Review

Artificial Intelligence in Glaucoma: A New Landscape of Diagnosis and Management

Patrick Xiang Ji 1, Vethushan Ramalingam 2, Michael Balas 1, Lauren Pickel 1 and David J. Mathew 1,3,4,*

1 Temerty Faculty of Medicine, University of Toronto, Toronto, ON M5S 1A8, Canada; x.ji@mail.utoronto.ca (P.X.J.)
2 York Finch Eye Associates, Toronto, ON M3N 2V6, Canada
3 Donald K. Johnson Eye Institute, Krembil Research Institute, University Health Network, Toronto, ON M5T 0S8, Canada
4 Department of Ophthalmology and Vision Sciences, University of Toronto, Toronto, ON M5T 2S8, Canada
* Correspondence: david.mathew@mail.utoronto.ca; Tel.: +1-416-603-5727

Abstract: Glaucoma refers to a spectrum of progressive optic neuropathies and remains the leading cause of irreversible blindness worldwide. Its insidious onset poses serious challenges to conventional diagnostic methods and clinicians striving to detect early-stage disease for timely and effective intervention. Artificial intelligence (AI) has demonstrated its ability to process and analyze large datasets which can help identify subtle changes in early glaucomatous clinical presentation. This study reviews the current state of AI utilization in glaucoma and elucidates the strengths and limitations of existing approaches. We dissect the role of AI in various domains: enhancing early detection and diagnosis, monitoring disease progression, and refining treatment strategies to optimize patient outcomes. Furthermore, we address the ethical, legal, and social implications, alongside the inherent limitations of AI in the clinical setting. Despite these challenges, AI holds transformative potential for glaucoma management. Future directions emphasize the need for interdisciplinary collaboration, advanced and explainable algorithm development, and equitable healthcare access to fully realize the promise of AI in combating this vision-threatening condition.

Keywords: glaucoma; artificial intelligence; machine learning; deep learning; neural networks; visual field; optical coherence tomography; electroretinogram; MIGS; trabeculectomy

1. Introduction

Glaucoma represents a spectrum of progressive optic neuropathies, often associated with elevated intraocular pressure (IOP), that lead to the degeneration of retinal ganglion cells and thinning of the retinal nerve fiber layer (RNFL), culminating in optic nerve damage, optic disc cupping, and the loss of peripheral vision [1,2]. Glaucoma remains the leading cause of irreversible blindness, affecting over 70 million people worldwide, with disease prevalence projected to increase to 112 million people in 2040 due to the aging population [3,4]. Patients are often asymptomatic until reaching advanced stages of the disease, suggesting that the true population of affected individuals may far outnumber those previously diagnosed. This trend signifies the critical need for early detection and timely intervention.

1.1. Current Challenges in Glaucoma Management

The insidious onset of glaucoma poses significant hurdles for both conventional diagnostic methods and healthcare professionals striving to detect early-stage diseases for timely and effective intervention [2,5–7]. For instance, the clinical evaluation of visual fields (VFs) can be difficult and inaccurate, particularly at the disease onset due to minimal or no detectable changes [8]. Monitoring glaucoma progression also demands frequent and comprehensive eye exams, which can present adherence challenges influenced by...
Artificial intelligence (AI) has emerged as a revolutionary tool in addressing some of these challenges. The utilization of AI in ocular conditions [12,13], particularly in diagnostic imaging, pattern recognition, and predictive analysis, offers a new frontier in glaucoma management (Figure 1). For instance, structural imaging modalities such as optical coherence tomography (OCT) and fundus photography have been utilized to improve the accuracy of early diagnosis, predict disease progression, and optimize treatment plans [14,15]. Moreover, AI can play a significant role in predicting individual response patterns and developing personalized treatment regimens.

1.2. Role of AI in Glaucoma

AI’s ability to efficiently process and learn from large datasets may enable the detection of complex patterns at a fine level of detail, making it ideal for VF analysis. A multi-center trial from mainland China trained convolutional neural networks (CNNs) on 4012 VFs from 1352 patients to differentiate glaucomatous from non-glaucomatous VFs [8]. The CNN achieved an accuracy of 87.6% on a validation set of 300 VFs, with a sensitivity of 0.826 and a specificity of 0.932. This study also evaluated the performance of residents, attending ophthalmologists, and glaucoma experts, who performed with accuracies of 60.7%, 58.5%, and 62.6%, respectively. The CNN also outperformed traditional scoring systems such as

Figure 1. Step-by-step process of AI algorithm used for glaucoma management.

1.3. Objectives

This study aimed to review the current state of AI utilization in glaucoma, exploring the various applications across diagnostic, monitoring, and predictive treatment models. By examining AI technologies in ophthalmology, we sought to elucidate the strengths and limitations of existing approaches. We highlighted current challenges associated with AI implementation in glaucoma management and identified future directions for research and clinical integration.

