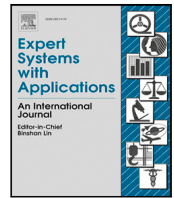




Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review



Glaucoma diagnosis in the era of deep learning: A survey

Mona Ashtari-Majlan^{a,d,*}, Mohammad Mahdi Dehshibi^{b,c}, David Masip^a^a Department of Computer Science, Multimedia, and Telecommunications, Universitat Oberta de Catalunya, Barcelona, Spain^b Escuela Politécnica Superior, Departamento de Informática, Universidad Carlos III de Madrid, Leganés, 28911, Spain^c Unconventional Computing Laboratory, University of the West of England, Bristol, BS16 1QY, UK^d Institute of Ophthalmology, University College London, London, UK

ARTICLE INFO

Keywords:

Glaucoma
Deep learning
Computer vision
Machine learning

ABSTRACT

Glaucoma, a leading cause of irreversible blindness worldwide, poses significant diagnostic challenges due to its reliance on subjective evaluation. Recent advances in computer vision and deep learning have demonstrated the potential for automated assessment. This paper provides a comprehensive survey of studies on AI-based glaucoma diagnosis using fundus, optical coherence tomography, and visual field images, with a focus on deep learning-based methods. We searched Web of Science, PubMed, IEEE Xplore, and Google Scholar, applying specific selection criteria to identify relevant studies published from 2017 to 2023. Our analysis provides a structured overview of architectural paradigms, including convolutional neural networks, autoencoders, attention networks, generative adversarial networks, and geometric deep learning models. Additionally, we discuss approaches for extracting informative features, such as structural, statistical, and hybrid techniques. Furthermore, we outline key research challenges and future directions, emphasizing the need for larger, more diverse datasets, strategies for early disease detection, multi-modal data integration, model explainability, and clinical translation. This survey is expected to be useful for Artificial Intelligence (AI) researchers seeking to translate advances into practice and ophthalmologists aiming to improve clinical workflows and diagnosis using the latest AI outcomes.

Contents

1. Introduction	2
2. Search protocol	2
3. Glaucoma: Definition and diagnosis	3
4. Datasets and evaluation metrics	5
5. Feature extraction	5
5.1. Structural features	5
5.2. Statistical features	6
5.3. Hybrid features	7
6. End-to-end glaucoma classification	8
6.1. Convolutional neural networks	8
6.2. Autoencoder-based networks	8
6.3. Attention-based networks	10
6.4. Generative adversarial networks	11
6.5. Geometric deep learning networks	11
6.6. Hybrid networks	12
7. Glaucoma prediction and progression detection	12
8. Challenges and future directions	13
8.1. Data challenges	13
8.2. Model development challenges	13
8.3. Clinical translation challenges	14
8.4. Future outlook	14

* Corresponding author.

E-mail addresses: mashtarimajlan@uoc.edu (M. Ashtari-Majlan), dmohamma@inf.uc3m.es (M.M. Dehshibi), dmasipr@uoc.edu (D. Masip).<https://doi.org/10.1016/j.eswa.2024.124888>

Received 16 April 2024; Received in revised form 9 July 2024; Accepted 23 July 2024

Available online 1 August 2024

0957-4174/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

9. Conclusion	14
CRediT authorship contribution statement	15
Declaration of competing interest	15
Data availability	15
Acknowledgments	15
References	15

1. Introduction

Glaucoma is the leading cause of irreversible blindness worldwide, affecting over 70 million people as of 2020 (Quigley & Broman, 2006). If left untreated, glaucoma leads to permanent vision loss due to damage to the optic nerve head and retinal nerve fiber layer (Jonas et al., 2017). However, despite improved understanding and management of glaucoma, it still accounts for approximately 10% of global blindness (Steinmetz et al., 2021; Tham et al., 2014). This high disease burden motivates the development of enhanced diagnostic techniques to enable early diagnosis and timely intervention to prevent or slow down the further deterioration of vision.

The application of Artificial Intelligence (AI) and deep learning in glaucoma diagnosis has advanced significantly in recent years (Mayro, Wang, Elze, & Pasquale, 2020). Early efforts focused on conventional image processing techniques to analyze retinal images (Lee et al., 2019; Thakur & Juneja, 2020). However, these approaches were limited by their reliance on hand-crafted features and lack of generalizability across diverse patient populations. The emergence of deep learning has transformed this field by enabling the development of algorithms capable of identifying subtle patterns in retinal images associated with glaucoma (Chai, Liu, & Xu, 2018; Hemelings et al., 2021; Raja, Hassan, Akram, & Werghi, 2021; Xue et al., 2022).

Accurately diagnosing glaucoma remains challenging for several reasons (Dervisevic, Pavljasevic, Dervisevic, & Kasumovic, 2016). Firstly, glaucoma is often asymptomatic in its early stages, impeding detection without comprehensive eye examinations. Secondly, current diagnostic modalities, such as imaging tests and functional assessments, have limitations in their sensitivity and specificity. Finally, the wide variability in glaucoma presentation, from subtle early symptoms to severe late-stage damage, makes definitive diagnoses difficult (Jonas et al., 2017). The complexity and subjectivity of evaluating diverse examination findings further complicate diagnosis. Given these challenges, the development of automated techniques for (1) glaucoma incidence prediction (identifying the transition from non-glaucoma to glaucoma), (2) glaucoma progression prediction (forecasting future changes in existing glaucoma cases), and (3) glaucoma progression detection (measuring whether glaucoma has progressed post-hoc) could enable earlier detection, more consistent quantification of progression, and ultimately preserving vision that would otherwise be lost to glaucoma (Muhammad et al., 2017).

This survey aims to provide a comprehensive overview of applying deep learning and computer vision algorithms to enhance glaucoma diagnosis,¹ by overcoming these challenges. Such automated techniques offer the potential for earlier detection, more consistent quantification of progression, and ultimately preserving vision that would otherwise be lost to glaucoma (Muhammad et al., 2017). Synthesizing recent techniques, results, and open problems can deliver value to both ophthalmology practitioners and AI researchers. For ophthalmology practitioners, it highlights cutting-edge research to improve diagnostics accuracy and integrate intelligent systems into clinical workflows. For

¹ When we use the term “diagnosis” in the context of papers that focus on deep learning and computer vision, we are referring to the use of AI technology to support medical diagnosis.

AI researchers, it provides a landscape analysis of the state-of-the-art, remaining gaps, and future opportunities to advance glaucoma diagnosis algorithms.

The remainder of the paper is organized as follows: Section 2 outlines the employed search protocol to select relevant papers for review, ensuring comprehensive topic coverage. Section 3 discusses clinical terminologies and definitions used in ophthalmology, enabling AI researchers to grasp the essential concepts necessary for interpreting the research in this field. The subsequent sections follow the general pipeline for glaucoma diagnosis, starting with Section 4, which delves into the datasets used for training and testing deep learning and computer vision models and the performance metrics employed to evaluate their effectiveness. Section 5 explores the various types of features, including structural, statistical, and hybrid ones, extracted for glaucoma diagnosis. Section 6 reviews the latest research on developing end-to-end deep learning models for glaucoma diagnosis, categorizing them based on the type and architecture of the models. This section highlights the ability of deep learning models to integrate multiple modalities and analyze complex data. These models provide accurate diagnosis and monitoring. Section 7 focuses on the methods proposed for early glaucoma prediction, shedding light on the advancements in prognostic techniques. In Section 8, challenges and potential future directions are discussed, addressing the limitations and paving the way for further research in the field. Finally, Section 9 serves as the conclusion, summarizing the key findings and contributions. It emphasizes the transformative potential of deep learning and computer vision in revolutionizing glaucoma diagnosis and management.

2. Search protocol

A thorough search was conducted to identify relevant studies on deep learning and computer vision techniques for glaucoma diagnosis published from the beginning of 2017 until the end of 2023. This date range was selected to capture the state-of-the-art advancements in this rapidly progressing field. Fig. 1 depicts the search protocol used to carry out this survey.

The following scholarly databases were searched: Web of Science, PubMed, IEEE Xplore, and Google Scholar. Targeted search terms included “glaucoma”, “deep learning”, “computer vision”, “machine learning”, “artificial intelligence”, and “medical diagnosis” keywords. These were combined using Boolean operators and customized search strings tailored to each database. When available, searches were restricted to titles, abstracts, and author keywords to filter potentially relevant papers efficiently. We also gave priority to leading conferences and journals in deep learning, computer vision, medical imaging, and ophthalmology.

After deduplication, the records retrieved underwent two-phase screening. Title/abstract screening evaluated relevance to glaucoma diagnosis using deep learning and computer vision approaches. The full-text review then confirmed papers met the inclusion criteria: (1) written in English; (2) primary focus on automated glaucoma diagnosis/screening using deep learning or computer vision; (3) rigorous machine learning experiments and substantial technical depth. Studies were excluded if they: (1) focused on other ocular diseases, even with glaucoma sub-analysis; (2) primarily contributed towards clinical insights. We also excluded review papers, case reports, and conference abstracts given limited methodological detail.

Additionally, the reference lists of included papers were manually searched to identify any additional relevant studies that might have

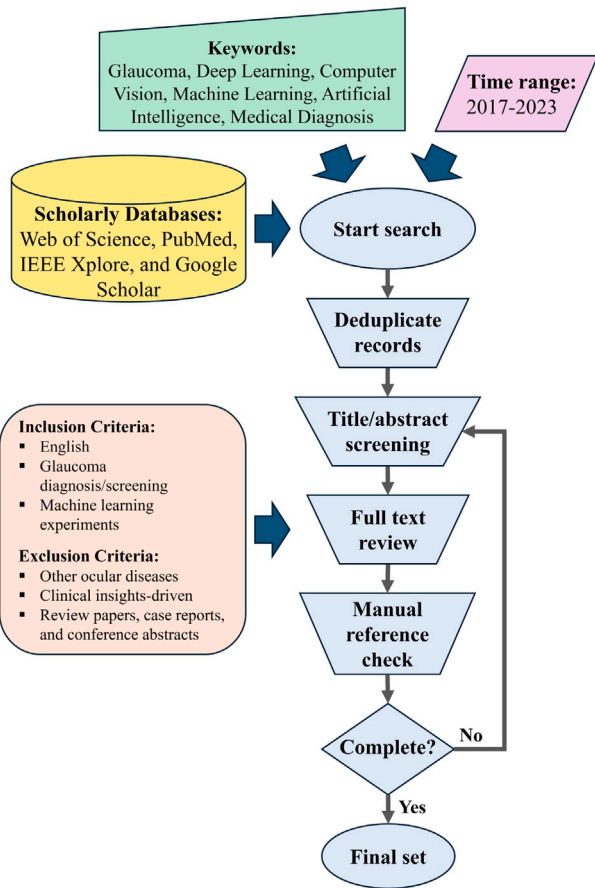


Fig. 1. Schematic representation of the search protocol.

been missed in the initial search. This way, we could include seminal papers even if published outside the target venues. Data extraction from the selected papers involved capturing key information such as study design, dataset characteristics, deep learning architectures, evaluation metrics, and major findings. This information serves as the foundation for synthesizing the current state of research in glaucoma diagnosis using deep learning and computer vision.

While this search protocol aimed to be comprehensive, some relevant studies may have been inadvertently omitted, given the rapid pace of research in this field. Nonetheless, the final selection aimed to identify a representative sample for assessing the state-of-the-art deep learning and computer vision techniques for glaucoma diagnosis.

3. Glaucoma: Definition and diagnosis

Familiarity with key ophthalmic concepts and terminologies (Tielsch et al., 1991) is necessary, specifically for AI researchers, to effectively interpret glaucoma diagnosis research. This section defines relevant terms and tests used in clinical glaucoma assessment, with a summary provided in Table 1.

Glaucoma is a condition characterized by the degeneration of Retinal Ganglion Cells (RGCs), leading to structural changes in the retina (Nouri-Mahdavi et al., 2021). These changes manifest as (1) the thinning of the Retinal Nerve Fiber Layer (RNFL), Ganglion Cell with the Inner Plexiform Layer (GCIPL), and Ganglion Cell Complex (GCC) profiles, (2) narrowing Neuroretinal Rim (NRR), and (3) cupping of the Optic Nerve Head (ONH) or enlargement of the Cup-to-Disc Ratio (CDR) (Jonas et al., 2017) (see Fig. 2). In addition to these structural changes, glaucoma also causes functional damage, resulting in defects in visual field sensitivity (Yousefi et al., 2018).

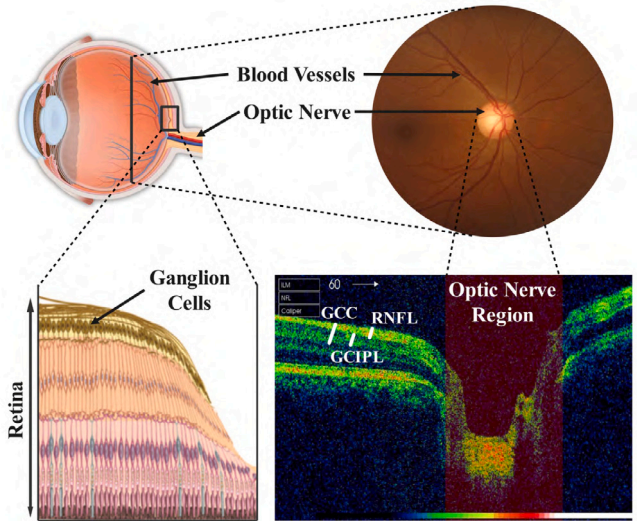


Fig. 2. Anatomical structures of the human eye and optic nerve relevant to glaucoma detection. [Left] Schematic views, [Right] Fundus and OCT views (Raja et al., 2020). This figure was created using images licensed under Creative Commons.

Table 1
Ophthalmic terminologies.

Terminology	Abbreviation	Definition
Retinal Ganglion Cells	RGCs	Neurons in the innermost layer of the retina which receive and transmit visual information to the brain.
Retinal Nerve Fiber Layer	RNFL	Layer of RGCs' nerve fibers (i.e., axons) that comprises the optic nerve and extends into the retina.
Ganglion Cell Layer	GCL	Layer of RGCs' cell bodies.
Inner Plexiform Layer	IPL	Layer of RGCs' dendrites.
Ganglion Cell Complex	GCC	Forms the combination of RNFL, GCL, and IPL layers.
Ganglion Cell with the Inner Plexiform Layer	GCIPL	Layer that consists of GCL and IPL.
Optic Nerve Head	ONH	The structure in the posterior section of the eye that enables the exit of axons of RGCs and the entry/exit of blood vessels.
Neuroretinal Rim	NRR	Area of the optic nerve head composed of retinal ganglion cell axons.
Optic Disc	OD	The optic disc and optic nerve head are interchangeable terms.
Optic Cup	OC	Central depression in the optic nerve head.
Cup-to-Disc Ratio	CDR	Quantitative measure comparing the size of the optic cup to the optic disc.
Intraocular Pressure	IOP	Fluid pressure inside the eye, influenced by the balance of aqueous humor production and drainage.

The process of detecting glaucoma is both complex and time-consuming (Gutierrez & Chen, 2023). To gain valuable insights into the structural and functional changes associated with the disease, both medical examinations and clinical expertise are used, where imaging techniques play a vital role.

Fundus imaging is a method that takes detailed images of the retina and OD. This helps with evaluating the appearance of the optic nerve, detecting vascular changes, and identifying any abnormalities. This imaging modality is a useful diagnostic tool for a variety of ocular conditions, including glaucoma. Optical Coherence Tomography (OCT) is another non-invasive imaging technique that produces highly detailed cross-sectional images of the retina, optic nerve, and other eye

Table 2

A review of the most commonly used datasets for glaucoma diagnosis. GT: Ground Truth, G: Glaucoma, H: Healthy.

