool.41 This high performance underscores the potential of AI to serve as a reliable tool for both clinical diagnosis and as a robust methodology for evaluating new therapies in research settings.

#### **2.2.2 Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME)**

Diabetic retinopathy is a major microvascular complication of diabetes and a leading cause of vision loss in the working-age population. Diabetic macular edema (DME), the accumulation of fluid in the macula, is the most common cause of vision loss in patients with DR.13 AI algorithms applied to OCT images excel at identifying and quantifying the hallmark signs of DME, including

**intraretinal fluid (IRF)**, **subretinal fluid (SRF)**, disruptions in the **external limiting membrane (ELM)** and **ellipsoid zone (EZ)**, and **hyperreflective foci (HRF)**.13

The accuracy of these AI systems has been rigorously validated against expert human graders. One multicenter validation study of an AI algorithm for DME biomarkers found "almost perfect agreement" (Cohen's kappa value > 0.81) for the presence of SRF and ELM/EZ disruption, and an excellent intraclass correlation coefficient (ICC) of 0.973 for the automated counting of HRF, demonstrating that the AI's performance was statistically indistinguishable from that of trained clinicians.13 Furthermore, AI models are capable of precisely segmenting individual retinal layers in patients with DR, which is crucial for detecting early neurodegenerative changes that may precede visible vascular signs. In this task, AI has achieved performance comparable to human graders, even in the presence of mild retinopathy.35

The high accuracy and efficiency of these systems have enabled their deployment in real-world clinical settings. For example, an OCT-AI-based telemedicine platform deployed in primary care stations for community screening demonstrated a sensitivity of 98.5% and a specificity of 96.2% for detecting referable retinal diseases, proving its practical value in identifying at-risk patients and facilitating timely referral to specialists.42

#### **2.2.3 Glaucoma**

Glaucoma is a progressive optic neuropathy characterized by the gradual loss of retinal ganglion cells and their axons, leading to irreversible vision loss. OCT is the primary imaging modality used to detect and monitor the structural damage associated with glaucoma, specifically the thinning of the **retinal nerve fiber layer (RNFL)** around the optic nerve head (ONH) and the loss of the **ganglion cell-inner plexiform layer (GC-IPL)** in the macula.33

AI models, particularly CNNs, have been extensively applied to OCT scans for glaucoma diagnosis. A review of published studies showed that these models consistently achieve high diagnostic accuracy, with Area Under the Receiver Operating Characteristic Curve (AUC) values ranging from 0.78 to as high as 0.99 in the task of differentiating glaucomatous eyes from healthy controls.34 Advanced

**3D CNNs** are capable of analyzing entire volumetric OCT scans without requiring prior segmentation of the retinal layers. These models learn to identify the most salient three-dimensional anatomical features associated with glaucomatous optic neuropathy, such as optic disc cupping and neuroretinal rim thinning, achieving AUCs as high as 0.94.34

A significant challenge in glaucoma management is tracking its slow progression over time. This requires the analysis of longitudinal data, a task for which static, single-visit images are insufficient. To address this, researchers are employing **Recurrent Neural Networks (RNNs)**, which are designed to process sequential data. By analyzing a series of OCT scans taken over multiple visits, RNNs can learn the temporal patterns of structural change, providing a dynamic understanding of disease trajectories and potentially predicting future rates of progression.33 This application highlights AI's evolving capability from static diagnosis to dynamic, predictive monitoring.

The consistently high performance of AI across these major diseases, as summarized in Table 2, provides compelling evidence of its clinical utility and transformative potential in ophthalmic imaging.

| **Table 2: Performance of AI Models in OCT-Based Disease Detection** |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Disease** | **AI Task** | **AI Model Type** | **Key Performance Metric** | **Source** |
| **AMD** | Exudative AMD Classification | CNN (OCT images) | Specificity: 0.97 (95% CI 0.93-0.98) | 41 |
| **AMD** | Clinical Trial Endpoint (GA) | AI-based Segmentation | Enabled EZ integrity as a new FDA-acknowledged endpoint | 40 |
| **DME** | Biomarker Identification (SRF) | Deep Learning (GAN-based) | Kappa: 0.831 (Almost Perfect Agreement) | 13 |
| **DME** | Biomarker Quantification (HRF) | Deep Learning (GAN-based) | ICC: 0.973 (Excellent Correlation) | 13 |
| **DR** | Retinal Layer Segmentation | FCN (DenseNet) + GP | Mean Unsigned Error: 1.06 pixels (vs. 1.10 inter-human) | 35 |
| **Glaucoma** | Disease Classification | 3D CNN (unsegmented volumes) | AUC: 0.94 | 34 |
| **Glaucoma** | Disease Classification | Various CNNs | AUC Range: 0.78 - 0.99 | 34 |
| **Glaucoma** | Progression Analysis | CNN + RNN | F-measure: 96.2% | 34 |
| **General Retinal Disease** | Community Screening/Referral | OCT-AI Platform | Sensitivity: 98.5%, Specificity: 96.2% | 42 |

## **Part III: The Multimodal Revolution: Integrating LLMs with Medical Imaging**

The evolution of AI in medical imaging is entering a new, more powerful phase, moving beyond the unimodal analysis of images toward a multimodal paradigm orchestrated by Large Language Models. This revolution is driven by the capacity of LMMs to understand, reason about, and generate information from disparate data types—including images, text, and structured data—simultaneously. This capability is unlocking novel applications that more closely mirror the complex, integrative reasoning of human clinicians, from automated report generation to fully interactive diagnostic dialogues.