2. AI in Glaucoma Detection and Diagnosis

2.1. AI in Functional Imaging

The VF test is one of the main functional diagnostic tests for glaucoma that evaluates the degree of peripheral and central visual impairment caused by the disease. AI’s ability to efficiently process and learn from large datasets may enable the detection of complex patterns at a fine level of detail, making it ideal for VF analysis. A multi-center trial from mainland China trained convolutional neural networks (CNNs) on 4012 VFs from 1352 patients to differentiate glaucomatous from non-glaucomatous VFs [8]. The CNN achieved an accuracy of 87.6% on a validation set of 300 VFs, with a sensitivity of 0.826 and a specificity of 0.932. This study also evaluated the performance of residents, attending ophthalmologists, and glaucoma experts, who performed with accuracies of 60.7%, 58.5%, and 62.6%, respectively. The CNN also outperformed traditional scoring systems such as
the Advanced Glaucoma Intervention Study (AGIS) and Glaucoma Severity Staging (GSS) system, which had accuracies of 45.9% and 52.3%, respectively.

In the evolving landscape of glaucoma diagnostics, AI offers a novel perspective on interpreting data from various imaging modalities. One particularly intriguing area of study involves the electroretinogram (ERG), a tool not traditionally associated with glaucoma detection due to its primary focus on retinal electrophysiology. ERG can capture functional deficits at the cellular level, specifically the loss of retinal ganglion cell function, introducing an innovative avenue for early glaucoma detection, even before conventional methods indicate disease presence [16]. Recent research has begun to explore the potential of ERGs in glaucoma diagnostics, propelled by AI’s capacity to analyze complex time-series data. For instance, an animal study applying dark-adapted Ganzfeld flash ERGs to mice employed AI algorithms trained on ERG data to perform both binary and multiclass classifications, distinguishing between healthy and glaucomatous states and differentiating stages of the disease. The models had accuracies reaching 91.7% and 80% for binary and multiclass classifications, respectively, showcasing the potential of AI in identifying early, subtle changes in ERG patterns indicative of glaucoma [17]. Building on these findings, a recent human study applied the continuous wavelet transform to analyze multifocal ERG signals from a cohort that included both healthy individuals and glaucoma patients [18]. Using a radial basis function neural network, the researchers trained the algorithm on morphological characteristics of ERG signals, achieving a sensitivity and specificity of 0.89 and 0.84, respectively, in distinguishing glaucomatous from non-glaucomatous eye sectors. This level of accuracy in detecting early glaucomatous changes suggests a promising role for AI in enhancing glaucoma diagnosis and understanding disease progression.

While ERG remains an underutilized tool in the routine assessment of glaucoma, these studies highlight its potential when combined with AI. This synergy could unlock new diagnostic markers and offer a deeper understanding of glaucoma’s early stages, ultimately leading to more timely and personalized interventions. As the field of ophthalmology continues to embrace technological advancements, integrating such innovative approaches could significantly improve our diagnostic capabilities and patient care strategies.

2.2. AI in Structural Imaging

The cornerstone of glaucoma detection lies in structural imaging modalities such as OCT and fundus photography. AI algorithms have shown remarkable success in identifying subtle, yet critical, changes in both OCT and fundus photography images that may help to describe clinical presentations of early-onset glaucoma. Studies have demonstrated the ability of AI to successfully identify clinical features of glaucoma, including optic disc cupping, changes in retinal microvasculature, and Bruch’s membrane opening-minimum rim width (BMO-MRW), among others. For example, models have been developed to segment the optic disc and cup using fundus photography to reveal retinal vessel kinking or BMO-MRW [19–21]. Antony et al. employed an iterative graph-based method for the automated segmentation of surfaces with a shared hole, which was applied to the two- and three-dimensional segmentation of BMO from Spectral Domain OCT (SD-OCT) volumes to compute BMO-MRW. This method demonstrated good accuracy, with an error of 55.29 μm in three-dimensional Euclidean space compared to manual annotations. Since BMO-MRW can estimate the remaining nerve fiber bundles in the retina, this can be a strong structural parameter for glaucoma diagnosis and monitoring [22].

AI algorithms have also been employed in OCT analysis to detect thinning of the RNFL and ganglion cell layer complex, which are key indicators of glaucomatous damage [23,24]. Koozekanani et al. employed a Markov boundary model trained on 330 images from 14 healthy individuals to detect retinal boundaries in OCT-B-scans [23]. Retinal thickness measurements derived from this algorithm were comparable to manually corrected methods for 1450 test images, with differences of <10 microns for 74% of the images and <25 microns for 98.4% of the images. These errors were reported to be near the resolution limit of OCT machines, and well below clinical significance, suggesting that quantitatively
accurate systems such as the model proposed in this study show promise to enhance the precision of glaucoma diagnostics and monitoring, facilitating early interventions and potentially altering the course of the disease.

