End-to-End Glaucoma Detection using Deep Learning: Recent Progress and Perspectives

Vaibhav Yadav[1], *Saumya Das*[2], *Barnali Dey*[1], *Udayan Baruah*[3], *Hashinur Islam*[4], *Manas Kanti Saha*[2]

[1]Department of Information Technology, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Majitar, Rangpo, Sikkim, India
[2]Department of Computer Science and Engineering – Artificial Intelligence, Brainware University, Kolkata, West Bengal, India
[3]Birangana Sati Sadhani Rajyik Vishwavidyalaya (A state university under Government of Assam), Golaghat, Assam, India
[4]Department of Electronics and Communication Engineering, HKBK College of Engineering, Bengaluru, Karnataka, India & VTU, Belagavi, India

Abstract. Glaucoma is the leading cause of irreversible blindness globally, with its diagnosis remaining a complex challenge due to the variability and intricacy of clinical assessments. Recent progress in deep learning-based artificial intelligence (AI) offers promising avenues for automating glaucoma detection. This review provides a comprehensive survey of state-of-the-art deep learning methods applied to glaucoma diagnosis using data from fundus photography, optical coherence tomography (OCT), and visual field tests. Relevant studies published between 2020 and 2026 were selected from databases such as ScienceDirect, Google Scholar, and IEEE Xplore, based on well-defined inclusion criteria and search keywords. The review categorizes key deep learning approaches into several groups: convolutional neural networks (CNNs), autoencoder-based models, attention mechanisms, generative adversarial networks (GANs), geometric deep learning frameworks, Transformer-based architectures, reinforcement learning models, self-supervised learning techniques, recurrent neural networks (RNNs) including LSTMs, diffusion models, and hybrid configurations. Notable challenges include the scarcity and lack of diversity in available datasets, difficulties in integrating multimodal data, and the limited interpretability of AI models from a clinical perspective. This article aims to support AI researchers in selecting appropriate deep learning frameworks for glaucoma detection by considering factors such as data characteristics, model architecture, and clinical applicability.

Keywords: Glaucoma Detection, Deep Learning, Computer-Aided Glaucoma Detection System

1 Introduction

Glaucoma is one of the leading causes of blindness worldwide. At present, globally about 80 million people are suffering from glaucoma [1-2] at present. If not treated in time, glaucoma can lead to permanent vision loss because of the damage of optic nerve head and retinal nerve fibre layer. Despite the advent of advanced treatments, many people lose vision because of late diagnosis. Fig. 1 illustrates a gradual loss of vision due to glaucoma [3]. Conventional glaucoma diagnosis is complex, time-consuming and expensive. Therefore, the time has come to develop a simple, less time consuming, inexpensive glaucoma detection method that can save the vision of many people. Recent advances in artificial intelligence are motivating researchers to use machine learning for glaucoma detection [4-6].

Fig. 1. Gradual loss of vision due to glaucoma [3]

But because of unreliable handcrafted feature extraction and lack of generalizability among the patient population, conventional machine learning is unsuccessful to deliver satisfactory outcome in glaucoma detection. But deep learning's automatic feature selection and natural pattern identification characteristics have inspired researchers to develop a reliable glaucoma detection method [7-10]. It is expected that these deep learning-based glaucoma detection methods will soon be used in clinical practice as well.

In this review, an overview of different deep learning-based glaucoma detection methods has been illustrated. Different categories of deep learning-based glaucoma detection methods such as Convolutional neural networks-based glaucoma diagnosis, Autoencoder based network for glaucoma diagnosis, Attention-based networks for glaucoma diagnosis, Generative adversarial networks for glaucoma diagnosis, Geometric deep learning networks for glaucoma diagnosis, Transformer architecture for glaucoma diagnosis, Reinforcement Learning (RL)-based Deep Models for glaucoma diagnosis, Self-supervised Learning (SSL) Architectures for glaucoma diagnosis, Recurrent Neural Networks (RNNs) / LSTMs for glaucoma diagnosis, Diffusion Models for glaucoma diagnosis, and Hybrid networks for glaucoma diagnosis have been illustrated citing the existing work in recent years. Finally, a conclusion has been drawn to display the uses and challenges of various deep learning-based glaucoma detection methods.

2 Article Searching Strategy

In this work, published articles in the duration 2020 to 2026 have been considered based on deep learning-based glaucoma detection. Scholarly databases such as IEEE Xplore, Science Direct, Google Scholar and Springer Nature have been explored for searching the existing research. Primarily keywords like ‘glaucoma detection’, machine learning’, ‘deep learning’, ‘artificial intelligence’ have been applied to find the relevant research in recent years. For final consideration two phase screening has been applied on preliminary collected articles. In the first phase, screening is done on Title/Abstract of the article. If it is found relevant then in the second phase, the methods and the outcome have been accepted and listed in comparison table. Fig. 2 shows the searching process diagrammatically for this review article.

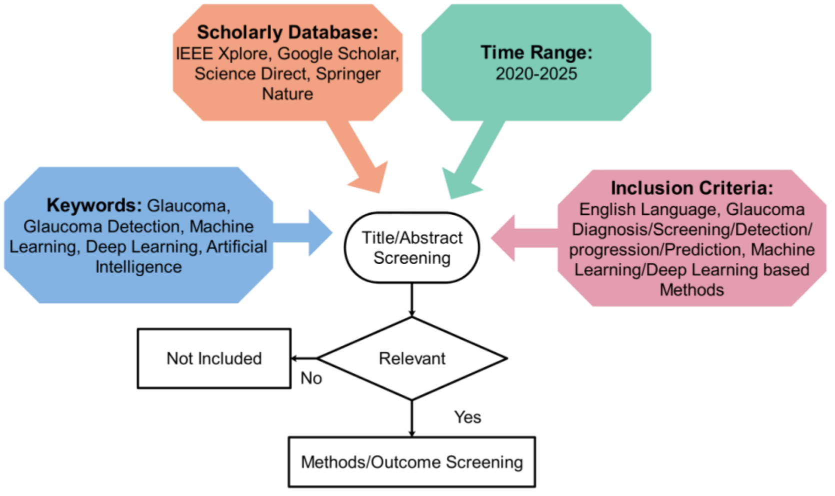

Fig. 2. Article Searching Strategy

3 Glaucoma: Definition, Anatomical features and Terminologies

Glaucoma is an eye disease caused by damage to retinal ganglion cells (RGCs), leading to morphological changes in the retina. These changes result in a thinner Retinal Nerve Fibre Layer (RNFL), narrower Neuroretinal Rim (NRR), and greater Cup-to-Disc ratio (CDR) [11]. Apart from these morphological changes, glaucoma also results in functional changes, which impairs visual field sensitivity. Relevant terminologies [12] used in glaucoma diagnosis are listed in Table 1. Medical imaging plays an important role in detecting both the structural and functional changes caused due to glaucoma. But for extracting insights from the images, it requires long term expertise.

Fundus imaging captures fine images of the retina and optic disc. These images identify any structural changes in OC, OD, RNFL and NRR. Fundus imaging can identify the morphological changes in optic nerve and blood vessels. Fundus imaging is an effective diagnostic tool for detecting a wide range of ocular diseases, including glaucoma. Another non-invasive image modality method is called optical coherence tomography (OCT) [13-14]. It delivers detailed cross-sectional imaging of the retina, optic nerve, and other internal eye structures [15].