Dataset	Number of images			Resolution	Ground truth description	Note	Address
	Glaucoma	Healthy	Total				
Fundus							
REFUGE (Orlando et al., 2020)	121	1079	1200	2124 × 2056, 1634 × 1634	Subject-level label, Segmentation GT (OC/OD)	–	Link
LAG (Li, Xu et al., 2020)	4878	6882	11,760	500 × 500	Subject-level label, Attention GT map	5824 images have Attention GT maps	Link
RIM-ONE DL (Batista et al., 2020)	172	313	485	Various	Subject-level label	Combination of RIM-ONE-r1, r2 and r3 (Fumero, Alayón, Sanchez, Sigut, & Gonzalez-Hernandez, 2011), Cropped at ON	Link
DRISHTI-GS1 (Sivaswamy et al., 2015)	70	31	101	2896 × 1944	Subject-level label, Segmentation GT (OC/OD), CDR Values, Notching	An extension of DRISHTI-GS (Sivaswamy, Krishnadas, Datt Joshi, Jain, & Syed Tabish, 2014)	Link
ACRIMA (Diaz-Pinto, Morales et al., 2019)	396	309	705	2048 × 1536	Subject-level label	Cropped at ON	Link
AIROGS (Lemij, Vente, Sánchez, & Vermeer, 2023)	3270	98,172	101,442	Various	Subject-level label	The training set is publicly available.	Link
PAPILA (Kovalyk et al., 2022)	155	333	488	2576 × 1934	Subject-level label, Segmentation GT (OC/OD)	Both eyes of the same patient	Link
Glaucoma Fundus (Ahn et al., 2018)	756	786	1542	240 × 240	Subject-level label	467 Advanced and 289 Early G, Cropped at ON	Link
SIGF (Li, Wang, Xu, Liu and Chen, 2020)	–	–	3671 ^a	Various	Subject-level label	405 Sequential images, avg 9 images per eye.	Link
INSPIRE-stereo (Tang et al., 2011)	30	–	30	768 × 1019	Subject-level label	Stereo color fundus images of the optic disc	Link
HRF (Budai, Bock, Maier, Hornegger, & Michelson, 2013)	15	15	45	3504 × 2336	Subject-level label, Vessel GT	Includes 15 images from DR patients	Link
DRIONS-DB (Carmona, Rincón, García-Feijóo, & de-la Casa, 2008)	–	–	110	600 × 400	Contour of ON	23.1% glaucoma and 76.9% eye hypertension	Link
ODIR-5K (Ocular Disease Intelligent Recognition, 2024)	307	1620	5000	Various	Subject-level label	Dataset is divided into eight categories	Link
JSIEC (Cen et al., 2021)	13	54	1087	Various	Subject-level label	Dataset is divided into 37 categories	Link
RIGA (Almazroa, Sun, Alodhayb, Raahemifar, & Lakshminarayanan, 2017)	–	–	750	Various	Segmentation GT (OC/OD)	No subject-level label	Link
OCT							
OCT Glaucoma Detection (Maetschke et al., 2019)	847	263	1110	200 × 200 × 1024	Subject-level label	–	Link
FDA (Soltanian-Zadeh et al., 2021)	10	6	16	297 × 259 × 450	Subject-level label	–	Link
AGE (Fu et al., 2020)	300	–	300	Various	Subject-level label ^b , Scleral Spur localization	1:4 ratio of Angle-closure to Open-angle glaucoma patients	Link
OCT & Fundus							
AFIO (Raja et al., 2020)	32	18	50	OCT: 951 × 456, Fundus: 2032 × 1934	Subject-level label, CDR values	The data is from 26 subjects	Link
Fundus & VF							
GRAPE (Huang et al., 2023)	1115	–	1115	Various	Subject-level label, Segmentation GT (OC/OD)	Longitudinal follow-ups of 263 eyes with 3-9 times of visits per eye	Link

(continued on next page)

structures. It enables clinicians to visualize the retinal layers, measure the thickness of the RNFL, and assess the integrity of the NRR (Bioussé, Danesh-Meyer, Saindane, Lamirel, & Newman, 2022; Weinreb et al., 2019).

While fundus and OCT imaging techniques are commonly used to capture structural changes associated with glaucoma, the Visual Field

analysis (VF) measures the sensitivity of the visual field. This test, based on standard automated perimetry, allows patients to respond to visual stimuli presented at different locations within their visual field. By mapping the patient's responses, clinicians can identify visual field defects associated with glaucoma and assess the extent of retinal sensitivity loss to provide better ground truth (Lee et al., 2019; Li,

Table 2 (continued).

Dataset	Number of images			Resolution	Ground truth description	Note	Address
	Glaucoma	Healthy	Total				
VF							
Rotterdam (Kucur, Holló, & Sznitman, 2018)	2279	244	2523	–	Subject-level label	24-2 VF test pattern	Link
Budapest (Kucur et al., 2018)	532	1735	2267	–	Subject-level label	24-2 VF test pattern	Link
UWHVF (Montesano, Chen, Lu, Lee, & Lee, 2022)	28,943	–	28,943	–	Subject-level label	Progression data for 2985 eyes, at least 4 VFs per eye	Link

^a 37 time-variant sequential images (convert from H to G), 368 time-invariant Sequential images (H).

^b Glaucoma can be classified into two broad categories: angle-closure and open-angle. The former is considered a more aggressive form of the disease compared to the latter (Fu et al., 2020).

Song et al., 2020; Tékouabou et al., 2022; Yousefi, Pasquale, Boland, & Johnson, 2022).

The integration of fundus, OCT, and VF modalities (Chen et al., 2019; Song et al., 2021), in addition to expert-level features (Akter et al., 2022; Chai et al., 2018) and biomarkers such as Intraocular Pressure (IOP) (Dixit, Yohannan, & Boland, 2021; Ibrahim, Hacibeyoglu, Agaoglu, & Ucar, 2022; Xue et al., 2022), can leverage a broader range of inputs to improve the accuracy and enhance the performance of deep learning models, enabling more precise detection, prediction, and management of glaucoma.

4. Datasets and evaluation metrics

In this section, we present a concise overview of the datasets used in the studies reviewed within this paper, as well as the evaluation metrics employed to assess the performance of trained models in glaucoma diagnosis. To facilitate ease of reference and enhance clarity, we have summarized this information in Table 2.²

Various evaluation metrics have been used to assess the performance of glaucoma diagnostic models. To account for the lack of standardization in metrics across datasets, researchers have used both subject-level and finer-grained spatial metrics in their studies. Eq. (1) illustrates the commonly used quantitative measures in the reviewed studies. These include accuracy (ACC), sensitivity (SEN), specificity (SPE), precision (PRC), and the F1-score. Here, TP , TN , FP , and FN represent true positive, true negative, false positive, and false negative, respectively (Yerushalmy, 1947). Additionally, the area under the receiver operating characteristic curve (AUC) metric has been employed to evaluate the performance of glaucoma diagnosis.

Table 2 highlights various glaucoma datasets with fundus images, OCT scans, and visual fields. These datasets have been crucial for developing deep learning techniques in glaucoma diagnosis. However, limitations exist. For instance, while some datasets (e.g., LAG (Li, Xu et al., 2020) and UWHVF (Montesano et al., 2022)) have many samples, others (e.g., DRISHTI-GS1 (Sivaswamy et al., 2015)) have relatively small sizes, hindering deep learning model training due to their scale. Another challenge is the imbalanced class distribution. Datasets like AIROGS (Lemij et al., 2023) have a significant imbalance. Such imbalances can bias the models towards the majority class, resulting in poor

² The links to the databases were active at the time of submitting this paper. In case of any inactive links, please contact the authors of the respective papers for updated information.

performance on the minority class, which is often the class of interest.

$$\begin{aligned}
 \text{Accuracy (ACC)} &= \frac{TP + TN}{TP + TN + FP + FN} \\
 \text{Specificity (SPE)} &= \frac{TN}{TN + FP} \\
 \text{Sensitivity (SEN)} &= \frac{TP}{TP + FN} \\
 \text{Precision (PRC)} &= \frac{TP}{TP + FP} \\
 \text{F1 - score} &= \frac{2 \times \text{SEN} \times \text{PRC}}{\text{SEN} + \text{PRC}}
 \end{aligned} \tag{1}$$

Furthermore, missing information in some datasets (e.g., subject-level labels in DRIONS-DB (Carmona et al., 2008)) restricts research possibilities.

5. Feature extraction

The feature extraction process is pivotal in transforming raw data into a more manageable and conducive format for analysis. It involves deriving a set of representative features from the input data, serving as the foundation for subsequent analytical processes. The taxonomy adopted in this survey categorizes the feature extraction techniques into structural features, statistical features, and hybrid features. This categorization aligns with the general pipeline for glaucoma diagnosis, where structural features capture physical characteristics of the optic nerve head, statistical features provide image-derived information, and hybrid features combine both types for a comprehensive assessment. The extracted informative features are subsequently utilized as inputs to train the model to enhance its accuracy and reliability.

5.1. Structural features

Structural image measurements, focusing on the physical characteristics of the optic nerve head, are essential for clinicians in quantifying retinal structures relevant to glaucoma. These measurements yield objective data for assessing the severity and monitoring glaucoma progression. Incorporating clinical knowledge can significantly enhance the precision and interpretability of machine learning algorithms (Xu, Hu et al., 2021). The CDR, an important structural measurement extractable from fundus images, is widely used by clinicians due to its relevance and significance in glaucoma assessment. An elevated CDR signifies a larger Cup, correlating to an increased risk of glaucoma (Jiang et al., 2020). Therefore, many glaucoma screening works focus on accurate OC and OD segmentation tasks (Bajwa et al., 2019; Fu, Cheng, Xu, Wong et al., 2018; Haleem et al., 2018; Juneja et al., 2020; Latif et al., 2022; Liu et al., 2019; P., J., & R., 2021; Tabassum et al., 2020; Wang et al., 2021).

M., Issac, and Dutta (2018) introduced an adaptive threshold framework for segmenting the OC and OD using geometrical features. Employing diagnostic methodologies analogous to those used by ophthalmologists, the algorithm extracts and tracks blood vessels in the

Disc region, identifying the initial bending locations of different vessels and connecting them to outline the contours of the OC. They also computed the vertical CDR, representing the distance between the uppermost and lowermost locations of the OC and OD, and established a threshold to classify a fundus image as normal, suspect glaucoma, or glaucoma. [Mvoulana, Kachouri, and Akil \(2019\)](#) proposed using the K-means clustering algorithm. This method incorporates an intensity-based proximity criterion for pixel classification and employs a model-based boundary-fitting methodology using the circular Hough transform for OC and OD segmentation. Subsequently, they calculated the CDR value and set a threshold to differentiate between healthy individuals and those with glaucoma, achieving a 98% accuracy on the DRISHTI-GS1 dataset in glaucoma diagnosis.

Building upon earlier studies in employing deep learning techniques for glaucoma diagnostics, [Jiang et al. \(2018\)](#) used two Faster R-CNN architectures ([Ren, He, Girshick, & Sun, 2015](#)) for the individual segmentation of the OC and OD, with the goal of identifying the smallest bounding boxes for each area. They further developed their approach in [Jiang et al. \(2020\)](#), focusing on the unified segmentation of OC and OD by introducing an end-to-end region-based convolutional neural network, namely JointRCNN. JointRCNN is composed of four key elements: a module for feature extraction, a network for Disc proposal, a module for Disc attention, and a network for Cup proposal. The feature extraction module is used for both OC and OD segmentation tasks, with atrous convolution ([Chen, Zhu, Papandreou, Schroff, & Adam, 2018](#)) enhancing the feature extraction performance in this module. The Disc and Cup proposal networks define bounding box proposals, and a Disc attention module is introduced to interconnect the two networks.

[Fu, Cheng, Xu, Wong et al. \(2018\)](#) introduced polar transformation coupled with a multi-label deep network (M-Net) for the concurrent segmentation of the OC and OD in retinal images. The proposed M-Net has several components, each serving a unique purpose in the image analysis process. It incorporates a multi-scale input layer responsible for constructing an image pyramid. Additionally, a U-shaped convolutional network is employed to capture rich hierarchical representations. A side-output layer acts as a preliminary classifier to create a corresponding local prediction map across different scales. Lastly, a multi-label loss function is used to derive the concluding segmentation map. In this work, the CDR values were calculated specifically for the screening for glaucoma. The effectiveness of the M-Net was then evaluated on the ORIGA ([Zhang et al., 2010](#)) and Singapore Chinese Eye Study (SCES) datasets, achieving an AUC of 85.08% and 89.98%, respectively.

[Liu et al. \(2019\)](#) introduced a method for joint OC and OD segmentation that leverages a semi-supervised model, benefiting from both labeled and unlabeled datasets to enhance segmentation results. The architecture, based on conditional Generative Adversarial Nets (cGAN), incorporates a segmentation, a generator, and a discriminator network aimed at understanding the relationship between fundus images and their respective segmentation maps. Both the segmentation network and the generator are focused on learning the conditional distributions between the fundus images and their corresponding segmentation maps. Concurrently, the discriminator evaluates if the pairs of images and labels are sourced from the empirical joint distribution. Furthermore, the proposed method performed better than its fully-supervised version on both ORIGA and REFUGE datasets, with AUC values of 86.22% and 90.11% and accuracy rates of 76.57% and 82.78%, respectively.

Although segmentation-based methods have demonstrated efficacy, obtaining accurate CDR measurements still poses challenges due to issues like overlapping regions, low contrast between regions, variability and inhomogeneity of the OD's shape, lack of sufficient labels, and errors occurring in intermediate steps. [Zhao, Chen, Liu et al. \(2020\)](#) introduced a method for direct CDR estimation, leveraging a semi-supervised learning model that eliminates the need for intermediate segmentation steps. This method is structured in a two-stage cascaded manner, first applying unsupervised feature representation of fundus

images with a Convolutional Neural Network (CNN), followed by CDR value regression utilizing a Random Forest (RF) regressor. The proposed method achieved an AUC of 90.50% for diagnosing glaucoma on a dataset of 421 fundus images. [Zhou, Yi, Bao, and Wang \(2019\)](#) introduced an adaptive weighted locality-constrained sparse coding methodology for glaucoma diagnosis. This methodology integrates locality and sparse constraints. It employs a weighted locality constraint developed by adaptively integrating multiple distance measurements. The effectiveness of the approach was tested on two datasets, DRISHTI-GS and RIM-ONE-r2, where it achieved an accuracy of 88.63% and 85.56%, respectively, in diagnosing CDR.

[Medeiros, Jammal, and Thompson \(2019\)](#) curated a private data set comprising 32,820 pairs of OD fundus images and SD-OCT.³ scans. They then developed a deep learning model to predict the average thickness of the RNFL from the fundus images and facilitate the quantification of neural damage. They evaluated the capability of both predicted and actual RNFL thickness values to distinguish between glaucomatous and healthy eyes, achieving AUCs of 94% and 94.40%, respectively. [Raja Raja et al. \(2021\)](#) introduced a method to objectively grade glaucoma, categorizing it as either an early suspect or advanced stage, focusing on the degeneration of RGCs. To this end, they segmented the RNFL, GCIPL, and GCC regions to extract information related to their thickness. Subsequently, they used a Support Vector Machine (SVM) to assess the severity of glaucoma, achieving an accuracy of 91.17% on the AFIO dataset. [Lee, Kim, Park, and Jeoung \(2020\)](#) developed a deep learning framework, using NASNet ([Zoph, Vasudevan, Shlens, & Le, 2018](#)) backbone, designed to diagnose glaucoma by analyzing SD-OCT images. They also extracted features from maps representing GCIPL and RNFL thicknesses and deviations, which were incorporated into the deep learning model for glaucoma diagnosis. The model achieved an AUC of 99%, a sensitivity of 94.70%, and a specificity of 100% when tested on a private dataset comprising 350 glaucomatous and 307 healthy SD-OCT image sets.

5.2. Statistical features

The diagnosis of glaucoma also relies on statistical features such as intensity-based, texture-based, morphological, and wavelet-based measurements ([Semmlow & Griffel, 2021](#)). These features enable differentiation between normal and glaucomatous eyes. In recent years, advancements in this field have integrated these features with deep learning-based methods to enhance the accuracy and precision of glaucoma detection. For instance, [Claro et al. \(2019\)](#) determined the optimal feature set for representing fundus images. The comprehensive set included Local Binary Pattern (LBP), Gray Level Co-occurrence Matrix (GLCM) ([Haralick, Shanmugam, & Dinstein, 1973](#)), Histogram of Oriented Gradients (HOG) ([Dalal & Triggs, 2005](#)), Tamura ([Tamura, Mori, & Yamawaki, 1978](#)), Gray Level Run Length Matrix (GLRM) ([Galloway, 1975](#)), morphology, and seven CNN architectures. These features were combined to create a feature vector with 30,682 dimensions. Subsequently, they employed the gain ratio algorithm for feature selection, concluding that integrating the GLCM and pre-trained CNNs not only bridged the gap between conventional and state-of-the-art methods but also achieved 92.78% accuracy on a test set of 1675 images across DRISHTI-GS, RIM-ONE, HRF, JSIEC, and ACRIMA datasets.