Other deep learning algorithms have been employed to assess neuroretinal rim loss and the three-dimensional structure of the optic nerve head [25,26]. In a joint cross-sectional study between the Singapore National Eye Centre and the University Hospital Santaros Klinikos, Braeu et al. developed a geometric deep learning model to explore structural landmarks around the optic nerve head and analyze differences in connective and neural tissues between various stages of glaucoma [26]. The model found structural differences in both neural and connective tissues across all stages of glaucoma and achieved AUROCs of 0.94 in binary classification tasks between normal and mild glaucoma, 0.68 between mild and moderate glaucoma, and 0.80 between moderate and advanced glaucoma. This demonstrates the potential utility of AI in uncovering complex three-dimensional structural changes in the retina as a function of glaucoma severity. Such models may be useful for both early glaucoma detection and the monitoring of disease severity.

Automated segmentation algorithms enhance the precision of retinal structure delineation, mitigating human error and streamlining the diagnostic process. AI systems, by learning from extensive datasets of manually annotated images, have demonstrated diagnostic capabilities that meet or exceed those of traditional methods. For example, one study built a deep neural network trained on 1364 fundus photographs exhibiting glaucomatous features, alongside 1768 non-glaucomatous images, and found that it achieved an area under the receiver operating characteristic (AUROC) of 0.965 for glaucoma diagnosis. It even outperformed three ophthalmology resident physicians, whose AUROC values ranged from 0.725 to 0.912 [27]. Similarly, another study trained an ensemble of five CNNs using 43,055 fundus images from 12 public datasets. This ensemble model was evaluated using a previously unseen dataset of 100 images to identify various retinal diseases, including diabetic retinopathy, age-related macular degeneration, glaucoma, and normal eyes [28]. It achieved a mean accuracy of 79.2% across all cases, which outperformed seven board-certified ophthalmologists who achieved a mean accuracy of 72.7%.

With novel imaging modalities such as SD-OCT with enhanced depth imaging (EDI) and adaptive optics (AO), glaucoma disease detection can be further improved. Belghith et al. leveraged SD-OCT with EDI to employ a sophisticated non-local Markov random field-based segmentation technique, focusing on measuring variations in the anterior lamina cribrosa surface depth over time. The model was able to closely match the segmentations of two expert annotators, the results of which were correlated significantly to IOP, indicating that the positioning of the lamina cribrosa could act as an early structural biomarker for glaucoma, reflecting changes via IOP fluctuations [29]. Furthermore, AO can significantly enhance the spatial resolution of photoreceptors and map blood flow within the eye [30,31]. This technology also facilitates the exploration of glaucoma pathophysiology, offering insights into disease progression and monitoring. For example, weakly supervised deep learning algorithms were trained on AO-OCT to quantitatively assess retinal ganglion cell and displaced amacrine cell features in a pilot study. Utilizing a K-Nearest Neighbor classifier, these retinal ganglion cell and amacrine cell features, such as their diameters, were used to effectively distinguish between four healthy and five glaucomatous eyes with a sensitivity and specificity of 1.0 and 0.75, respectively [32,33].

2.3. Integrating Multiple Modalities Using AI for Detection and Diagnosis

By training on extensive datasets, AI algorithms can learn to analyze various forms of historical patient data all at once, including demographics, clinical findings, ocular imaging, and IOP readings, among many other forms of data. Research indicates that multimodal models outperform those restricted to singular data types in predicting glaucomatous optic neuropathy [15,23,34–36]. A study from Taiwan developed a multimodal model called Xception using fundus photographs, predicting RNFL thickness values, vertical cup-to-disc ratios, and numerical data from past clinical exams to diagnose glaucoma [37]. The model
classified normal, pre-perimetric glaucoma, and glaucoma with an AUROC of 93.9% in a population of patients with a high incidence of myopia [37]. The integration of not only multimodal inputs but also multiple classifiers as seen with ensemble techniques has emerged as an effective strategy for enhancing AI performance. For example, an ensemble model that utilized a best–worst weighted voting system, combining various classifiers trained on both color and grayscale images to identify glaucomatous changes, demonstrated superior performance with an accuracy of 99.78% and AUROC of 0.992 compared to each of its individual component models, which had an average accuracy of 99.21% and an AUROC of 0.981 [38].

2.4. Telemedicine and Remote Monitoring

The integration of AI into telemedicine has opened new avenues in glaucoma management, particularly benefiting patients in remote regions or those with mobility concerns, facilitating screening and referrals from primary care settings, and offering solutions for individuals unable to access traditional clinical services [34,39,40]. Telemedicine technologies are now capable of distinguishing between various eye diseases, with successful applications including the detection of glaucoma within diabetic populations [41]. AI-powered applications can analyze patient-reported data or measurements from take-home devices, such as a portable tonometer, VF assessment (Eyecatcher), or ERG measurements (RETeval), the best models of which were able to achieve an accuracy of 93% in classifying glaucomatous versus non-glaucomatous eyes [42–44]. These data can be used to monitor disease progression or treatment efficacy, reducing the need for frequent in-person visits. Moreover, AI-driven telemedicine platforms can provide patients with personalized feedback and reminders for medication adherence, appointment scheduling, and lifestyle modifications [45]. Such interactive systems may enhance patient engagement and motivate self-management of the disease. For physicians, on the other hand, AI can serve as a decision support tool in telemedicine settings, offering recommendations based on remote data. Screening tools can be applied in telemedicine contexts to help identify candidates for referral to ophthalmology [34,39].