4 Anatomical Terminology used in Glaucoma Detection

Table 1. Anatomical terminologies related to glaucoma detection

Term	Abbreviation	Description
Retinal Ganglion Cells	RGCs	Neurons located in the innermost retinal layer that transmit visual signals from the eye to the brain.
Inferior Superior Nasal Temporal	ISNT	Represents the typical order of neuroretinal rim thickness around the optic disc: Inferior (I), Superior (S), Nasal (N), and Temporal (T).
Ganglion Cell Layer	GCL	A retinal layer containing ganglion cells that carry visual data to the brain through the optic nerve.
Inner Plexiform Layer	IPL	A region in the retina where bipolar, amacrine, and ganglion

		cells form synapses to facilitate signal transmission.
Ganglion Cell Complex	GCC	A combination of the retinal nerve fiber layer, ganglion cell layer, and inner plexiform layer, essential for visual signal processing.
Ganglion Cell–Inner Plexiform Layer	GCIPL	Refers to the ganglion cells and adjacent inner plexiform layer, involved in conveying visual input to the brain.
Optic Nerve Head	ONH	The junction where the optic nerve exits the eye, channelling visual data to the brain.
Neuroretinal Rim	NRR	Tissue encircling the optic disc made up of ganglion cell axons, responsible for signal transmission from the retina.
Optic Disc	OD	A circular region on the retina where the optic nerve emerges.
Optic Cup	OC	The central depressed area within the optic nerve head.
Cup-to-Disc Ratio	CDR	The proportion of the optic cup's diameter relative to the optic disc's diameter, often used in glaucoma assessment.
Intraocular Pressure	IOP	The fluid pressure maintained within the eyeball.
Retinal Nerve Fiber Layer	RNFL	Comprised of axons from retinal ganglion cells, forming part of the optic nerve and retina.
Retinal Blood Vessels	RBVs	Vessels that deliver nutrients and oxygen to the retina and remove metabolic waste.
Cornea	–	The transparent front part of the eye that helps focus light and provides a protective barrier.

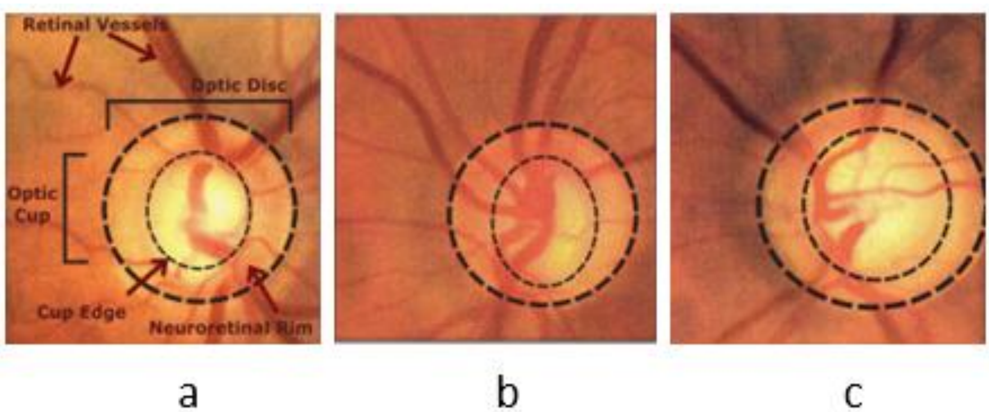

Fig. 3. Optic Disc and Optic Cup in (a) healthy eye (b) early stage of glaucomatous eye (c) advance stage of glaucomatous eye [15]

Although fundus and OCT imaging techniques are commonly used to record morphological changes associated with glaucoma, they cannot detect functional changes induced by glaucoma. Therefore, another parameter called Visual Field (VF) is used to detect the functional changes caused by glaucoma. It measures the sensitivity of the vision. For measuring the VF sensitivity of a patient, a visual stimulus is applied to patient's eye following the process of standard automated perimetry. Fundus, OCT and VF modalities are integrated [16-17] along with expert level features and biomarkers like Intraocular pressure (IOP) to achieve higher accuracy and better performance from deep learning models.

5 Available datasets and performance metrics

In this section, datasets used in this article have been listed in Table 2. Datasets are created with information taken from fundus, OCT or VF modalities. Sometimes they have been combined to prepare a more robust dataset.
The performance of the deep learning model is mainly evaluated through metrics like accuracy, specificity, sensitivity, precision and F1 score. Formal definitions of the above-mentioned metrics are displayed in Eq. (1-5) [18-21].

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Instances} \quad (1)$$

$$Specificity = \frac{True\ negatives}{True\ negatives + False\ Positives} \quad (2)$$

$$Sensitivity = \frac{Total\ Positives}{Total\ Positives\ +\ False\ Negatives} \quad (3)$$

$$Precision = \frac{Total\ Positives}{Total\ Positives + False\ Positives} \quad (4)$$

$$F1\ Score = \frac{2\ \times\ Sensitivity\ \times\ Precision}{Sensitivity\ +\ Precision} \quad (5)$$

True positives are events that the model accurately identifies as positive. False negatives are occasions where the model wrongly classifies something as negative when it is actually positive. True negatives are instances that the model accurately identifies as negative. False Positives are events that are genuinely negative but are misclassified as positive by the model. Higher accuracy, specificity, sensitivity, precision, and F1 Score values suggest a better classification model.

Table 2. Available Datasets

Dataset Information	**Number of images: Glaucomatous (G), Healthy (H), Total (T)**	**Data Type**	**Resolution**	**Ref.**
REFUGE	G: 121, H: 1079, T: 1200	Fundus	2124 × 2056, 1634 × 1634	[22]
LAG	G: 4878, H: 6882, T: 11760	Fundus	500 × 500	[23]
RIM-ONE DL	G: 172, H: 313, T: 485	Fundus	Many types	[24]
AIROGS	G: 3270, H: 98172, T: 101442	Fundus	Many types	[25]
PAPILA	G: 155, H: 333, T: 488	Fundus	2576 × 1934	[26]
ODIR-5K	G: 307, H: 1620, T: 5000	Fundus	Many types	[27]
JSIEC	G: 13, H: 54, T: 1087	Fundus	Many types	[28]
FDA	G: 10, H: 6, T: 16	OCT	297 × 259 × 450	[29]
AGE	G: 300, T: 300	OCT	Many types	[30]
Drishti -GS1	G: 70, H: 31, T: 101	Fundus	2896 × 1944	[31]
GRAPE	G: 1115, T: 1115	Fundus & VF	Many types	[32]

6 End to end deep learning-based processes for glaucoma detection

Deep learning-based methods have shown promising outcomes for the purpose of end-to-end glaucoma detection. They have consistently outperformed feature-based method of detection [33-40]. The characteristics of acquiring comprehensive contextual information, minimizing information loss and promoting generalizability, made deep learning models to achieve better accuracy and reliability in glaucoma detection. Deep learning-based glaucoma detection architecture mainly divided into the following categories:

1. Convolutional neural networks-based glaucoma diagnosis
2. Autoencoder based network for glaucoma diagnosis
3. Attention-based networks for glaucoma diagnosis
4. Generative adversarial networks for glaucoma diagnosis
5. Geometric deep learning networks for glaucoma diagnosis
6. Transformer architecture for glaucoma diagnosis
7. Reinforcement Learning (RL)-based Deep Models for glaucoma diagnosis
8. Self-supervised Learning (SSL) Architectures for glaucoma diagnosis
9. Recurrent Neural Networks (RNNs) / LSTMs for glaucoma diagnosis
10. Diffusion Models for glaucoma diagnosis
11. Hybrid networks for glaucoma diagnosis