[Juneja et al. \(2022\)](#) extracted GLRM and GLCM features from wavelet-filtered OCT images, along with 18 other statistical features. They used gain ratio, information gain, and correlation statistical methods to isolate discriminative features. In addition, they adopted a 3D-CNN architecture for feature extraction and classification. Majority voting and weighted decision fusion strategies were used to consolidate

³ SD-OCT is a variant of OCT that employs accelerated scanning speeds with higher resolution to produce more intricate images of the eye's internal structures compared to conventional time-domain OCT.

the final classification results, considering K-nearest neighbor (k-NN), RF, SVM, and the probabilities provided by the 3D-CNN model. The proposed framework achieved promising results on a publicly available dataset of 1110 OCT scans, with 847 glaucoma cases and 263 normal cases. The framework obtained 95% precision, 97% sensitivity, and an F1-score of 96%. Maheshwari, Pachori, and Acharya (2017) proposed a new method for diagnosing glaucoma using the Empirical Wavelet Transform (EWT) to decompose fundus images. They extracted correntropy features from the decomposed EWT components and used the t -value feature selection algorithm to rank them. These features were then fed into a least-squares SVM to classify normal and glaucoma images. The approach achieved an accuracy of 98.33% and a specificity of 96.67% on both a private and a public database. Nayak, Nayak, Das, Majhi, Bhandary, and Acharya (2021) conducted a study where they used a meta-heuristic optimization technique called the real-coded genetic algorithm to perform feature extraction. The goal was to extract high-level features from fundus images by maximizing inter-class distance and minimizing intra-class variability. They used an SVM classifier to diagnose glaucoma based on the derived features. The experiments were conducted on a dataset consisting of 1426 fundus images, which included 589 normal cases and 837 cases of glaucoma. The results demonstrated an accuracy of 97.20%.

Vinícius dos Santos Ferreira, Oseas de Carvalho Filho, Dalívia de Sousa, Corrêa Silva, and Gattass (2018) developed a framework that uses the U-Net model to perform semantic segmentation of the optic disc, which is a crucial Region of Interest (ROI) in the diagnosis of glaucoma. This segmentation allows for faster processing by excluding irrelevant image regions. Then, they extracted texture features from both the RGB channels and gray levels of the segmented region using phylogenetic diversity indexes. Finally, they classified images using a CNN-based architecture on RIM-ONE, DRIONS-DB, and DRISHTI-GS datasets. Experiments demonstrated promising results with a precision of 98.50%, sensitivity of 98%, specificity of 100%, F1-score of 96%, and AUC of 98.10% in the diagnosis of glaucoma. In a similar way, Bisneto, de Carvalho Filho, and Magalhães (2020) developed a cGAN framework that employed a U-Net generator and a PatchGAN discriminator for segmenting the optic disc. The cGAN, in combination with taxonomic indexes, facilitated the extraction of textural attributes from the segmented OD area, which were crucial in glaucoma classification. They utilized various classifiers, such as Multilayer Perceptron (MLP), Sequential Minimal Optimization, and RF, and attained 100% accuracy and AUC in diagnosing glaucoma using the RIM-ONE and DRISHTI-GS datasets.

5.3. Hybrid features

The integration of structural and statistical features can significantly improve the performance of diagnostic models for glaucoma. This approach provides a comprehensive perspective on the condition, revealing clinical measurements and anatomical changes in the retina through structural features, while statistical features provide valuable image-derived information. Combining both types of features makes the glaucoma screening process more effective and detailed, taking advantage of the strengths of each type of feature.

Balasubramanian and N.P. (2022) proposed a method that integrates structural and statistical features extracted from fundus images to improve the diagnosis of glaucoma. They used correlation-based feature selection algorithms with a Kernel-Extreme Learning Machine classifier. To derive structural features, they segmented the optic disc and optic cup using a Fuzzy C-Means Clustering algorithm to compute CDR and extract cup shape features. Additionally, they extracted several statistical features using different techniques, including GLCM, Anisotropic Dual-Tree Complex Wavelet Transform (ADT-CWT) (Swamidoss et al., 2013), Fractal Texture Analysis (Cheung et al., 2009), SURF (Bay, Tuytelaars, & Van Gool, 2006), Pyramid

HOG (Bosch, Zisserman, & Munoz, 2007), Mean Gray-Level, Color Intensity, and Super-pixels. Their proposed method achieved remarkable results, with 99.61% accuracy, 99.89% sensitivity, and 100% specificity when tested on a dataset consisting of 7280 images from both public and private retinal fundus datasets. Guo, Li, Tang, Zou, and Fan (2020) developed a model called Increasing Field of View (IFOV) that meticulously extracts various textural and statistical features from images. The model operates on four different image scales: the OC region, the OD region, an ROI that is a cropped image surrounding the OD, and the entire global fundus image. The Gabor transform and GLCM were applied to images at different scales to extract CDR and statistical features. An adaptive synthetic sampling approach (He, Bai, Garcia, & Li, 2008) was used for feature selection. The extracted features were then used to train a gradient-boosting decision tree classifier, which showed an accuracy rate of 84.30% on the ORIGA dataset and 83.70% on the DRISHTI-GS1 dataset for screening for glaucoma.

Thakur and Juneja (2020) proposed a hybrid feature set that combined structural and statistical features for classifying retinal fundus images. The features included CDR, disc damage likelihood scale, GLRM, GLCM, first-order measures, higher-order spectra, higher-order cumulant, and wavelets. They used the wrapper approach to select the best features and trained k-NN, Neural network (NN), RF, SVM, and Naïve Bayes (NB) classifiers for glaucoma diagnosis. Among them, SVM showed the highest accuracy rate of 97.20% on the DRISHTI-GS and RIM-ONE datasets, with a specificity of 96%, a precision of 97%, and a sensitivity of 97%. Kausu, Gopi, Wahid, Doma, and Niwas (2018) introduced a unique method for glaucoma identification using time-invariant feature CDR and ADT-CWT features. The initial step involved segmenting the OD using the Fuzzy C-Means method, followed by OC segmentation through Otsu's thresholding. The classification of glaucoma was then performed using an MLP model, achieving an accuracy of 97.6% and a sensitivity of 98% on a dataset collected from 51 healthy subjects and 35 glaucoma patients at the Venu Eye Institute & Research Centre.

The Inferior Superior Nasal Temporal (ISNT) rule is a widely recognized technique in ophthalmology to evaluate structural damage to the optic nerve head (Chan et al., 2013). According to this rule, the thickness of the NRR in healthy eyes reduces in a specific order, where the inferior, superior, nasal, and temporal regions are subject to this decrease, respectively. However, in the case of glaucomatous optic discs, changes in the thickness of NRR violate this rule. Pathan, Kumar, Pai, and Bhandary (2021) extracted clinical features from the segmented OD and OC regions and calculated CDR. They also employed ISNT rule verification in the NRR area. They used RGB, CIEL*a*b, and HSV color spaces to extract color-related features and applied the GLCM method to extract texture features from the NRR area to examine glaucoma-related changes in fundus images. They used an SVM, a three-layered NN, and an AdaBoost classifier with dynamic ensemble selection for classifying normal and glaucoma samples. Their experiments demonstrated that the SVM algorithm could achieve 95% accuracy on the DRISHTI-GS1 dataset and 90% on a private dataset. Singh, Garg, Khanna, and Bhadoria (2021) also focused on the ISNT rule and CDR, integrating 20 statistical features such as Homogeneity, Contrast, and Correlation. They used the combination of SVM, K-NN, NB, and a 3-layered MLP, achieving 95.82% sensitivity, 98.59% specificity, and 98.60% accuracy on the DRIONS-DB dataset.

Martins, Cardoso, and Soares (2020) proposed a method that involved the segmentation of the OC and OD as well as morphological features extraction. The extracted features included the CDR, vertical length CDR, and rim-to-disc area ratio, providing an interpretation of the ONH shape in compliance with the ISNT rule. The proposed pipeline also incorporated a MobileNetV2 (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2018) as the feature extractor backbone to assess the glaucoma diagnosis confidence level. This approach combined the calculated morphological features with the confidence level to render the final decision. The methodology demonstrated an accuracy of 87%, a sensitivity of 85%, a specificity of 88%, and an AUC of 93% on a consolidated dataset from various publicly available sources.

6. End-to-end glaucoma classification

End-to-end deep learning models have demonstrated promising results in classifying glaucoma, outperforming feature-based methods (Dixit et al., 2021; Guo, Peng, Sun, Li and Zhang, 2022; Hervella, Rouco, Novo, & Ortega, 2022; Li, Xu, Wang, Jiang, & Liu, 2019; Raghavendra et al., 2018, 2019; Thompson, Jammal, Berchuck, Martiottoni, & Medeiros, 2020; Xu, Hu et al., 2021; Zhao, Chen, Chen, & Li, 2022). These models excel because they can incorporate comprehensive contextual information throughout training, potentially minimizing information loss and enhancing generalizability. This section reviews studies concentrating on the use of these models for glaucoma diagnosis, categorizing them based on their architectures into Convolutional Neural Networks, Autoencoder-based Networks, Attention-based Networks, Generative Adversarial Networks, Geometric Deep Learning Networks, and Hybrid Networks.

6.1. Convolutional neural networks

Convolutional Neural Networks (CNNs) have been frequently used for diagnosing glaucoma, serving to distill higher-level features from input data (Deperlioglu et al., 2022; Jun et al., 2021; Kucur et al., 2018; Li et al., 2018; Maetschke et al., 2019; Ran et al., 2019; Younesi et al., 2024; Zhao, Liao, Zou, Chen, & Li, 2019). The studies referenced herein have either used state-of-the-art CNN architectures such as Inception-v3 (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016), ResNet (He, Zhang, Ren, & Sun, 2016), EfficientNet (Tan & Le, 2019), and DenseNet (Huang, Liu, Van Der Maaten, & Weinberger, 2017), or designed optimized CNN architectures from scratch for glaucoma classification. In the case of using state-of-the-art CNN architectures, researchers have usually fine-tuned the pre-trained architectures on ImageNet (Russakovsky et al., 2015) using glaucoma datasets. Table 3 summarizes the studies that used CNN for glaucoma classification.

Wang, Chen et al. (2020) proposed an end-to-end semi-supervised CNN architecture designed for multi-task learning. This model can classify OCT B-scan images as either normal or glaucomatous. It also explores the correlation between structural and functional changes in glaucomatous eyes. The proposed model had three key components, including a glaucoma classification module, a VF measurement regression module, and a shared feature extraction module with a ResNet-18 backbone. To validate the proposed method, they curated the HK dataset, one of the largest glaucoma OCT image datasets, containing 975,400 B-scans gathered from 4877 volumes.

Zhao et al. (2019) introduced the Weakly-Supervised Multi-Task Learning (WSMTL) method, designed to accurately generate evidence maps, segment the OD, and automate glaucoma diagnosis. The WSMTL framework incorporates a densely connected CNN with skip connections, facilitating a multi-scale representation of fundus structure and a pyramid integration architecture for generating high-resolution evidence maps. Additionally, it employs a constrained clustering branch for precise OD segmentation and a fully connected discriminator for glaucoma diagnosis. They trained the model on weakly labeled data with binary diagnostic labels (normal/glaucoma) to produce a pixel-level segmentation mask and predict glaucoma diagnosis labels. Liao et al. (2020) developed EAMNet, a clinically interpretable CNN architecture for glaucoma diagnosis. This model could aggregate features extracted from a CNN backbone with ResBlock at multiple scales, harmonizing semantic and localization information to enhance the accuracy of glaucoma diagnosis. EAMNet could also generate refined evidence activation maps to highlight the discriminative regions specific to glaucoma, providing transparent model interpretations.

Xue et al. (2022) proposed a structured framework with three distinct phases: (1) initial screening, (2) distinguishing between normal and glaucomatous conditions, and (3) classifying the severity of glaucoma into three levels. The first phase employs IOP measurements to identify potential glaucoma patients. Subsequently, two distinct ResNet

architectures, DetectionNet and ClassificationNet, are independently trained during the second and third phases. DetectionNet uses a fusion of fundus and Voronoi VF images to classify cases as normal or glaucomatous, while ClassificationNet uses Voronoi VF images to determine the severity of glaucoma, categorizing it as mild, moderate, or severe. Jun et al. (2021) developed a Transferable Ranking Convolutional Neural Network (TRk-CNN) designed for the multi-class classification of conditions as normal, glaucoma suspect, or glaucomatous. TRk-CNN uses DenseNet as its backbone CNN model and combines the weights of the initial classification model to integrate inter-class information with the final classification stage, optimizing the process where high correlation exists between classes.

Hemelings et al. (2021) explored an alternative approach to glaucoma diagnosis by formulating it as a regression task rather than a binary classification problem. They developed a deep learning model that estimates the CDR directly from fundus images without requiring explicit segmentation of the optic cup and disc. This regression-based approach leverages the continuity of disease severity and may be better suited for detecting early glaucomatous changes than binary classification models. The authors also found that their model effectively used contextual information beyond the optic nerve head region to improve CDR estimation accuracy.

6.2. Autoencoder-based networks

An autoencoder functions as an unsupervised neural network, using an encoder to map input data into a lower-dimensional latent space and a decoder to convert the latent features back to their original input dimensionality. Raghavendra et al. (2019) proposed a two-layer sparse autoencoder to extract pivotal and effective features from fundus images for glaucoma detection. This network incorporates two cascaded autoencoders for unsupervised feature learning and a softmax layer for glaucoma classification. Table 4 summarizes the studies that have used autoencoder-based architectures for diagnosing glaucoma.

Pal et al. (2018) introduced G-EyeNet, a deep autoencoder designed for glaucoma detection. G-EyeNet consists of three modules: an encoder, a decoder, and a classifier. The encoder and decoder are used for image reconstruction and unsupervised feature learning, while the classifier uses the salient features to classify glaucoma. The model is trained using multi-task learning, minimizing both the reconstruction and classification losses simultaneously. The experimental results, obtained by combining the HRF, RIM ONE v.3, DRISHTI-GS, and DRIONS-DB datasets, showed that G-EyeNet outperformed existing state-of-the-art networks with an AUC of 92.30%. Similarly, Raja et al. (2021) developed RAG-Net_{v2}, a deep convolutional-based autoencoder network that is capable of segmenting specific regions in the retinal ganglion cell and diagnosing glaucoma. These regions include the RNFL, ganglion cell-inner plexiform layer, and ganglion cell complex. In this study, the encoder end of RAG-Net_{v2} is used to assign RGC-aware labels to normal and glaucomatous samples.

Hervella et al. (2022) developed an end-to-end multi-task learning approach designed for simultaneous glaucoma classification and segmenting of the optic disc and cup in fundus images. The proposed architecture employs a shared encoder-decoder structure between tasks, using both pixel-level and image-level labels throughout the training process. To balance the contribution of both tasks to parameter updates during training, they implemented a multi-adaptive optimization strategy, eliminating the need for loss-weighting hyperparameters. They conducted experiments on the REFUGE and DRISHTI-GS datasets. The ablation study indicated that the best performance is only achieved using the multi-adaptive optimization and the proposed network architecture together.

Pascal et al. (2022) introduced MTL-IO, a deep learning model for detecting glaucoma in retinal fundus images, segmenting the OD and OC, and locating the fovea. The model employs a U-Net encoder-

Table 3

Comparison of papers utilizing CNN architectures for glaucoma classification. N: Normal, G: Glaucoma, S: Glaucoma Suspect, A: Advanced glaucoma, E: Early glaucoma, Avg: Average, T: Total.

Study	Model	Data type	Dataset	Classification performance measures (%)						Code
				ACC	SEN	SPE	PRC	F1-score	AUC	
Raghavendra et al. (2018)	18-layer CNN	Fundus	Private (589 N, 837 G)	98.13	98.00	98.30	-	-	-	-
Wu et al. (2020) ^c	EfficientNet-based	Fundus	Private (1586 G, 2244 N), RIGA LAG	93.29	96.03	91.42	-	-	98.29	Link
Li et al. (2018)	Inception-v3	Fundus	Private (29,466 N, 2620 S, 7659 G)	95.81	98.40	94.22	-	-	99.49	-
Jun et al. (2021)	TRk-CNN	Fundus	Private (403 N, 208 S, 381 G)	88.94 (Avg)	85.37 (S Vs. N), 90.36 (G Vs. N)	89.33 (Avg)	74.47 (S Vs. N), 94.94 (G Vs. N)	79.55 (S Vs. N), 92.59 (G Vs. N)	-	-
Hemelings et al. (2023) ^c	ResNet-50	Fundus	Private (6927 T, 2.14% G) Private (23,318 T, 1.08% G) 11 publicly available datasets ^a	-	81.00	95.00	-	-	97.60	-
Shibata et al. (2018)	ResNet-18	Fundus	Private (Training: 1424 G, 1818 N)	-	94.00	94.00	-	-	98.40	-
Ahn et al. (2018)	3-layer CNN	Fundus	Glaucoma Fundus	87.90	-	-	-	-	94.00	-
Hemelings et al. (2021) ^c	ResNet-50	Fundus	Private (13,551 T, 55% G)	-	-	-	-	-	94.00	-
Zhao et al. (2019)	WSMTL	Fundus	REFUGES	-	-	-	-	-	87.00	-
Liao et al. (2020)	EAMNet	Fundus	ORIGA	-	-	-	-	-	92.00	-
Xue et al. (2022)	ResNet	Fundus, VF	ORIGA	-	-	-	-	-	88.00	-
Kucur et al. (2018) ^c	7-layer CNN	VF	Private (1695 N, 1201 mild, 1607 moderate, 1868 severe)	85.50 (G Vs. N), 82.50 (Multi-class ^b)	-	-	-	-	88.10 (G Vs. N), 95.60 (Multi-class)	-
Wang, Chen et al. (2020) ^c	CNN	OCT	Rotterdam	-	-	-	87.40, 98.60	-	-	Link
Noury et al. (2022) ^c	3D-CNN	OCT	Budapest HK (2926 G, 1951 N)	92.70	-	-	98.60	-	-	-
Ran et al. (2019)	ResNet-34	OCT	Stanford (806 G, 425 N)	86.00	-	-	-	88.90	93.30	-
Thompson et al. (2020)	ResNet-34	OCT	Stanford (157 G, 241 N)	-	86.00	78.00	-	87.00	91.00	-
Thakoor, Koorathota, Hood, and Sajda (2021) ^c	Inception-v3	OCT	HK (959 G, 666 N)	-	73.00	73.00	-	76.00	80.00	-
Asaoka et al. (2019a)	6-layer CNN	OCT	India (461 G, 211 N)	-	93.00	71.00	-	91.00	94.00	-
			Nepal (461 G, 211 N)	-	79.00	79.00	-	80.00	87.00	-
			Private (2926 G, 1951 N)	91.00	89.00	96.00	-	-	96.90	-
			Private (612 G, 542 N)	-	-	-	-	-	96.00	-
			Private (192 G, 545 N)	90.40	-	-	-	-	-	Link
			Private (57 G, 44 N)	91.10	-	-	-	-	-	-

^a AIROGS, ORIGA, REFUGES1, LAG, ODIR, REFUGES2, GAMMA, RIM-ONEr3, RIM-ONE DL, ACRIMA, PAPILA.

