

TOPICAL REVIEW

Retinal Health Screening Using Artificial Intelligence With Digital Fundus Images: A Review of the Last Decade (2012–2023)

SAAD ISLAM¹, RAVINESH C. DEO¹, (Senior Member, IEEE), PRABAL DATTA BARUA^{2,3}, JEFFREY SOAR², PING YU⁴, (Senior Member, IEEE), AND U. RAJENDRA ACHARYA¹

¹School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia

²School of Business, University of Southern Queensland, Springfield, QLD 4300, Australia

³Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia

⁴School of Computing and Information Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, NSW 2516, Australia

Corresponding authors: Saad Islam (saad.islam@unisq.edu.au) and Ravinesh C. Deo (ravinesh.deo@unisq.edu.au)

This work was supported in part by the University of Southern Queensland International Ph.D. Fee Scholarship, and in part by the University of Southern Queensland International Ph.D. Stipend Scholarship.

ABSTRACT Prolonged diabetic retinopathy (DR), glaucoma, and age-related macular degeneration (AMD) may lead to vision loss. Hence, early detection and treatment are crucial to prevent irreversible vision loss. Fundus retinal images have been widely used to help detect these diseases. Manual screening is susceptible to human errors, tedious, and expensive. Hence, artificial intelligence (AI) techniques have been widely employed to overcome these constraints. This paper reviewed the work published on automated retinal health detection models using various machine learning (ML) and deep learning (DL) techniques. We reviewed 142 papers and 262 studies (124 on glaucoma, 60 on AMD, and 78 on DR) from January 2012 to June 2024 using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We found that Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) models were widely used in DL and ML techniques, respectively. To the best of our knowledge, this is the first review developed for detecting AMD, DR, and glaucoma using AI techniques over the last decade. We have discussed the limitations of the present methods and also suggested future directions for accurately detecting eye diseases.

INDEX TERMS Retinal health, automated detection, deep learning, machine learning, glaucoma, fundus.

I. INTRODUCTION

2.2 billion people worldwide have near- or far-sightedness, and 1 billion cases may have been avoidable [1]. It is estimated that only 36% of people worldwide with distance vision impairment due to refractive errors and 17% of persons with cataract-related vision impairment have access to the right intervention. A large financial burden is associated with the projected \$411 billion yearly global cost of productivity loss due to eyesight impairment. The US\$25 billion cost gap to address the unmet requirement of vision impairment

is far higher than the productivity cost [1]. A decline in eyesight quality harms one's efficiency and standard of living. Millions of people globally are impacted by these illnesses, which have the potential to cause limited vision if they are not detected and treated on time. While vision loss can affect persons of any age, most of those who have limited vision or have impaired vision are over 50. The population in this age group will increase from 900 million to 2 billion between 2015 and 2050 [2].

The elderly have a higher risk of illness and experience a quicker decline in health [3]. The hardship of a visual impairment extends beyond the affected person to include family members and carers [4]. The elderly must take

The associate editor coordinating the review of this manuscript and approving it for publication was Asadullah Shaikh¹.

preventative measures, such as going to an annual eye test, to identify various eye illnesses early on. It is possible to better limit or control how eye illnesses (glaucoma, DR, cataract, and AMD) proceed. Ophthalmic diseases such as Glaucoma, DR, and AMD are the leading sources of vision loss worldwide [5]. These disorders can proceed more slowly or have a better chance of responding well to therapy if they are identified in their early stages.

Between 2004 and 2006, the Singapore Malay Eye Study looked at 3,280 randomly chosen Malay people. The results showed that 5.6% of participants had AMD, 12.9% had DR, and 4.6% had glaucoma [6]. Globally, AMD, DR, Cataracts, and Glaucoma are the leading causes of limited vision and visual impairment, contributing to one-third of all eye diseases. The WHO lists these four eye conditions as priority conditions that, if caught early enough, can be treated to avoid vision loss [7]. The elderly are more likely to get AMD, cataracts, DR, and glaucoma [8]. Artificial intelligence models can solve or slow down many health-related problems. AI algorithms can identify eye diseases from fundus images. By analysing these images, they can extract characteristics and classify patients' diseases. Using digital fundus images, image mining can be used to train algorithms to identify eye conditions like glaucoma [9]. They can also identify early signs of disorders like glaucoma by looking for abnormal enlargements in the eyes' optic cups [10].

Recently, deep learning (DL), a branch of machine learning (ML), has demonstrated potential in the identification of retinal disorders [11]. For retinal image analysis, fundus cameras and Optical Coherence Tomography (OCT) can be used to take images of the eye. The most widely used methods for capturing changes in retinal morphology, including the optic disc, blood vessels, and macula, are funduscopy and OCT imaging [12]. These images can be analysed for diseases like glaucoma, DR, and AMD. Numerous studies have been published that identify these specific eye diseases [13], [14]. Many DL algorithms have been successfully used to create AI systems for automated detection, utilising sizable databases [14], [15], [16], [17], [18].

A. RESEARCH MOTIVATION

The best-performing ML and DL algorithms for eye disorders such as glaucoma, DR, and AMD over the last decade are examined for the first time in this review. This study highlights potential research directions to show the role of ML and DL approaches for automatically detecting ophthalmic diseases using fundus images. These illnesses can be slowed down or have a better chance of recovery if identified in their early stages [19]. Many countries can benefit from utilising AI in clinical settings to diagnose eye diseases due to not having enough access to ophthalmologists [20]. To reduce the burden on the low numbers of ophthalmologists, particularly in developing nations such as

Bangladesh, where only one ophthalmologist exists for every 162,494 people, automated and accurate detection of retinal diseases may prove useful [20].

B. RESEARCH QUESTIONS

Our systematic review aims to cover the following research questions:

- 1) What are some of the changes that have occurred in the AI methods used for retinal health detection over the last decade?
- 2) How accurate are the current ML and DL systems in detecting retinal diseases for the following diseases: Glaucoma, Diabetic Retinopathy (DR), and Age-related macular degeneration (AMD)?
- 3) How can the performance of AI models in detecting Glaucoma, DR, and AMD be further improved and translated in real-world clinical settings?

C. STRUCTURE OF THE PAPER

The paper is structured as follows: Section II gives a summary of different ophthalmic diseases such as DR, AMD, and Glaucoma. Section III explores how different AI techniques detect these diseases and Section IV discusses the current publicly available datasets of fundus images. Section V addresses the progress of AI techniques in retinal diseases and their evolution over the last decade. Section VI addresses the methodology of our review including following the PRISMA guideline in the design phase. Section VII addresses the results and analysis of the studies of the last decade of DR, AMD, and Glaucoma detection using ML and DL. The different stages (Segmentation, Classification, Segmentation followed by classification) in the analysis process of automated retinal health detection are discussed in Sections VIII, IX, and X respectively. Section XII addresses the current limitations. Future directions are discussed in Section XIII and Section XIV is the conclusion.

D. NOVEL CONTRIBUTION OF THIS REVIEW

This review makes several novel contributions to the field of AI-based retinal health screening. This is the first review to comprehensively cover the use of both ML and DL techniques for the screening of glaucoma, DR, and AMD. By addressing all three major retinal diseases, this review provides a holistic view of AI applications in retinal health. Unlike previous reviews that focus exclusively on either ML or DL, this review examines both techniques. This approach highlights how combining ML and DL can enhance diagnostic accuracy and efficiency. In this review, we will address the studies done till now on DR, AMD, and Glaucoma in the last decade. We will also give clear and detailed discussions on future research directions, emphasising the potential for multi-modality approaches. We will provide insights into the challenges and successes of deploying AI systems in clinical settings, offering valuable guidance for future implementations.

II. BACKGROUND

A. RETINAL IMAGING OF THE EYE

Retinal imaging has evolved through decades of ongoing research to become the cornerstone of clinical management and treatment for patients with eye diseases [21]. Today's most common retinal imaging modalities are fundoscopy and OCT [22], each offering unique advantages. Advancements in fundus imaging have greatly increased its accessibility; one of these advancements is the switch from film-based to digital imaging. Fundus imaging has become more user-friendly with the introduction of more standardised imaging procedures [21]. Color Fundus Photography (CFP) captures images of the retina using a specialised fundus camera, resulting in two-dimensional images that detail the retinal surface, including blood vessels, optic disc, and macula [23]. CFP is non-invasive in nature and provides fast, clear retinal surface images. It offers a wide field of view, enabling the detection of peripheral retinal abnormalities, and its standardised imaging facilitates comparative studies and longitudinal monitoring [23]. However, CFP is limited in that it provides only surface views without depth information, making it challenging to assess the layers of the retina. Additionally, optimal image quality often requires pupil dilation, which can be uncomfortable for patients and time-consuming.

Optical Coherence Tomography (OCT) is the most common imaging modality in ophthalmology and is widely used for diagnosing and monitoring various eye conditions [18]. It employs low-coherence interferometry to produce high-resolution, three-dimensional cross-sectional images of the retina. Accurate segmentation of retinal layers in OCT data provides essential information for clinical diagnosis [24]. This modality excels in providing detailed visualisation of the retinal layers, offering micrometre-scale resolution that enables precise imaging of these layers [25]. OCT is particularly valuable for its depth information, which is essential for assessing retinal thickness and structural changes. It also allows for quantitative analysis, making it useful for monitoring disease progression [25]. Nevertheless, OCT typically covers a smaller retinal area compared to CFP, which can result in missing peripheral abnormalities. The high cost and limited availability of OCT machines also pose challenges, potentially restricting access in some clinical settings.

Comparing CFP with other imaging modalities, such as fluorescein angiography (FA) and scanning laser ophthalmoscopy (SLO), highlights further differences [26]. CFP is safer than FA as it does not require dye injection, which can cause allergic reactions, and is simpler and quicker, making it suitable for routine screening [27]. Compared to SLO, CFP is generally more cost-effective and widely available, providing straightforward images that are easier to interpret for general screening [28]. When evaluating OCT against other modalities like ultrasound biomicroscopy (UBM) and magnetic resonance imaging (MRI), the distinctions become clear. OCT offers much higher resolution than UBM,

providing finer details of retinal structures and being a non-contact method that reduces the risk of discomfort and infection [29]. Compared to magnetic resonance imaging (MRI), OCT is more practical for routine use due to its quick acquisition time and lower cost [30]. While MRI can provide a comprehensive view of the entire orbit and surrounding structures, OCT excels in resolution for retinal imaging.