6.1 Convolutional neural networks-based glaucoma diagnosis

A convolutional neural network (CNN) is a mathematical model that employs filters to process images and detect characteristics. CNNs are composed of layers that perform specialized functions, such as feature extraction and pattern recognition. Convolutional Neural Networks (CNNs) are often used to diagnose glaucoma by extracting higher-level features from input data. In this section, some cutting-edge CNN architectures such as ResNet, EfficientNet, DenseNet etc. used for glaucoma detection, are illustrated. In addition to that specially designed optimized CNN applied in glaucoma detection has also been cited here. Table 3 summarizes research efforts that utilized CNN-based architectures for glaucoma detection. In [41], W. Liao and colleagues proposed a robust semi-supervised CNN framework designed for multitask learning. This model is capable of classifying OCT B-scan images as either normal or glaucomatous, while also analyzing the association between structural and functional changes in glaucoma patients. The architecture is composed of three main modules: a shared feature extraction unit using ResNet-18, a visual field (VF) regression component, and a glaucoma classification segment. The model was validated using the HK dataset, which contains 975,400 B-scan images derived from 4877 volume scans. In another notable contribution [42], researchers introduced EAMNet, a CNN model designed to provide medical interpretability for glaucoma diagnosis. EAMNet improves detection performance by combining semantic and spatial information through ResBlock features extracted from different stages of a CNN backbone. Additionally, the model produces clear evidence activation maps that pinpoint the key regions associated with glaucomatous changes. As a result, it offers clear model interpretations. A systematic framework comprising three distinct phases was presented in [43] (1) preliminary screening; (2) differentiating normal and glaucomatous eye; and (3) categorization of glaucoma in terms of severity. IOP measurements are used in the initial phase to find possible glaucoma patients. The second and third stages then involve the independent training of two different ResNet designs, DetectionNet and ClassificationNet. Identification of glaucoma as mild, moderate or severe is done by ClassificationNet exploring Voronoi VF images. DetectionNet uses a combination of fundus and Voronoi VF images to identify normal eye or glaucomatous eye. Glaucoma severity categorization is also done in [44]. They created a Transferable Ranking Convolutional Neural Network (TRk-CNN) to categorize the cases as normal eye, glaucoma suspect eye, or glaucomatous eye. TRk-CNN optimizes the procedure for closely related classes by integrating inter-class information into the final classification phase and using DenseNet as the backbone CNN model. Article [45] considered glaucoma diagnosis as a regression problem instead of classification problem. They used a deep learning algorithm for calculating the CDR directly from the fundus images. This process eliminates the need for manual segmentation of the optic cup and optic disc. This regression-based strategy, which takes advantage of disease severity continuity, may be more effective than classification models in detecting early glaucomatous alterations. Researchers also discovered that their model effectively improved CDR estimation accuracy by incorporating contextual information outside the ONH region.

Islami et.al. [50] developed customized architecture, using the Convolutional Neural Network (CNN) method. The customized CNN architecture is designed to address the complex characteristics of glaucoma changes in the identification process. This research works on a dataset that consists of 506 retinal images. Out of which 117 glaucoma images are categorized as glaucoma, 19 as suspected glaucoma, and 370 as healthy cases. The model yields an accuracy of 80.87%, sensitivity of 85.65%, and specificity of 71.26% in the test phase. The customized CNN architecture has shown potential in the glaucoma detection process and clinical acceptance.

Table 3. Use of CNN for glaucoma diagnosis (G : Glaucomatous, H : Healthy, T : Total, S : Glaucoma Suspect, AUC :Area under curve, Avg : Average)

Ref.	Model	Type of data	Dataset	Performance metrics for classification (%)					
				Accuracy	Sensitivity	Specificity	Precision	F1 score	AUC
[42]	EAMNet	Fundus	ORIGA	-	-	-	-	-	88
[44]	TRk-CNN	Fundus	G:381, H:403 S:208	88.94 (Avg)	90.36 (G Vs. H)	89.33 (Avg)	94.94 (G Vs. H)	92.59 (G Vs. H)	-
[46]	EfficientNet	Fundus	G:1586, H: 2244, RIGA	93.29	96.03	91.42	-	-	98.29
[47]	ResNet-50	Fundus	6927 T, 2.14% G) 23,318 T, 1.08% G) 11 publicly available dataset	-	81,94, 70 to 99	95, 94, 68 to 94	-	-	97, 98 85.40 to 98.80
[48]	3D CNN	OCT	G:363, H:291	-	86	90	-	76	91
[49]	ResNet18 and ResNet50	Fundus	G:1590, H:1865	98.03	99		97.61	-	-

6.2 Autoencoder based network for glaucoma diagnosis

An autoencoder is an unsupervised neural network that converts the complex input data into a lower dimensional to identify complicated patterns and representations. After processing, a decoder is used to convert the low dimensional data back to their original input form. In order to diagnose glaucoma, the authors of the paper [7] developed a deep convolutional autoencoder network known as RAG-Netv2. The retinal ganglion cell can be divided into three portions by this network: the ganglion cell complex, the ganglion cell-inner plexiform layer, and the RNFL. The encoder can identify glaucoma with the aid of this segmentation. In another work, [51] proposed a complete multitasking model for classifying glaucoma as well as segmenting the optic disc and cup in fundus images. The proposed model uses a common encoder-decoder structure between tasks. In addition to image, it also utilizes pixel labels during the complete training process. Model performance has been evaluated using REFUGE and DRISHTI-GS datasets. Another multitasking deep learning model autoencoder model was created by Pascal et.al. [52]. The model can segment the OD and OC and locate the fovea in retinal fundus images. The multitasking model uses lesser parameters than single tasking models and consumes almost similar computational time. The model performance is evaluated with REFUGE dataset. It is found that combination of multitasking and transfer learning work well even in limited resources and small dataset. Table 4 outlines research that has employed autoencoder-based architectures to diagnose glaucoma.

Table 4. Autoencoder based glaucoma diagnosis

Ref.`	**Model**	**Data type**	**Dataset Information**	**Performance metrics for classification (%)**					
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
[7]	Retinal Analysis and Grading Network	OCT	AFIO	94.91	97.14	91.66	94.44	95.77	98.71
[51]	Multitasking & Multiadaptive optimization	Fundus image	REFUGE, DRISHTI-GS	-	-	-	-	-	97.60, 94.74
[52]	Multitask Learning & Independent Optimizer	Fundus image	REFUGE	-	-	-	-	-	96.76
[53]	VGG16, VGG19, ResNet50, and MobileNetv2	Fundus image	AIROGS, RIM-ONE DL	95.21, 94.52, 93.15, 95.21	98.08, 96.15, 100, 92.31				95.85, 94.89, 94.68, 94.56

6.3 Attention-based networks for glaucoma diagnosis

The attention-based network in deep learning mainly focuses on important areas of input data rather than less significant parts. This concept enhances model performance in terms of accuracy and efficiency. For identifying the important areas in the input data, judgements are taken from domain experts [23]. Also, heatmaps based on Grad-CAM method has been considered to emphasize the selective part that influence most in the decision-making process [51,54,7]. In the architecture of attention-based network like Vision Transformer [55], important parts have been extracted based on the weights of the various parts. This allowed the model to focus on the parts which carries more weight than others. This has improved the performance of image classification task of the model. A summary of the reviewed papers on attention-based network can be found in Table 5.