^b Severity classification between mild, moderate and severe glaucoma.

^c Studies tested their model on an auxiliary dataset.

decoder convolutional network with independent optimizers for each task. By leveraging the similarities between related tasks, the multi-task model performs better than single-task models in detecting glaucoma,

using almost 3.5 times fewer parameters and comparable computational time. They evaluated the model on the REFUGES dataset, surpassing both single and multi-task models. The research also demonstrated

Table 4

Comparison of papers utilizing autoencoder-based architectures for glaucoma classification. SAE: Sparse Autoencoder, N: Normal, G: Glaucoma.

Study	Model	Data type	Dataset	Classification performance measures (%)					Code	
				ACC	SEN	SPE	PRC	F1-score		AUC
Hervella et al. (2022)^a	MT + M-Ada	Fundus	REFUGE	–	–	–	–	–	97.60	–
			DRISHTI-GS	–	–	–	–	–	94.74	–
Pascal et al. (2022)	MTL-IO	Fundus	REFUGE	–	–	–	–	–	96.76	–
Raghavendra et al. (2019)	SAE	Fundus	Private (589 N, 837 G)	95.30	95.20	–	96.80	95.00	–	–
Pal, Moorthy, and Shahina (2018)	G-EyeNet	Fundus	DRIONS-DB	–	–	–	–	–	92.30	–
Raja et al. (2021)	RAG-Net _{v2}	OCT	AFIO	94.91	97.14	91.66	94.44	95.77	98.71	Link

^a Studies tested their model on an auxiliary dataset.

that combining multi-task and transfer learning approaches can enhance generalization even with limited resources and small sample sizes.

[Ren et al. \(2020\)](#) presented a task decomposition algorithm in which an encoder–decoder model was designed to perform glaucoma classification and the semantic segmentation of OC and OD. The proposed method decomposes the single segmentation task into three subsequent sub-tasks. Each sub-task is trained to optimize individual loss functions at different levels of detail. To minimize the difference between the outputs of pixel-wise semantic segmentation and instance class prediction tasks, the framework uses sync-regularization. The model achieved remarkable results, with a Dice score of 80.88% and an HD score of 9.74 mm in the BRATS18 segmentation challenge, and a Dice score of 83.49% and an IoU of 91.29% in the REFUGE18 segmentation challenge.

6.3. Attention-based networks

The attention mechanism in deep learning models allows for the selective focus on salient parts of the input data, improving the precision and efficiency of predictions. Implementing attention mechanisms in glaucoma screening models is particularly intuitive, enabling the model to prioritize key pathological areas over irrelevant information.

The term “attention” in the reviewed literature is attributed to two distinct concepts. The first concept refers to models focusing on particular areas within the input data. This selective attention is informed by various types of auxiliary information. For instance, expert-derived attention maps guide the model’s focus towards regions deemed significant by domain specialists ([Li, Xu et al., 2020](#); [Li et al., 2019](#)). Similarly, heatmaps, produced by approaches like Grad-CAM, visually highlight the areas within the input that have the most impact on the model’s decision-making process ([George et al., 2020](#)). Additionally, models may focus on segments correlating with disease-related regions for more precise pathology identification ([Hervella et al., 2022](#); [Jiang et al., 2020](#); [Raja et al., 2021](#)). The second concept is integral to the design of attention-based architectures, such as Vision Transformers (ViT) ([Dosovitskiy et al., 2021](#)), where attention mechanisms are employed to weigh the importance of different parts of the input, allowing the model to focus on more relevant features and improve its performance in tasks like image classification. We will elaborate on these concepts subsequently, where a summary of the reviewed paper can be found in [Table 5](#).

[Li et al. \(2019\)](#) introduced AG-CNN, an Attention-based convolutional neural network designed for glaucoma detection. AG-CNN identifies salient regions using attention maps to filter out redundant information in fundus images. The model consists of three sub-nets that predict attention, localize pathological areas, and classify glaucoma. They curated the large-scale attention-based glaucoma (LAG) database, which contains 5824 fundus images labeled as either glaucoma-positive or negative. To assist ophthalmologists in highlighting the salient regions in the fundus images, they simulated an eye-tracking experiment to obtain the ground truth for attention maps. Their subsequent study ([Li, Xu et al., 2020](#)) expanded the LAG database to 11,760 fundus images and introduced a weakly supervised learning approach to integrate attention maps for glaucoma detection.

[George et al. \(2020\)](#) developed an end-to-end, attention-guided 3D CNN model to detect glaucoma and estimate Visual Field Index (VFI) in high-resolution 3D OCT volumes. This model has three streams, each with identical architecture. The first stream processes the raw 3D-OCT cube to learn global retinal structures. The inputs for the other two streams are determined during training, guided by 3D Grad-CAM ([Selvaraju et al., 2017](#)) attention heatmaps. The entire model is trained simultaneously to minimize the cumulative losses from the three streams. The final output is derived by fusing the predictions from all streams. [García, Del Amor et al. \(2021\)](#) developed a hybrid neural network, incorporating hand-crafted features with deep learning ones. This network used skip connections to incorporate tailored residual and attention modules. These modules served the purpose of enhancing and fine-tuning the representative features within the latent space. The backbone model was specifically designed to process raw OCT B-scans, and a feature descriptor was used to extract RNFL thickness as the hand-crafted features. The model was trained using a few-shot learning technique to classify healthy, early, and advanced glaucoma cases.

[Zhao et al. \(2022\)](#) proposed a teacher-student framework where both models have an identical architecture, including spatial and channel attention modules. The spatial attention module is designed to identify salient image regions directing the pruning of feature responses. The channel attention module selectively preserves feature channels crucial for the classification task. This framework is especially beneficial when dealing with imbalanced datasets, as it promotes a balanced feature distribution through feature distilling and re-weighting. [Guo, Li et al. \(2022\)](#) introduced a multitask teacher-student framework designed for unbiased glaucoma screenings and for visualizing areas pivotal for model decision-making. The teacher network, utilizing a ResNet-34 backbone, extracts semantic feature maps at different depths to construct a multi-scale discrimination module. It also incorporates a self-attention mechanism, enabling the network to concurrently focus on spatial and channel information. This approach aims to ensure the reliability of the generated evidence map and to provide dependable preliminary results for glaucoma classification. Meanwhile, the student network, featuring a dual-branch CNN structure and a collaborative learning module, performs glaucoma diagnosis and generates the corresponding evidence map.

Vision Transformer is an advanced deep learning architecture that employs a self-attention mechanism to comprehend the global attributes of an image, extending the transformer model’s capabilities to computer vision tasks ([Dosovitskiy et al., 2021](#)). [Wassel et al. \(2022\)](#) evaluated the effectiveness of eight ViT baseline models for glaucoma detection using a dataset compiled from six publicly available sets of fundus images. Additionally, they implemented voting and snapshot ensemble techniques. The former combined the 3, 5, and 7 best-performing standalone models, and the latter integrated 3 and 5 checkpoints of the top-performing ViT models.

[Xu, Guan et al. \(2021\)](#) proposed TIA-Net, a transfer-induced attention network that leverages features learned from similar ophthalmic data to extract holistic features. TIA-Net employs channel-wise attention and maximum mean discrepancy to derive deep features that capture distinctive patterns related to glaucoma. This method facilitates feature transferability by seamlessly transitioning between holistic and

Table 5

Comparison of papers utilizing attention-based architectures for glaucoma classification. N: Normal, G: Glaucoma, S: Glaucoma Suspect, A: Advanced glaucoma, E: Early glaucoma, Avg: Average, Att: Attention module.

Study	Model	Data type	Dataset	Classification performance measures (%)					
				ACC	SEN	SPE	PRC	F1-score	AUC
Zhao et al. (2022) ^c	CNN + Att	Fundus	LAG	97.12	95.20	98.16	–	95.47	99.28
			REFUGE	95.25	80.00	96.94	–	78.82	93.32
			RIM-ONE	93.96	89.74	97.12	–	90.91	97.19
Zhao, Chen, Chen and Li (2020)	CNN + Att	Fundus	LAG	97.12	97.21	97.07	–	96.65	99.31
Guo, Li, Shen and Shi (2022)	CNN + Att	Fundus	LAG	96.70	96.10	97.00	–	95.00	99.60
			6 publicly available datasets ^a	93.20	92.57	93.43	–	–	97.77
Wassel, Hamdi, Adly, and Torki (2022)	Swin (Liu et al., 2021) Cait (Touvron, Cord, Sablayrolles, Synnaeve and Jégou, 2021) CrossViT (Chen, Fan, & Panda, 2021) XciT (Ali et al., 2021) ResMlp (Touvron et al., 2023) DeiT (Touvron, Cord, Douze et al., 2021) ViT (Dosovitskiy et al., 2021) BeiT (Bao, Dong, Piao, & Wei, 2022)	Fundus	6 publicly available datasets ^a	94.50	89.92	96.04	–	–	97.90
			94.30	86.73	96.94	–	–	96.87	
			93.55	88.6	95.23	–	–	97.2	
			91.50	85.94	93.43	–	–	96.00	
			88.00	81.70	90.10	–	–	94.60	
			87.40	77.20	90.60	–	–	92.60	
			85.50	83.82	86.15	–	–	92.70	
Xu, Guan et al. (2021) ^c	TIA-Net	Fundus	Private (1005 N, 877 G)	85.70	84.90	86.90	–	–	92.90
			ORIGA	76.60	75.30	77.20	–	–	83.50
Li et al. (2019) ^c	AG-CNN	Fundus	LAG	95.30	95.40	95.20	–	95.10	97.50
			RIM-ONE	85.20	84.80	85.50	–	83.70	91.60
Li, Xu et al. (2020)	AG-CNN	Fundus	LAG	92.20	95.40	96.70	–	95.40	98.30
García, Del Amor et al. (2021)	CNN + Att	OCT	Private (90 N, 72 E, 57 A)	87.88	81.82	90.91	81.82	81.82	–
			(Avg ^b)	(Avg)	(Avg)	(Avg)	(Avg)	(Avg)	
Song et al. (2021)	DRT	OCT, VF	Private (697 G, 698 N)	88.30	93.70	82.40	–	88.90	93.90
George et al. (2020)	CNN + Grad-CAM	OCT	Private (427 N, 3355 G)	91.07	95.12	–	94.73	94.88	93.77

^a LAG, ODIR-5K, ORIGA, REFUGE, DRISHTI-GS1, HRF.

^b Average categorical results for discriminating between healthy, early, and advanced glaucoma samples.

^c Studies tested their model on an auxiliary dataset.

specific features. Song et al. (2021) introduced the Deep Relation Transformer (DRT) model, which used both OCT and VF modalities. This model comprises three sequential modules, including global relation, guided regional relation, and interaction transformer modules. These modules use deep reasoning and transformer mechanisms to examine implicit pairwise relations between OCT and VF data, enriching the representation with complementary information. The proposed DRT model surpasses existing methods and holds promise for precise glaucoma diagnosis using multimodal data.

6.4. Generative adversarial networks

Generative Adversarial Networks (GANs) are a subclass of deep neural networks designed to generate new samples that adhere to a given probability distribution (Goodfellow et al., 2014). GANs use a generator and a discriminator module. The generator trains to generate samples aligned with a given probability distribution, while the discriminator trains to differentiate the generated samples from real data. Diaz-Pinto, Colomer et al. (2019) explored synthesizing retinal images using a deep convolutional GAN. Additionally, they integrated a semi-supervised learning approach, modifying the GAN's discriminator to function as a 3-class classifier. This classifier can distinguish between normal, glaucomatous, and real/fake classes. Employing the semi-supervised learning approach allowed for enhanced performance of the classifier, even with a simplified architecture. The models, trained on an extensive dataset of 86,926 retinal images from 14 public databases, demonstrated an AUC of 90.17% in differentiating between normal and glaucomatous images.

Guo, Peng et al. (2022) trained a CycleGAN (Zhu, Park, Isola, & Efros, 2017) with a teacher-student approach to potentially minimize the appearance disparities between labeled images from source and target domains and to enhance the precision of glaucoma detection across different racial backgrounds. This framework has inter- and intra-image teacher models, a student model, and a backbone network to integrate the benefits of domain adaptation and semi-supervised learning. The inter-image teacher model uses CycleGAN to transfer knowledge to the student model by reducing the distillation loss. This strategy could potentially mitigate the domain shift issue and improve glaucoma detection efficacy. The intra-image teacher model uses the exponential moving average to leverage the unlabeled target domain and transfers knowledge to the student model by reducing prediction consistency loss. Additionally, the backbone network is a modified ResNet-50 that integrates the contextual features of the local OD region and the global fundus image. Table 6 shows the comparison between experimental outcomes on various datasets.

6.5. Geometric deep learning networks

Geometric deep learning (GDL) focuses on developing algorithms and network architectures capable of analyzing non-Euclidean structured information, such as graphs, manifolds, and point clouds, by integrating geometric principles and deep learning techniques (Bronstein, Bruna, LeCun, Szlam, & Vandergheynst, 2017). Thiéry, Braeu, Tun, Aung, and Girard (2023) proposed using GDL to diagnose glaucoma from an OCT image of the ONH. They experimentally validated it against a 3D CNN and RNFL thickness classification outputs. Using

Table 6
Comparison of studies in which GAN-based architectures were used for glaucoma classification.

Study	Model	Data type	Dataset	Classification performance measures (%)					Code
				ACC	SEN	SPE	F1-score	AUC	
Diaz-Pinto, Colomer et al. (2019)	DCGAN	Fundus	14 publicly available datasets ^a	–	82.90	79.86	84.29	90.17	Link
			LAG	98.14	98.62	98.17	–	96.41	
			REFUGE	98.74	96.85	96.57	–	97.06	
			ORIGA	97.64	97.13	97.62	–	96.93	
			DRISHTI-GS	98.45	97.66	96.84	–	97.04	
Guo, Peng et al. (2022) ^b	CycleGAN-based	Fundus	ACRIMA	96.52	97.62	96.72	–	97.26	–
			RIM-ONE-r1	97.48	98.31	97.14	–	97.60	
			RIM-ONE-r2	97.46	98.23	96.57	–	96.40	
			RIM-ONE-r3	96.46	97.18	96.82	–	96.56	

^a ORIGA-light, DRISHTI-GS1, RIM-ONE, s3choi86-HRF, HRF, DRIVE (Staal, Abramoff, Niemeijer, Viergever, & van Ginneken, 2004), MESSIDOR (Wang, Liu et al., 2019), DR KAGGLE (Diabetic Retinopathy Detection, 2024), STARE (Hoover & Goldbaum, 2003), e-optha (Decenci ere et al., 2013), ONHSD (Lowell et al., 2004), CHASEDB1 (Owen et al., 2011), DRIONS-DB, SASTRA (Narasimhan, Vijayarekha, JogiNarayana, SivaPrasad, & SatishKumar, 2012), ACRIMA.