3. AI in Monitoring Glaucoma Progression and Prediction

Effective glaucoma management hinges on the vigilant monitoring of disease progression to avert permanent vision impairment. AI introduces a promising avenue for augmenting the monitoring process, notably by streamlining the analysis of imaging data within time-constrained clinical environments. This section delves into how AI technologies can transform the analysis of glaucoma progression.

3.1. AI in Functional Imaging for Progression Monitoring and Prediction

AI models can detect subtle patterns of VF loss and predict future deterioration more accurately and earlier than traditional methods. A longitudinal study analyzing VFs from 2085 eyes of 1214 individuals sought to identify glaucomatous progression, revealing that AI models required the least amount of time to detect progression—identifying changes in 25% of eyes within 3.5 years. This surpasses traditional methods including global mean deviation, region-wise analyses that take 10 sections of the VF and point-wise analyses that track changes in individual points on the VF, with the most effective of these established methods taking 3.9 years for progression detection [46].

AI algorithms extend their capabilities to analyzing serial VF tests as well, offering insight into disease trajectory beyond the mere prediction of VF progression. For example, Elze et al. applied an archetypal analysis to characterize 17 distinct glaucomatous archetypes [47]. Building on this research by using the Ocular Hypertension Treatment Study database [48], Singh et al. employed the archetypal analysis to identify archetypes with increased risk of developing primary open angle glaucoma with a correlation coefficient of 0.75. Moreover, they discovered archetypes that were strongly associated with accelerated disease progression [49]. By recognizing and classifying the nuances of
glaucoma progression, AI can contribute to a more personalized, effective approach to glaucoma care.

3.2. AI in Structural Imaging for Progression Monitoring and Prediction

AI has shown remarkable proficiency in analyzing OCT images and fundus photography for signs of glaucoma progression. Similar to their success in diagnosing glaucoma, CNNs have been particularly effective in analyzing OCT images and fundus photography to predict glaucomatous progression [50]. These models have achieved accuracies of up to 93.7% with sensitivities and specificities of 0.91 and 0.97, respectively, in binary glaucoma diagnosis tasks using either OCT or fundus photography [51–53]. One longitudinal study developed models to predict glaucoma incidence and progression using color fundus photographs from 17,497 eyes of 9346 patients [54]. The model demonstrated AUROCs of 0.90 and 0.91 in predicting glaucoma incidence and progression, respectively. A similar study investigating disease progression in patients with normal tension glaucoma developed AI models to automatically measure retinal–vessel calibers from fundus photographs. The authors found that a decrease in vessel caliber by one standard deviation was associated with a greater than 30% increase in RNFL thinning, as well as a 90% risk of VF deterioration during the follow-up period (an average of 34.4 months), and subsequent glaucoma progression [55].

3.3. Integrating Multiple Modalities Using AI into Progression Monitoring and Prediction

Given that glaucoma typically progresses slowly, accurately predicting its rate of advancement can significantly improve patient outcomes. Conventional diagnostic tools such as OCT and VF tests, when used in isolation, suffer from patient-dependency and variability, posing challenges with reliable disease progression tracking. Addressing these challenges, Hussain et al. designed an ensemble of CNN and long short-term memory (LSTM) models trained on multimodal data including OCTs, VF data, patient demographics, and clinical history data over a 12-month period [56]. This ensemble model was able to predict glaucomatous progression with an AUROC of 0.83 in just five visits worth of data 6 months prior to VF progression. Furthermore, by utilizing synthetic images created by a generative adversarial network (GAN), the model enhanced its predictive performance, forecasting vision loss nine months in advance with an AUROC of 0.81. This marked a significant improvement over models that relied on a single modality, whether by structural (AUROC = 0.68) or functional (AUROC = 0.72) tests. The integration of diverse data sources also appears to enhance the predictive capabilities of transformer architectures. This was demonstrated in a study by Herbert et al., who developed a model that incorporated a transformer-based feature extractor with a linear regression model to merge VF, OCT, and clinical data. Their approach successfully identified rapidly progressing glaucoma with a high degree of accuracy, achieving an AUROC of 0.87 [57].

These predictive AI models, especially those that are multimodal in nature, can be beneficial in identifying patients who are at high risk of glaucoma disease progression. By detecting the potential trajectory of the disease earlier, ophthalmologists can personalize monitoring frequency and treatment plans to match individual patient needs.