While CFP is invaluable for wide-field imaging and detecting surface-level abnormalities, making it ideal for initial screenings and monitoring conditions like diabetic retinopathy and age-related macular degeneration, it lacks depth information and may require pupil dilation. OCT, on the other hand, provides high-resolution cross-sectional images essential for diagnosing and managing conditions that affect the integrity of retinal layers, such as glaucoma, diabetic macular edema, and age-related macular degeneration [25]. Its limitations include a limited field of view and higher costs. The choice of modality depends on clinical requirements, with CFP preferred for its simplicity and wide coverage, and OCT for its detailed structural insights.

Our research focuses on papers that only used fundus images in their studies. The primary use of fundoscopy is to detect DR, glaucoma, and AMD [22]. The photograph of a healthy eye captured by a fundus camera is shown in Figure 1a.

B. DIABETIC RETINOPATHY

Diabetic retinopathy (DR) refers to problems with the retina caused by damage to the retinal vessel walls [31]. One of the primary causes of adult vision loss is DR (Figure 1b). A patient is diagnosed with diabetic mellitus if their plasma glucose level is above 7mmol/L [32], [33]. Hyperglycemia (high blood sugar) has been linked to kidney, heart, brain, and eye damage because it can harm blood vessels and nerve cells [34]. Diabetic macular edema can result from hyperglycemia-induced harm to the retinal vessel walls [35]. New blood vessels that emerge during ischemia may later rupture due to their fragility and cause serious haemorrhages that can impair vision or even result in permanent limited vision [31]. Microaneurysms appear during the early stages of DR [31]. Neovascularisation, another name for this condition, causes proliferative diabetic retinopathy [35]. Proliferative DR, and diabetic macular edema, are two examples of severe stages of DR [36], [37]. The current treatment options for DR include surgery, intravitreal injections of steroid and anti-VEGF medications, and laser photocoagulation. DR is a significant public health issue and the leading cause of vision loss among the working-age population [38]. One-third of diabetes patients have DR [32]. Early diagnosis of DR is critical to preventing severe damage to the retina and avoiding loss of vision [39].

C. GLAUCOMA

Glaucoma arises from injury to the optic nerve, and visual field loss follows [10]. Glaucoma damages the retina's axons

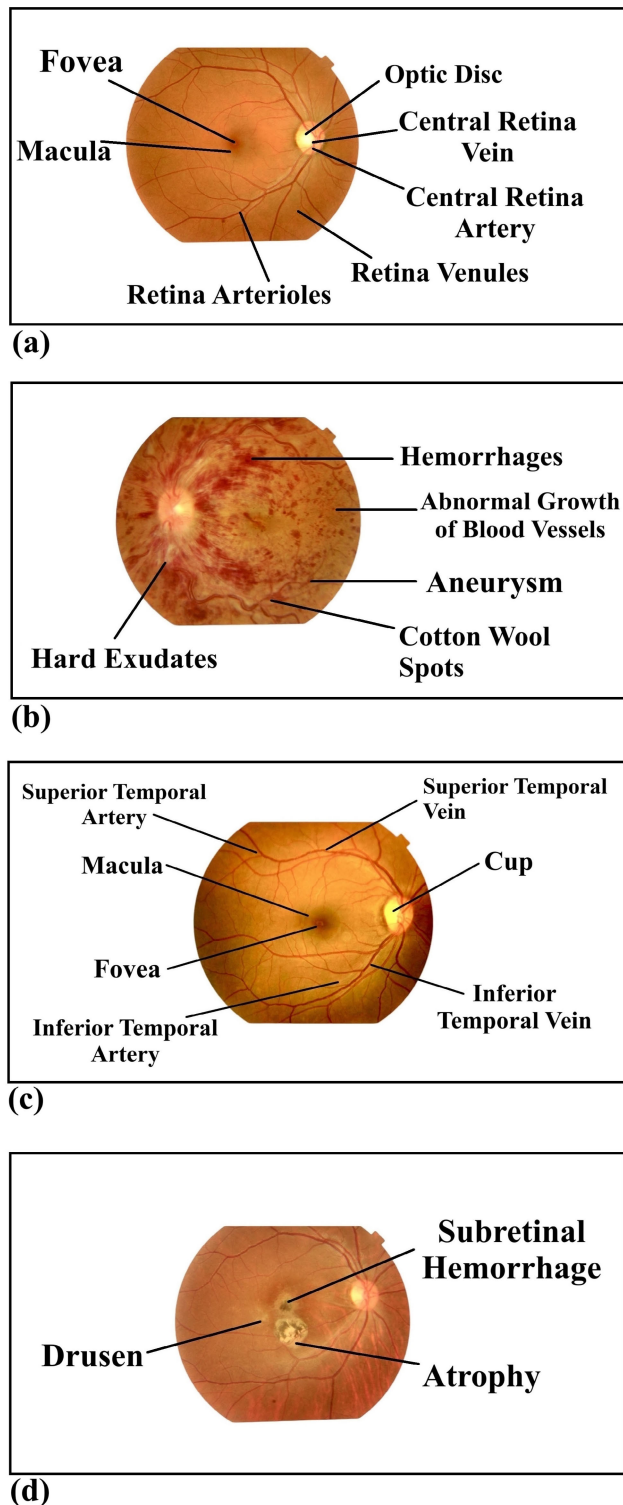


FIGURE 1. Fundus images of an eye with no diseases (a), DR (b), Glaucoma (c), AMD (d). Images provided by Bangladesh Eye Hospital and Institute Ltd.

and ganglion cells (Figure 1c). This occurs when the aqueous humour, or eye fluid, does not properly circulate in the front of the eye [40]. There are numerous glaucoma types, each

with its own set of pathogenic factors. However, they are all distinguished by virtually universal modifications to the optic nerve's structure and function [41]. The cupping of the optic disc characterizes glaucoma [42]. Tamim et al. [43] predicted that 111.8 million people globally will have glaucoma by 2040.

D. AGE-RELATED MACULAR DEGENERATION

AMD is caused by aging-related degeneration to the macula, the part of the eye that regulates accurate, straight-ahead vision [44], [45]. AMD, a prevalent condition in individuals above 50, generally manifests before Drusen, (Figure 1d) which are microscopic yellow fragments of fatty protein beneath the retina [46]. The two primary forms of AMD are wet and dry AMD. With dry AMD, vision loss or impairment typically occurs gradually [47]. Since it usually exhibits no symptoms in the intermediate stage, AMD is difficult to detect. OCT is currently the gold standard for assessing individuals for initial AMD diagnosis. Grading of AMD is crucial for detecting the early stages of the disease and preventing patients from progressing to advanced AMD [48]. Early detection allows for timely intervention, which can slow the progression and reduce the risk of severe vision loss [48]. Performing the traditional detection of this disease can take time and requires specialists with the necessary skills [49]. AMD affects 6.2 million people worldwide [50].

III. ARTIFICIAL INTELLIGENCE-BASED RETINAL SCREENING

Initially, traditional image processing techniques for analysing fundus images were used, yielding encouraging results, albeit on limited datasets and providing only partial clinical information [51]. ML methods subsequently enhanced the performance of automated analysis but still lacked robustness [52]. Recently, DL methods have demonstrated excellent performance on large datasets, showing significant potential for clinical applications [53]. In general, these models based on AI work by getting an input (e.g., fundus eye images) and if it is an ML model, it extracts features first, then performs classification and then gives an output of OD detected or not detected.

However, DL models do not need to extract features, which makes them better for automatic detection, as the user does not need to define each feature to detect the diseases. Many DL models can be developed to screen, classify and detect retinal diseases. Figure 2 shows a diagram of the different approaches of ML (Figure 2a) and DL (Figure 2b) to detect eye diseases from fundus images.

Expert ophthalmologists use fundus images from fundus cameras or OCT to identify whether an ophthalmic illness is present. Fundus eye imaging, a non-invasive technique capturing retina images, has emerged as a valuable diagnostic tool for detecting various retinal pathologies.

In recent years, the application of transformer-based methods, such as Vision Transformers (ViTs), has shown

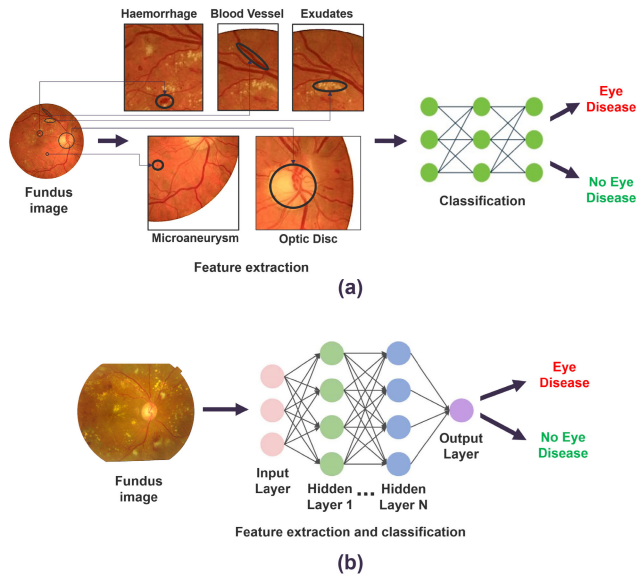


FIGURE 2. Diagram of the different approaches of (a) ML and (b) DL models to detect eye diseases from fundus images.

significant promise in the diagnosis of retinal diseases like DR, AMD, and glaucoma [54]. These advanced methodologies leverage DL techniques to enhance the accuracy and efficiency of retinal imaging analysis [55].