### **Section 3.1: Automated Report Generation and Summarization**

A significant bottleneck in clinical workflows is the manual composition of medical reports. This process is not only time-consuming and labor-intensive but is also subject to inter-reader variability in style, structure, and terminology.44 Furthermore, the specialized language used in ophthalmology and radiology reports often creates a "comprehension gap" for non-specialist clinicians and patients, potentially hindering interdisciplinary communication and shared decision-making.46 LLMs are now being applied to address these challenges directly.

Advanced frameworks are being developed to automate the generation of reports from medical images. Many of these systems adopt a two-stage process that emulates the clinician's workflow. For example, models like **RadAlign** and **KARGEN** first employ a specialized vision model (like a CNN) to analyze an image and extract key clinical findings or medical concepts. These extracted concepts are then translated into a structured text prompt that is fed to an LLM, which uses its powerful generative capabilities to weave the findings into a fluent, coherent narrative report.47 This approach grounds the LLM's output in the visual evidence, aiming to improve clinical accuracy.

To combat the variability of free-text reports, researchers have proposed the task of **Structured Radiology Report Generation (SRRG)**. This involves reformulating reports into a standardized format with predefined sections and anatomical headers, ensuring clarity and consistency.44 LLMs are being used both to convert existing free-text reports into this structured format and to generate structured reports directly from images. To evaluate the clinical accuracy of these generated reports, specialized models like

**SRR-BERT**, which can classify a fine-grained set of 55 disease labels, have been developed to serve as an objective evaluation metric.44

Beyond generating full reports for clinicians, LLMs are proving highly effective at creating **plain language summaries** for a broader audience. In a large quality improvement study, adding LLM-generated plain language summaries to ophthalmology notes was shown to significantly improve diagnostic understanding, satisfaction with note detail, and clarity for non-ophthalmology professionals.46 The participating ophthalmologists confirmed that the summaries were highly accurate (90% reported "a great deal" of accuracy) and required minimal review time (95% took less than one minute), demonstrating a practical path toward improving interdisciplinary communication and patient engagement.46

### **Section 3.2: Integrating Imaging Findings with Electronic Health Records (EHRs)**

A complete clinical diagnosis is rarely made from an image alone. It requires the synthesis of information from multiple sources, including the patient's clinical history, lab results, and prior treatments, much of which is stored as unstructured text within the Electronic Health Record (EHR).49 Manually sifting through potentially hundreds of pages of clinical notes to find relevant information is often practically infeasible for a busy clinician. Large Multimodal Models are uniquely positioned to automate this critical process of data integration.

The proposed clinical workflow of the future involves MLLMs that can access and process data from multiple hospital information systems concurrently. An MLLM could analyze a new OCT scan from the Picture Archiving and Communication System (PACS) while simultaneously querying the EHR database for relevant patient history.30 This enables a more holistic and contextualized analysis. For example, a proposed strategy uses an LLM to perform

**zero-shot evidence retrieval** from the EHR. The model can be prompted with a clinical question relevant to the image, such as, "Based on this patient's clinical notes, is there any evidence of poorly controlled diabetes or hypertension that would increase their risk for the retinal findings seen in this OCT?" The LLM can then retrieve and summarize the specific notes that support its answer, presenting the clinician with concise, relevant evidence from the patient's history.49

LLMs also demonstrate remarkable flexibility in handling the diverse data formats found within healthcare systems. They can, for instance, convert structured data like ICD-10 billing codes into natural language narratives (e.g., converting "E11.311" to "Type 2 diabetes mellitus with nonproliferative diabetic retinopathy with macular edema"). This allows the LLM to better leverage its vast internal knowledge base to perform higher-level tasks like predicting future disease risk based on a patient's complete visit history.50 By integrating imaging findings, EHR data, and even embedded clinical guidelines, LLMs can provide clinicians with a comprehensive diagnostic summary, flagging potential discrepancies between the imaging and the patient's history that might otherwise be missed.30

### **Section 3.3: Conversational AI and Clinical Decision Support**

Perhaps the most forward-looking application of LLMs in medical imaging is the transition from generating static reports to enabling interactive, conversational dialogues with clinicians. This transforms the AI from a passive tool into an active, collaborative partner in the diagnostic process.