4. AI in Glaucoma Treatment

The management of glaucoma encompasses a range of pharmacological treatments and procedural interventions. Selecting the optimal treatment strategy is complex, owing to the diverse factors unique to each patient. Decisions regarding treatment necessitate a deep understanding of underlying disease mechanisms and how various interventions alter ocular physiology [58–60]. AI may offer valuable assistance in this domain, enhancing surgical precision and predicting patient responses to treatments. This section delves into the multifaceted roles AI plays in the treatment of glaucoma, illustrating its potential to refine and personalize therapeutic approaches.
4.1. AI in Surgical Interventions

Glaucoma procedures, such as trabeculoplasty, trabeculectomy, or microinvasive glaucoma surgery (MIGS), demand high precision to achieve optimal patient outcomes. AI can assist surgeons in the preoperative planning phase, intraoperatively, and post-operatively. Preoperatively, AI algorithms can use a comprehensive array of clinical parameters to predict the success, risks, and potential complications of various surgical approaches [60]. This assessment can include factors such as optic nerve damage severity, VF loss, patient history of treatment responses, ocular anatomical parameters, as well as electronic health record data, enabling personalized surgical strategy selection and prognostic predictions [61,62]. For example, Baxter et al. used machine learning-based predictive modeling to forecast the need for surgical intervention in 385 patients with primary open-angle glaucoma using systemic health data and medications from electronic health records data [63]. The model achieved an AUROC of 0.67 and found that the most predictive feature from health records that associated with significantly increased odds of requiring glaucoma surgery was higher mean systolic blood pressure (odds ratio: 1.09, 95% CI: 1.06–1.13), whereas other parameters, such as ophthalmic medications, were associated with decreased odds of needing surgery.

Intraoperatively, AI has the potential to provide real-time feedback using live-feed video transmitted directly from the microscope. For example, Lin et al. used a neural encoder–decoder CNN (VGG16 U-Net) to identify the trabecular meshwork during gonioscopy [64]. This model was trained using 229 gonioscopic photographs and 149 laser trabeculoplasty video frames. Their model demonstrated remarkable accuracy, achieving a mean deviation of 0.8% compared to expert-derived annotations, and significantly surpassing manual identification by ophthalmologists and residents in various conditions, including changes in magnification, contrast, and lighting, as well as in the presence of instruments such as a trabectome. Such precision in localizing the TM in real-time can offer benefits for intraoperative surgical assistance and surgical training, showcasing the practical utility and clinical applicability of AI in enhancing glaucoma surgery outcomes.

In another study, a region-based CNN model trained on 600 extracted frames from ten stereoscopic videos of cataract surgery performed by attending physicians and surgical trainees demonstrated its abilities in providing real-time surgical guidance for phacoemulsification cataract surgery [65]. The model was able to perform real-time pupil tracking, segmentation, and identification of the current surgical phase with great accuracy (AUROC = 0.95), while also providing live intraoperative visual feedback such as harmful turbulence and instrument movements to improve tissue visualization. Although this work focused on cataract and not glaucoma surgery, their findings still help to elucidate that an AI-based surgical guidance platform has the potential to enhance intraoperative surgeon experience.

4.2. Treatment Response Prediction

In the era of personalized medicine, AI stands as a pivotal tool in tailoring healthcare to the individual needs of patients. Leveraging its capability to process vast datasets, AI enables the tailored prediction of patient responses to various glaucoma treatments, encompassing medications, laser therapy, and surgical interventions. For example, one innovative surgical paradigm, MIGS, offers less invasive alternatives to traditional IOP-lowering surgeries. Given the variability in patient responses due to individual characteristics, selecting the optimal MIGS approach presents a significant clinical challenge. However, one trial from the United Kingdom harnessed an Adaptive Neuro-Fuzzy Inference System (ANFIS) to navigate this complexity [58]. This system, designed to offer treatment recommendations, analyzed baseline clinical data—such as baseline demographics, IOP, VF scores, visual acuity, and glaucoma subtype—to guide ophthalmologists in choosing the most suitable MIGS procedure for each patient. Drawing on a dataset from 372 patients who underwent one of four MIGS options, the ANFIS model achieved an impressive 91% prediction accuracy of surgical success, with a sensitivity of 0.80 and a specificity of 0.90. These results...
underscore the potential of AI to significantly enhance clinical decision-making, providing a data-driven basis for personalized treatment strategies in glaucoma care.

AI can also predict postoperative outcomes in patients with glaucoma. One study developed a random forest model using preoperative data extracted from electronic health records to predict patient outcomes after trabeculectomy [66]. It was trained on health records (including demographic, systemic, and ocular health data) from 230 trabeculectomy procedures performed on 184 patients. It achieved an accuracy of 68% with an AUROC of 0.74, supporting the notion that machine learning models may in the future support physicians and patients with surgical decision-making.