ViTs represent a breakthrough in image analysis, transferring the success of transformer models in natural language processing to the field of computer vision [56]. ViTs process images as sequences of patches, enabling them to capture global contextual information more effectively than traditional convolutional neural networks (CNNs) [57]. The primary advantages of ViTs include their ability to understand global context, which is particularly useful in identifying complex retinal patterns and abnormalities [58]. ViTs are highly scalable, improving performance with larger datasets and more extensive training. They also benefit from transfer learning, where pre-trained transformers can be fine-tuned on specific retinal disease datasets, enhancing their diagnostic capabilities [58].

In applications, ViTs can detect subtle changes in the retinal vasculature, such as microaneurysms and hemorrhages in diabetic retinopathy [56], identify early signs of AMD, including drusen and pigmentary changes [59], and measure retinal nerve fibre layer (RNFL) thickness and optic nerve head morphology for early glaucoma detection [54]. However, ViTs require substantial computational power and resources for training and deployment, which may be a barrier in some clinical settings. ViTs also need large annotated datasets for effective training, which can be challenging to obtain in the medical field.

According to Qummar et al., the manual method is subjective, time-consuming, and arduous, making it difficult for such diagnoses to be repeated [13]. The growing interest in leveraging ML and DL to analyse fundus images has

the potential to revolutionize retinal health diagnostics. This automatic retinal health screening system may help clinicians detect these diseases in their early stages and have a higher chance of saving patients' vision. There would also be no subjective bias on the part of the clinicians. If applied properly, these systems would produce results faster and more consistently than manual or human processes.

IV. PUBLIC DATASETS OF FUNDUS EYE IMAGES

The main retinal image databases that are publicly available and have recently been used to gauge algorithm performance in literature are listed in this section. These databases are appropriate for assessing algorithm performance because they have a clearly defined standard. The demand for validating or training models has increased, thus research teams have created and made their own datasets public [72]. The databases contain retinal images that show, among other things, DR, AMD, and glaucoma.

The accessibility of these datasets is critical for the development and evaluation of ML and DL models. A fully annotated database, MESSIDOR, displays the DR grade for all its 1200 fundus photos [60]. MESSIDOR-2 has 1748 photos, one for each eye and two for each subject. There are 40 photos in DRIVE, 33 of which are DR-free, and 7 of which show only minor DR symptoms [61].

RIM-ONE contains 159 images, each with an optic cup and disc label. Of the images, 74 exhibit glaucoma symptoms, and 85 are normal. STARE has 400 images, with 40 that are manually segmented and annotated. Specialists label all images [62]. DIARET DB1 contains 89 fundus images, 84 with at least mild DR. The images were acquired from Kuopio University Hospital in Finland [36]. KAGGLE by EyePACS has 88,702 images, 35,126 from the training set and 53,576 for the test set. LAG-DB has 11,760 images, 4878 with glaucoma and the rest as normal [68]. IDRiD has 597 images showing DR and its severity and images with normal retinal structures [71].

CHASEDB1 (Child Heart and Health Study in England Database 1) is a public database created as part of the Child Heart and Health Study in England [72]. It contains retinal images used for research into the correlation between retinal vessel characteristics and cardiovascular disease risk factors in children. The CHASEDB1 database comprises 28 manually segmented monochrome ground-truth images with a resolution of 1280×960 pixels. Retinal imaging was conducted on over 1000 children. Expert ophthalmologists performed the image segmentation.

The ACRIMA dataset emerged from a project funded by Spain's Ministerio de Economía y Competitividad, which focused on the development of algorithms for detecting ocular diseases [73]. This database comprises 705 images, including 396 glaucomatous and 309 normal ones. Images were obtained using the Topcon TRC retinal camera from previously dilated left and right eyes. Two glaucoma experts performed the image annotation.

TABLE 1. List of public databases with fundus images. Note: MESSIDOR = methods to evaluate segmentation and indexing techniques in the field of Retinal Ophthalmology RIM-ONE = retinal image database for optic nerve evaluation; DRIVE = digital retinal image for vessel extraction; STARE = structured analysis of the retina; DIARET DB1 = standard diabetic retinopathy database calibration level 1; LAG DB = large-scale attention Glaucoma Database; IDRiD = indian diabetic retinopathy image dataset.

Database Name	Images	Specific Eye Disease.	Ref	Web Link
MESSIDOR	1200	DR	[60]	www.adcis.net/en/third-party/messidor
MESSIDOR-2	1748	DR	[61]	www.adcis.net/en/third-party/messidor2
RIM-ONE	159	Glaucoma	[62]	https://github.com/miag-ull/rim-one-dl
HRF	45	AMD, Glaucoma, DR	[63]	https://www5.cs.fau.de/research/data/fundus-images
DRIVE	40	DR	[64]	www.kaggle.com/datasets/andrewmvd/drive-digital-retinal-images-for-vessel-extraction
STARE	400	AMD, DR	[65]	https://cecas.clemson.edu/\$sim\$ahoover/stare
APTOS	5590	DR	[66]	https://www.kaggle.com/c/aptos2019-blindness-detection/data
E-optha	463	Glaucoma, AMD, DR	[67]	http://www.adcis.net/en/third-party/e-optha
DIARETDB1	89	DR	[36]	www.it.lut.fi/project/imageret/diaretdb1_v2_1
KAGGLE by EyePACS	88702	DR	[68]	www.kaggle.com/c/diabetic-retinopathy-detection/data
LAG DB	11760	Glaucoma	[69]	https://github.com/smilell/AG-CNN
HEIMED	169	DR	[70]	https://github.com/lgiancaUTH/HEI-MED
IDRiD	597	DR	[71]	https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid
CHASEDB1	28	DR	[72]	https://blogs.kingston.ac.uk/retinal/chasedb1
ACRIMA	705	Glaucoma	[73]	https://figshare.com/s/c2d31f850af14c5b5232
ORIGA	650	Glaucoma	[74]	https://www.kaggle.com/datasets/arnavjain1/glaucoma-datasets?select=ORIGA
DIARETDB0	130	DR	[75]	www.it.lut.fi/project/imageret/
ODIR	6426	AMD, DR, Glaucoma	[76]	https://odir2019.grand-challenge.org/dataset/
ARIA	450	AMD, DR	[77]	https://paperswithcode.com/dataset/aria/
AREDS	120656	AMD	[78]	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000001.v3.p1
IOSTAR	30	DR	[79]	https://www.retinacheck.org/download-iostar-retinal-vessel-segmentation-dataset

The Online Retinal Fundus Image Dataset for Glaucoma Analysis and Research (ORIGA) was developed by the Singapore Malay Eye Research Institute (SERI) for segmenting the optic cup and optic disc [74]. This publicly accessible database contains 650 retinal images intended for benchmarking segmentation and classification algorithms. It includes 168 glaucomatous images and 482 healthy images, each with a resolution of 3072×2048 pixels. The images were collected between 2004 and 2007 and were annotated by highly trained professionals. The study's subjects ranged in age from 40 to 80 years.

The Ocular Disease Intelligence Recognition (ODIR) dataset is a structured collection of data from 5,000 patients, curated by the Peking University National Institute of Health Sciences [76]. It includes multiple label annotations for retinal diseases including DR, AMD, Glaucoma, and others. The images are stored in various sizes in JPEG format. The distribution of image class labels is as follows: Normal: 3098, DR: 1406, Glaucoma: 224, AMD: 293, and others. Expert ophthalmologists participated in the annotation process.

The ARIA database consists of 450 images in JPEG format [77]. These images are divided into three categories: a healthy control group, a group with AMD, and a group with DR. Two expert ophthalmologists were responsible for annotating the images.

The Age-Related Eye Disease Study (AREDS) was a longitudinal study spanning up to 12 years, during which the AMD conditions of numerous patients were monitored [78]. The study included cases of geographic atrophy, neovascular AMD, and control patients. Retinal images of both the left and right eyes of each patient were taken throughout the study. These images were graded for AMD severity by various eye specialists. Over the course of the study, some patients who initially showed mild AMD symptoms progressed to

more severe stages. The database is divided into training, validation, and test sets, consisting of 86,770, 21,867, and 12,019 images, respectively.

The IOSTAR database contains 30 retinal photos captured using a laser fundus camera. These images were edited and annotated by two specialists [79]. The IOSTAR database includes annotations for the optic disc and the images have a resolution of 1024×1024 pixels. DIARETDB0 consists of 130 images, with 20 normal images and 110 images showing signs of diabetic retinopathy (DR), saved in PNG format.

Among other things, the quantity of photos, the pre-processing tasks, and the quality of the images affect ML and DL performance. Publicly available datasets such as these have had an important role in advancing retinal disease detection systems. Links to these databases are given in Table 1, which shows the details of the datasets and relevant links to these public data.

V. RELATED RESEARCH

Despite the considerable progress in AI-based retinal health screening, many existing reviews in this field have several limitations. Many reviews focus only on one specific disease such as DR or AMD, or a particular AI model between ML or DL, rather than providing a holistic view of the entire field. This limits the understanding of the broader impact and potential of AI in retinal health screening.

Few reviews offer a comprehensive longitudinal analysis that tracks the evolution of AI technologies over an extended period. This makes it challenging to appreciate the incremental improvements and significant breakthroughs achieved over the years. The existing reviews often emphasize theoretical developments and laboratory results, neglecting the practical challenges and successes of implementing AI systems in real-world clinical settings. Rapid advancements

TABLE 2. A table that shows a comparison with other published systematic review articles on detecting eye diseases such as DR, AMD, and Glaucoma from fundus images.