^b Studies tested their model on an auxiliary dataset.

a deep learning model, they first segmented seven relevant neural and connective tissues from OCT images. After that, each ONH was represented as a cloud containing nearly 1000 3D points. Then they used GDL (PointNet (Charles, Su, Kaichun, & Guibas, 2017)) to diagnose glaucoma from a single 3D point cloud. The proposed geometric deep learning model achieved an AUC of 95% on a private dataset consisting of 873 glaucomatous and 3897 non-glaucomatous OCT scans, outperforming the results achieved with a 3D CNN (AUC of 87% on raw OCT images and 91% on segmented OCT scans) and the ones obtained only from RNFL thickness (AUC = 80%). In another study, Braeu et al. (2023) proposed to evaluate the experimental results of PointNet against dynamic graph convolutional neural network (DGCNN) (Wang, Sun et al., 2019) for glaucoma diagnosis. Following the same procedure as their previous paper (Thi ery et al., 2023), each ONH was transformed to a 3D point cloud and utilized for glaucoma diagnosis. They demonstrated that both the DGCNN and PointNet could accurately classify 2259 glaucomatous from 2247 Non-glaucomatous OCT scans based on 3D ONH point clouds with an AUC of 97% and 95%, respectively. Furthermore, they identified an hourglass pattern, constituting relevant 3D structural characteristics for glaucoma diagnosis. These patterns were localized in the lower and upper quadrants of the ONH.

6.6. Hybrid networks

Hybrid networks, which integrate diverse neural network architectures to leverage their unique strengths, have been used in glaucoma diagnosis. These networks can effectively process various types of data and tasks to discern intricate patterns and relationships within the data. Chai et al. (2018) developed a multi-branch neural network (MB-NN) to leverage domain knowledge and derive features from complex patterns within retinal fundus images for glaucoma diagnosis. This model uses the entire fundus image and the OD region image as inputs to the first and second branches, respectively, with the latter employing a FasterRCNN (Ren et al., 2015) trained on a distinct dataset. The domain knowledge extracted from these branches is then integrated into a 1D flattened vector and entered into the third branch of the model, enhancing its diagnostic proficiency.

Fu, Cheng, Xu, Zhang et al. (2018) proposed DENet, an ensemble network for automated glaucoma diagnosis. DENet consists of four streams: a global image stream, a segmentation-guided network, a local disc region stream, and a disc polar transformation stream. All streams, except the segmentation-guided network, have a ResNet-50 backbone. The segmentation-guided network uses a U-shape convolutional network to detect the optic disc region and guide glaucoma screening on the entire fundus image. These streams provide complementary information, and their output probabilities are fused for the final classification result.

Yu et al. (2020) recognized that critical information contained in the raw multi-rater grading is discarded when the final ground-truth label is only used for model training. To address this problem,

they proposed a multi-branch model that employed raw multi-rater grading to improve the performance of glaucoma classification. In this multi-branch model, one branch optimizes sensitivity, one optimizes specificity, and one provides balanced fused result for the input images. They introduced a consensus loss to promote uniformity in results from the sensitivity and specificity branches for consensus labels and to encourage divergent results for disagreement labels. In this way, the grade of consistency or inconsistency between the predictions of the two branches implies the difficulty level of an image. This information then directed the balanced fusion branch to focus more on difficult cases.

Garc a, Colomer and Naranjo (2021) introduced a fusion of CNN and LSTM architectures to leverage the spatial dependencies of the features extracted from raw SD-OCT volumes. The model is structured with a feature extractor at the slide level and a volume-based predictor. This feature extractor employs residual and attention convolutional modules, enhanced with fine-tuning techniques, to capture intricate patterns and details. The model incorporates LSTM architecture with a sequential weighting module to integrate spatial information with three-dimensional data effectively. This module helps to optimize the LSTM outputs, resulting in a more steady and effective model training. A summary of results is presented in Table 7.

7. Glaucoma prediction and progression detection

Glaucoma is a degenerative disease that often exhibits minimal symptoms during its initial stages (Li et al., 2022). Therefore, predicting the risk of developing glaucoma is essential for timely detection and intervention to prevent vision loss. Recent approaches have shown the applicability of deep learning models in predicting the transition from non-glaucoma to glaucoma through the analysis of diverse demographic, clinical, and imaging data (Dixit et al., 2021; Li et al., 2022; Thakur, Goldbaum, & Yousefi, 2020b). Li et al. (2022), for example, suggested the use of a CNN-based network to predict and stratify the risk of glaucoma onset and progression using fundus images of 17,497 eyes in 9346 patients. The model demonstrated the ability to predict patients who may develop glaucoma within a five-year period with an AUC of 90%. Thakur et al. (2020b) also used deep learning models to predict the development of glaucoma from fundus images in a prospective longitudinal study several years before disease onset. The study reported an AUC of 77% for predicting the development of glaucoma 4 to 7 years before its onset, 88% for predicting the development of glaucoma nearly 1 to 3 years before the onset, and 95% for detecting glaucoma after the onset.

Li, Wang et al. (2020) introduced the DeepGF algorithm using sequential fundus images for glaucoma prediction. They first established a database of sequential fundus images for glaucoma prediction (SIGF), which included an average of 9 images per eye, for a total of 3671 images. The proposed DeepGF consists of an attention-polar

Table 7
Comparison of papers utilizing hybrid architectures for glaucoma classification. N: Normal, G: Glaucoma, SINDI: Singapore Indian Eye Study.

Study	Model	Data type	Dataset	Classification performance measures (%)					Code
				ACC	SEN	SPE	F1-score	AUC	
Fu, Cheng, Xu, Zhang et al. (2018) ^a	DENet	Fundus	ORIGA, SCES (46 G, 1630 N)	84.29	84.78	83.80	–	91.83	Code
			SINDI (113 G, 5670 N)	74.95	78.76	71.15	–	83.80	
Chai et al. (2018)	MB-NN	Fundus	Private (1023 G, 1531 N)	91.51	92.33	90.90	–	–	–
			Private (2952 G, 3366 N)	92.54	90.96	93.92	91.89	97.94	–
Yu et al. (2020) ^a	Multi-rater deep model	Fundus	DRISHTI-GS1	86.14	91.43	74.19	90.14	89.63	–
			REFUGE	98.00	82.50	99.72	–	96.83	–
García, Colomer et al. (2021)	CNN + LSTM	OCT	Private (144 G, 176 N)	81.25	75.86	85.71	78.57	80.79	–

^a Studies tested their model on an auxiliary dataset.

CNN alongside a variable time interval long short-term memory (VTI-LSTM) network. The architecture was designed with the objective of capturing spatio-temporal transitions within sequential medical images at different time intervals. Furthermore, active convergence training was introduced to address the imbalanced sample distribution issue inherent in glaucoma forecasting. The authors reported an accuracy of 80.7%, a sensitivity of 85.70%, a specificity of 80.60%, and an AUC of 87% for glaucoma prediction. Thakur, Goldbaum, and Yousefi (2020a) developed a deep archetypal algorithm capable of predicting the onset of glaucoma approximately four years in advance. They used simplex projections to derive unsupervised convex representations of VFs, proving clinically insightful and more discriminative than raw VFs or conventional VF analysis methods. They evaluated the proposed algorithm on a dataset with 7248 VF tests collected at baseline, achieving an AUC of 71%.

Dixit et al. (2021) proposed using a convolutional LSTM model to detect glaucoma progression, leveraging a longitudinal dataset consisting of VF and clinical data from patients diagnosed with glaucoma. The database comprises 11,242 eyes with at least four VF tests and the associated clinical data, including CDR, central corneal thickness, and IOP for each sample. The experimental results demonstrated an AUC of 93.93% for distinguishing progressing glaucoma from stable cases.

While predicting the onset and future progression of glaucoma is crucial, accurately detecting whether glaucoma has progressed post-hoc is also important for monitoring and managing the disease. Several studies have focused on developing deep learning models for this task. For instance, Chen et al. (2019) proposed a method for early glaucoma screening that combines information from enhanced depth imaging optical coherence tomography (EDI-OCT) and fundus images. In addition to glaucoma detection, their approach could measure changes in the lamina cribrosa depth and deformation over time, potentially enabling the detection of glaucoma progression using longitudinal EDI-OCT and fundus image data. These methods typically leverage longitudinal data, such as sequential fundus images, OCT scans, or visual field tests, to quantify changes over time and determine if glaucoma has progressed beyond a certain threshold. The ability to reliably detect progression can inform treatment decisions and enable timely interventions to slow or halt further vision loss.

8. Challenges and future directions

The adoption of deep learning for glaucoma diagnosis faces several key challenges that need to be addressed through continued research and cross-disciplinary collaboration. We structure these challenges and promising future directions as follows.

8.1. Data challenges

Developing effective deep learning models for glaucoma diagnosis relies on the availability of comprehensive training datasets. However, assembling high-quality annotated datasets presents several key challenges.

One major challenge is the limited size of publicly available glaucoma datasets compared to common computer vision benchmarks like

ImageNet. The prohibitive expertise and cost required for reliable manual annotations further constrains dataset development. Data diversity is another issue, with many datasets lacking varied representations across different demographics, ethnicities, and imaging equipment (De Fauw et al., 2018; Orlando et al., 2020). These limitations have caused existing public datasets to remain relatively small in scale, which can hinder model performance and restrict real-world applicability. However, alternative training methods such as transfer learning (Latif et al., 2022; Xu, Guan et al., 2021), zero/few-shot learning (García, Del Amor et al., 2021; Wang, Yao, Kwok & Ni, 2020), and knowledge distillation (Chelaramani, Gupta, Agarwal, Gupta, & Habash, 2021; Guo, Li et al., 2022; Zhao et al., 2022) offer potential solutions to mitigate the impact of limited dataset size. Additionally, active learning techniques can optimize the annotation process by selectively identifying the most useful and ambiguous samples for labeling. Incorporating uncertainty estimation into the sample selection process (Adhane, Dehshibi, & Masip, 2022) further focuses active learning on improving model performance on underrepresented classes.

Imbalanced class distribution poses another persistent problem, with normal cases dominating many current datasets. This can skew model performance toward low sensitivity and high false negative rates (Thabtah, Hammoud, Kamalov, & Gonsalves, 2020). Techniques like data augmentation (Asaoka, Tanito et al., 2019; Shorten & Khoshgoftaar, 2019), weighted sampling (Cui, Jia, Lin, Song, & Belongie, 2019), and generation of synthetic minority over-samples (Li, Wang, Liu, Chen, & Chen, 2023) may help mitigate imbalance. However, augmentation requires judicious implementation with clinician input to prevent medical inaccuracies and ensure fidelity (Finlayson et al., 2019).

Healthcare data also carries ethical and legal obligations around privacy that researchers must proactively address (Kondylakis et al., 2023). Laws and cultural norms around medical data sharing vary geographically, necessitating localized considerations. Policymakers have a role in developing balanced frameworks that promote research and innovation while protecting patient privacy. Techniques like federated learning (Zhang, Qu, Singh, Kalpathy-Cramer, & Rubin, 2022) can also enable collaborative model development without raw data sharing. However, the onus remains on researchers to be privacy stewards and ensure compliance with applicable national and regional privacy laws in their research studies (Vollmer et al., 2020).

8.2. Model development challenges

Deep learning models need to detect subtle signs of early-stage glaucoma, which is challenging even for experienced clinicians (Asaoka et al., 2019b). Heterogeneity in early disease appearance further complicates this task. Researchers are exploring various model architecture techniques to improve performance for early detection. For instance, attention mechanisms have shown promise in focusing models on salient retinal regions, enhancing informative feature extraction, and representation learning (Li, Xu et al., 2020). Regression models have also been explored for glaucoma diagnosis rather than binary classification (Hemelings et al., 2021). This regression-based approach aims to

estimate continuous biomarkers, such as the cup-to-disc ratio, directly from fundus or OCT images. Regression models may be better equipped to detect subtle early changes associated with glaucoma progression by leveraging the continuity of disease severity. Furthermore, these models can effectively use contextual information beyond the optic nerve head region, potentially improving their diagnostic performance. However, careful validation and comparison against traditional classification approaches are still needed to assess the relative strengths and limitations of regression-based methods for glaucoma diagnosis.

Integrating multi-modal data and multi-label learning may further improve performance by leveraging interrelated tasks like segmentation and diagnosis (Song et al., 2021). However, this poses challenges such as optimal data fusion, balancing diverse tasks, and continual learning as new data emerges. To address these issues, researchers are exploring specialized multi-modal architectures, adaptive optimization strategies, dynamic network expansion, and meta-learning (Gao et al., 2023).

Despite the potential benefits, the opaque nature of deep learning models remains a challenge. The importance of employing or proposing explainable AI (XAI) techniques to explain predictions and feature importance from deep learning models using different image modalities like OCT, Fundus, and VF has been recognized by researchers (Hemelings et al., 2021; Shibata et al., 2018). These studies employed post-hoc explanation techniques like SHAP values (Lundberg & Lee, 2017), locally interpretable model-agnostic explanations (Ribeiro, Singh, & Guestrin, 2016), integrated gradients (Sundararajan, Taly, & Yan, 2017), occlusion sensitivity (Zeiler & Fergus, 2014), saliency maps (Vasu & Long, 2020), and contrastive explanations (Dhurandhar et al., 2018) to identify influential segments, patterns, and abnormalities in fundus images that lead the models to predict glaucoma. They found that XAI methods can highlight relevant regions like the optic disc, retinal nerve fiber layer defects, and areas of hemorrhaging that most inform the models' predictions (Mehta et al., 2021; Ting et al., 2019). However, further development of XAI techniques tailored for glaucoma is still necessary, as most methods are model-agnostic and not optimized for glaucomatous features. Most work has focused on post-hoc methods rather than real-time explainability during model development and evaluation. To enhance understanding and trust in AI-assisted glaucoma screening, the field needs to create glaucoma-specific XAI solutions that provide human-interpretable explanations in real time. Another pragmatic approach involves clinician-researcher collaboration to assess XAI techniques and identify medically relevant insights (Liao et al., 2020).

8.3. Clinical translation challenges

For clinical adoption, glaucoma diagnosis models must demonstrate consistent performance across diverse populations and imaging equipment. However, demographic and acquisition differences can affect model generalizability. Wider collaborations between academia, healthcare providers, and industry could facilitate large-scale external validation across multiple centers and populations to identify failure modes. The models should integrate seamlessly into ophthalmic workflows with efficient computations on edge devices (Rahmani et al., 2018). Physician trust is essential for uptake, requiring initiatives to improve model explainability and transparency. Frameworks for uncertainty estimation could also indicate situations where clinician oversight is necessary.

Careful design of human-AI interaction mechanisms will be vital, allowing physicians to accept, reject, or modify model recommendations. Regulatory agencies play a crucial role in establishing standards for rigorous clinical validation of deep learning systems before approval (Kelly, Karthikesalingam, Suleyman, Corrado, & King, 2019). Beyond accuracy, aspects like usability, interoperability, cybersecurity, and patient privacy must be addressed to ensure safety and effectiveness (Matheny, Whicher, & Israni, 2020). Overall, a patient-centered approach with clinical partnership throughout the life cycle of model development will be key to clinical translation.

8.4. Future outlook

Advancements in the diagnosis of glaucoma will require synergistic progress across multiple disciplines. On the data front, innovations in sensor and imaging technologies can enable the acquisition of informative multi-modal datasets. In parallel, advances in deep learning architectures, optimization algorithms, and computing hardware will allow more sophisticated analysis of these rich datasets.

The recent wave of interest in Large Language Models (LLMs) has presented exciting possibilities for the future of glaucoma diagnosis. Studies like (Bernstein et al., 2023; Delsoz et al., 2023) have demonstrated the potential for LLMs to analyze vast amounts of medical literature and patient data, potentially aiding in identifying risk factors, predicting disease progression, and even suggesting personalized treatment plans. However, further research is necessary to explore the integration of LLMs' ability to process complex information and deep learning's proficiency in image analysis. It is crucial to ensure responsible development and clinical validation of such combined models before real-world implementation.

Ultimately, sustained collaborative efforts across medicine, engineering, and computer science will be the key to realizing AI's potential to improve glaucoma outcomes. Initiatives to promote open datasets, model repositories, and evaluation benchmarks will facilitate collective progress. Interdisciplinary teams should lead technological development, ensuring clinical applicability and integration into practice. With concerted efforts, AI-enabled glaucoma care could soon transition from promise to reality.

9. Conclusion

In this paper, we have provided a comprehensive overview of the state-of-the-art research that applies deep learning and computer vision techniques for the diagnosis of glaucoma. We synthesized the existing work across diverse architectural categories, including CNNs, autoencoders, attention networks, GANs, and geometric deep learning models.

The review highlighted promising capabilities demonstrated by these techniques in analyzing fundus, OCT, and visual field data. Tasks like classification, segmentation, and prediction of glaucoma have shown strong results across a wide range of experiments. We also discussed different feature extraction approaches, covering structural, statistical, and hybrid techniques for identifying informative glaucoma biomarkers from retinal imaging data. However, key challenges remain around limited dataset size and diversity, class imbalance, optimizing models for early disease detection, integrating multi-modal data, and translating solutions to clinical practice.