Several frameworks are pioneering this new paradigm. **RaDialog** is a vision-language model designed specifically for interactive radiology dialogue. After generating an initial report from a chest X-ray, it allows the radiologist to engage with it conversationally: they can ask follow-up questions ("Which findings support the diagnosis of pneumonia?"), request corrections ("The patient has no history of smoking, please remove that from the report"), ask for reformatting ("Summarize the key findings in bullet points for the referring physician"), or request a simplified explanation for the patient.52

Taking this concept a step further, the proposed **MedChat** framework for glaucoma diagnosis employs a **multi-agent** approach to emulate the collaborative reasoning of a clinical team. Rather than relying on a single LLM, MedChat assigns distinct roles to multiple LLM "agents" (e.g., an "Image Analysis Agent," a "Clinical History Agent," a "Differential Diagnosis Agent"). A "Director Agent" then coordinates and synthesizes the outputs from these specialized agents to produce a more robust and comprehensive final report. This multi-agent design is intended to mitigate the risk of errors or hallucinations that can arise from a single model's perspective and to encourage more nuanced, multi-faceted reasoning.53

Ultimately, these conversational systems aim to function as true Clinical Decision Support Systems (CDSS). The goal is not just to answer questions, but to provide a structured rationale for their conclusions, weighing evidence from the image and EHR, and even referencing external medical literature or clinical guidelines to support their reasoning.54 This ability to ground explanations in verifiable evidence is absolutely critical for building the clinician trust necessary for widespread adoption.

This evolution marks a fundamental change in the "user interface" of medical AI. The first generation of AI systems provided a discrete, static output—a segmentation map or a classification score—and the clinician's interaction was limited to simply accepting or rejecting that result. The inherently conversational nature of LLMs is being leveraged to create a dynamic, interactive, and collaborative workspace. This new paradigm transforms the clinician's role from that of a passive reviewer of an AI's conclusion to an active collaborator with an AI assistant. This has profound implications for improving workflow efficiency through rapid report refinement, enhancing medical education by allowing trainees to probe the AI's reasoning, and, most importantly, fostering trust and accelerating adoption by making the AI's decision-making process more transparent, interpretable, and controllable.

## **Part IV: Critical Challenges and Limitations in Clinical Implementation**

Despite the immense potential and rapid progress of AI and LLMs in ophthalmic imaging, the path to widespread, responsible clinical implementation is fraught with significant challenges. These hurdles span technical, ethical, and practical domains and must be critically addressed to ensure that these powerful technologies are safe, effective, and equitable. Acknowledging these limitations is essential for all stakeholders, from developers and researchers to clinicians and regulators, to navigate the future of this field responsibly.

### **Section 4.1: Technical and Algorithmic Hurdles**

* **Hallucinations:** The tendency of generative models, including LLMs, to produce outputs that are factually incorrect, nonsensical, or inconsistent with source data is arguably the most critical technical barrier.56 Termed "hallucinations," these fabrications are particularly perilous in a medical context, where a confident-sounding but incorrect diagnosis or a fictitious reference to a non-existent clinical study could lead to severe patient harm.57 Several strategies are being actively researched to mitigate this risk.  
  **Retrieval-Augmented Generation (RAG)** is a key technique that grounds the LLM's response in a trusted external knowledge base, such as clinical practice guidelines or a patient's specific EHR, before generating an answer.56  
  **Reinforcement Learning from Human Feedback (RLHF)**, which involves using human experts to rate and rank model outputs to better align the model with human values and expectations of accuracy, has also been widely applied to reduce hallucinations.56
* **Interpretability (The "Black Box" Problem):** Deep learning models are often described as "black boxes" because their internal decision-making processes are not inherently transparent.57 For clinicians trained in evidence-based medicine, trusting a diagnostic recommendation without understanding the underlying rationale is a major challenge.59 This lack of interpretability complicates error analysis and raises profound questions about accountability. The field of  
  **Explainable AI (XAI)** is dedicated to developing methods that can provide insights into a model's reasoning, for instance, by highlighting the specific regions of an image that most influenced its classification decision. Progress in XAI is vital for building clinician trust and facilitating responsible adoption.59
* **Robustness and Generalizability:** An AI model may demonstrate stellar performance on the curated dataset it was trained and validated on, but its performance can degrade significantly when deployed in a real-world clinical setting. This is because of variations in patient populations, imaging hardware from different manufacturers, and differing image acquisition protocols.62 This "domain shift" problem highlights the critical need for rigorous  
  **external validation** of AI models on diverse, multi-center, real-world datasets before they can be considered for clinical use. Many published studies lack this level of validation, limiting the generalizability of their findings.41
* **Multimodal Integration Challenges:** While the integration of image and text data is a key advantage of LMMs, it is also a formidable technical challenge. Visual information is continuous, high-dimensional, and spatially structured, whereas language is discrete and sequential. Developing architectures that can effectively and meaningfully align these fundamentally different data types remains an active area of research. Naive approaches can lose critical information or be computationally prohibitive.63