Year	Author(s)	Eye disease(s)	Technique(s)	Features focused on in the review
2024	This review	DR, AMD, Glaucoma	ML, DL	Trend of different AI techniques used over the last decade for detecting DR, AMD, and Glaucoma. Comparison of different AI methods. ML, DL techniques used for segmentation of optic disc, optic cup, retinal blood vessels, microaneurysms, etc. Classification, Segmentation followed by classification methods. Comparison of performance of different ML and DL models on fundus images, over the last decade.
2024	[80]	DR	ML, DL	Grading of DR, Analyse blood vessels, microaneurysms, lesion appearance variability
2023	[81]	Glaucoma	DL	Optic disc and cup segmentation, classification of glaucoma stages, evaluating performance
2023	[82]	AMD	DL	Diagnostic accuracy of different DL algorithms to identify AMD
2023	[83]	Glaucoma	DL	Fundus image analysis techniques, dataset overview, and image processing methodologies
2020	[84]	Glaucoma	ML	Image segmentation, optic disc, and optic cup segmentation, feature reduction
2020	[85]	DR	DL	Degree of DR severity, grading of fundus images of DR
2020	[86]	AMD, DR, Glaucoma	DL	Retinal blood vessel segmentations, Optic disc and optic nerve segmentations
2020	[68]	DR	DL	Lesion detection methods, hierarchical layer processing, and unsupervised feature learning
2019	[87]	AMD	ML, DL	Severity of AMD, Detection of AMD using ML and DL classifiers
2018	[88]	Glaucoma	ML	Progression and detection measures of glaucoma, pattern classification, ML methods
2017	[89]	DR	ML	Analysing retinal blood vessels to assess DR, image processing, pattern recognition
2016	[90]	Glaucoma	ML	Localisation and segmentation of the optic nerve head, pixel level glaucomatous changes
2014	[91]	AMD	ML	Automated AMD grading techniques, drusen detection, Texture based segmentation
2013	[92]	DR	ML	Detection of microaneurysms, exudates, optic disk and cup, retinal blood vessels

in AI mean that emerging trends and technologies may be underrepresented in reviews. This includes the latest innovations in deep learning architectures, transfer learning, and federated learning, which are crucial for the future of retinal health screening.

In Table 2, we compared several reviews related to our research work to the automated detection of ophthalmic diseases such as DR, AMD, and Glaucoma. Initially, many earlier studies used computer-aided diagnosis (CAD) systems [90], [92]. These systems used image processing techniques and handcrafted features to assist ophthalmologists in diagnosing eye diseases from fundus images. CAD systems primarily served as tools to aid ophthalmologists by highlighting areas of concern in fundus images, rather than providing definitive diagnoses [93]. Early ML systems often used rule-based algorithms to classify images [93]. These rules were derived from clinical expertise and predefined criteria, which limited their adaptability and accuracy.

In 2013, a review by Mookiah et al. showed how ML was used to detect DR by extracting different features such as microaneurysms, exudates and blood vessels [92]. ML systems relied heavily on the manual extraction of features such as blood vessel patterns, microaneurysms, exudates, and other retinal abnormalities. These features were then used to identify potential signs of diseases like DR and AMD [92].

Bhuiyan et al. [91] reviewed AMD detection techniques using ML in 2014. Different techniques, such as drusen detection techniques, and texture-based segmentation, were discussed [91]. As the field progressed, ML techniques began to be employed more frequently for retinal health screening [94]. ML algorithms improved upon CAD systems by learning from data, reducing the need for handcrafted features [95]. ML models, particularly supervised learning algorithms like support vector machines (SVMs) and random forests, were trained on labelled datasets of fundus images. These

models learned to distinguish between healthy and diseased retinas based on patterns in the data [96].

In 2016, Mary et al. [90] reviewed detecting glaucoma using ML models such as SVM. While ML reduced reliance on manual feature extraction, feature selection remained an important step. Techniques like principal component analysis (PCA) were used to identify the most relevant features for classification [97]. ML algorithms achieved higher accuracy compared to traditional CAD systems by leveraging larger datasets and more sophisticated learning techniques [98]. However, they still required significant human intervention for feature engineering and data pre-processing.

In recent years, deep learning (DL) has emerged as the dominant approach for retinal health screening [99]. DL models, particularly convolutional neural networks (CNNs), have transformed the field by automating feature extraction and improving diagnostic performance [100]. Unlike ML, DL models automatically learn relevant features from raw image data [101]. CNNs, with their multiple layers of convolutional and pooling operations, can identify complex patterns in fundus images without manual intervention [102].

In 2019, Pead et al. [87] reviewed DL methods, particularly CNNs, in detecting AMD. In the review, CNNs are highlighted for their high performance in detection from fundus images. Transfer learning from pre-trained networks and ensemble learning are also discussed [87]. Examples include a 14-layer CNN achieving a high accuracy of 95.45%, sensitivity of 96.43%, and specificity of 93.75%.

In 2020, Islam et al. [85] reviewed DL methods for detecting DR. Their findings demonstrated that DL algorithms exhibited high sensitivity and specificity in detecting DR from fundus images. They concluded that implementing a DL-based automated tool to assess DR from colour fundus images can offer an alternative solution to reduce misdiagnosis and enhance workflow. It can also have significant advantages, including lowering screening costs,

increasing healthcare accessibility, and facilitating earlier treatments [85].

In 2023, Soofi [81] reviewed DL methods for detecting glaucoma. Using CNNs, Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks for glaucoma detection are highlighted in their review. Their review also focused on the different techniques including U-Net for segmentation of the optic disc, MobileNet V2 for classification, and attention-based mechanisms to improve focus on relevant image regions [81].

Bhulakshmi and Rajput [80] reviewed ML and DL methods to detect DR in 2024. They reviewed different methods such as CNNs, RNNs, GANs. The review concluded that CNNs were effective in classifying DR severity, RNNs were useful for sequential data, tracking disease progression over time, and GANs were used in many studies to generate synthetic retinal images for training DL models [80].

DL enables end-to-end learning, where the entire process from image input to disease classification is handled by the model [103]. This reduces the need for intermediate steps like feature selection and allows for more streamlined workflows. DL models have achieved state-of-the-art performance in detecting retinal diseases. Studies have shown that CNNs can match the accuracy of expert ophthalmologists in detecting conditions like DR, AMD, and glaucoma [69].

DL models are highly scalable and can be trained on large datasets, enabling them to generalise well to diverse populations and varying image quality [69]. This makes them suitable for widespread clinical use. In general, earlier studies focused on documenting the performance of different ML techniques for automatic ophthalmic disease detection. In contrast, more recent studies have reviewed DL techniques for automatic ophthalmic disease detection.

VI. MATERIALS AND METHOD

During the design phase of this systematic review, we applied the PRISMA guideline to evaluate relevant research on AI in retinal screening using fundus images. We focused on the following conditions: Glaucoma, AMD, and DR. We targeted articles that incorporated ML or DL techniques.

We followed the PRISMA guidelines to perform a systemic search of studies related to AI approaches in retinal health screening published from January 2012 to June 2024. We identified, screened, and selected 142 papers that satisfied the criteria of this review. These particular databases were picked because they had many excellent research papers. The database queries were created using the topic-related specific Boolean strings, as shown in Table 3.

Using the Boolean search queries outlined in Table 3, a systematic search was performed across four databases: Institute of Electrical and Electronics Engineers (IEEE) Xplore Digital Library, PubMed, Science Direct, and Google Scholar, covering all publications. The initial identification phase of the PRISMA method, shown in Figure 3,

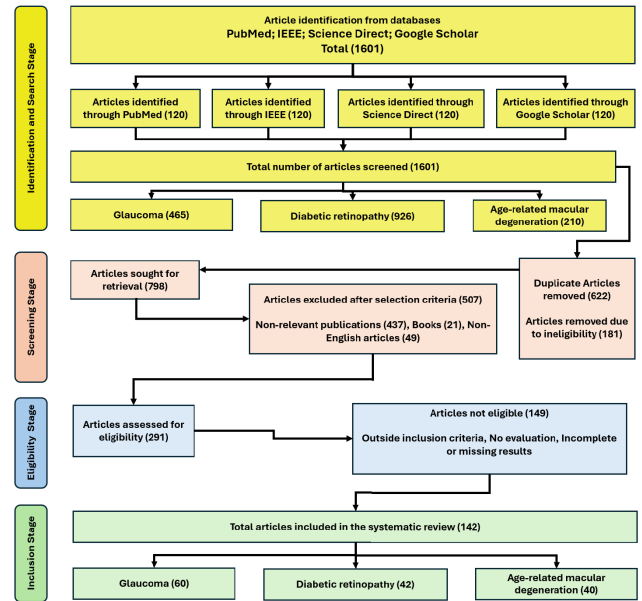


FIGURE 3. Flow diagram of PRISMA approach used for this systematic review.

spanned from January 2012 to June 2024 and resulted in 1601 publications.

The distribution of studies across databases was as follows: 531 studies from Google Scholar, 176 studies from Science Direct, 429 studies from PubMed, and 465 studies from IEEE Xplore Digital Library.

In the screening phase, we removed 622 duplicates and 181 ineligible studies from the initial search results. After removing the 803 articles, 798 publications remained. We further examined the titles and abstracts of the remaining publications. This process led to the systematic elimination of 507 articles, including 437 non-relevant publications, 21 books, and 49 non-English language publications, which resulted in 291 articles remaining.

In the inclusion stage, we thoroughly read and assessed the eligibility of the remaining 291 articles. We checked for completeness of information and whether they contained the necessary details pertinent to AI approaches in retinal health screening. After meticulous review, an additional 149 articles were excluded for not being in our inclusion criteria, missing or irrelevant information, or no evaluation. Finally, 142 journal articles met the eligibility criteria for inclusion in this review. Our detailed inclusion and exclusion criteria are presented in Table 4. In Figure 3, the PRISMA workflow diagram of this process is shown.