Table 5. Attention based network for glaucoma diagnosis

Ref.	Model	Data type	Dataset information	Performance metrics for classification (%)					
				Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC
[23]	Attention-based CNN	Fundus	LAG	96.2	95.4	96.7	-	95.40	98.30
[56]	Self-ensemble dual curriculum Learning framework	Fundus	LAG, REFUGE, RIM-ONE	97.12, 95.25, 93.96	95.20 80.00 89.74	98.16 96.94 97.12	- - -	95.47 78.82 90.91	99.28 93.32 97.19
[57]	Self-attention student network	Fundus	LAG (G:1711, H:3143)	97.12	97.21	97.07	-	96.65	99.31
[58]	4-layer CNN	OCT scans centered on the ONH	Two private datasets	87.88	81.82	90.91	81.82	81.81	-
[59]	Attention-guided DL model (AG-OCT)	3782 OCT scans from both eyes of 555 individuals	Private dataset (G:3355, H:427)	93.07	95.11	-	94.73	94.88	93.76
[60]	CNN-based multi-input model	Fundus	ODIR (Ocular Disease Intelligent Recognition)	-	77	-	83	80	-

6.4 Generative adversarial networks for glaucoma diagnosis

Generative adversarial networks (GANs) are deep learning models that combine two neural networks to generate new data that looks like training data. GANs are a form of generative model, which implies they produce new data instances. New data generation is based on the probability distribution [61]. GANs includes two modules: generator module and discriminator module. A generator is a neural network that learns to create new data. A discriminator is a neural network that learns to distinguish between real data and the data created by the generator. In the work [62] authors trained a GAN following a teacher-student approach. This helped the model to identify the differences between glaucomatous eye and healthy eye. This approach got significant success across different racial backgrounds. A comparison of different Generative adversarial networks used for glaucoma detection is illustrated in Table 6.

Table 6. Use of Generative adversarial networks for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**		**Performance metrics for classification (%)**				
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 score**	**AUC**
[62]	CycleGAN	Fundus	DRISHTI-GS	98.45	97.66	96.84		-	97.04
[63]	CycleGAN	Fundus	ACRIMA	96.8	97.4	-	82.1	89.1	-

6.5 Geometric deep learning networks for glaucoma diagnosis

Geometric deep learning (GDL) is a way to apply neural networks to non-Euclidean data, where traditional deep learning methods may struggle. It is a subfield of machine learning that focuses on creating algorithms to analyze and process data with a geometric structure, such as graphs, point clouds, and meshes. This is accomplished by combining deep learning techniques with geometry and topology concepts, enabling machines to learn and reason about complex 3D objects and environments. In the work authors mentioned the use of GDL for diagnosing glaucoma from OCT images of optical nerve head. It achieved an AUC of 95% on a private dataset that includes 477 glaucomatous and 2296 non-glaucomatous OCT images. A summary of different GDL network-based glaucoma diagnosis is presented in Table 7.

Table 7. Use of GDL networks for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**		**Performance metrics for classification (%)**				
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 score**	**AUC**
[64]	Geometric deep learning (PointNet)	OCT	477 glaucoma and 2296 nonglaucoma		-	-	-	-	95
[65]	DGCNN	Fundus	2,247 non-glaucoma & 2,259 glaucoma	-	-	-	-	-	97.0

6.6 Transformer architecture for glaucoma diagnosis

Transformer-based deep learning models, originally developed for natural language processing, have recently been adapted for medical image analysis, including glaucoma diagnosis. These models utilize self-attention mechanisms to capture global contextual relationships within data, making them particularly effective for analyzing complex ocular images like fundus photographs and OCT scans. In glaucoma detection, Transformers can model spatial dependencies across retinal regions more effectively than traditional CNNs, which rely on local receptive fields. Vision Transformers (ViTs), for instance, divide an image into patches and process them as a sequence, allowing the model to learn both local and global patterns associated with glaucomatous changes. Some studies also explore hybrid models combining CNN backbones with Transformer layers to enhance feature representation and interpretability. Overall, Transformer-based approaches offer promising improvements in diagnostic accuracy, generalization, and the ability to handle multimodal or high-dimensional data in glaucoma screening and progression analysis. Few existing research work related to this category are illustrated in Table 8.

Table 8. Transformer architecture for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**		**Performance metrics for classification (%)**				
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
[66]	Vision Transformer based Hybrid Model	Fundus	Private Dataset 2874 Cases	99.4	-	-	-	99.3	-
[67]	Vision Transformer architecture	Fundus	Standardized Multi-Channel Dataset for Glaucoma (SMDG),	90.48	-	-	-	-	91.0

6.7 Reinforcement Learning (RL)-based Deep Models for glaucoma diagnosis

Reinforcement Learning (RL), traditionally applied in control and decision-making tasks, has recently been explored in the context of medical image analysis, including glaucoma diagnosis. In RL-based deep models, an agent learns to make sequential decisions by interacting with a visual environment—such as fundus or OCT images—to optimize a reward signal tied to diagnostic accuracy. These models can learn optimal policies for tasks like localizing the optic disc or cup, refining segmentation boundaries, or prioritizing regions of interest for analysis. By integrating RL with deep neural networks, such models can adaptively focus on clinically relevant features, enabling more interpretable and efficient diagnosis. Although research in this area is still emerging, RL offers promising potential for enhancing end-to-end glaucoma analysis pipelines through active and adaptive learning strategies. Table 9 includes few researches on applying reinforcement learning based deep models for glaucoma diagnosis.

Table 9. Reinforcement Learning (RL)-based Deep Models for glaucoma diagnosis

Ref.	Model	Data type	Dataset Information		Performance metrics for classification (%)				
				Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC
[68]	Deep reinforcement learning	Fundus	520 Fundus images	-	-	-	80	-	-

[69]	Reinforcement-Based Leveraging Transfer Learning	OCT	83,484 images of different categories	99.20	-	-	-	-	-

6.8 Self-supervised Learning (SSL) Architectures for glaucoma diagnosis

Self-supervised learning (SSL) has gained significant attention in medical imaging due to its ability to learn rich and generalizable feature representations from unlabeled data—a key advantage in domains like glaucoma diagnosis where annotated datasets are limited. SSL architectures pre-train deep models using pretext tasks (e.g., image inpainting, jigsaw puzzles, contrastive learning) that do not require manual labels, enabling the network to understand structural and anatomical patterns in retinal images. These pre-trained models can then be fine-tuned on small labeled datasets for downstream tasks such as classification, segmentation, or severity grading of glaucoma. By leveraging the abundant availability of unlabeled fundus or OCT images, SSL improves model robustness, reduces reliance on costly annotations, and enhances performance in low-data regimes. Its integration into end-to-end glaucoma analysis frameworks represents a promising direction for building scalable and data-efficient diagnostic tools. Table 10 includes few researches on using Self-supervised Learning (SSL) Architectures for glaucoma diagnosis.

Table 10. Self-supervised Learning (SSL) Architectures for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**		**Performance metrics for classification (%)**				
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
[70]	Self-supervised clustering network	2397 images in total, with 956 glaucoma cases	REFUGE-2	96.32				68.32	
[71]	Vision Transformer (ViT) with a self-supervised volume contrast (VoCo) learning framework.	SD-OCT	Private (G:93, H:156)	95.56	83.55	95.94	-		97.11

6.9 Recurrent Neural Networks (RNNs) / LSTMs for glaucoma diagnosis

Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) models, are designed to capture sequential dependencies in data and have shown potential in glaucoma diagnosis where temporal or spatial sequences are involved. In ophthalmic imaging, RNNs and LSTMs can be used to model progressive changes in retinal structures over time, enabling longitudinal analysis of disease progression from time-series patient records or sequential OCT slices. These architectures are especially useful for integrating multi-modal or volumetric data, where the temporal or ordered nature of slices provides contextual information critical for accurate diagnosis. By retaining long-range dependencies and mitigating the vanishing gradient problem, LSTMs can learn patterns of progression or subtle changes in structure that are often indicative of early glaucoma. Although less common than convolutional models, RNN-based approaches offer valuable tools for dynamic and progression-aware glaucoma assessment. Few recent works on Recurrent Neural Networks (RNNs) / LSTMs for glaucoma diagnosis can be seen in Table 11.