### **Section 4.2: Data Governance and Ethical Considerations**

* **Data Privacy and Security:** The development of medical AI models requires access to vast quantities of patient data, which is highly sensitive and protected by stringent regulations like the Health Insurance Portability and Accountability Act (HIPAA) in the US and the General Data Protection Regulation (GDPR) in Europe.57 Ensuring the privacy and security of this data is paramount. Uploading patient information to third-party cloud-based AI systems poses significant risks.57 Solutions include robust data encryption, secure on-premise deployments, and privacy-preserving techniques like  
  **federated learning**, where the model is trained locally at different institutions without the raw data ever leaving its source.65
* **Algorithmic Bias:** AI models learn from the data they are trained on. If this training data is not representative of the diverse patient population—for example, if it underrepresents certain racial, ethnic, or socioeconomic groups—the resulting algorithm will inherit and may even amplify these biases.59 A biased model could be less accurate for underrepresented groups, exacerbating existing health disparities. Mitigating bias requires careful auditing of datasets for diversity, inclusivity, and fairness, as well as post-deployment monitoring to ensure equitable performance across all demographic groups.
* **Accountability and Liability:** The "black box" nature of AI creates a complex legal and ethical dilemma regarding accountability. If an AI system contributes to a diagnostic error that results in patient harm, who is liable? Is it the clinician who acted on the AI's recommendation, the hospital that deployed the system, or the company that developed the software? The current legal and regulatory frameworks are still evolving and do not yet provide clear answers to these critical questions, creating a climate of uncertainty for all parties involved.57

### **Section 4.3: Clinical Workflow and Regulatory Pathways**

* **Integration with Existing Systems:** A significant practical barrier to AI adoption is the challenge of integrating new tools with the complex and often fragmented IT infrastructure of hospitals and clinics. AI systems must be able to seamlessly communicate with existing Electronic Health Record (EHR) and Picture Archiving and Communication System (PACS) platforms. A lack of interoperability and data standardization can make deployment a costly and resource-intensive endeavor.64
* **Clinician and Staff Adoption:** The successful implementation of AI is not just a technical problem; it is also a human one. Healthcare professionals may resist adopting new technologies due to a lack of familiarity, a distrust of AI, fears of job displacement, or concerns that the tools will increase their administrative workload rather than decrease it.65 Overcoming this resistance requires early and continuous engagement with clinical staff, comprehensive training programs, and a clear demonstration that the AI tool provides tangible value and fits smoothly into their established workflows.64
* **Cost and Resource Allocation:** The financial investment required for AI implementation—including software licenses, necessary hardware upgrades, IT support, and staff training—can be substantial and prohibitive, particularly for smaller clinics or healthcare systems in low-resource settings.64
* **Evolving Regulatory Landscape:** As medical devices, AI algorithms are subject to oversight by regulatory bodies like the U.S. Food and Drug Administration (FDA). The unprecedented pace of AI innovation, especially with the advent of adaptable, generative models like LLMs, presents a significant challenge for regulators. They are tasked with developing new frameworks and guidelines that can ensure these technologies are proven to be safe and effective for clinical use without stifling innovation. This evolving regulatory landscape requires developers to engage proactively with regulators to navigate the pathway to market.11

A structured approach to addressing these multifaceted issues is necessary for progress, as outlined in Table 3.

| **Table 3: Challenges and Mitigation Strategies for AI/LLM Implementation in Ophthalmology** |  |  |  |
| --- | --- | --- | --- |
| **Challenge Category** | **Specific Challenge** | **Description of Challenge** | **Proposed Mitigation Strategies** |
| **Technical** | **Hallucinations** | LLMs generate factually incorrect or fabricated information with high confidence, posing a direct risk to patient safety. | Implement Retrieval-Augmented Generation (RAG) to ground outputs in trusted sources; use Reinforcement Learning from Human Feedback (RLHF) to align models with accuracy; design multi-agent systems to cross-validate outputs.53 |
| **Technical** | **Interpretability ("Black Box")** | The inability to understand the reasoning behind an AI's decision erodes clinician trust and complicates error analysis and accountability. | Develop and integrate Explainable AI (XAI) techniques to provide visual or textual explanations for model outputs; promote research into inherently more transparent model architectures.59 |
| **Technical** | **Robustness & Generalizability** | Models perform poorly when deployed on data from different devices or patient populations than those they were trained on. | Mandate rigorous external validation on diverse, multi-center, real-world datasets; develop methods for domain adaptation and continuous model monitoring post-deployment.41 |
| **Ethical/Regulatory** | **Data Privacy & Security** | Training models requires large amounts of sensitive patient data, creating risks of privacy breaches under HIPAA/GDPR. | Utilize privacy-preserving techniques like federated learning; ensure end-to-end encryption and secure on-premise or HIPAA-compliant cloud solutions; de-identify data rigorously.57 |
| **Ethical/Regulatory** | **Algorithmic Bias** | Training on non-representative data can lead to models that perform inequitably across different demographic groups, worsening health disparities. | Audit training datasets for diversity and representation; implement fairness metrics in model evaluation; conduct post-market surveillance for performance disparities.59 |
| **Ethical/Regulatory** | **Accountability & Liability** | Unclear legal frameworks for assigning responsibility when an AI-related error causes harm. | Develop clear institutional policies on the use of AI; professional societies and regulatory bodies must collaborate to establish legal and ethical guidelines for AI in clinical practice.57 |
| **Practical/Workflow** | **EHR/PACS Integration** | Lack of interoperability and standardization makes it difficult to integrate AI tools into existing hospital IT systems. | Prioritize AI tools designed with interoperability standards (e.g., HL7 FHIR); work closely with vendors and IT departments to plan for seamless integration.64 |
| **Practical/Workflow** | **Clinician & Staff Adoption** | Resistance from staff due to distrust, fear of job replacement, or perceived increase in workload. | Involve clinicians and staff early in the selection and implementation process; provide comprehensive training and ongoing support; design tools that demonstrably reduce administrative burden.65 |