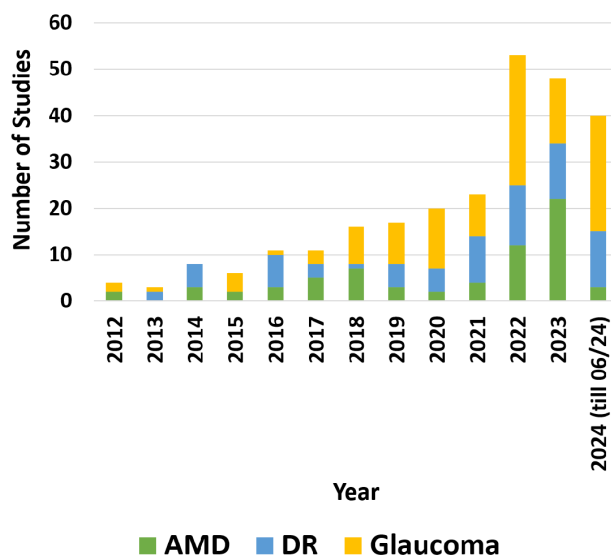
Figure 4 illustrates the number of ML and DL studies per year from January 2012 to June 2024 that focused on detecting AMD, DR, and glaucoma. The graph shows an exponential increase in research output over this period. Early on from 2012-2017 there were just a handful of studies but from 2018-2021 the number of studies rose steadily for each disease area. There was a sharp increase in the year

TABLE 3. Boolean search terms and the quantity of chosen papers from the corresponding databases.

Boolean	AND Disease Title	Databases				Total No. of studies
		PubMed	IEEE	Science Direct	Google Scholar	
fundus image + deep learning OR fundus image + automated detection	Glaucoma	26	93	41	89	249
	Diabetic Retinopathy	73	201	39	227	540
	Age-related macular degeneration	18	26	23	40	107
OR fundus image + machine learning	Glaucoma	81	38	36	61	216
	Diabetic Retinopathy	164	95	24	103	386
	Age-related macular degeneration	67	12	13	11	103

TABLE 4. Inclusion and exclusion criteria of our search.

Inclusion Criteria	Exclusion Criteria
<ul style="list-style-type: none"> - Studies that used only digital fundus eye images. - Detection of the following eye diseases: DR, AMD, and Glaucoma. - Studies that used Accuracy, Sensitivity, Specificity, Recall, and AUC as metrics. - Studies that used ML, DL, or hybrid techniques in classifying the diseases. - Studies that used a specific number of fundus eye images. - Papers taken only from January 2012 till June 2024 - Only papers published in English - Only from IEEE, Google Scholar, Science Direct, PubMed 	<ul style="list-style-type: none"> - Studies conducted with eye images other than colour fundus photography. - Studies that did not use Accuracy, Sensitivity, Specificity, Recall, and AUC as performance metrics. - Studies focused on using ML or DL to classify other eye diseases than the ones we are focusing on (DR, AMD, and Glaucoma). - Studies where the dataset is unknown, or the number of images are not provided. - Articles that do not employ ML or DL techniques in the classification or prediction phase.

**FIGURE 4.** Bar graph illustrating the number of annual studies that used ML and DL methods to identify AMD, DR, and glaucoma from 2012 to 2024 (till June 2024).

2022, which was just after the COVID-19 pandemic period. Numerous studies followed it in the year 2023 and up until June 2024. This reflects the steady increase of interest in applying ML and DL for the detection of AMD, DR, and glaucoma.

VII. RESULTS, ANALYSIS, SYNTHESIS AND INTERPRETATION

In this paper, we reviewed 142 papers, which had 262 studies using different ML and DL models to identify ophthalmic diseases. Out of 262 studies, 47.33% (124) were on

Glaucoma, 22.90% (60) were on AMD and 29.77% (78) were on DR. Different studies used different performance metrics such as F1-Score and precision, but our review focused only on studies that used accuracy, specificity, and sensitivity/recall or AUC as metrics. As demonstrated by the bar charts in Figures 5 to 7, and 11 to 14, the models utilised in all of these studies had a generally strong performance in terms of accuracy, specificity, and sensitivity.

A. MACHINE LEARNING

If we consider the 104 ML-based studies that attempted to identify ODs, we note that around 36.5% (38) utilised a support vector machine (SVM) as the classifier. The discussed 3 ODs (Table 14-19) achieved greater performance with SVM despite the results varying from 73.3% to 100% accuracy, 53.16% to 100% specificity, and 82.6% to 100% sensitivity.

With the SVM classifier, the highest accuracy, specificity, and sensitivity of 100% were achieved in Glaucoma [105]. For AMD, the highest accuracy of 93.7%, specificity of 96.3%, and sensitivity of 91.11% were achieved using SVM [119]. For DR, the highest accuracy of 100% [120], specificity of 96.88% [121], and sensitivity of 100% [98] were achieved using ML classifiers.

SVM was still the most commonly used ML classifier. For Glaucoma, 37.8% (17 of the 45 classifiers) were SVM (Figure 10). Two studies used the least-square support vector machine (LS-SVM), one of the SVM variants [122], [123]. Large data processing and computational time reduction are common uses for the LS-SVM model [124].

Many models for OD detection have recently been developed using ML methods. To find patterns and forecast the presence of ODs, ML algorithms can analyse large

TABLE 5. Summary of some high-performing ML methods to detect Glaucoma, DR, and AMD on fundus images over the last decade.

Authors (Year)	Disease detected	No. of images	Features Extracted	Performance
[104] (2022)	Glaucoma	DRISHTI-GS (101), ACRIMA (705)	Homogeneity, Contrast, Entropy Correlation and Energy	NB: Acc = 88.86% Sens = 90.25%
[105] (2021)	Glaucoma	DRISHTI-GS (101), RIM-One (169), NIO (118) DRIONS-DB (110), JSIEC (124), DRIVE (44). Total: 666	CDR, NRR ISNT ratio, Blood Vessels ISNT ratio	Adaboost: Acc = 100% Spec = 90% Spec = 100%
[106] (2020)	Glaucoma	DRISHTI-GS (101), RIM-One (159)	DDLS, CDR, GLRM, GLCM, FoS, HOS, HOC and Wavelets	KNN: Acc = 91.6% Spec = 90% Sens = 91%
[107] (2019)	DR, Glaucoma	ONHSD (99), DRIONS-DB (110), MESSIDOR (1200)	Intensity-based statistical features and texton map histogram	SVM: Acc = 99.3% Spec = 99.4% Sens = 96.9%
[108] (2018)	DR, AMD, Glaucoma	KMC: Normal: 404 Abnormal: 1082 (381 AMD, 170 DR, 115 Glaucoma)	Scalogram of 2-D CWT, PSO, entropy, energy, ADASYN	RF: Acc = 92.48% Spec = 95.58% Sens = 89.37%
[109] (2018)	DR, AMD, Glaucoma	KMC: Normal: 790, Abnormal: 1430 (531 AMD, 346 DR, 533 Glaucoma)	Pyramid histogram of visual words (PHOW), Fisher vectors	RF: Acc = 96.79%
[110] (2017)	Glaucoma	Private (702) (143 Normal, 559 Glaucoma)	Local configuration pattern	KNN : Acc = 95.7% Spec = 93.7% Sens = 96.2%
[111] (2016)	DR, AMD, Glaucoma	KMC: Normal: 400, Abnormal: 400 (115 AMD, 170 DR, 115 Glaucoma)	Bi-dimensional empirical mode decomposition	SVM: Acc = 88.63% Spec = 91% Sen = 86.25%
[112] (2016)	Glaucoma	Venu Eye Research Centre, India (63 images)	Wavelet feature extraction	KNN : Acc = 94.8% Spec = 100% Sens = 90.9%
[98] (2014)	DR	MESSIDOR (1200)	Various	GMM: Spec = 53.16% Sens = 100% AUC = 0.904
[9] (2012)	AMD	ARIA + STARE: 258 (160 AMD and 98 normal)	Weighted Frequent Subgraph (WFSG) Mining: This technique extracts frequent subgraphs (image features) from the graph representation of the images.	SVM : Acc = 99.6% Spec = 100% Sens = 99.4%

TABLE 6. Summary of some high-performing DL methods to detect Glaucoma, DR, and AMD on fundus images over the last decade.

Authors (Year)	Disease detected	No. of images	Features Extracted	Performance
[113] (2024)	Glaucoma	Private (1426) (837 glaucoma and 589 normal)	Class-specific features	CDAM-Net: Acc = 99.30% Spec = 100% Sens = 98.80% AUC = 0.99
[114] (2023)	AMD	Private (7049)	Drusen: Small yellow or white deposits under the retina, Retinal blood vessels.	GAN : Acc = 88.5% Spec = 94% Sens = 88.5% AUC = 0.967
[115] (2022)	DR	STARE (60)	Various deep features	VSUL-Net : Acc = 97.27%, Spec = 98.13 Sens = 86.64%, AUC = 0.99
[116] (2021)	Glaucoma	UK Biobank Eye and Vision Consortium, UK (1193+1283)	IOP, corneal hysteresis, corneal resistance factor	Densenet: Sens = 81%
[117] (2018)	Glaucoma	Private (304)	RNFL thickness and deviation map, GCC thickness and deviation map	CNN: Sens = 96%
[118] (2015)	Glaucoma	ORIGA (168 glaucoma, 482 normal), SCES (46 glaucoma, 1630 normal)	Various deep features	CNN : AUC = 0.831

amounts of data. From the literature, we can see that several studies have employed traditional ML techniques, such as

SVM, decision trees, and random forests, to detect retinal diseases from fundus images. Reference [21] proposed an

automated system based on SVM for detecting DR. The system showed high sensitivity and specificity in identifying DR-related lesions.

Similar studies have used other ML techniques, such as KNN and decision trees, to detect and classify ODs such as glaucoma [110]. Retinal nerve fibre layer thickness was the basis for Bock et al. [125] application of SVM to distinguish between retinas in good condition and those with glaucoma. Their approach demonstrated promising results in detecting glaucomatous damage.

García-Florian et al. [126] developed an automated system using ML techniques, including SVM, to detect AMD. Their approach achieved high classification accuracy in distinguishing between different AMD stages. These methods for medical diagnosis, such as decision trees [127] and the Gaussian mixture model [128] were able to match the accuracy levels of human experts, as noted by Jain et al. [129]. Still, their disadvantage was that they heavily relied on knowledge of the disease-specific features and required a lot of work to be able to extract and analyse the features.