Ongoing efforts are beginning to address these gaps through transfer learning approaches, data augmentation techniques, attention mechanisms, multitask learning, model explainability, and physician collaboration. Additionally, AI-based glaucoma diagnosis systems present notable economic benefits by automating routine tasks, enhancing diagnostic efficiency, and improving scalability in underserved areas. In addition, AI tools could identify glaucoma in patients undergoing retinography for other conditions (e.g., age-related macular degeneration or diabetic retinopathy) and raise alarms to consult with an ophthalmologist. It is important to note that AI is not intended to substitute ophthalmologists but to complement them, serving as a tool to enhance diagnosis and patient care. Nonetheless, realizing AI's full potential in transforming glaucoma care will require sustained cross-disciplinary teamwork.

From our perspective, open datasets, model repositories, and evaluation benchmarks will be critical to accelerate collective progress. Ultimately, an integrated approach spanning medicine, engineering, and computer science will be essential for developing and validating solutions ready for real-world clinical deployment. This review highlights the tremendous opportunities for us at the intersection of ophthalmology and artificial intelligence. Overall, we aimed to provide a comprehensive overview and analysis of the state-of-the-art in this exciting and high-impact emerging field.

CRedit authorship contribution statement

Mona Ashtari-Majlan: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Mohammad Mahdi Dehshibi:** Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. **David Masip:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data presented in this study is available within the main text.

Acknowledgments

We are grateful for the support Provided to this study. Mona Ashtari-Majlan and David Masip receive financial support from Spanish Ministry of Science and Innovation, which is funded by the FEDER initiative (PID2022-138721NB-I00). Mohammad Mahdi Dehshibi receives financial support from the European Research Council (ERC) through the Horizon 2020 research and innovation program, with grant agreement number 101002711 for the BODYinTRANSIT project.

References

- Adhane, G., Dehshibi, M. M., & Masip, D. (2022). On the use of uncertainty in classifying aedes albopictus mosquitoes. *IEEE Journal of Selected Topics in Signal Processing*, 16(2), 224–233. <http://dx.doi.org/10.1109/JSTSP.2021.3122886>.
- Ahn, J. M., Kim, S., Ahn, K.-S., Cho, S.-H., Lee, K. B., & Kim, U. S. (2018). A deep learning model for the detection of both advanced and early glaucoma using fundus photography. *PLoS One*, 13(11), 1–8. <http://dx.doi.org/10.1371/journal.pone.0207982>.
- Akter, N., Fletcher, J., Perry, S., Simunovic, M. P., Briggs, N., & Roy, M. (2022). Glaucoma diagnosis using multi-feature analysis and a deep learning technique. *Scientific Reports*, 12(1), 8064. <http://dx.doi.org/10.1038/s41598-022-12147-y>.
- Ali, A., Touvron, H., Caron, M., Bojanowski, P., Douze, M., Joulin, A., et al. (2021). Xcit: Cross-covariance image transformers. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, J. W. Vaughan (Eds.), *Vol. 34, Advances in neural information processing systems* (pp. 20014–20027). Curran Associates, Inc..
- Almazroa, A., Sun, W., Alodhayb, S., Raahemifar, K., & Lakshminarayanan, V. (2017). Optic disc segmentation for glaucoma screening system using fundus images. *Clinical Ophthalmology*, 11, 2017–2029. <http://dx.doi.org/10.2147/OPTH.S140061>.
- Asaoka, R., Murata, H., Hirasawa, K., Fujino, Y., Matsuura, M., Miki, A., et al. (2019a). Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *American Journal of Ophthalmology*, 198, 136–145. <http://dx.doi.org/10.1016/j.ajo.2018.10.007>.
- Asaoka, R., Murata, H., Hirasawa, K., Fujino, Y., Matsuura, M., Miki, A., et al. (2019b). Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images. *American Journal of Ophthalmology*, 198, 136–145. <http://dx.doi.org/10.1016/j.ajo.2018.10.007>.
- Asaoka, R., Tanito, M., Shibata, N., Mitsuhashi, K., Nakahara, K., Fujino, Y., et al. (2019). Validation of a deep learning model to screen for glaucoma using images from different fundus cameras and data augmentation. *Ophthalmology Glaucoma*, 2(4), 224–231. <http://dx.doi.org/10.1016/j.ogla.2019.03.008>.
- Bajwa, M. N., Malik, M. I., Siddiqui, S. A., Dengel, A., Shafait, F., Neumeier, W., et al. (2019). Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. *BMC Medical Informatics and Decision Making*, 19(1), 1–16. <http://dx.doi.org/10.1186/s12911-019-0842-8>.
- Balasubramanian, K., & N.P., A. (2022). Correlation-based feature selection using bio-inspired algorithms and optimized KELM classifier for glaucoma diagnosis. *Applied Soft Computing*, 128, Article 109432. <http://dx.doi.org/10.1016/j.asoc.2022.109432>.
- Bao, H., Dong, L., Piao, S., & Wei, F. (2022). BEit: BERT pre-training of image transformers. In *International conference on learning representations* (pp. 1–18). URL <https://openreview.net/forum?id=p-BhZSz59o4>.
- 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 & Stereology*, 39(3), 161–167. <http://dx.doi.org/10.5566/ias.2346>.
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. *Lecture Notes in Computer Science*, 3951, 404–417. http://dx.doi.org/10.1007/11744023_32.
- Bernstein, I. A., Zhang, Y. V., Govil, D., Majid, I., Chang, R. T., Sun, Y., et al. (2023). Comparison of ophthalmologist and large language model chatbot responses to online patient eye care questions. *JAMA Network Open*, 6(8), <http://dx.doi.org/10.1001/jamanetworkopen.2023.30320>, e2330320–e2330320.
- 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, 1135–1150. [http://dx.doi.org/10.1016/S1474-4422\(22\)00173-9](http://dx.doi.org/10.1016/S1474-4422(22)00173-9).
- Bisneto, T. R. V., de Carvalho Filho, A. O., & Magalhães, D. M. V. (2020). Generative adversarial network and texture features applied to automatic glaucoma detection. *Applied Soft Computing*, 90, Article 106165. <http://dx.doi.org/10.1016/j.asoc.2020.106165>.
- Bosch, A., Zisserman, A., & Munoz, X. (2007). Representing shape with a spatial pyramid kernel. In *Proceedings of the 6th ACM international conference on image and video retrieval* (pp. 401–408). Association for Computing Machinery, <http://dx.doi.org/10.1145/1282280.1282340>.
- Braeu, F. A., Thiéry, A. H., Tun, T. A., Kadziauskiene, A., Barbastathis, G., Aung, T., et al. (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. <http://dx.doi.org/10.1016/j.ajo.2023.01.008>.
- Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4), 18–42. <http://dx.doi.org/10.1109/MSP.2017.2693418>.
- Budai, A., Bock, R., Maier, A., Hornegger, J., & Michelson, G. (2013). Robust vessel segmentation in fundus images. *International Journal of Biomedical Imaging*, 2013, <http://dx.doi.org/10.1155/2013/154860>.
- Carmona, E. J., Rincón, M., García-Feijó, J., & de-la Casa, J. M. M. (2008). Identification of the optic nerve head with genetic algorithms. *Artificial Intelligence in Medicine*, 43(3), 243–259. <http://dx.doi.org/10.1016/j.artmed.2008.04.005>.
- Cen, L.-P., Ji, J., Lin, J.-W., Ju, S.-T., Lin, H.-J., Li, T.-P., et al. (2021). Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. *Nature Communications*, 12(1), 4828. <http://dx.doi.org/10.1038/s41467-021-25138-w>.
- 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. <http://dx.doi.org/10.1016/j.knsys.2018.07.043>.
- Chan, E. W., Liao, J., Chao Ming Foo, R., Loon, S. C., Aung, T., Wong, T. Y., et al. (2013). Diagnostic performance of the ISNT rule for glaucoma based on the heidelberg retinal tomograph. *Translational Vision Science & Technology*, 2(5), <http://dx.doi.org/10.1167/tvst.2.5.2>, 2–2.
- Charles, R. Q., Su, H., Kaichun, M., & Guibas, L. J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation. In *2017 IEEE conference on computer vision and pattern recognition* (pp. 77–85). IEEE, <http://dx.doi.org/10.1109/CVPR.2017.16>.
- Chelaramani, S., Gupta, M., Agarwal, V., Gupta, P., & Habash, R. (2021). Multi-task knowledge distillation for eye disease prediction. In *2021 IEEE winter conference on applications of computer vision* (pp. 3983–3993). IEEE, <http://dx.doi.org/10.1109/WACV48630.2021.00403>.
- Chen, C.-F. R., Fan, Q., & Panda, R. (2021). CrossViT: Cross-attention multi-scale vision transformer for image classification. In *2021 IEEE/CVF international conference on computer vision* (pp. 347–356). <http://dx.doi.org/10.1109/ICCV48922.2021.00041>.
- Chen, Z., Zheng, X., Shen, H., Zeng, Z., Liu, Q., & Li, Z. (2019). Combination of enhanced depth imaging optical coherence tomography and fundus images for glaucoma screening. *Journal of Medical Systems*, 43, 1–12. <http://dx.doi.org/10.1007/s10916-019-1303-8>.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Computer vision – ECCV 2018* (pp. 833–851). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-01234-2_49.
- Cheung, N., Donaghue, K. C., Liew, G., Rogers, S. L., Wang, J. J., Lim, S.-W., et al. (2009). Quantitative assessment of early diabetic retinopathy using fractal analysis. *Diabetes Care*, 32(1), 106–110. <http://dx.doi.org/10.2337/dc08-1233>.
- Claro, M., Veras, R., Santana, A., Araújo, F., Silva, R., Almeida, J., et al. (2019). An hybrid feature space from texture information and transfer learning for glaucoma classification. *Journal of Visual Communication and Image Representation*, 64, Article 102597. <http://dx.doi.org/10.1016/j.jvcir.2019.102597>.
- Cui, Y., Jia, M., Lin, T.-Y., Song, Y., & Belongie, S. (2019). Class-balanced loss based on effective number of samples. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9268–9277). IEEE, <http://dx.doi.org/10.1109/CVPR.2019.00949>.
- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *Vol. 1, In 2005 IEEE computer society conference on computer vision and pattern recognition* (pp. 886–893). <http://dx.doi.org/10.1109/CVPR.2005.177>.

- De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., et al. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature Medicine*, 24(9), 1342–1350. <http://dx.doi.org/10.1038/s41591-018-0107-6>.
- Decenci re, E., Cazuguel, G., Zhang, X., Thibault, G., Klein, J.-C., Meyer, F., et al. (2013). TeleOphta: Machine learning and image processing methods for teleophthalmology. *IRBM*, 34(2), 196–203. <http://dx.doi.org/10.1016/j.irbm.2013.01.010>, Special issue : ANR TECSAN : Technologies for Health and Autonomy.
- Delsoz, M., Raja, H., Madadi, Y., Tang, A. A., Wirosko, B. M., Kahook, M. Y., et al. (2023). The use of ChatGPT to assist in diagnosing glaucoma based on clinical case reports. *Ophthalmology and Therapy*, 12(6), 3121–3132. <http://dx.doi.org/10.1007/s40123-023-00805-x>.
- Deperlioglu, O., Kose, U., Gupta, D., Khanna, A., Giampaolo, F., & Fortino, G. (2022). Explainable framework for glaucoma diagnosis by image processing and convolutional neural network synergy: Analysis with doctor evaluation. *Future Generation Computer Systems*, 129, 152–169. <http://dx.doi.org/10.1016/j.future.2021.11.018>.
- Dervisevic, E., Pavljasevic, S., Dervisevic, A., & Kasumovic, S. S. (2016). Challenges in early glaucoma detection. *Medical Archives*, 70(3), 203–207. <http://dx.doi.org/10.5455/medarh.2016.70.203-207>.
- Dhurandhar, A., Chen, P.-Y., Luss, R., Tu, C.-C., Ting, P., Shanmugam, K., et al. (2018). Explanations based on the missing: Towards contrastive explanations with pertinent negatives. In *Advances in neural information processing systems* (pp. 1–12). Curran Associates, Inc..
- Diabetic Retinopathy Detection (2024). Diabetic retinopathy detection. <https://www.kaggle.com/competitions/diabetic-retinopathy-detection>. (Accessed 9 July 2024).
- Diaz-Pinto, A., Colomer, A., Naranjo, V., Morales, S., Xu, Y., & Frangi, A. F. (2019). Retinal image synthesis and semi-supervised learning for glaucoma assessment. *IEEE Transactions on Medical Imaging*, 38(9), 2211–2218. <http://dx.doi.org/10.1109/TMI.2019.2903434>.
- Diaz-Pinto, A., Morales, S., Naranjo, V., K hler, T., Mossi, J. M., & Navea, A. (2019). CNNs for automatic glaucoma assessment using fundus images: an extensive validation. *BioMedical Engineering Online*, 18, 1–19. <http://dx.doi.org/10.1186/s12938-019-0649-y>.
- 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. <http://dx.doi.org/10.1016/j.ophtha.2020.12.020>.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. In *International conference on learning representations* (pp. 1–21). URL <https://openreview.net/forum?id=YicbFdNTTy>.
- Finlayson, S. G., Bowers, J. D., Ito, J., Zittrain, J. L., Beam, A. L., & Kohane, I. S. (2019). Adversarial attacks on medical machine learning. *Science*, 363(6433), 1287–1289. <http://dx.doi.org/10.1126/science.aaw4399>.
- Fu, H., Cheng, J., Xu, Y., Wong, D. W. K., Liu, J., & Cao, X. (2018). Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. *IEEE Transactions on Medical Imaging*, 37(7), 1597–1605. <http://dx.doi.org/10.1109/TMI.2018.2791488>.
- Fu, H., Cheng, J., Xu, Y., Zhang, C., Wong, D. W. K., Liu, J., et al. (2018). Disc-aware ensemble network for glaucoma screening from fundus image. *IEEE Transactions on Medical Imaging*, 37(11), 2493–2501. <http://dx.doi.org/10.1109/TMI.2018.2837012>.
- Fu, H., Li, F., Sun, X., Cao, X., Liao, J., Orlando, J. I., et al. (2020). AGE challenge: Angle closure glaucoma evaluation in anterior segment optical coherence tomography. *Medical Image Analysis*, 66, Article 101798. <http://dx.doi.org/10.1016/j.media.2020.101798>.
- Fumero, F., Alay n, S., Sanchez, J. L., Sigut, J., & Gonzalez-Hernandez, M. (2011). RIM-ONE: An open retinal image database for optic nerve evaluation. In *2011 24th international symposium on computer-based medical systems* (pp. 1–6). IEEE, <http://dx.doi.org/10.1109/CBMS.2011.5999143>.
- Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, 4(2), 172–179. [http://dx.doi.org/10.1016/S0146-664X\(75\)80008-6](http://dx.doi.org/10.1016/S0146-664X(75)80008-6).
- Gao, M., Jiang, H., Zhu, L., Jiang, Z., Geng, M., Ren, Q., et al. (2023). Discriminative ensemble meta-learning with co-regularization for rare fundus diseases diagnosis. *Medical Image Analysis*, 89, Article 102884. <http://dx.doi.org/10.1016/j.media.2023.102884>.
- Garc a, 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, Article 105855. <http://dx.doi.org/10.1016/j.cmpb.2020.105855>.
- Garc a, G., Del Amor, R., Colomer, A., Verd -Monedero, R., Morales-S nchez, J., & Naranjo, V. (2021). Circumpapillary OCT-focused hybrid learning for glaucoma grading using tailored prototypical neural networks. *Artificial Intelligence in Medicine*, 118, Article 102132. <http://dx.doi.org/10.1016/j.artmed.2021.102132>.
- 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. <http://dx.doi.org/10.1109/JBHI.2020.3001019>.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, K. Weinberger (Eds.), *Vol. 27, Advances in neural information processing systems* (pp. 2672–2680). Curran Associates, Inc..
- Guo, F., Li, W., Shen, Z., & Shi, X. (2022). MTCLF: A multitask curriculum learning framework for unbiased glaucoma screenings. *Computer Methods and Programs in Biomedicine*, 221, Article 106910. <http://dx.doi.org/10.1016/j.cmpb.2022.106910>.
- Guo, F., Li, W., Tang, J., Zou, B., & Fan, Z. (2020). Automated glaucoma screening method based on image segmentation and feature extraction. *Medical & Biological Engineering & Computing*, 58, 2567–2586. <http://dx.doi.org/10.1007/s11517-020-02237-2>.
- 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. <http://dx.doi.org/10.1007/s00521-022-07078-8>.
- Gutierrez, A., & Chen, T. C. (2023). Artificial intelligence in glaucoma: posterior segment optical coherence tomography. *Current Opinion in Ophthalmology*, 34(3), 245. <http://dx.doi.org/10.1097/ICU.0000000000000934>.
- Haleem, M. S., Han, L., Hemert, J. v., Li, B., Fleming, A., Pasquale, L. R., et al. (2018). A novel adaptive deformable model for glaucoma optic disc and cup segmentation to aid glaucoma diagnosis. *Journal of Medical Systems*, 42, 1–18. <http://dx.doi.org/10.1007/s10916-017-0859-4>.
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6), 610–621. <http://dx.doi.org/10.1109/TSMC.1973.4309314>.
- He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)* (pp. 1322–1328). <http://dx.doi.org/10.1109/IJCNN.2008.4633969>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE conference on computer vision and pattern recognition* (pp. 770–778). <http://dx.doi.org/10.1109/CVPR.2016.90>.
- 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. <http://dx.doi.org/10.1038/s41598-021-99605-1>.
- Hemelings, R., Elen, B., Schuster, A. K., Blaschko, M. B., Barbosa-Breda, J., Hujanen, P., et al. (2023). A generalizable deep learning regression model for automated glaucoma screening from fundus images. *NPJ Digital Medicine*, 6(1), 112. <http://dx.doi.org/10.1038/s41746-023-00857-0>.
- Hervella,  . 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, Article 108347. <http://dx.doi.org/10.1016/j.asoc.2021.108347>.
- Hoover, A., & Goldbaum, M. (2003). Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22(8), 951–958. <http://dx.doi.org/10.1109/TMI.2003.815900>.
- Huang, X., Kong, X., Shen, Z., Ouyang, J., Li, Y., Jin, K., et al. (2023). GRAPE: A multi-modal dataset of longitudinal follow-up visual field and fundus images for glaucoma management. *Scientific Data*, 10(1), 520. <http://dx.doi.org/10.1038/s41597-023-02424-4>.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE conference on computer vision and pattern recognition* (pp. 2261–2269). <http://dx.doi.org/10.1109/CVPR.2017.243>.
- Ibrahim, M. H., Hacıbeyoglu, M., Agaoglu, A., & Ucar, F. (2022). Glaucoma disease diagnosis with an artificial algae-based deep learning algorithm. *Medical & Biological Engineering & Computing*, 60(3), 785–796. <http://dx.doi.org/10.1007/s11517-022-02510-6>.
- Jiang, Y., Duan, L., Cheng, J., Gu, Z., Xia, H., Fu, H., et al. (2020). JointRCNN: A region-based convolutional neural network for optic disc and cup segmentation. *IEEE Transactions on Biomedical Engineering*, 67(2), 335–343. <http://dx.doi.org/10.1109/TBME.2019.2913211>.
- Jiang, Y., Xia, H., Xu, Y., Cheng, J., Fu, H., Duan, L., et al. (2018). Optic disc and cup segmentation with blood vessel removal from fundus images for glaucoma detection. In *2018 40th annual international conference of the IEEE engineering in medicine and biology society* (pp. 862–865). <http://dx.doi.org/10.1109/EMBC.2018.8512400>.
- Jonas, J. B., Aung, T., Bourne, R. R., Bron, A. M., Ritch, R., & Panda-Jonas, S. (2017). Glaucoma. *The Lancet*, 390(10108), 2183–2193. [http://dx.doi.org/10.1016/S0140-6736\(17\)31469-1](http://dx.doi.org/10.1016/S0140-6736(17)31469-1).
- Jun, T. J., Eom, Y., Kim, D., Kim, C., Park, J.-H., Nguyen, H. M., et al. (2021). TRK-CNN: Transferable ranking-CNN for image classification of glaucoma, glaucoma suspect, and normal eyes. *Expert Systems with Applications*, 182, Article 115211. <http://dx.doi.org/10.1016/j.eswa.2021.115211>.
- Juneja, M., Minhas, J. S., Singla, N., Thakur, S., Thakur, N., & Jindal, P. (2022). Fused framework for glaucoma diagnosis using optical coherence tomography (OCT) images. *Expert Systems with Applications*, 201, Article 117202. <http://dx.doi.org/10.1016/j.eswa.2022.117202>.
- Juneja, M., Singh, S., Agarwal, N., Bali, S., Gupta, S., Thakur, N., et al. (2020). Automated detection of glaucoma using deep learning convolution network (G-net). *Multimedia Tools and Applications*, 79, 15531–15553. <http://dx.doi.org/10.1007/s11042-019-7460-4>.