## **Part V: The Future of Ophthalmic AI: Synthesis and Strategic Recommendations**

The convergence of artificial intelligence, particularly deep learning and large language models, with advanced imaging modalities like Optical Coherence Tomography is not merely an incremental technological advancement but the beginning of a fundamental reshaping of ophthalmology and, potentially, medicine at large. By synthesizing the current evidence, applications, and challenges, a clear trajectory emerges: a move away from reactive diagnosis toward proactive, predictive, and highly personalized eye care.

### **Section 5.1: The Trajectory Towards Predictive and Personalized Eye Care**

The current paradigm of ophthalmic AI is largely focused on *detecting existing disease* with a proficiency that can match or exceed human experts. However, the ultimate promise of this technology lies in its potential to shift the clinical focus from detection to **prediction**. By leveraging AI to analyze longitudinal data—such as series of OCT scans captured over many years—models can learn the subtle, subclinical patterns of change that precede a clinical diagnosis. When this temporal imaging data is integrated by LMMs with a patient's genomic data, EHR history, and lifestyle factors, the potential emerges to create powerful predictive models.33 These models could forecast an individual's lifetime risk of developing a condition like glaucoma or AMD, predict the likely rate of disease progression if it occurs, and even estimate their probable response to different therapeutic interventions, allowing for truly personalized treatment strategies.40

Furthermore, the eye is increasingly being recognized as a unique "window" to systemic health. The retina's vasculature and neural tissue share developmental and physiological properties with the brain and cardiovascular system. AI algorithms are already showing promise in detecting biomarkers of neurodegenerative conditions like Alzheimer's disease and cardiovascular risk factors from retinal fundus and OCT images.69 In the future, this could fundamentally reposition the role of the ophthalmologist and the routine eye exam, transforming it into a non-invasive, cost-effective screening hub for a wide range of systemic diseases.

Looking further ahead, the convergence of these powerful multimodal AI systems could enable the creation of a **"digital twin"** for each patient's eye. This would be a dynamic, computational model that integrates all of a patient's relevant data—imaging, clinical, genomic, and more. This digital twin could be used to simulate the natural course of the patient's disease and, crucially, to test the potential efficacy and side effects of various treatments *in silico* before they are ever administered to the patient. This represents the ultimate realization of precision medicine, moving from population-level evidence to individualized prediction and intervention.

### **Section 5.2: Recommendations for Stakeholders**

To navigate this complex but promising future, key stakeholders must adopt strategic approaches tailored to their roles within the healthcare ecosystem.