Traditional ML algorithms like SVMs do not have the same feature learning capabilities as deep nets. In traditional ML, having more fundus images does not meaningfully improve the model as there are diminishing returns - at some point, the performance may plateau or even degrade with too much data. They cannot take full advantage of very large image datasets in the same way as DL. In DL, the performance typically continues to improve as more images are added to the dataset without hitting the same diminishing returns as ML. This is because DL models have a higher capacity to take advantage of large datasets.

ML models often need help in capturing complex patterns in large datasets. If a dataset is too large and diverse, it can lead to overfitting, where the model memorizes the data rather than generalising from it, potentially resulting in lower performance.

B. DEEP LEARNING

Most of the models in our review that used DL have used CNN. CNN is a subclass of multilayer neural networks. It is a model based on the biological neural networks found in the human brain [101]. There have already been some publications of CNN applications for analysing retinal images. As an illustration, Van Grinsven et al. [130] used CNN to find haemorrhages in fundus photographs.

In their work, they used a CNN model with nine layers. A deep CNN was also used to simultaneously locate and segment the vasculature, optic disc, and fovea [131]. The proposed CNN model's high accuracy performance demonstrates its potential for use in CAD systems. In visual recognition tasks, CNN models have demonstrated exceptional recognition capacity [132], [133].

A Convolutional Neural Network (CNN) or one of its variants was used as the classifier in about 58.9%

(93 out of 158 DL studies) that attempted to identify ODs (Appendix A).

The discussed 3 ODs (see Table 14 to Table 19) achieved superior performance with CNN despite the results varying from 63.3% to 100% accuracy, 66.6% to 100% specificity, and 51.5% to 100% sensitivity.

AMD achieved 100% accuracy, specificity, and sensitivity of 100% with the CNN classifier or one of its variants [134].

CNN achieved the highest accuracy of 99% [135], specificity of 96.7% [136], and sensitivity of 95.6% [22] for glaucoma.

For DR, the highest accuracy of 99.62% [137], specificity of 96.37% [138], and sensitivity of 96.87% [139] were achieved using CNN or one of its variants.

DL models like CNN are specifically designed to extract features from image data. The more images they are trained on, the better they recognise patterns and features, leading to continued gains in performance. DL models with many layers can be overfitted on small datasets. More training images help prevent overfitting and improve generalisation, so, performance improves with more data.

DL can leverage large image datasets more effectively to steadily improve performance, while traditional ML sees diminishing returns after a certain threshold of data size. Due to the advantages of DL over ML, future work should focus on DL for image classification tasks. Using deep learning (DL) over traditional machine learning (ML) for detecting eye diseases offers several advantages:

- **Automatic Feature Extraction:** DL models, particularly Convolutional Neural Networks (CNNs), are adept at automatically identifying and learning relevant features from raw data, such as images [140]. This contrasts with ML, where feature extraction requires manual intervention and expert knowledge, making DL more efficient and scalable for complex tasks like detecting a wide range of retinal diseases [84].
- **Handling Complex Patterns:** DL models can capture and model complex patterns in data that are often missed by traditional ML algorithms [141]. This capability is crucial for detecting eye diseases, where subtle variations in retinal images can indicate different conditions [9]. DL can more accurately identify these nuances, leading to better diagnosis and treatment strategies.
- **Versatility and Adaptability:** DL models can be designed to handle multiple tasks simultaneously, such as screening, categorisation, and detection of various eye conditions, including AMD, DR, and glaucoma [142]. This multipurpose nature makes DL more versatile than ML, which might require different models or feature sets for each task.
- **Improved Accuracy and Efficiency:** Due to their ability to learn from large datasets and improve over time, DL models can achieve higher accuracy in disease detection than ML models [101]. This is particularly beneficial in medical imaging, where precision is

TABLE 7. A comparison of ML methods for optic disc and optic segmentation in fundus eye images.

Year	Author	ML Method for optic disc and optic cup segmentation	Database
2021	[143]	Clustering	Drishti-GS
2019	[144]	Clustering and level set combination	RIM-ONE, Drishti-GS
2018	[145]	Active contour model	RIM-ONE, Drishti-GS
2017	[146]	Conditional random field formulation	Drishti-GS
2015	[147]	Adaptive histogram equalisation and Gabor filter	Private
2013	[148]	K-means clustering and Gabor wavelet transform	Private
2013	[149]	Superpixel classification	ORIGA, SCES
2013	[150]	Gradient method	Private
2012	[151]	K-means and openCV code	AEH

TABLE 8. A comparison of DL methods for optic disc and optic segmentation in fundus eye images.

Year	Author	DL Method for optic disc and optic cup segmentation	Database
2024	[152]	MRSNet	RIM-ONE, Drishti-GS, REFUGE
2023	[153]	Transformer and U-Net	Drishti-GS, REFUGE
2022	[154]	U-Net with (Inception-v3)	RIM-ONE, ACRIMA
2022	[154]	U-Net (with VGG19)	RIM-ONE, ACRIMA
2022	[154]	U-Net (with ResNet50)	RIM-ONE, ACRIMA
2021	[155]	GAN	Drishti-GS, REFUGE
2021	[156]	U-Net (with GAN)	RIM-ONE, REFUGE
2019	[157]	Patch GAN	RIM-ONE, Drishti-GS
2018	[158]	FCN	Drishti-GS
2017	[159]	Modified U-Net	RIM-ONE, Drishti-GS, DRIONS-DB
2016	[160]	Ensemble-based CNN	Drishti-GS

critical. DL models can also process and analyse data faster once trained, offering real-time diagnostic capabilities that are essential in clinical settings.

- **Potential for Novel Discoveries:** The DL approach can uncover new patterns or biomarkers for diseases that were previously unknown [37]. By learning from comprehensive datasets, DL models might identify new indicators of eye diseases, leading to breakthroughs in how these conditions are understood and treated.

In summary, DL offers significant advantages over traditional ML in the context of detecting eye diseases, including the ability to automatically learn from data, handle complex patterns, adapt to various tasks, improve diagnostic accuracy and efficiency, and potentially lead to new medical insights.

VIII. SEGMENTATION

In the field of retinal health screening, segmentation involves partitioning a fundus image into multiple segments (sets of pixels) to simplify or change the representation of an image into something more meaningful and easier to analyse [161]. Segmentation is a fundamental process, particularly for delineating key structures such as retinal lesions, the optic cup, and the optic disc [161]. Segmentation aims to isolate and highlight specific structures or regions of interest, such as blood vessels, the optic disc, or exudates. Although traditional segmentation techniques involved image processing techniques, it was improved upon by ML techniques. Recently, DL techniques like U-Net, Fully Convolutional Networks (FCNs), and Region-Based Convolutional Neural Networks (R-CNN) have been commonly used for the segmentation of fundus images [162].

A. OPTIC DISC AND OPTIC CUP SEGMENTATION

The segmentation of the optic disc and optic cup is essential in detecting glaucoma [163]. The optic cup-to-disc ratio (CDR) is a key metric analysed in this process. An increased CDR is indicative of glaucomatous changes, reflecting the loss of retinal nerve fibres and the associated excavation of the optic nerve head [164]. Accurate segmentation of the optic disc and optic cup allows for precise measurement of the CDR, facilitating early detection and monitoring of glaucoma [165]. Additionally, the segmentation of the optic cup helps in assessing the neuroretinal rim, which is crucial for evaluating glaucomatous damage [118].

Optic disc segmentation was first accomplished using techniques such as mathematical morphology, thresholding, and template matching [166], [167], [168]. Later, many studies incorporated the Hough transform into mathematical morphology [169]. Another approach involved tracing blood vessels first and then locating the optic disc by identifying the point where the vessels converged [170].

Active contours, ellipse fitting, and thresholding were commonly used techniques in early works for optic disc and optic cup and segmentation [171]. These algorithms utilised colour/intensity variations or vessel bends (or a combination of both) within the optic disc.

ML techniques have significantly enhanced the segmentation of the optic disc and optic cup, aiding in the estimation of clinically relevant parameters [171]. Image pre-processing, feature selection, and classification techniques are commonly emphasized in many ML studies. The results from these automated methods have shown their effectiveness in detecting glaucoma, often producing results comparable to those obtained through manual analysis by expert clinicians.

However, these techniques tend to be computationally intensive, have been validated on limited datasets, and may favour specific types of images while struggling with others, such as those with very large optic discs and optic cups.

DL has also been applied in optic disc and optic cup segmentation and can be particularly useful in identifying specific structures in the eye, such as the optic nerve head in glaucoma [92].

Al-Bander et al. [74] utilised DenseNet for optic disc and optic cup segmentation, providing precise cup-to-disc ratio (CDR) measurements crucial for glaucoma detection. Similarly, Tan et al. developed a CNN model to segment the optic disc from fundus images, resulting in high classification accuracy [131].

Advanced segmentation techniques have been developed using U-Nets. For instance, Sevastopolsky proposed a method using U-Net for segmenting the optic disc and optic cup, achieving high intersection-over-union (IOU) scores across several databases [159].

Shyamalee and Meedeniya [154] proposed a design of attention U-Net architectures with different CNN backbones to detect glaucoma. Two datasets, RIM-ONE and ACRIMA, were used. The attention U-Net model incorporates attention gates at each skip connection to enhance feature retention and spatial data.

The encoder part of the U-Net was replaced with pre-trained networks (Inception-v3, VGG19, and ResNet50) to identify the best segmentation performance. The three U-Net models with different CNN architectures were trained and evaluated using multiple metrics [154].

The attention U-Net with ResNet50 backbone achieved the highest accuracy of 99.53% in segmenting the optic disc on the RIM-ONE dataset [154]. The study demonstrates the superior performance of the ResNet50-based attention U-Net in accurately segmenting the optic disc and optic cup, which is crucial for glaucoma identification.

The study successfully demonstrates the effectiveness of the attention U-Net model with ResNet50 backbone for optic disc and optic cup segmentation in fundus images, achieving high accuracy and sensitivity. This approach can significantly aid in the early diagnosis of glaucoma, potentially preventing vision loss.