Table 11. Recurrent Neural Networks (RNNs) / LSTMs for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**		**Performance metrics for classification (%)**				
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
[72]	Deep learning-based RNNLSTM	Fundus	DRISHTI-GS	97.4	97	97.9	-	-	-
[73]	CNN and RNN	Fundus Images and Videos	1810 fundus images and 295 fundus videos	-	95	96.0	-	96.2	99.0

6.10 Diffusion Models for glaucoma diagnosis

Diffusion models, a class of generative deep learning models, have recently emerged as powerful tools for high-fidelity image synthesis and denoising, with growing interest in medical imaging applications, including glaucoma diagnosis. These models work by learning to reverse a gradual noising process, effectively reconstructing high-quality images from noisy inputs. In the context of glaucoma, diffusion models can be leveraged for tasks such as synthetic retinal image generation to augment limited datasets, enhancing training diversity and generalization. Additionally, they show promise in denoising low-quality fundus or OCT images, thereby improving diagnostic clarity for downstream deep learning models. Their ability to generate anatomically plausible and high-resolution samples also aids in explainable AI by enabling visualizations of disease-relevant features. Although still in the early stages of adoption in ophthalmology, diffusion models represent a promising frontier for augmenting and improving end-to-end deep learning pipelines for glaucoma analysis. Table 12 shows some significant contributions on diffusion model based glaucoma detection.

Table 12. Diffusion Models for glaucoma diagnosis

Ref.	**Model**	**Data type**	**Dataset Information**	**Performance metrics for classification (%)**					
				Accuracy	**Sensitivity**	**Specificity**	**Precision**	**F1 Score**	**AUC**
[74]	Diffusion models	Fundus	AIROGS Dataset	-	94.44	97.19	-	-	98.98
[75]	Coarse-to-Fine Latent Diffusion Model	SD-OCT	LAG and ACRIMA	94.4	93.8	94.6	-	-	95.5

6.11 Hybrid networks for glaucoma diagnosis

A hybrid network refers to a neural network model that integrates multiple architectures—such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs)—to leverage the unique strengths of each. This approach is particularly effective for handling complex tasks and heterogeneous data types. In [76], the authors implemented a hybrid model that combines CNN and LSTM components to extract features directly from raw spectral-domain optical coherence tomography (SD-OCT) data. The architecture consists of two main components: a feature extraction module and a volume-based prediction unit. The feature extractor employs residual connections and attention-based CNNs to capture critical patterns from the input data. To incorporate spatial and three-dimensional context, the model integrates an LSTM network along with a sequential weighting mechanism, which refines the LSTM outputs and enhances training consistency. Table 13 provides an overview of various hybrid network approaches applied to glaucoma diagnosis.

Table 13. Use of hybrid networks for glaucoma diagnosis

Ref.	Model	Data type	Dataset Information		Performance metrics for classification (%)				
				Accuracy	Sensitivity	Specificity	Precision	F1 Score	AUC
[76]	Multi-Rater	Fundus	REFUGE, DRISHTI	98.00 86.14	82.50 91.43	99.72 74.19	-	-	96.83

	Consensus Model		-GS1						89.63
[77]	CNN and LSTM	SD-OCT	Private (G:93, H:156)	81.25	75.86	85.71	-	78.57	80.79
[78]	CNNs (ResNet50, VGG-16) and Random Forest	Fundus RGB images	ACRIMA, G1020, ORIGA and REFUGE	95.41	88.37	-	99.37	93.52	

7 Summary and Conclusions

This paper provides an overview of current research using deep learning and computer vision approaches to diagnose glaucoma. The existing research have been compiled on many architectural categories, including convolutional neural networks (CNNs), autoencoder-based networks, attention-based models, generative adversarial networks (GANs), geometric deep learning frameworks, Transformer architecture, Reinforcement Learning (RL)-based Deep Models, Self-supervised Learning (SSL), Recurrent Neural Networks (RNNs) / LSTMs, Diffusion Models for glaucoma diagnosis and hybrid architectures The review found that fundus, OCT, and visual field data are most frequently used for glaucoma detection using deep learning. Many neural network-based detection methods have shown significantly positive results. However, the main problem lies on the dataset size and diversity, pattern finding, feature extraction, model preparation and clinical acceptance. Research is going on to address these issues. Many state-of-the-art learning techniques such as transfer learning, multitask learning are getting implemented. It also requires support from medical experts of glaucoma detection.

Deep learning-based AI is paving the way for accurate glaucoma detection. This will drastically minimize the cost and duration of the detecting process. Automatic glaucoma detection systems can simplify the process so that patients can do it at their own convenience. Medical practitioners could perform this even during a routine eyesight examination. To use AI in glaucoma diagnosis, continual cross-disciplinary collaboration is required. Dataset availability, model selection and evaluation metrics are crucial in this research. Finally, a multidisciplinary approach integrating engineering, medicine, and computer science will be critical for developing and validating solutions that are ready for clinical adoption. This article may assist AI researchers in selecting a good deep learning technique to generate models for glaucoma diagnosis based on dataset dimension, dataset variety, information extraction process, and clinical acceptance. A summary diagram as shown in Fig. 4 can help researchers quickly grasp the various aspects of this research domain.