* **For Clinicians and Healthcare Organizations:**
  + **Embrace a Collaborative Mindset:** It is imperative to view AI not as a competitor or replacement, but as a powerful collaborative tool. AI's strength lies in its ability to perform rapid, quantitative analysis and reduce administrative burdens, which augments and elevates the clinician's expertise, freeing them to focus on complex reasoning, patient communication, and care management.11
  + **Invest in Training and Workflow Planning:** Successful AI implementation is contingent on human factors. Organizations must invest in comprehensive training for all staff, from technicians to physicians, to ensure they understand the capabilities and limitations of these tools. Careful, collaborative planning is required to integrate AI seamlessly into existing clinical workflows to maximize efficiency and avoid disruption.64
  + **Prioritize Validated, Interpretable Systems:** When procuring AI systems, preference should be given to those that have undergone rigorous, independent, external validation on diverse populations. Furthermore, clinicians should advocate for and prioritize systems that offer some degree of explainability, as transparency is fundamental to maintaining clinical trust, accountability, and the principles of evidence-based practice.59
* **For Researchers and Academics:**
  + **Focus on Robustness and Fairness:** The field must move beyond proof-of-concept studies on limited, curated datasets. Research priority should be given to the external validation of algorithms on large-scale, multi-center, real-world data to ensure robustness. Concurrently, developing and implementing robust methods to detect, measure, and mitigate algorithmic bias is a critical ethical and scientific imperative.41
  + **Advance Explainable AI (XAI):** A concerted research effort is needed to "open the black box." Developing and validating novel XAI techniques that can provide clinically meaningful explanations for the outputs of complex models like LMMs is essential for their translation into high-stakes clinical environments.
  + **Develop Standardized Benchmarks:** The research community should collaborate to create and adopt standardized, public benchmarks and datasets for evaluating medical AI models, especially multimodal systems. This will enable fair, reproducible comparisons between different approaches and accelerate progress in the field.56
* **For Technology Developers and Industry:**
  + **Design for the Clinical Workflow:** The most successful AI tools will be those that solve a tangible, real-world clinical problem. This requires deep collaboration with end-users—clinicians, technicians, and administrators—to ensure that products are not only technically powerful but also intuitive, efficient, and seamlessly integrated into the complex realities of daily practice.66
  + **Prioritize Safety and Reliability Above All Else:** For generative AI and LLMs, mitigating the risk of hallucinations and ensuring clinical accuracy is non-negotiable. Developers must invest heavily in safety measures, including grounding models in verifiable evidence (RAG), extensive expert-led fine-tuning (RLHF), and building in robust "human-in-the-loop" validation steps where a clinician must confirm critical outputs.54
  + **Navigate the Regulatory Landscape Collaboratively:** The regulatory environment for AI is in flux. Developers should engage proactively and transparently with regulatory bodies like the FDA. This collaborative approach can help shape clear, effective, and predictable regulatory pathways that ensure patient safety while fostering continued innovation in this transformative field.68