5. Ethical, Legal, and Social Implications

5.1. Ethical Considerations

As AI integrates deeper into the realm of glaucoma diagnostics and treatment, ethical considerations become paramount. The primary concern revolves around patient privacy and data security, given the extensive collection and processing of personal health information. Ensuring that patient data remains confidential and protected against breaches is a fundamental ethical obligation. Additionally, there is the matter of algorithmic bias, where AI models might inadvertently reflect or amplify existing disparities in healthcare access or outcomes among different populations. This raises the necessity for transparent, equitable AI systems that are rigorously validated across diverse demographic groups to prevent any form of discrimination or inequality in patient care. A study on algorithmic bias evaluated an AI model developed from Ocular Hypertension Treatment Study data [48] to diagnose primary open-angle glaucoma. The validation process involved stratifying a large-scale, longitudinal dataset by sex, race, and age, revealing that the model under-diagnosed females under 60 years of age and over-diagnosed Black females [67]. This was measured by comparing false negative rates, which were higher in females, non-Black individuals, and younger individuals under the age of 60. The study also examined intersectional groups, finding that biases were amplified in individuals who belonged to two or more subpopulations, such as Black females and younger individuals, who showed higher underdiagnosis rates compared to their counterparts. Such biases pose serious risks, as underdiagnosis can deprive patients of necessary treatment, whereas overdiagnosis can lead to the administration of unwarranted therapies. This underscores the necessity for bias mitigation strategies that consider multiple characteristics influencing bias, rather than a focus on a single identifying factor. Given the demographic variations in glaucoma prevalence and response to treatment, implementing bias mitigation testing as a routine component of AI tool development and validation is crucial to ensure equitable healthcare outcomes.

Furthermore, the advent of AI in healthcare introduces questions about accountability, particularly in the face of diagnostic errors or adverse treatment outcomes. It is essential to maintain clear lines of responsibility among healthcare providers, AI developers, and other stakeholders, emphasizing the supplementary role of AI in enhancing, not replacing, the clinician’s judgment. Preserving the patient–physician relationship and ensuring informed patient consent are paramount in this new technological landscape, reinforcing the importance of human oversight in AI-assisted clinical decision-making.

5.2. Legal Implications

The regulatory landscape of AI in healthcare is continually evolving, with significant variations across jurisdictions. Ensuring that AI applications in glaucoma management comply with medical device regulations, data protection laws, and other relevant statutes is critical. The regulation of Software as a Medical Device (SaMD) encompasses various requirements tailored to ensure patient safety and product efficacy. Among these requirements is the stipulation for a “locked” algorithm, which mandates that the algorithm’s output remains consistent across its applications. This characteristic, while ensuring stability and reliability, poses a constraint on the inherent adaptive capabilities of AI technologies, which are designed to refine their responses as they encounter new data [68]. The FDA

has approved several AI models with locked algorithms, highlighting the potential for regulatory acceptance. However, this necessitates a delicate balance between harnessing the adaptive nature of AI and adhering to rigorous regulatory standards that ensure these technologies do not deviate from their intended purpose or pose unforeseen risks to patients.

Moreover, the legal frameworks concerning liability in the use of AI in medical practice are yet to be fully developed. There is a pressing need for consensus and clarity on protocols that safeguard patient interests, ensuring accountability and protection in the event of adverse outcomes [69]. Such focused efforts are crucial for fostering trust and confidence among healthcare professionals and patients alike, paving the way for the widespread adoption of AI in improving patient care and outcomes in glaucoma care and beyond.

5.3. Social Implications

The deployment of AI in glaucoma management also entails profound social implications. On one hand, AI has the potential to democratize access to high-quality eye care, particularly for underserved populations or those in remote areas through telemedicine platforms. By automating certain diagnostic and monitoring tasks, AI can help bridge gaps in healthcare access, reduce wait times for appointments, and alleviate the burden on healthcare systems [70,71]. On the other hand, there exists the risk of exacerbating healthcare disparities if AI technologies are primarily available only in wealthier, urban centers, or if they require resources that are out of reach for lower-income individuals [72]. Additionally, the shift towards AI-driven care necessitates digital literacy among patients and healthcare providers, highlighting the need for comprehensive education and training programs to ensure widespread adoption and effective use.

In summary, while AI presents transformative opportunities for enhancing glaucoma diagnosis and management, navigating the ethical, legal, and social landscapes requires concerted efforts from all stakeholders to realize its full potential in a manner that is equitable, responsible, and beneficial to society at large.

6. Limitations of AI in Glaucoma

While AI holds significant promise for revolutionizing glaucoma management, it is not without its limitations. Understanding these challenges is crucial for refining AI applications and ensuring their effective integration into clinical practice.

The efficacy of AI systems in diagnosing and managing glaucoma heavily relies on the quantity and quality of data on which they are trained [15,32,33]. These systems require large datasets of annotated images and patient records to learn and make accurate predictions [73,74]. However, the acquisition of high-quality, comprehensive datasets poses several challenges. Firstly, privacy regulations and ethical considerations surrounding patient data significantly limit the availability of datasets for training AI models. Institutions must navigate complex legal frameworks to share clinical data, often resulting in fragmented datasets that lack the diversity necessary for training robust AI systems. Moreover, the inconsistency in data collection standards and practices across different healthcare settings can introduce variability in the data. AI systems learn from the data provided to them, such that their training and validation are highly dependent on the diversity and representativeness of the datasets used; if the data contain biases—whether due to demographic skews, socioeconomic factors, or differences in clinical presentations of glaucoma or in imaging techniques—the AI model may replicate or even amplify these biases in its predictions and recommendations. For instance, deep learning algorithms may exhibit different accuracies when presented with different ethnic groups depending on fundus pigmentation and optic disc sizes [75]. Moreover, most datasets used for AI model training include images or data from European or Asian backgrounds, with lower African representation [76]. A lack of dataset diversity may cause the AI models to perform poorly when presented with newly acquired data from a different population in a real-world clinical setting, potentially leading to underdiagnosis or overdiagnosis of patients.
belonging to specific demographic backgrounds. This situation worsens the generalizability of AI models and can lead to disparities in the accuracy and effectiveness of AI-driven diagnostics and treatment plans across different patient groups or clinical settings [67].