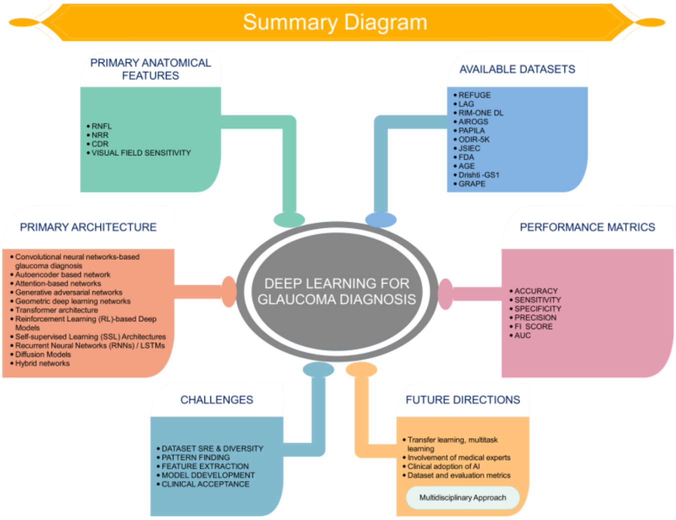

Fig. 4. Summary Diagram

References

1. Steinmetz, J. D., Bourne, R. R., Briant, P. S., Flaxman, S. R., Taylor, H. R., Jonas, J. B., ... & Morse, A. R. F. (2021). Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. *The Lancet Global Health*, *9*(2), e144-e160. https://doi.org/10.1016/S2214-109X(20)30489-7
2. Alghamdi, M., & Abdel-Mottaleb, M. (2021). A comparative study of deep learning models for diagnosing glaucoma from fundus images. *IEEE access*, *9*, 23894-23906. https://doi.org/10.1109/ACCESS.2021.3056641
3. Kim, M., Han, J. C., Hyun, S. H., Janssens, O., Van Hoecke, S., Kee, C., & De Neve, W. (2019). Medinoid: computer-aided diagnosis and localization of glaucoma using deep learning. *Applied Sciences*, *9*(15), 3064. https://doi.org/10.3390/app9153064
4. Yadav, V., Dey, B., Baruah, U., Das, S., & Prakash, O. (2025). Machine learning-assisted image analysis techniques for glaucoma detection. EURASIP Journal on Image and Video Processing, 2025(1), 1-28. https://doi.org/10.1186/s13640-025-00668-1
5. Y. Chai, Y., Liu, H., & Xu, J. (2018). Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. *Knowledge-Based Systems*, *161*, 147-156. https://doi.org/10.1016/j.knosys.2018.07.043
6. Ashtari-Majlan, M., Dehshibi, M. M., & Masip, D. (2024). Glaucoma diagnosis in the era of deep learning: A survey. *Expert Systems with Applications*, *256*, 124888. https://doi.org/10.1016/j.eswa.2024.124888
7. Raja, H., Hassan, T., Akram, M. U., & Werghi, N. (2020). Clinically verified hybrid deep learning system for retinal ganglion cells aware grading of glaucomatous progression. *IEEE Transactions on Biomedical Engineering*, *68*(7), 2140-2151. https://doi.org/10.1109/TBME.2020.3030085
8. Chen, X., Xu, Y., Yan, S., Wong, D. W. K., Wong, T. Y., & Liu, J. (2015). Automatic feature learning for glaucoma detection based on deep learning. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International*

Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 669-677). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_80
9. Eye Diseases Prevalence Research Group. (2004). Prevalence of open-angle glaucoma among adults in the United States. *Archives of ophthalmology*, *122*(4), 532.
10. Abdullah, F., Imtiaz, R., Madni, H. A., Khan, H. A., Khan, T. M., Khan, M. A., & Naqvi, S. S. (2021). A review on glaucoma disease detection using computerized techniques. *IEEE Access*, *9*, 37311-37333. https://doi.org/10.1109/ACCESS.2021.3061451
11. Biousse, V., Danesh-Meyer, H. V., Saindane, A. M., Lamirel, C., & Newman, N. J. (2022). Imaging of the optic nerve: technological advances and future prospects. *The Lancet Neurology*, *21*(12), 1135-1150. https://doi.org/10.1016/S1474-4422(22)00173-9
12. Varma, R., Ying-Lai, M., Francis, B. A., Nguyen, B. B. T., Deneen, J., Wilson, M. R., ... & Los Angeles Latino Eye Study Group. (2004). Prevalence of open-angle glaucoma and ocular hypertension in Latinos: the Los Angeles Latino Eye Study. *Ophthalmology*, *111*(8), 1439-1448. https://doi.org/10.1016/j.ophtha.2004.01.025
13. Weinreb, R. N., Bowd, C., Moghimi, S., Tafreshi, A., Rausch, S., & Zangwill, L. M. (2019). Ophthalmic diagnostic imaging: glaucoma. *High Resolution Imaging in Microscopy and Ophthalmology: New Frontiers in Biomedical Optics*, 107-134. https://doi.org/10.1007/978-3-030-16638-0_5
14. Watanabe, T., Hiratsuka, Y., Kita, Y., Tamura, H., Kawasaki, R., Yokoyama, T., ... & Yamada, M. (2022). Combining optical coherence tomography and fundus photography to improve glaucoma screening. *Diagnostics*, *12*(5), 1100. https://doi.org/10.3390/diagnostics12051100
15. Li, F., Yan, L., Wang, Y., Shi, J., Chen, H., Zhang, X., ... & Zhou, K. (2020). Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *258*, 851-867. https://doi.org/10.1007/s00417-020-04609-8
16. Raghunathan, T., Mishra, A., Mahur, A. K., & Balaji, B. (2024, March). Multi-Modal AI/ML Integration for Precision Glaucoma Detection: A Comprehensive Analysis using Optical Coherence Tomography, Fundus Imaging, RNFL, and Vessel Density. In *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)* (pp. 1-7). IEEE. https://doi.org/10.1109/AIMLA59606.2024.10531451
17. Song, D., Fu, B., Li, F., Xiong, J., He, J., Zhang, X., & Qiao, Y. (2021). Deep relation transformer for diagnosing glaucoma with optical coherence tomography and visual field function. *IEEE Transactions on Medical Imaging*, *40*(9), 2392-2402. https://doi.org/10.1109/TMI.2021.3077484
18. Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: sensitivity and specificity. *Continuing education in anaesthesia, critical care & pain*, *8*(6), 221-223. https://doi.org/10.1093/bjaceaccp/mkn041
19. Van Stralen, K. J., Stel, V. S., Reitsma, J. B., Dekker, F. W., Zoccali, C., & Jager, K. J. (2009). Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney international*, *75*(12), 1257-1263. https://doi.org/10.1038/ki.2009.92
20. Kumar, R., & Indrayan, A. (2011). Receiver operating characteristic (ROC) curve for medical researchers. *Indian pediatrics*, *48*, 277-287. https://doi.org/10.1007/s13312-011-0055-4
21. Hung, K. H., Kao, Y. C., Tang, Y. H., Chen, Y. T., Wang, C. H., Wang, Y. C., & Lee, O. K. S. (2022). Application of a deep learning system in glaucoma screening and further classification with colour fundus photographs: a case control study. *BMC ophthalmology*, *22*(1), 483. https://doi.org/10.1186/s12886-022-02730-2