- Kausu, T., Gopi, V. P., Wahid, K. A., Doma, W., & Niwas, S. I. (2018). Combination of clinical and multiresolution features for glaucoma detection and its classification using fundus images. *BioCybernetics and Biomedical Engineering*, 38(2), 329–341. <http://dx.doi.org/10.1016/j.bbe.2018.02.003>.
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 195. <http://dx.doi.org/10.1186/s12916-019-1426-2>.
- Kondylakis, H., Kalokyri, V., Sfakianakis, S., Marias, K., Tsiknakis, M., Jimenez-Pastor, A., et al. (2023). Data infrastructures for AI in medical imaging: a report on the experiences of five EU projects. *European Radiology Experimental*, 7(1), 20. <http://dx.doi.org/10.1186/s41747-023-00336-x>.
- 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. <http://dx.doi.org/10.1038/s41597-022-01388-1>.
- Kucur, Ş. S., Holló, G., & Sznitman, R. (2018). A deep learning approach to automatic detection of early glaucoma from visual fields. *PLoS One*, 13(11), 1–18. <http://dx.doi.org/10.1371/journal.pone.0206081>.
- Latif, J., Tu, S., Xiao, C., Ur Rehman, S., Imran, A., & Latif, Y. (2022). Odgnet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images. *SN Applied Sciences*, 4(4), 98. <http://dx.doi.org/10.1007/s42452-022-04984-3>.
- Lee, J., Kim, Y. K., Park, K. H., & Jeoung, J. W. (2020). Diagnosing glaucoma with spectral-domain optical coherence tomography using deep learning classifier. *Journal of Glaucoma*, 29(4), 287–294. <http://dx.doi.org/10.1097/JG.0000000000001458>.
- Lee, S.-D., Lee, J.-H., Choi, Y.-G., You, H.-C., Kang, J.-H., & Jun, C.-H. (2019). Machine learning models based on the dimensionality reduction of standard automated perimetry data for glaucoma diagnosis. *Artificial Intelligence in Medicine*, 94, 110–116. <http://dx.doi.org/10.1016/j.artmed.2019.02.006>.
- Lemij, H. G., Vente, C. d., 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), <http://dx.doi.org/10.1016/j.xops.2023.100300>.
- Li, Z., He, Y., Keel, S., Meng, W., Chang, R. T., & He, M. (2018). Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. *Ophthalmology*, 125(8), 1199–1206. <http://dx.doi.org/10.1016/j.ophtha.2018.01.023>.
- Li, F., Song, D., Chen, H., Xiong, J., Li, X., Zhong, H., et al. (2020). Development and clinical deployment of a smartphone-based visual field deep learning system for glaucoma detection. *NPJ Digital Medicine*, 3(1), 123. <http://dx.doi.org/10.1038/s41746-020-00329-9>.
- Li, F., Su, Y., Lin, F., Li, Z., Song, Y., Nie, S., et al. (2022). A deep-learning system predicts glaucoma incidence and progression using retinal photographs. *The Journal of Clinical Investigation*, 132(11), <http://dx.doi.org/10.1172/JCI157968>.
- Li, T., Wang, Y., Liu, L., Chen, L., & Chen, C. P. (2023). Subspace-based minority oversampling for imbalance classification. *Information Sciences*, 621, 371–388. <http://dx.doi.org/10.1016/j.ins.2022.11.108>.
- Li, L., Wang, X., Xu, M., Liu, H., & Chen, X. (2020). DeepGF: Glaucoma forecast using the sequential fundus images. In A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. acoceanu, & L. Joskowicz (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2020* (pp. 626–635). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-59722-1_60.
- Li, L., Xu, M., Liu, H., Li, Y., Wang, X., Jiang, L., et al. (2020). A large-scale database and a CNN model for attention-based glaucoma detection. *IEEE Transactions on Medical Imaging*, 39(2), 413–424. <http://dx.doi.org/10.1109/TMI.2019.2927226>.
- Li, L., Xu, M., Wang, X., Jiang, L., & Liu, H. (2019). Attention based glaucoma detection: A large-scale database and cnn model. In *2019 IEEE/CVF conference on computer vision and pattern recognition* (pp. 10563–10572). <http://dx.doi.org/10.1109/CVPR.2019.01082>.
- Liao, W., Zou, B., Zhao, R., Chen, Y., He, Z., & Zhou, M. (2020). Clinical interpretable deep learning model for glaucoma diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 24(5), 1405–1412. <http://dx.doi.org/10.1109/JBHI.2019.2949075>.
- Liu, S., Hong, J., Lu, X., Jia, X., Lin, Z., Zhou, Y., et al. (2019). Joint optic disc and cup segmentation using semi-supervised conditional GANs. *Computers in Biology and Medicine*, 115, Article 103485. <http://dx.doi.org/10.1016/j.combiomed.2019.103485>.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In *IEEE/CVF international conference on computer vision* (pp. 9992–10002). IEEE, <http://dx.doi.org/10.1109/ICCV48922.2021.00986>.
- Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., et al. (2004). Optic nerve head segmentation. *IEEE Transactions on Medical Imaging*, 23(2), 256–264. <http://dx.doi.org/10.1109/TMI.2003.823261>.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Proceedings of the 31st international conference on neural information processing systems* (pp. 4768–4777). Curran Associates Inc., <http://dx.doi.org/10.5555/3295222.3295230>.
- M., S., Issac, A., & Dutta, M. K. (2018). An automated and robust image processing algorithm for glaucoma diagnosis from fundus images using novel blood vessel tracking and bend point detection. *International Journal of Medical Informatics*, 110, 52–70. <http://dx.doi.org/10.1016/j.ijmedinf.2017.11.015>.
- Maetschke, S., Antony, B., Ishikawa, H., Wollstein, G., Schuman, J., & Garnavi, R. (2019). A feature agnostic approach for glaucoma detection in OCT volumes. *PLoS One*, 14(7), 1–12. <http://dx.doi.org/10.1371/journal.pone.0219126>.
- Maheshwari, S., Pachori, R. B., & Acharya, U. R. (2017). Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. *IEEE Journal of Biomedical and Health Informatics*, 21(3), 803–813. <http://dx.doi.org/10.1109/JBHI.2016.2544961>.
- Martins, J., Cardoso, J. S., & Soares, F. (2020). Offline computer-aided diagnosis for glaucoma detection using fundus images targeted at mobile devices. *Computer Methods and Programs in Biomedicine*, 192, Article 105341. <http://dx.doi.org/10.1016/j.cmpb.2020.105341>.
- Matheny, M. E., Whicher, D., & Israni, S. T. (2020). Artificial intelligence in health care: A report from the national academy of medicine. *JAMA*, 323(6), 509–510. <http://dx.doi.org/10.1001/jama.2019.21579>.
- Mayro, E. L., Wang, M., Elze, T., & Pasquale, L. R. (2020). The impact of artificial intelligence in the diagnosis and management of glaucoma. *Eye*, 34(1), 1–11. <http://dx.doi.org/10.1038/s41433-019-0577-x>.
- Medeiros, F. A., Jammal, A. A., & Thompson, A. C. (2019). From machine to machine: An OCT-trained deep learning algorithm for objective quantification of glaucomatous damage in fundus photographs. *Ophthalmology*, 126(4), 513–521. <http://dx.doi.org/10.1016/j.ophtha.2018.12.033>.
- Mehta, P., Petersen, C. A., Wen, J. C., Banitt, M. R., Chen, P. P., Bojkian, K. D., et al. (2021). Automated detection of glaucoma with interpretable machine learning using clinical data and multimodal retinal images. *American Journal of Ophthalmology*, 231, 154–169. <http://dx.doi.org/10.1016/j.ajo.2021.04.021>.
- Montesano, G., Chen, A., Lu, R., Lee, C. S., & Lee, A. Y. (2022). UWHVF: A real-world, open source dataset of perimetry tests from the humphrey field analyzer at the university of washington. *Translational Vision Science & Technology*, 11(1), <http://dx.doi.org/10.1167/tvst.11.1.1>, 2–2.
- Muhammad, H., Fuchs, T. J., De Cuir, N., De Moraes, C. G., Blumberg, D. M., Liebmann, J. M., et al. (2017). Hybrid deep learning on single wide-field optical coherence tomography scans accurately classifies glaucoma suspects. *Journal of Glaucoma*, 26(12), 1086. <http://dx.doi.org/10.1097/JG.0000000000000765>.
- Mvoulana, A., Kachouri, R., & Akil, M. (2019). Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images. *Computerized Medical Imaging and Graphics*, 77, Article 101643. <http://dx.doi.org/10.1016/j.compedimag.2019.101643>.
- Narasimhan, K., Vijayarekha, K., JogiNarayana, K. A., SivaPrasad, P., & SatishKumar, V. (2012). Glaucoma detection from fundus image using opencv. *Research Journal of Applied Sciences, Engineering and Technology*, 4(24), 5459–5463.
- Nayak, D. R., Das, D., Majhi, B., Bhandary, S. V., & Acharya, U. R. (2021). ECNet: An evolutionary convolutional network for automated glaucoma detection using fundus images. *Biomedical Signal Processing and Control*, 67, Article 102559. <http://dx.doi.org/10.1016/j.bspc.2021.102559>.
- Nouri-Mahdavi, K., Mohammadzadeh, V., Rabiolo, A., Edalati, K., Caprioli, J., & Yousefi, S. (2021). Prediction of visual field progression from OCT structural measures in moderate to advanced glaucoma. *American Journal of Ophthalmology*, 226, 172–181. <http://dx.doi.org/10.1016/j.ajo.2021.01.023>.
- Noury, E., Mannil, S. S., Chang, R. T., Ran, A. R., Cheung, C. Y., Thapa, S. S., et al. (2022). Deep learning for glaucoma detection and identification of novel diagnostic areas in diverse real-world datasets. *Translational Vision Science & Technology*, 11(5), <http://dx.doi.org/10.1167/tvst.11.5.11>, 11–11.
- Ocular disease intelligent recognition. (2024). <https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k>. (Accessed 9 July 2024).
- Orlando, J. I., Fu, H., Barbosa Breda, J., van Keer, K., Bathula, D. R., Diaz-Pinto, A., et al. (2020). REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. *Medical Image Analysis*, 59, Article 101570.
- Owen, C. G., Rudnicka, A. R., Nightingale, C. M., Mullen, R., Barman, S. A., Sattar, N., et al. (2011). Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; the child heart and health study in England (CHASE). *Arteriosclerosis, Thrombosis, and Vascular Biology*, 31(8), 1933–1938. <http://dx.doi.org/10.1161/ATVBAHA.111.225219>.
- P., S., J. R., & R., P. (2021). An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier. *Applied Soft Computing*, 109, Article 107512. <http://dx.doi.org/10.1016/j.asoc.2021.107512>.
- Pal, A., Moorthy, M. R., & Shahina, A. (2018). G-eyenet: A convolutional autoencoding classifier framework for the detection of glaucoma from retinal fundus images. In *2018 25th IEEE international conference on image processing* (pp. 2775–2779). <http://dx.doi.org/10.1109/ICIP.2018.8451029>.
- 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), 1–10. <http://dx.doi.org/10.1038/s41598-022-16262-8>.