Integrating AI into the glaucoma workflow requires thoughtful consideration to ensure it complements rather than complicates existing practices. Seamless integration demands not only the technical compatibility of AI systems with existing healthcare IT infrastructure but also the adaptation of clinical processes to fully leverage AI’s capabilities. This transition often necessitates targeted training for healthcare professionals, who must understand how to interpret AI-generated insights and incorporate them into their decision-making processes [77,78]. Additionally, the integration must be designed to enhance, not hinder, the efficiency of clinical operations, ensuring that the use of AI technologies does not introduce unnecessary steps or delays in patient care. Achieving this balance is critical for maintaining the trust and confidence of both clinicians and patients in AI as a valuable tool for glaucoma management [78]. The ultimate goal is to foster an environment where AI acts as a silent partner, augmenting the clinician’s expertise and facilitating a more personalized and effective approach to patient care.

Studies on the real-world applications of AI models in glaucoma care, specifically those providing evidence of their contributions in the context of improving glaucoma outcomes, are very limited. Most studies in the literature and those included in this review only focused on reporting the performances of various AI models and did not proceed further to report on the clinical outcomes following their implementation. For example, Lin et al. implemented a real-time CNN model to identify the trabecular meshwork in gonioscopy surgical videos, and it achieved remarkable accuracy, surpassing manual identification by ophthalmologists and residents in various conditions [64]. While not yet tested in real-world surgical settings, such precision in localizing the trabecular meshwork in real-time shows the potential to offer benefits for intraoperative surgical assistance and surgical training, especially for training ophthalmologists.

Furthermore, the financial aspects of implementing AI in clinical practice are non-trivial [79], encompassing not just the initial outlay for technology acquisition but also the expenses associated with integrating these systems into existing healthcare infrastructures. Training staff, updating IT systems, and the ongoing maintenance of AI technologies further contribute to the total cost of ownership. For many healthcare providers, particularly those in under-resourced settings, these expenses pose a formidable barrier to adoption. This financial burden may limit access to AI-enhanced care, potentially widening the gap in healthcare quality between well-funded urban centers and rural or economically disadvantaged areas. Overcoming this hurdle necessitates evidence of clear long-term cost benefits, such as reduced diagnostic errors, more efficient management, and ultimately, decreased healthcare costs through early intervention and personalized treatment strategies.

Finally, the issue of algorithmic transparency and interpretability in AI applications presents a significant challenge for clinical integration. AI models often lack understandable explanations for their outputs [80,81]. This opacity can hinder clinician trust, as healthcare professionals rely on clear, logical reasoning to make informed decisions about patient care. The complexity of these algorithms means that even their creators can struggle to elucidate how specific conclusions are drawn. This “black box” nature of AI complicates efforts to validate and trust AI-assisted diagnostic or treatment recommendations. To foster a collaborative environment where AI and clinicians work in tandem to improve patient outcomes, efforts must be directed towards developing more interpretable AI systems. This could involve creating models that not only predict with high accuracy but also provide insights into the features and reasoning behind their predictions, thereby enhancing clinician understanding and confidence in utilizing AI as a decision support tool for glaucoma.
7. Future Directions

The journey of integrating AI into glaucoma management is ongoing, with the potential to transform patient care in profound ways. As we look to the future, several key areas emerge where efforts should be concentrated to overcome current limitations and harness the full capabilities of AI (Figure 2).

A critical step forward involves the enhancement of data quality and the diversification of datasets used to train AI models. Future work should aim to gather large, high-quality, and expert-annotated datasets that are truly representative of the population of glaucoma patients that will be served. This initiative requires overcoming privacy concerns through stringent data protection measures and anonymization techniques ensuring patient confidentiality is never compromised. Moreover, establishing universal standards for data collection, annotation, and storage will enable AI systems to learn from a broader spectrum of patient demographics and disease manifestations. By embracing a standardized approach, the development of AI models becomes more scalable and inclusive, allowing for the creation of tools that are not only more accurate but also universally applicable across different clinical settings. This evolution is essential for paving the way toward a future where AI-driven insights can be seamlessly integrated into personalized glaucoma care.