22. Orlando, J. I., Fu, H., Breda, J. B., Van Keer, K., Bathula, D. R., Diaz-Pinto, A., ... & Bogunović, H. (2020). Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical image analysis*, *59*, 101570. https://doi.org/10.1016/j.media.2019.101570
23. Li, L., Xu, M., Liu, H., Li, Y., Wang, X., Jiang, L., ... & Wang, N. (2019). A large-scale database and a CNN model for attention-based glaucoma detection. *IEEE transactions on medical imaging*, *39*(2), 413-424. https://doi.org/10.1109/TMI.2019.2927226
24. Batista, F. J. F., Diaz-Aleman, T., Sigut, J., Alayon, S., Arnay, R., & Angel-Pereira, D. (2020). Rim-one dl: A unified retinal image database for assessing glaucoma using deep learning. *Image Analysis and Stereology*, *39*(3), 161-167. https://orcid.org/0000-0001-9806-2477
25. Lemij, H. G., de Vente, C., Sánchez, C. I., & Vermeer, K. A. (2023). Characteristics of a large, labeled data set for the training of artificial intelligence for glaucoma screening with fundus photographs. *Ophthalmology Science*, *3*(3), 100300. https://doi.org/10.1016/j.xops.2023.100300
26. Kovalyk, O., Morales-Sánchez, J., Verdú-Monedero, R., Sellés-Navarro, I., Palazón-Cabanes, A., & Sancho-Gómez, J. L. (2022). PAPILA: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. *Scientific Data*, *9*(1), 291. https://doi.org/10.1038/s41597-022-01388-1
27. Khan, S. D., Basalamah, S., & Lbath, A. (2025). A novel deep learning framework for retinal disease detection leveraging contextual and local features cues from retinal images. *Medical & Biological Engineering & Computing*, 1-18. https://doi.org/10.1007/s11517-025-03314-0
28. Cen, L. P., Ji, J., Lin, J. W., Ju, S. T., Lin, H. J., Li, T. P., ... & Zhang, M. (2021). Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nature communications*, *12*(1), 4828. https://doi.org/10.1038/s41467-021-25138-w
29. Soltanian-Zadeh, S., Kurokawa, K., Liu, Z., Zhang, F., Saeedi, O., Hammer, D. X., ... & Farsiu, S. (2021). Weakly supervised individual ganglion cell segmentation from adaptive optics OCT images for glaucomatous damage assessment. *Optica*, *8*(5), 642-651. https://doi.org/10.1364/OPTICA.418274
30. Fu, H., Li, F., Sun, X., Cao, X., Liao, J., Orlando, J. I., ... & iChallenge-PACG study group. (2020). Age challenge: angle closure glaucoma evaluation in anterior segment optical coherence tomography. *Medical Image Analysis*, *66*, 101798. https://doi.org/10.1016/j.media.2020.101798
31. Sivaswamy, J., Krishnadas, S. R., Joshi, G. D., Jain, M., & Tabish, A. U. S. (2014, April). Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In *2014 IEEE 11th international symposium on biomedical imaging (ISBI)* (pp. 53-56). IEEE. https://doi.org/10.1109/ISBI.2014.6867807
32. Gao, M., Jiang, H., Zhu, L., Jiang, Z., Geng, M., Ren, Q., & Lu, Y. (2023). Discriminative ensemble meta-learning with co-regularization for rare fundus diseases diagnosis. *Medical Image Analysis*, *89*, 102884. https://doi.org/10.1016/j.media.2023.102884
33. Dixit, A., Yohannan, J., & Boland, M. V. (2021). Assessing glaucoma progression using machine learning trained on longitudinal visual field and clinical data. *Ophthalmology*, *128*(7), 1016-1026. https://doi.org/10.1016/j.ophtha.2020.12.020
34. Al-Bander, B., Al-Nuaimy, W., Al-Taee, M. A., & Zheng, Y. (2017, March). Automated glaucoma diagnosis using deep learning approach. In *2017 14th International Multi-Conference on Systems, Signals & Devices (SSD)* (pp. 207-210). IEEE. https://doi.org/10.1109/SSD.2017.8166974

35. Kashyap, R., Nair, R., Gangadharan, S. M. P., Botto-Tobar, M., Farooq, S., & Rizwan, A. (2022, December). Glaucoma detection and classification using improved U-Net Deep Learning Model. In *Healthcare* (Vol. 10, No. 12, p. 2497). MDPI. https://doi.org/10.3390/healthcare10122497
36. Li, L., Xu, M., Wang, X., Jiang, L., & Liu, H. (2019). Attention based glaucoma detection: A large-scale database and CNN model. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10571-10580).
37. Raghavendra, U., Fujita, H., Bhandary, S. V., Gudigar, A., Tan, J. H., & Acharya, U. R. (2018). Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Information Sciences*, *441*, 41-49. https://doi.org/10.1016/j.ins.2018.01.051
38. Thompson, A. C., Jammal, A. A., Berchuck, S. I., Mariottoni, E. B., & Medeiros, F. A. (2020). Assessment of a segmentation-free deep learning algorithm for diagnosing glaucoma from optical coherence tomography scans. *JAMA ophthalmology*, *138*(4), 333-339. doi:10.1001/jamaophthalmol.2019.5983
39. Xu, Y., Hu, M., Liu, H., Yang, H., Wang, H., Lu, S., ... & Wang, N. (2021). A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis. *NPJ digital medicine*, *4*(1), 48. https://doi.org/10.1038/s41746-021-00417-4
40. Schottenhamml, J., Würfl, T., Ploner, S., Husvogt, L., Lämmer, R., Hohberger, B., ... & Mardin, C. (2024). Impact of acquisition area on deep-learning-based glaucoma detection in different plexuses in OCTA. *Scientific Reports*, *14*(1), 20414. https://doi.org/10.1038/s41598-024-71235-3
41. Wang, X., Chen, H., Ran, A. R., Luo, L., Chan, P. P., Tham, C. C., ... & Heng, P. A. (2020). Towards multi-center glaucoma OCT image screening with semi-supervised joint structure and function multi-task learning. *Medical Image Analysis*, *63*, 101695. https://doi.org/10.1016/j.media.2020.101695
42. Liao, W., Zou, B., Zhao, R., Chen, Y., He, Z., & Zhou, M. (2019). Clinical interpretable deep learning model for glaucoma diagnosis. *IEEE journal of biomedical and health informatics*, *24*(5), 1405-1412. https://doi.org/10.1109/JBHI.2019.2949075
43. Xue, Y., Zhu, J., Huang, X., Xu, X., Li, X., Zheng, Y., ... & Si, K. (2022). A multi-feature deep learning system to enhance glaucoma severity diagnosis with high accuracy and fast speed. *Journal of biomedical informatics*, *136*, 104233. https://doi.org/10.1016/j.jbi.2022.104233
44. Jun, T. J., Eom, Y., Kim, D., Kim, C., Park, J. H., Nguyen, H. M., ... & Kim, D. (2021). TRk-CNN: transferable ranking-CNN for image classification of glaucoma, glaucoma suspect, and normal eyes. *Expert Systems with Applications*, *182*, 115211. https://doi.org/10.1016/j.eswa.2021.115211
45. Hemelings, R., Elen, B., Barbosa-Breda, J., Blaschko, M. B., De Boever, P., & Stalmans, I. (2021). Deep learning on fundus images detects glaucoma beyond the optic disc. *Scientific Reports*, *11*(1), 20313. https://doi.org/10.1038/s41598-021-99605-1
46. Wu, J., Yu, S., Chen, W., Ma, K., Fu, R., Liu, H., ... & Zheng, Y. (2020). Leveraging undiagnosed data for glaucoma classification with teacher-student learning. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23* (pp. 731-740). Springer International Publishing. https://doi.org/10.1007/978-3-030-59710-8_71