B. RETINAL BLOOD VESSEL SEGMENTATION

Changes and abnormalities in retinal blood vessels, such as neovascularisation, are crucial for DR detection. Segmentation of retinal blood vessels allows for the identification of these changes. Neovascularisation, or the formation of new, fragile blood vessels, is a severe complication of DR that can lead to vision loss if not promptly treated [216]. Vessel segmentation also facilitates the assessment of vessel density and tortuosity, which are important biomarkers for DR progression [92].

The literature in this retinal blood vessel segmentation has seen many works published in the last decade [217].

Supervised methods for blood vessel segmentation use a classifier that requires a training stage with pre-labelled pixel information to adjust parameters, whereas unsupervised methods tackle the segmentation problem directly using various image processing techniques such as vessel tracking, matched filtering, morphological transformations, or model-based algorithms, among others [218], [219].

Two main categories of supervised methods can be distinguished: those based on conventional ML models and those based on DL using CNNs. Supervised methods require a set of mathematical descriptors to characterise and differentiate pixels as either part of the vascular structure or not. ML Classifiers such as SVM then use this mathematical representation to determine the class of each pixel. Some recent ML methods for blood vessel segmentation are given in Table 9.

Mehidi et al. [172] proposed a vessel segmentation method that used CLAHE and bottom-hat filtering to increase the contrast between the vascular and fundus, followed by a Jerman filter. The proposed segmentation model has been evaluated on the STARE and DRIVE databases, reaching an accuracy of 96.18% and 95.86%, and a specificity of 98.10% and 98.74%, respectively.

In recent years, research on blood vessel segmentation in retinal images has increasingly focused on DL methods, as shown in Table 10. Traditional image processing techniques often fail to detect all vessels accurately. Unlike conventional methods, DL approaches internally generate the most appropriate mathematical representation of the vascular structure.

A study [208] proposed a novel deep learning method based on a convolutional neural network (CNN) with a dice loss function for retinal vessel segmentation. The proposed method was tested on the DRIVE and STARE databases and showed superior performance compared to existing methods. Specifically, it achieved a sensitivity of 73.9% and an accuracy of 94.8% on the DRIVE database, and a sensitivity of 74.8% and an accuracy of 94.7% on the STARE database [208].

In Li et al. [214], blood vessel segmentation is treated as a cross-modality data transformation problem, utilising a broad and deep NN to model the relationship between the input image and the output vessel map. The characteristics are extracted in the intermediate layers of the network. DL methods typically employ a CNN architecture that uses a multi-layered cascade (including convolutional, pooling, and activation layers) to extract hierarchical descriptors for the classification stage. Although this final stage can be performed by any trainable classifier (e.g., a Random Forest ensemble as used by a study by Wang et al. [215]), a typical CNN architecture ends with a fully connected neural network structure to make the final classification decision.

In a study by Soomro et al. [220], different DL-based retinal blood vessel segmentation was reviewed. The study showed that methods like the self-organising map (SOM)

TABLE 9. A comparison of ML methods for blood vessel segmentation in fundus eye images.

Year	Author(s)	ML Method used for blood vessel segmentation	Database
2024	[172]	CLAHE, Bottom-hat and Jerman filter	STARE, DRIVE
2023	[173]	Grasshopper Optimisation with Fuzzy Edge Detection	Private
2022	[174]	PCA, RF	Private
2020	[175]	Hessian-based multiscale filter	AEH
2020	[176]	PCA, CLAHE and Fréchet matched filter	STARE, DRIVE
2020	[177]	HMM	STARE, DRIVE
2019	[178]	Cascade classification model	DRIVE, STARE, CHASEDB1
2019	[179]	Probabilistic Patch-Based (PPB) denoiser	STARE, DRIVE
2018	[180]	Patches clustering through K-means algorithm	STARE, DRIVE
2017	[181]	Discriminative dictionary learning and sparse representation	STARE, DRIVE
2016	[182]	Fully connected conditional random field model	DRIVE, STARE, CHASEDB1, HRF
2015	[183]	Iterative process based on adaptive thresholding	DRIVE, STARE, CHASEDB1
2015	[184]	MCA	DRIVE, STARE
2013	[185]	Multiwavelet kernels and multiscale hierarchical decomposition	DRIVE, STARE
2013	[186]	Multilayered thresholding technique and 2-D Gabor wavelet	DRIVE, STARE
2012	[187]	Ensemble system of bagged and boosted DT	DRIVE, STARE, CHASEDB1

TABLE 10. A comparison of DL methods for blood vessel segmentation in fundus eye images.

Year	Author(s)	DL Method used for blood vessel segmentation	Database
2024	[188]	U-Net, DenseU-Net, LadderNet, R2U-Net, ATTU-Net	STARE
2023	[189]	U-Net	DRIVE
2022	[190]	CNN	DRIVE
2021	[191]	U-Net with pre-processing	HRF, CHASEDB1, DIARETDB0, DIARETDB1, STARE, MESSIDOR, DRIVE, IDRiD
2020	[192]	CNN	DRIVE, STARE
2020	[193]	CNN	DRIVE, STARE
2020	[194]	Lightweight CNN	DRIVE, STARE, CHASEDB1
2020	[195]	U-Net network and semi-supervised learning	DRIVE
2020	[196]	HANet	DRIVE, STARE, CHASEDB1, HRF
2020	[197]	Multi-Scale Feature Fusion U-Net	DRIVE, STARE
2020	[198]	U-Net	ORIGA, REFUGE, DRIVE, CHASEDB1
2020	[199]	U-Net following with a U-Net with residual blocks	DRIVE, CHASEDB1
2020	[200]	Custom U-Net	DRIVE, STARE, CHASEDB1, HRF, ARIA
2020	[201]	GANs and Dense U-Net	DRIVE
2020	[202]	U-Net	DRIVE, STARE, CHASEDB1
2019	[203]	Multi-scale CNN	DRIVE, STARE, CHASEDB1, HRF
2019	[204]	Recurrent U-Net	DRIVE, STARE, CHASEDB1
2019	[205]	U-Net and residual multi-kernel pooling block	DRIVE, MESSIDOR
2019	[206]	Cascade CNN	DRIVE, STARE
2019	[207]	DenseNet extended to fully convolutional network	DRIVE, STARE
2018	[208]	CNN with dice loss function	DRIVE, STARE
2018	[209]	Multi-scale FCN	DRIVE, STARE, CHASEDB1
2018	[102]	CNN (3 convolutional layers, 1 fully connected layer)	DRIVE, DIARETDB0, DIARETDB1, AEH
2018	[210]	FCN	DRIVE, STARE, CHASEDB1, HRF
2018	[211]	MCNN	DRIVE, STARE
2016	[212]	CNN (Multiple convolutional layers followed by 3 fully connected layers)	DRIVE, STARE, CHASEDB1
2016	[213]	Multi-scale and multi-level CNN	DRIVE, STARE, CHASEDB1
2016	[214]	5-layer NN	DRIVE, STARE, CHASEDB1
2015	[215]	CNN (feature extraction), RF (classification)	DRIVE, STARE

and ensemble learning have shown good results but often struggle with tiny vessel detection [220]. Sangeethaa and Maheswari [102] proposed a CNN that learns from pre-processed retinal images instead of raw image data.

Another study proposed an ICA-based image enhancement technique that significantly improves retinal vessel segmentation, achieving better performance than existing methods [221]. The study's findings suggest that this approach can be extended to other medical imaging applications requiring low-contrast feature detection [221].

Jiang et al. [210] used a fully convolutional neural network (FCN) pre-trained on a natural image dataset,

using transfer learning for vascular tree segmentation. Feng et al. [192] suggested a cross-connected CNN, where all convolutional layers of the primary and secondary paths are connected to facilitate multi-level feature fusion.

The primary benefit of unsupervised vessel segmentation methods is that they do not require manual annotation. These methods utilise or identify image properties to classify pixels as either vessel or non-vessels. The GMM-expectation maximisation (EM) algorithm has also been employed for vessel segmentation. The EM algorithm provides a maximum-likelihood classification of a vessel and non-vessel

TABLE 11. A comparison of ML methods for microaneurysm segmentation in fundus eye images.

Year	Author	ML Method for microaneurysm segmentation segmentation	Database
2020	[225]	RUSBoost, RF, AdaBoost, SVM, KNN, etc.	e-Ophtha, DIARETDB1, ROC
2020	[226]	ANN	ROC, AGAR300
2020	[227]	SVM, MLP	e-Ophtha
2019	[228]	SVM	ROC
2019	[229]	KNN, Ensemble	MESSIDOR, DIARETDB0
2017	[230]	RF, SVM, NN	DIARETDB1
2017	[231]	KNN, SVM	DIARETDB0, DIARETDB1
2016	[232]	KNN	MESSIDOR
2016	[233]	SVM	DIARETDB1
2013	[234]	NB	Private
2013	[235]	NB, KNN, SVM	ROC

TABLE 12. A comparison of DL methods for microaneurysm segmentation in fundus eye images.

Year	Author	DL Method for microaneurysm segmentation segmentation	Database
2021	[236]	Sparse auto-encoder based feature transfer network	ROC, DIARETDB1, e-Ophtha, RetinaCheck
2019	[237]	Mask R-CNN with ResNet101	e-Ophtha, IDRiD
2019	[238]	U-Net	IDRiD
2019	[239]	L-Seg	DDR, IDRiD
2019	[240]	YOLO	Kaggle EyePACS
2019	[241]	CNN	ROC, e-Ophtha
2018	[242]	Deep U-Net	e-Ophtha, DIARETDB1, ROC
2017	[243]	CNN	Private (CLEOPATRA)
2016	[244]	Stacked sparse auto-encoder	DIARETDB

pixels, with vessel enhancement achieved through high-pass filtering and the top-hat transform [222].

Recently, many works have adopted the U-Net DL model such as the one proposed by Ronneberger [223]. It has proven effective in medical image segmentation, especially for problems involving class imbalance and limited sample sizes, as with blood vessel segmentation in retinal images. The conventional U-Net structure has been used as a network model in studies by Darmo et al. [189], Chen et al. [195], and Yin et al. [198] for blood vessel segmentation.