In terms of dataset diversification, there have been several attempts to address these challenges, and one such example was the prospective study by Kang et al., which implemented a machine-learning-based archetypal analysis that could identify the ethnic differences in the risks and regional patterns of visual field loss [82]. When compared to the non-Hispanic Caucasian group, Black participants showed higher risks of incident primary open-angle glaucoma with central and advanced visual field loss, suggesting that glaucoma subtyping using machine learning approaches can identify unique risk factors that may help to improve the generalizability of AI models across different populations. In another study by Shi et al., techniques aimed at adjusting AI models to reduce bias, such as fair identity normalization (FIN), were shown to improve the performance in glaucomatous status detection amongst vulnerable populations [83]. In comparison to the standalone deep learning model, the model combined with FIN enhanced the overall performance and screening accuracy in terms of AUROC by 0.03 and equity-scaled AUROC by 0.05 for racial groups, with noticeable improvements in AUROCs when stratified by race, ethnicity, and...
gender. Nevertheless, future studies are required to validate these methods and to continue improving on novel approaches with hopes to fully address the challenges associated with lack of dataset diversity and move towards the development of more robust and equitable AI solutions.

Advancements in explainable AI (XAI) represent a pivotal frontier in making AI more accessible and trusted within the medical community. By developing algorithms that not only make predictions but also provide insights into their reasoning, XAI aims to bridge the gap between AI’s computational complexity and the practical needs of clinical decision-making [84,85]. Many techniques have been developed in attempts to provide explanations for how AI models reach their decisions, and these include Shapley additive explanations, local interpretable model-agnostic explanations, and gradient-weighted class activation maps [85–89]. While these methods have demonstrated clinical utility in helping users determine how AI models arrive at certain diagnostic decisions, no single technique independently provides the level of interpretability clinicians require to fully understand the algorithms. Therefore, it is likely necessary to employ multiple explainability algorithms that can work synergistically [90]. In the future, explainable models may eventually be able to facilitate informed discussions with patients about their care plans, enhancing patient trust and engagement. As research in XAI progresses, it will play a crucial role in demystifying AI processes, ensuring that AI tools become integral, understandable, and trusted components of glaucoma care.

For AI to realize its potential in glaucoma care, its integration into clinical workflows must be seamless and intuitive. Key to this integration is designing AI tools that complement existing healthcare practices without requiring extensive modifications to current procedures. This involves developing user-friendly interfaces, tailored training programs, and AI solutions that are adaptable to the diverse needs and contexts of different healthcare settings. Prioritizing these aspects can help ensure AI integration complements rather than complicates patient care.

In addition to integration into clinical workflows, healthcare professionals should also recognize the importance of interdisciplinary collaboration in the development and implementation of AI technologies in glaucoma care. Collaboration between ophthalmologists, AI researchers, ethicists, patients, and policymakers is crucial for ensuring that AI tools are not only technologically advanced but are also ethically sound, clinically relevant, tailored to meet the needs of patients, and accessible across diverse healthcare settings. Prioritizing these aspects can help ensure AI integration complements rather than complicates patient care.

To fully realize the benefits of AI in glaucoma, conducting in-depth cost-effectiveness and accessibility studies will be essential. These studies should aim to quantify the long-term economic benefits of integrating AI tools, such as reduced healthcare costs due to earlier detection and more efficient disease management, against the initial investment and operational expenses. Highlighting AI’s potential for cost savings and improved patient outcomes can encourage healthcare systems to adopt these technologies more widely. Additionally, focusing on making AI solutions accessible in resource-limited environments is crucial for mitigating healthcare disparities. By developing scalable and affordable AI applications, we can extend the reach of advanced glaucoma care to underserved populations, ensuring that the advantages of AI in healthcare are universally accessible. This approach not only enhances patient care globally but also aligns with the broader goal of achieving equitable healthcare access.

8. Conclusions

This review explored various applications of AI in glaucoma practice, from enhancing diagnostic accuracy and monitoring disease progression to optimizing management pathways and predicting treatment responses. The ability of AI to study complex ocular tests and predict disease progression demonstrates its promising potential as a clinical tool in early detection and personalized treatment planning. AI’s ability to process and learn through large volumes of datasets and detect subtle disease patterns that are virtually unnoticeable to the eyes of experts showcases a more accurate and effective glaucoma diagnostic
future. However, the limitations highlighted within this review, including training data accessibility, selection bias, lack of qualitative or quantitative descriptors of improved care, and the challenges of healthcare integration, remind us that lots of work still lies ahead for the reliable implementation of AI in glaucoma management.

**Author Contributions:** Conceptualization, P.X.J., M.B. and D.J.M.; methodology, P.X.J., M.B. and D.J.M.; data curation, P.X.J., V.R. and L.P.; writing—original draft preparation, P.X.J., V.R., L.P. and M.B.; writing—review and editing, P.X.J., V.R., L.P., M.B. and D.J.M.; supervision, D.J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

**References**


86. Li, C. Glaucoma Detection Based on Optical Coherence Tomography Imaging. Master’s Thesis, Nanyang Technological University, Singapore, 2023. [CrossRef]


90. Hossain, I.; Zamzmi, G.; Mouton, P.R.; Salekin, S.; Sun, Y.; Goldgof, D. Explainable AI for Medical Data: Current Methods, Limitations, and Future Directions. *ACM Comput. Surv.* 2023. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.