47. Hemelings, R., Elen, B., Schuster, A. K., Blaschko, M. B., Barbosa-Breda, J., Hujanen, P., ... & Stalmans, I. (2023). A generalizable deep learning regression model for automated glaucoma screening from fundus images. *NPJ digital medicine*, *6*(1), 112. https://doi.org/10.1038/s41746-023-00857-0
48. Noury, E., Mannil, S. S., Chang, R. T., Ran, A. R., Cheung, C. Y., Thapa, S. S., ... & Zadeh, R. (2022). Deep learning for glaucoma detection and identification of novel diagnostic areas in diverse real-world datasets. *Translational Vision Science & Technology*, *11*(5), 11-11. https://doi.org/10.1167/tvst.11.5.11
49. Velpula, V. K., Sharma, D., Sharma, L. D., Roy, A., Bhuyan, M. K., Alfarhood, S., & Safran, M. (2024). Glaucoma detection with explainable AI using convolutional neural networks based feature extraction and machine learning classifiers. *IET Image Processing*, *18*(13), 3827-3853. https://doi.org/10.1049/ipr2.13211
50. Islami, F., & Defit, S. (2024). Customized Convolutional Neural Network for Glaucoma Detection in Retinal Fundus Images. *Jurnal Penelitian Pendidikan IPA*, *10*(8), 4606-4613. https://doi.org/10.29303/jppipa.v10i8.7614
51. Hervella, Á. S., Rouco, J., Novo, J., & Ortega, M. (2022). End-to-end multi-task learning for simultaneous optic disc and cup segmentation and glaucoma classification in eye fundus images. *Applied Soft Computing*, *116*, 108347. https://doi.org/10.1016/j.asoc.2021.108347
52. Pascal, L., Perdomo, O. J., Bost, X., Huet, B., Otálora, S., & Zuluaga, M. A. (2022). Multi-task deep learning for glaucoma detection from color fundus images. *Scientific Reports*, *12*(1), 12361. https://doi.org/10.1038/s41598-022-16262-8
53. Singh, M., Vivek, B. S., Gubbi, J., & Pal, A. (2024). Prototype-based Interpretable Model for Glaucoma Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5056-5065).
54. Jiang, Y., Duan, L., Cheng, J., Gu, Z., Xia, H., Fu, H., ... & Liu, J. (2019). JointRCNN: a region-based convolutional neural network for optic disc and cup segmentation. *IEEE Transactions on Biomedical Engineering*, *67*(2), 335-343. https://doi.org/10.1109/TBME.2019.2913211
55. Dosovitskiy, A. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.
56. Zhao, R., Chen, X., Chen, Z., & Li, S. (2022). Diagnosing glaucoma on imbalanced data with self-ensemble dual-curriculum learning. *Medical image analysis*, *75*, 102295. https://doi.org/10.1016/j.media.2021.102295
57. Zhao, R., Chen, X., Chen, Z., & Li, S. (2020). EGDCL: An adaptive curriculum learning framework for unbiased glaucoma diagnosis. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16* (pp. 190-205). Springer International Publishing. https://doi.org/10.1007/978-3-030-58589-1_12
58. García, G., Del Amor, R., Colomer, A., Verdú-Monedero, R., Morales-Sánchez, J., & Naranjo, V. (2021). Circumpapillary OCT-focused hybrid learning for glaucoma grading using tailored prototypical neural networks. *Artificial Intelligence in Medicine*, *118*, 102132. https://doi.org/10.1016/j.artmed.2021.102132
59. George, Y., Antony, B. J., Ishikawa, H., Wollstein, G., Schuman, J. S., & Garnavi, R. (2020). Attention-guided 3D-CNN framework for glaucoma detection and structural-

functional association using volumetric images. *IEEE journal of biomedical and health informatics*, *24*(12), 3421-3430. https://doi.org/10.1109/JBHI.2020.3001019

60. Cho, Y. S., Song, H. J., Han, J. H., & Kim, Y. S. (2024). Attention Mechanism-Based Glaucoma Classification Model Using Retinal Fundus Images. *Sensors*, *24*(14), 4684. https://doi.org/10.3390/s24144684
61. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.
62. Guo, Y., Peng, Y., Sun, J., Li, D., & Zhang, B. (2022). DSLN: Dual-tutor student learning network for multiracial glaucoma detection. *Neural Computing and Applications*, *34*(14), 11885-11910. https://doi.org/10.1007/s00521-022-07078-8
63. Lenka, S., Mayaluri, Z. L., & Panda, G. (2024). Retinal fundus image enhancement using an ensemble framework for accurate glaucoma detection. *Neural Computing and Applications*, 1-19. https://doi.org/10.1007/s00521-024-10500-y
64. Thiéry, A. H., Braeu, F., Tun, T. A., Aung, T., & Girard, M. J. (2023). Medical application of geometric deep learning for the diagnosis of glaucoma. *Translational Vision Science & Technology*, *12*(2), 23-23. https://doi.org/10.1167/tvst.12.2.23
65. Braeu, F. A., Thiéry, A. H., Tun, T. A., Kadziauskiene, A., Barbastathis, G., Aung, T., & Girard, M. J. (2023). Geometric deep learning to identify the critical 3D structural features of the optic nerve head for glaucoma diagnosis. *American Journal of Ophthalmology*, *250*, 38-48. https://doi.org/10.1016/j.ajo.2023.01.008
66. Yu, S., Zhou, H. Y., Ma, K., Bian, C., Chu, C., Liu, H., & Zheng, Y. (2020). Difficulty-aware glaucoma classification with multi-rater consensus modeling. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23* (pp. 741-750). Springer International Publishing. https://doi.org/10.1007/978-3-030-59710-8_72
67. Sivakumar, R., & Penkova, A. (2025). Enhancing glaucoma detection through multi-modal integration of retinal images and clinical biomarkers. Engineering Applications of Artificial Intelligence, 143, 110010.
68. Musthafa, N. (2023, July). Precise Localization of Optic Disc Region for Accurate Glaucoma Diagnosis Using Deep Reinforcement Learning. In 2023 International Conference on Innovations in Engineering and Technology (ICIET) (pp. 1-7). IEEE. https://doi.org/10.1109/ICIET57285.2023.10220844
69. Abbas, Y., Hadi, H. J., Aziz, K., Ahmed, N., Akhtar, M. U., Alshara, M. A., & Chakrabarti, P. (2025). Reinforcement-based leveraging transfer learning for multiclass optical coherence tomography images classification. Scientific Reports, 15(1), 6193. https://doi.org/10.1038/s41598-025-89831-2
70. Rezaei, M., Vahidi, A., Bischl, B., Wang, M., Elze, T., & Eslami, M. (2023). Self-supervised learning and self-labeling framework for retina glaucoma detection. Investigative Ophthalmology & Visual Science, 64(8), 352-352.
71. Liao, C., Todo, Y., & Tang, Z. (2025). Integrating self-supervised learning with vision transformers for glaucoma detection. Journal of Electronic Imaging, 34(2), 023016-023016. https://doi.org/10.1117/1.JEI.34.2.023016

72. Veena, H.N., Patil, K.K., Vanajakshi, P. et al. An Enhanced RNN-LSTM Model for Fundus Image Classification to Diagnose Glaucoma. SN COMPUT. SCI. 5, 514 (2024). https://doi.org/10.1007/s42979-024-02867-5
73. Gheisari, S., Shariflou, S., Phu, J., Kennedy, P. J., Agar, A., Kalloniatis, M., & Golzan, S. M. (2021). A combined convolutional and recurrent neural network for enhanced glaucoma detection. Scientific reports, 11(1), 1945. https://doi.org/10.1038/s41598-021-81554-4
74. Nawar, Y., Soliman, N., Wassel, M., ElHabebe, M., Adly, N., Torki, M., ... & Ahmed, I. (2025, February). DiffuPT: Class Imbalance Mitigation for Glaucoma Detection via Diffusion Based Generation and Model Pretraining. In 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) (pp. 4098-4107). IEEE. https://doi.org/10.1109/WACV61041.2025.00403
75. Zhang, Y., Huang, K., Yang, X., Ma, X., Wu, J., Wang, N., ... & Heng, P. A. (2024, October). Coarse-to-Fine Latent Diffusion Model for Glaucoma Forecast on Sequential Fundus Images. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 166-176). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72086-4_16
76. Chincholi, F., & Koestler, H. (2024). Transforming glaucoma diagnosis: transformers at the forefront. Frontiers in Artificial Intelligence, 7, 1324109. https://doi.org/10.3389/frai.2024.1324109
77. García, G., Colomer, A., & Naranjo, V. (2021). Glaucoma detection from raw SD-OCT volumes: a novel approach focused on spatial dependencies. *Computer methods and programs in biomedicine*, *200*, 105855. https://doi.org/10.1016/j.cmpb.2020.105855
78. Aljohani, A., & Aburasain, R. Y. (2024). A hybrid framework for glaucoma detection through federated machine learning and deep learning models. *BMC Medical Informatics and Decision Making*, *24*(1), 115. https://doi.org/10.1186/s12911-024-02518-y