- Pathan, S., Kumar, P., Pai, R. M., & Bhandary, S. V. (2021). Automated segmentation and classification of retinal features for glaucoma diagnosis. *Biomedical Signal Processing and Control*, 63, Article 102244. <http://dx.doi.org/10.1016/j.bspc.2020.102244>.
- Quigley, H. A., & Broman, A. T. (2006). The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology*, 90(3), 262–267. <http://dx.doi.org/10.1136/bjo.2005.081224>.
- 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. <http://dx.doi.org/10.1016/j.ins.2018.01.051>.
- Raghavendra, U., Gudigar, A., Bhandary, S. V., Rao, T. N., Ciaccio, E. J., & Acharya, U. R. (2019). A two layer sparse autoencoder for glaucoma identification with fundus images. *Journal of Medical Systems*, 43, 1–9. <http://dx.doi.org/10.1007/s10916-019-1427-x>.
- Rahmani, A. M., Gia, T. N., Negash, B., Anzanpour, A., Azimi, I., Jiang, M., et al. (2018). Exploiting smart e-health gateways at the edge of healthcare internet-of-things: A fog computing approach. *Future Generation Computer Systems*, 78, 641–658. <http://dx.doi.org/10.1016/j.future.2017.02.014>.
- Raja, H., Akram, M. U., Khawaja, S. G., Arslan, M., Ramzan, A., & Nazir, N. (2020). Data on OCT and fundus images for the detection of glaucoma. *Data in Brief*, 29, Article 105342. <http://dx.doi.org/10.1016/j.dib.2020.105342>.
- Raja, H., Hassan, T., Akram, M. U., & Werghi, N. (2021). Clinically verified hybrid deep learning system for retinal ganglion cells aware grading of glaucomatous progression. *IEEE Transactions on Biomedical Engineering*, 68(7), 2140–2151. <http://dx.doi.org/10.1109/TBME.2020.3030085>.
- Ran, A. R., Cheung, C. Y., Wang, X., Chen, H., yang Luo, L., Chan, P. P., et al. (2019). Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis. *The Lancet Digital Health*, 1(4), e172–e182. [http://dx.doi.org/10.1016/S2589-7500\(19\)30085-8](http://dx.doi.org/10.1016/S2589-7500(19)30085-8).
- Ren, X., Ahmad, S., Zhang, L., Xiang, L., Nie, D., Yang, F., et al. (2020). Task decomposition and synchronization for semantic biomedical image segmentation. *IEEE Transactions on Image Processing*, 29, 7497–7510. <http://dx.doi.org/10.1109/TIP.2020.3003735>.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Vol. 28, Advances in neural information processing systems* (pp. 1–9). Curran Associates, Inc..
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135–1144). Association for Computing Machinery. <http://dx.doi.org/10.1145/2939672.2939778>.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252. <http://dx.doi.org/10.1007/s11263-015-0816-y>.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. In *2018 IEEE/CVF conference on computer vision and pattern recognition* (pp. 4510–4520). <http://dx.doi.org/10.1109/CVPR.2018.00474>.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE international conference on computer vision* (pp. 618–626). <http://dx.doi.org/10.1109/ICCV.2017.74>.
- Semmlow, J. L., & Griffel, B. (2021). *Biosignal and Medical Image Processing*. CRC Press.
- Shibata, N., Tanito, M., Mitsuhashi, K., Fujino, Y., Matsuura, M., Murata, H., et al. (2018). Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. *Scientific Reports*, 8(1), 14665. <http://dx.doi.org/10.1038/s41598-018-33013-w>.
- Shorten, C., & Khoshgoftar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60: 1–48. <http://dx.doi.org/10.1186/s40537-019-0197-0>.
- Singh, L. K., Garg, H., Khanna, M., & Bhadoria, R. S. (2021). An enhanced deep image model for glaucoma diagnosis using feature-based detection in retinal fundus. *Medical & Biological Engineering & Computing*, 59, 333–353. <http://dx.doi.org/10.1007/s11517-020-02307-5>.
- Sivaswamy, J., Krishnadas, S., Chakravarty, A., Joshi, G., Tabish, A. S., et al. (2015). A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis. *JSM Biomedical Imaging Data Papers*, 2(1), 1004.
- Sivaswamy, J., Krishnadas, S. R., Datt Joshi, G., Jain, M., & Syed Tabish, A. U. (2014). Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation. In *2014 IEEE 11th international symposium on biomedical imaging* (pp. 53–56). <http://dx.doi.org/10.1109/ISBI.2014.6867807>.
- Soltanian-Zadeh, S., Kurokawa, K., Liu, Z., Zhang, F., Saeedi, O., Hammer, D. X., et al. (2021). Weakly supervised individual ganglion cell segmentation from adaptiveoptics OCT images for glaucomatous damage assessment. *Optica*, 8(5), 642–651. <http://dx.doi.org/10.1364/OPTICA.418274>.
- Song, D., Fu, B., Li, F., Xiong, J., He, J., Zhang, X., et al. (2021). Deep relation transformer for diagnosing glaucoma with optical coherence tomography and visual field function. *IEEE Transactions on Medical Imaging*, 40(9), 2392–2402. <http://dx.doi.org/10.1109/TMI.2021.3077484>.
- Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., & van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4), 501–509. <http://dx.doi.org/10.1109/TMI.2004.825627>.
- Steinmetz, J. D., et al. (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. [http://dx.doi.org/10.1016/S2214-109X\(20\)30489-7](http://dx.doi.org/10.1016/S2214-109X(20)30489-7).
- Sundararajan, M., Taly, A., & Yan, Q. (2017). Axiomatic attribution for deep networks. In *proceedings of the 34th international conference on machine learning* (pp. 3319–3328). PMLR.
- Swamidoss, I. N., arsnäs, A. K., Uhlmann, V., Ponnusamy, P., Kampf, C., Simonsson, M., et al. (2013). Automated classification of immunostaining patterns in breast tissue from the human protein atlas. *Journal of Pathology Informatics*, 4(2), 14. <http://dx.doi.org/10.4103/2153-3539.109881>.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *2016 IEEE conference on computer vision and pattern recognition* (pp. 2818–2826). <http://dx.doi.org/10.1109/CVPR.2016.308>.
- Tabassum, M., Khan, T. M., Arsalan, M., Naqvi, S. S., Ahmed, M., Madni, H. A., et al. (2020). CDED-net: Joint segmentation of optic disc and optic cup for glaucoma screening. *IEEE Access*, 8, 102733–102747. <http://dx.doi.org/10.1109/ACCESS.2020.2998635>.
- Tamura, H., Mori, S., & Yamawaki, T. (1978). Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6), 460–473. <http://dx.doi.org/10.1109/TSMC.1978.4309999>.
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In K. Chaudhuri, & R. Salakhutdinov (Eds.), *Proceedings of machine learning research: Vol. 97, Proceedings of the 36th international conference on machine learning* (pp. 6105–6114). MLResearchPress.
- Tang, L., Garvin, M. K., Lee, K., Alward, W. L., Kwon, Y. H., & Abramoff, M. D. (2011). Robust multiscale stereo matching from fundus images with radiometric differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(11), 2245–2258. <http://dx.doi.org/10.1109/TPAMI.2011.69>.
- Tékouabou, S. C. K., Alaoui, E. A. A., Chabbar, I., Toulmi, H., Cherif, W., & Silkan, H. (2022). Optimizing the early glaucoma detection from visual fields by combining preprocessing techniques and ensemble classifier with selection strategies. *Expert Systems with Applications*, 189, Article 115975. <http://dx.doi.org/10.1016/j.eswa.2021.115975>.
- Thabtah, F., Hammoud, S., Kamalov, F., & Gonsalves, A. (2020). Data imbalance in classification: Experimental evaluation. *Information Sciences*, 513, 429–441. <http://dx.doi.org/10.1016/j.ins.2019.11.004>.
- Thakoor, K. A., Koorathota, S. C., Hood, D. C., & Sajda, P. (2021). Robust and interpretable convolutional neural networks to detect glaucoma in optical coherence tomography images. *IEEE Transactions on Biomedical Engineering*, 68(8), 2456–2466. <http://dx.doi.org/10.1109/TBME.2020.3043215>.
- Thakur, A., Goldbaum, M., & Yousefi, S. (2020a). Convex representations using deep archetypal analysis for predicting glaucoma. *IEEE Journal of Translational Engineering in Health and Medicine*, 8, 1–7. <http://dx.doi.org/10.1109/JTEHM.2020.2982150>.
- Thakur, A., Goldbaum, M., & Yousefi, S. (2020b). Predicting glaucoma before onset using deep learning. *Ophthalmology Glaucoma*, 3(4), 262–268. <http://dx.doi.org/10.1016/j.jgla.2020.04.012>.
- Thakur, N., & Juneja, M. (2020). Classification of glaucoma using hybrid features with machine learning approaches. *Biomedical Signal Processing and Control*, 62, Article 102137. <http://dx.doi.org/10.1016/j.bspc.2020.102137>.
- Tham, Y.-C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., & Cheng, C.-Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040. *Ophthalmology*, 121(11), 2081–2090. <http://dx.doi.org/10.1016/j.ophtha.2014.05.013>.
- 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), <http://dx.doi.org/10.1167/tvst.12.2.23>, 23–23.
- 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. <http://dx.doi.org/10.1001/jamaophthol.2019.5983>.
- Tielsch, J. M., Sommer, A., Katz, J., Royall, R. M., Quigley, H. A., & Javitt, J. (1991). Racial variations in the prevalence of primary open-angle glaucoma: The baltimore eye survey. *Journal of American Medical Association*, 266(3), 369–374. <http://dx.doi.org/10.1001/jama.1991.03470030069026>.
- Ting, D. S., Peng, L., Varadarajan, A. V., Keane, P. A., Burlina, P. M., Chiang, M. F., et al. (2019). Deep learning in ophthalmology: The technical and clinical considerations. *Progress in Retinal and Eye Research*, 72, Article 100759. <http://dx.doi.org/10.1016/j.preteyeres.2019.04.003>.

- Touvron, H., Bojanowski, P., Caron, M., Cord, M., El-Nouby, A., Grave, E., et al. (2023). ResMLP: Feedforward networks for image classification with data-efficient training. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(4), 5314–5321. <http://dx.doi.org/10.1109/TPAMI.2022.3206148>.
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. In *Proceedings of the 38th international conference on machine learning* (p. 139). MLResearchPress, 10347–10357.
- Touvron, H., Cord, M., Sablayrolles, A., Synnaeve, G., & Jégou, H. (2021). Going deeper with image transformers. In *2021 IEEE/CVF international conference on computer vision* (pp. 32–42). <http://dx.doi.org/10.1109/ICCV48922.2021.00010>.
- Vasu, B., & Long, C. (2020). Iterative and adaptive sampling with spatial attention for black-box model explanations. In *IEEE winter conference on applications of computer vision* (pp. 2949–2958). IEEE, <http://dx.doi.org/10.1109/WACV45572.2020.9093576>.
- Vinícius dos Santos Ferreira, M., Oseas de Carvalho Filho, A., Dalíia de Sousa, A., Corrêa Silva, A., & Gattass, M. (2018). Convolutional neural network and texture descriptor-based automatic detection and diagnosis of glaucoma. *Expert Systems with Applications*, 110, 250–263. <http://dx.doi.org/10.1016/j.eswa.2018.06.010>.
- Vollmer, S., Mateen, B. A., Bohner, G., Király, F. J., Ghani, R., Jonsson, P., et al. (2020). Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ*, 368, l6927. <http://dx.doi.org/10.1136/bmj.l6927>.
- Wang, X., Chen, H., Ran, A.-R., Luo, L., Chan, P. P., Tham, C. C., et al. (2020). Towards multi-center glaucoma OCT image screening with semi-supervised joint structure and function multi-task learning. *Medical Image Analysis*, 63, Article 101695. <http://dx.doi.org/10.1016/j.media.2020.101695>.
- Wang, L., Gu, J., Chen, Y., Liang, Y., Zhang, W., Pu, J., et al. (2021). Automated segmentation of the optic disc from fundus images using an asymmetric deep learning network. *Pattern Recognition*, 112, Article 107810. <http://dx.doi.org/10.1016/j.patcog.2020.107810>.
- Wang, L., Liu, H., Lu, Y., Chen, H., Zhang, J., & Pu, J. (2019). A coarse-to-fine deep learning framework for optic disc segmentation in fundus images. *Biomedical Signal Processing and Control*, 51, 82–89. <http://dx.doi.org/10.1016/j.bspc.2019.01.022>.
- Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., & Solomon, J. M. (2019). Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 38(5), <http://dx.doi.org/10.1145/3326362>.
- Wang, Y., Yao, Q., Kwok, J. T., & Ni, L. M. (2020). Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys (CSUR)*, 53(3), 63: 1–34. <http://dx.doi.org/10.1145/3386252>.
- Wassel, M., Hamdi, A. M., Adly, N., & Torki, M. (2022). Vision transformers based classification for glaucomatous eye condition. In *2022 26th international conference on pattern recognition* (pp. 5082–5088). <http://dx.doi.org/10.1109/ICPR56361.2022.9956086>.
- Weinreb, R. N., Bowd, C., Moghimi, S., Tafreshi, A., Rausch, S., & Zangwill, L. M. (2019). Ophthalmic diagnostic imaging: glaucoma. In *High resolution imaging in microscopy and ophthalmology: new frontiers in biomedical optics* (pp. 107–134). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-16638-0_5.
- Wu, J., Yu, S., Chen, W., Ma, K., Fu, R., Liu, H., et al. (2020). Leveraging undiagnosed data for glaucoma classification with teacher-student learning. In A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, L. Joskowicz (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2020* (pp. 731–740). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-59710-8_71.
- Xu, X., Guan, Y., Li, J., Ma, Z., Zhang, L., & Li, L. (2021). Automatic glaucoma detection based on transfer induced attention network. *BioMedical Engineering Online*, 20(1), 1–19. <http://dx.doi.org/10.1186/s12938-021-00877-5>.
- Xu, Y., Hu, M., Liu, H., Yang, H., Wang, H., Lu, S., et al. (2021). A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis. *NPJ Digital Medicine*, 4(1), 48. <http://dx.doi.org/10.1038/s41746-021-00417-4>.
- Xue, Y., Zhu, J., Huang, X., Xu, X., Li, X., Zheng, Y., et al. (2022). A multi-feature deep learning system to enhance glaucoma severity diagnosis with high accuracy and fast speed. *Journal of Biomedical Informatics*, 136, Article 104233. <http://dx.doi.org/10.1016/j.jbi.2022.104233>.
- Yerushalmy, J. (1947). Statistical problems in assessing methods of medical diagnosis, with special reference to X-Ray techniques. *Public Health Reports (1896-1970)*, 62(40), 1432–1449. <http://dx.doi.org/10.2307/4586294>.
- Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends. *IEEE Access*, 12, 41180–41218. <http://dx.doi.org/10.1109/ACCESS.2024.3376441>.
- Yousefi, S., Kiwaki, T., Zheng, Y., Sugiura, H., Asaoka, R., Murata, H., et al. (2018). Detection of longitudinal visual field progression in glaucoma using machine learning. *American Journal of Ophthalmology*, 193, 71–79. <http://dx.doi.org/10.1016/j.ajo.2018.06.007>.
- Yousefi, S., Pasquale, L. R., Boland, M. V., & Johnson, C. A. (2022). Machine-identified patterns of visual field loss and an association with rapid progression in the ocular hypertension treatment study. *Ophthalmology*, 129(12), 1402–1411. <http://dx.doi.org/10.1016/j.ophtha.2022.07.001>.
- Yu, S., Zhou, H.-Y., Ma, K., Bian, C., Chu, C., Liu, H., et al. (2020). Difficulty-aware glaucoma classification with multi-rater consensus modeling. In A. L. Martel, P. Abolmaesumi, D. Stoyanov, D. Mateus, M. A. Zuluaga, S. K. Zhou, D. Racoceanu, L. Joskowicz (Eds.), *Medical image computing and computer assisted intervention – MICCAI 2020* (pp. 741–750). Springer International Publishing, http://dx.doi.org/10.1007/978-3-030-59710-8_72.
- Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Computer vision – ECCV 2014* (pp. 818–833). Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-10590-1_53.
- Zhang, M., Qu, L., Singh, P., Kalpathy-Cramer, J., & Rubin, D. L. (2022). SplitAVG: A heterogeneity-aware federated deep learning method for medical imaging. *IEEE Journal of Biomedical and Health Informatics*, 26(9), 4635–4644. <http://dx.doi.org/10.1109/JBHI.2022.3185956>.
- Zhang, Z., Yin, F. S., Liu, J., Wong, W. K., Tan, N. M., Lee, B. H., et al. (2010). ORIGA-light: An online retinal fundus image database for glaucoma analysis and research. In *Annual international conference of the IEEE engineering in medicine and biology* (pp. 3065–3068). IEEE, <http://dx.doi.org/10.1109/IEMBS.2010.5626137>.
- Zhao, R., Chen, X., Chen, Z., & Li, S. (2020). EGDCL: An adaptive curriculum learning framework for unbiased glaucoma diagnosis. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer vision–ECCV 2020* (pp. 190–205). Springer International Publishing.
- Zhao, R., Chen, X., Chen, Z., & Li, S. (2022). Diagnosing glaucoma on imbalanced data with self-ensemble dual-curriculum learning. *Medical Image Analysis*, 75, Article 102295. <http://dx.doi.org/10.1016/j.media.2021.102295>.
- Zhao, R., Chen, X., Liu, X., Chen, Z., Guo, F., & Li, S. (2020). Direct cup-to-disc ratio estimation for glaucoma screening via semi-supervised learning. *IEEE Journal of Biomedical and Health Informatics*, 24(4), 1104–1113. <http://dx.doi.org/10.1109/JBHI.2019.2934477>.
- Zhao, R., Liao, W., Zou, B., Chen, Z., & Li, S. (2019). Weakly-supervised simultaneous evidence identification and segmentation for automated glaucoma diagnosis. Vol. 33, In *Proceedings of the AAAI conference on artificial intelligence* (pp. 809–816). <http://dx.doi.org/10.1609/aaai.v33i01.3301809>.
- Zhou, W., Yi, Y., Bao, J., & Wang, W. (2019). Adaptive weighted locality-constrained sparse coding for glaucoma diagnosis. *Medical & Biological Engineering & Computing*, 57, 2055–2067. <http://dx.doi.org/10.1007/s11517-019-02011-z>.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE international conference on computer vision* (pp. 2242–2251). <http://dx.doi.org/10.1109/ICCV.2017.244>.
- Zoph, B., Vasudevan, V., Shlens, J., & Le, Q. V. (2018). Learning transferable architectures for scalable image recognition. In *2018 IEEE/CVF conference on computer vision and pattern recognition* (pp. 8697–8710). <http://dx.doi.org/10.1109/CVPR.2018.00907>.