#### Works cited

1. Artificial Intelligence in Medical Imaging | AI in Imaging - Spectral AI, accessed August 7, 2025, <https://www.spectral-ai.com/blog/artificial-intelligence-in-medical-imaging/>
2. Artificial Intelligence (AI) in Healthcare & Medical Field, accessed August 7, 2025, <https://www.foreseemed.com/artificial-intelligence-in-healthcare>
3. What is Artificial Intelligence in Medicine? | IBM, accessed August 7, 2025, <https://www.ibm.com/think/topics/artificial-intelligence-medicine>
4. Medical Image Segmentation: A Comprehensive Review of Deep Learning-Based Methods, accessed August 7, 2025, <https://www.mdpi.com/2379-139X/11/5/52>
5. Medical image analysis using deep learning algorithms - PMC - PubMed Central, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10662291/>
6. Machine Learning for Medical Image Analysis: A Survey - ResearchGate, accessed August 7, 2025, <https://www.researchgate.net/publication/371258367_Machine_Learning_for_Medical_Image_Analysis_A_Survey>
7. Advancements and Challenges in Medical Image Segmentation: A Comprehensive Survey-Scilight, accessed August 7, 2025, <https://ojs.sciltp.com/journals/sai/article/view/658>
8. A Survey on Medical Image Segmentation Based on Deep Learning ..., accessed August 7, 2025, <https://www.mdpi.com/2504-2289/6/4/117>
9. Deep Convolutional Neural Networks in Medical Image Analysis: A Review - MDPI, accessed August 7, 2025, <https://www.mdpi.com/2078-2489/16/3/195>
10. A review of the application of deep learning in medical image classification and segmentation - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC7327346/>
11. The Good, the Bad, and the Ugly of AI in Medical Imaging - EMJ, accessed August 7, 2025, <https://www.emjreviews.com/radiology/article/the-good-the-bad-and-the-ugly-of-ai-in-medical-imaging-j140125/>
12. Rajalakshmi Institute of Technology adopts avant-garde tech in healthcare, engineering, accessed August 7, 2025, <https://timesofindia.indiatimes.com/education/rajalakshmi-institute-of-technology-adopts-avant-garde-tech-in-healthcare-engineering/articleshow/123088475.cms>
13. Validation of an Automated Artificial Intelligence Algorithm for the ..., accessed August 7, 2025, <https://www.mdpi.com/2077-0383/12/6/2134>
14. Optical Coherence Tomography (OCT) Eye Test: What It Is - Cleveland Clinic, accessed August 7, 2025, <https://my.clevelandclinic.org/health/diagnostics/optical-coherence-tomography-oct>
15. en.wikipedia.org, accessed August 7, 2025, <https://en.wikipedia.org/wiki/Optical_coherence_tomography>
16. Optical Coherence Tomography - StatPearls - NCBI Bookshelf, accessed August 7, 2025, <https://www.ncbi.nlm.nih.gov/books/NBK606122/>
17. Optical Coherence Tomography - EyeWiki, accessed August 7, 2025, <https://eyewiki.org/Optical_Coherence_Tomography>
18. What Is Optical Coherence Tomography? - American Academy of Ophthalmology, accessed August 7, 2025, <https://www.aao.org/eye-health/treatments/what-is-optical-coherence-tomography>
19. Optical Coherence Tomography (OCT) - Wills Eye Hospital, accessed August 7, 2025, <https://www.willseye.org/treatment/optical-coherence-tomography-oct/>
20. What Are Large Language Models (LLMs)? - IBM, accessed August 7, 2025, <https://www.ibm.com/think/topics/large-language-models>
21. What is a large language model (LLM)? - University of Arizona Libraries, accessed August 7, 2025, <https://ask.library.arizona.edu/faq/407985>
22. What is an LLM (large language model)? - Cloudflare, accessed August 7, 2025, <https://www.cloudflare.com/learning/ai/what-is-large-language-model/>
23. What is LLM? - Large Language Models Explained - AWS, accessed August 7, 2025, <https://aws.amazon.com/what-is/large-language-model/>
24. What Are Large Language Models Used For? - NVIDIA Blog, accessed August 7, 2025, <https://blogs.nvidia.com/blog/what-are-large-language-models-used-for/>
25. Understanding the Capabilities and Limitations of LLMs in Radiology - RSNA, accessed August 7, 2025, <https://www.rsna.org/news/2025/july/using-llms-in-radiology>
26. Large Language Models and Large Multimodal Models in Medical Imaging: A Primer for Physicians | Journal of Nuclear Medicine, accessed August 7, 2025, <https://jnm.snmjournals.org/content/66/2/173.abstract>
27. Multimodal Large Language Models in Health Care: Applications, Challenges, and Future Outlook - Journal of Medical Internet Research, accessed August 7, 2025, <https://www.jmir.org/2024/1/e59505/>
28. Large Language Models and Large Multimodal Models in Medical Imaging: A Primer for Physicians - PubMed, accessed August 7, 2025, <https://pubmed.ncbi.nlm.nih.gov/39819692/>
29. The role of large language models in medical image processing: a narrative review - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10784029/>
30. Implementing Large Language Models in Health Care: Clinician ..., accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12299950/>
31. From CNN to Transformer: A Review of Medical Image Segmentation Models - arXiv, accessed August 7, 2025, <https://arxiv.org/abs/2308.05305>
32. Deep Learning Medical Image Segmentation Methods: A Survey - EKB Journal Management System, accessed August 7, 2025, <https://journals.ekb.eg/article_338478_cc0d01882da9b93c5d40ea1ccf0a353e.pdf>
33. Artificial Intelligence for Optical Coherence Tomography in ... - TVST, accessed August 7, 2025, <https://tvst.arvojournals.org/article.aspx?articleid=2802509>
34. Artificial Intelligence for Optical Coherence Tomography in Glaucoma - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11760759/>
35. Deep Learning based Retinal OCT Segmentation, accessed August 7, 2025, <https://arxiv.org/abs/1801.09749>
36. AI Can Enhance OCT Analysis for Retinal Disease Monitoring, accessed August 7, 2025, <https://www.ajmc.com/view/ai-can-enhance-oct-analysis-for-retinal-disease-monitoring>
37. Application and progress of artificial intelligence technology in the segmentation of hyperreflective foci in OCT images for ophthalmic disease research - PubMed Central, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11144766/>
38. Deep learning-based classification of eye diseases using ... - Frontiers, accessed August 7, 2025, <https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2023.1252295/full>
39. AI-based support for optical coherence tomography in age-related macular degeneration, accessed August 7, 2025, <https://www.researchgate.net/publication/379686210_AI-based_support_for_optical_coherence_tomography_in_age-related_macular_degeneration>
40. Rethinking Clinical Trials in Age-Related Macular Degeneration ..., accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11945239/>
41. Comparison of deep learning approaches for OCT diagnostics of ..., accessed August 7, 2025, <https://www.researchgate.net/publication/374466289_Comparison_of_deep_learning_approaches_for_OCT_diagnostics_of_age-related_macular_degeneration>
42. Evaluation of an OCT-AI–Based Telemedicine Platform for Retinal Disease Screening and Referral in a Primary Care Setting | TVST, accessed August 7, 2025, <https://tvst.arvojournals.org/article.aspx?articleid=2778636>
43. Artificial Intelligence for Optical Coherence Tomography in Glaucoma - PubMed, accessed August 7, 2025, <https://pubmed.ncbi.nlm.nih.gov/39854198/>
44. Automated Structured Radiology Report Generation - arXiv, accessed August 7, 2025, <https://arxiv.org/abs/2505.24223>
45. [2410.15642] Resource-Efficient Medical Report Generation using Large Language Models, accessed August 7, 2025, <https://arxiv.org/abs/2410.15642>
46. Large language model-generated summaries improve interdisciplinary communication in ophthalmology, accessed August 7, 2025, <https://www.ophthalmologytimes.com/view/large-language-model-generated-summaries-improve-interdisciplinary-communication-in-ophthalmology>
47. [2501.07525] RadAlign: Advancing Radiology Report Generation with Vision-Language Concept Alignment - arXiv, accessed August 7, 2025, <https://arxiv.org/abs/2501.07525>
48. [2409.05370] KARGEN: Knowledge-enhanced Automated Radiology Report Generation Using Large Language Models - arXiv, accessed August 7, 2025, <https://arxiv.org/abs/2409.05370>
49. Retrieving Evidence from EHRs with LLMs: Possibilities and Challenges - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC11368037/>
50. LLMs-based Few-Shot Disease Predictions using EHR: A Novel Approach Combining Predictive Agent Reasoning and Critical Agent Instruction - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12099430/>
51. Harnessing the Power of Large Language Models (LLMs) for Electronic Health Records (EHRs) Optimization - PubMed, accessed August 7, 2025, <https://pubmed.ncbi.nlm.nih.gov/37644945/>
52. RaDialog: A Large Vision-Language Model for Radiology Report Generation and Conversational Assistance - arXiv, accessed August 7, 2025, <https://arxiv.org/html/2311.18681v1>
53. MedChat: A Multi-Agent Framework for Multimodal Diagnosis with Large Language Models, accessed August 7, 2025, <https://arxiv.org/html/2506.07400v2>
54. Conversational AI for Diagnostics and Patient Support - John Snow Labs, accessed August 7, 2025, <https://www.johnsnowlabs.com/conversational-ai-for-diagnostics-and-patient-support/>
55. Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System - ResearchGate, accessed August 7, 2025, <https://www.researchgate.net/publication/380525352_Human-Algorithmic_Interaction_Using_a_Large_Language_Model-Augmented_Artificial_Intelligence_Clinical_Decision_Support_System>
56. Large Language Models in Medicine: Applications, Challenges, and Future Directions, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12163604/>
57. EDITORIAL - Opportunities and Challenges of Large Language ..., accessed August 7, 2025, <https://fbt.tums.ac.ir/index.php/fbt/article/view/1182/488>
58. Large Language Models in Radiology Reporting—A Systematic ..., accessed August 7, 2025, <https://www.medrxiv.org/content/10.1101/2025.03.18.25324193v1.full-text>
59. Ophthalmology balances the promises and challenges of AI, accessed August 7, 2025, <https://www.ophthalmologytimes.com/view/ophthalmology-balances-the-promises-and-challenges-of-ai>
60. Artificial Intelligence Applications in Ophthalmology - PubMed, accessed August 7, 2025, <https://pubmed.ncbi.nlm.nih.gov/39926073/>
61. Artificial Intelligence in Ophthalmology: Advantages and Limits - MDPI, accessed August 7, 2025, <https://www.mdpi.com/2076-3417/15/4/1913>
62. Artificial intelligence in ophthalmological practice: when ideal meets reality, accessed August 7, 2025, <https://bmjophth.bmj.com/content/8/1/e001129>
63. A Review of the Opportunities and Challenges with Large Language Models in Radiology: The Road Ahead | American Journal of Neuroradiology, accessed August 7, 2025, <https://www.ajnr.org/content/46/7/1292>
64. Integrating AI Technologies into Ophthalmology: Challenges and ..., accessed August 7, 2025, <https://www.simbo.ai/blog/integrating-ai-technologies-into-ophthalmology-challenges-and-opportunities-for-future-practice-management-and-patient-outcomes-2941929/>
65. Identifying the Challenges Healthcare Professionals Face When Integrating AI Technologies in Ophthalmology Settings and Potential Solutions | Simbo AI - Blogs, accessed August 7, 2025, <https://www.simbo.ai/blog/identifying-the-challenges-healthcare-professionals-face-when-integrating-ai-technologies-in-ophthalmology-settings-and-potential-solutions-3995847/>
66. Generative artificial intelligence in ophthalmology: current innovations, future applications and challenges, accessed August 7, 2025, <https://bjo.bmj.com/content/108/10/1335>
67. Artificial intelligence technology in ophthalmology public health: current applications and future directions - Frontiers, accessed August 7, 2025, <https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2025.1576465/full>
68. Exploring the future of AI in ophthalmology with Mark Daly, accessed August 7, 2025, <https://www.ophthalmologytimes.com/view/exploring-the-future-of-ai-in-ophthalmology-with-mark-daly>
69. AI-assisted ophthalmic imaging for early detection of neurodegenerative diseases - PMC, accessed August 7, 2025, <https://pmc.ncbi.nlm.nih.gov/articles/PMC12054287/>
70. Artificial Intelligence in Biomedical Imaging | NYU Langone Health, accessed August 7, 2025, <https://med.nyu.edu/departments-institutes/radiology/research/ai-biomedical-imaging>
71. How Conversational AI is Transforming Diagnostics and Patient Support - John Snow Labs, accessed August 7, 2025, <https://www.johnsnowlabs.com/how-conversational-ai-is-transforming-diagnostics-and-patient-support/>