Moreover, Bayesian U-Net and weakly supervised learning approaches have been employed to enhance segmentation efficiency and reduce manual annotation efforts, as demonstrated by Xiong et al. [224]. These methods address inter-subject variability and improve model performance by optimising the segmentation process through innovative techniques.

C. MICROANEURYSM AND EXUDATES SEGMENTATION

Microaneurysms, appearing as small red dots on the retina, are one of the earliest signs of DR. AI algorithms segment these microaneurysms to detect the onset of DR. The presence and number of microaneurysms are critical indicators of the severity of DR, with more advanced stages showing an increased number of these lesions [32].

Early detection through the segmentation of microaneurysms enables timely intervention and can prevent progression to more severe stages [245]. Exudates are lipid residues that appear as yellow spots on the retina and are another hallmark of DR [246]. Their segmentation is vital for diagnosing DR, as the presence of exudates signifies leakage from damaged blood vessels in the retina [246]. Detecting and quantifying exudates help in assessing the extent of

retinal damage and the progression of DR [53]. Advanced segmentation techniques enable precise localisation and classification of exudates, contributing to more accurate diagnosis [247].

Microaneurysm and exudates segmentation from fundus images have seen significant advancements over the past decade. The methods for segmentation can be categorized into traditional image-processing techniques and modern DL approaches. For microaneurysms segmentation, traditional methods include morphological processing, wavelet transformation, and hybrid classifier approaches.

One of the earliest studies by Spencer et al. [248] used morphological operations to eliminate vasculature in fluorescein angiograms, isolating small structures like microaneurysms. This approach relied heavily on manual feature engineering and traditional image processing techniques.

Quelleg et al. [247] introduced an adaptive wavelet method using local template matching in the wavelet domain to detect microaneurysms. Akram et al. [249] developed a hybrid classifier combining a Gaussian mixture model, and support vector machine (SVM) to identify microaneurysms.

A study by Sreng et al. [250] presented an effective method for the segmentation of microaneurysms from fundus images. Initially, they pre-processed the fundus images to reduce noise and enhance contrast. They then segmented the images using Canny edge detection and maximum entropy thresholding. Microaneurysms were distinguished from other lesions and anatomical structures in the fundus image using area and eccentricity methods [250]. Finally, morphological operations were applied to highlight these symptoms. Ophthalmologists analysed the results to assess the system's accuracy and precision. Their comparative analysis showed a 90% accuracy rate [250].

In contrast to traditional methods, DL utilise deep networks to perform segmentation tasks, automatically extracting useful image features. With the advancement of DL, neural networks have become prevalent in microaneurysms and exudates segmentation.

Haloi [251] employed a deep neural network with three convolutional layers and two fully connected layers for automatic microaneurysms segmentation. Kou et al. [252] proposed a deep residual U-Net, combining a deep residual model and recurrent convolutional operations into a U-Net for microaneurysms segmentation.

Exudates segmentation has seen the development of traditional methods such as thresholding, and morphological processing.

A study by Phillips et al. [253] used global threshold techniques for fundus images for exudates. Walter et al. [51] applied morphological reconstruction to locate exudates. For exudates segmentation, Perdomo et al. [254] applied LeNet, a CNN.

In AMD, particularly the wet form, exudates indicate fluid leakage and neovascularisation [255]. Segmentation of exudates in AMD patients helps in identifying the presence of abnormal blood vessels and fluid accumulation, which are critical factors in diagnosing and managing wet AMD [18]. Automated exudate segmentation supports early detection and monitoring, improving patient outcomes [256].

Microaneurysm and exudates segmentation has evolved from traditional image-processing techniques to sophisticated DL models. Early methods relied on morphological operations, wavelet transformations, and hybrid classifiers. The transition to DL introduced CNNs, FCNNs, and U-Net variants, significantly enhancing segmentation accuracy. Recent advancements include enhanced residual U-Nets, attention mechanisms, and transformer-based models.

A study by Kou et al. [257] introduced an enhanced residual U-Net (ERU-Net), which featured one downsampling path and three upsampling paths. Unlike the original U-Net, the three upsampling paths in ERU-Net enhanced the fusion feature maps and captured more details of fundus images. Additionally, a residual block in ERU-Net was designed to extract more representative features. The study showed that ERU-Net performs well in segmenting microaneurysms and exudates. Compared to other U-Net variants, ERU-Net achieved the best performance across three publicly available fundus image segmentation datasets.

D. HAEMORRHAGE SEGMENTATION

Retinal haemorrhages, which are bleeding spots, play a significant role in DR detection [258]. Segmentation of these haemorrhages enables the identification of AMD and more advanced stages of DR. The presence of retinal haemorrhages indicates significant vascular damage and warrants immediate medical attention [258].

Automated segmentation of haemorrhages aids in comprehensive retinal screening and monitoring [246]. The initial

attempts at haemorrhage segmentation relied heavily on traditional ML approaches, which focused on the extraction of handcrafted features from fundus images.

Kande et al. [259] employed pixel classification and mathematical morphology to detect haemorrhages. They utilised the red and green channels of the images to determine the presence of red lesions. Subsequently, the SVM algorithm was applied to classify candidate areas for red lesion containment. This approach achieved a specificity of 91% and a sensitivity of 100%.

A study by Tan et al. [243] used a 10-layer multiclass neural network for segmenting haemorrhages in retinal fundus images. Their method achieved a haemorrhage segmentation sensitivity of 62.57% and a specificity of 98.93%.

Orlando et al. [70] developed a method that combines a CNN with an RF for segmenting hemorrhages and microaneurysms. The RF algorithm generates probability maps of hemorrhages and microaneurysms at the image level, utilising features from the green layer of patches extracted by the CNN architecture. This approach achieved a sensitivity of 48.83% for detecting hemorrhages and microaneurysms.

Badar et al. [260] introduced an encoder-decoder model for the simultaneous segmentation of hemorrhages and exudates, based on a CNN classifier. When trained and tested on the Messidor dataset, this model achieved a hemorrhage segmentation accuracy of 97.86%.

E. DRUSEN SEGMENTATION

Drusen, yellow deposits under the retina, are the main symptoms of AMD [99]. Segmentation of drusen is important in detecting AMD, as their presence and size correlate with the severity of the disease [261]. Early detection of drusen can help in monitoring AMD progression and initiating timely interventions. AI-based segmentation of drusen enables accurate quantification and characterisation, aiding in personalised treatment plans ([261]).

The literature shows, that drusen segmentation methods follow two main approaches. The first approach relies on traditional image processing techniques, where various local features are extracted and then classified using an ML, such as SVM [262]. The primary objective is either to directly detect the drusen region or to delineate its boundaries. Traditional methods relied on handcrafted features and were typically limited by their inability to generalise across varying image conditions.

For instance, Kim and Kim [263] applied multiple filters to candidate regions to detect drusen. They used traditional image processing techniques for drusen segmentation. However, by using only local features and handcrafted filters, it will not work as well in early AMD, where the drusens are not as prominent in fundus images.

In 2011, Mora et al. [264] used a gradient-based segmentation algorithm to isolate drusen and provide basic drusen characterisation. The approach had a maximum sensitivity of 74% and a specificity of 97%. Mohaimin et al. [265]

introduced a colour normalisation method to address the issue of colour variations in fundus images for detecting drusen.

Ren et al. [262] used SVM to classify drusen from fundus images from the STARE and DRIVE datasets. The method achieved sensitivity, specificity, and accuracy of 90.03%, 97.06% and 96.92% on the STARE dataset and a sensitivity, specificity, and accuracy of 87.41%, 94.93%, and 94.81% on the DRIVE dataset [262].

Sbeh et al. [266] proposed a method for drusen segmentation from fundus images using an adaptive algorithm based on mathematical morphology transforms.

Rapantzikos et al. [267] developed the histogram-teased adaptive local thresholding technique for drusen detection in fundus images, efficiently extracting useful information while ignoring other pathological structures. Various fuzzy logic-based techniques and texture-based methods have been proposed for drusen detection and segmentation from fundus images [268], [269] [270].

Brandon and Hoover [271] employed a multi-level approach, beginning with pixel-level classification and progressing to region-level, area-level, and finally image-level analysis, which enabled the detection of drusen with an accuracy of 87%.

Recently, DL models such as U-Nets have been utilised in drusen segmentation. Yan et al. [238] utilised two U-Nets to capture both global and local information. In this approach, feature maps are treated as global information and are merged in the final layer. However, this configuration necessitates limiting the number of channels in the feature maps to prevent excessive computational demands [238].

Pham et al. [272] proposed a multi-scale DL model to make a fine drusen segmentation prediction. Their method is suitable for high-resolution fundus images. Whereas previous studies on drusen segmentation analysed a cropped image to solve the high-resolution problem, the method by Pham et al. combined both global and local information, by which the model is able to predict more accurate drusen segmentation. Additionally, by utilising the pre-trained model and the combination of different loss functions, the performance of detecting drusen in the early stages of AMD is improved [272].

IX. CLASSIFICATION

Classification involves assigning a label to an image based on its content, such as identifying the presence or absence of specific eye diseases from retinal images [137]. This process helps in diagnosing and categorising retinal conditions.

Classification approaches in retinal health screening have evolved significantly with the advent of ML and DL. ML techniques such as SVM, KNN, NB, DT, RF and others have been extensively used to classify eye diseases like AMD, DR, and Glaucoma. by analysing various features from fundus images to detect and classify these conditions with high accuracy. Recently, DL models, particularly CNNs

and their variations have been used in the detection of fundus images. DL models have automated feature extraction from raw image data and enhanced classification accuracy [102]. These models enable end-to-end learning, streamlining workflows by handling the entire process from image input to disease classification.

In the literature, one of the more popular ML models, Support Vector Machines (SVM) has been extensively used for classifying DR, AMD, and glaucoma. SVM is used due to their ability to handle high-dimensional data and create decision boundaries that maximize the margin between classes [273]. It involves extracting relevant features from fundus images and then selecting important features to reduce dimensionality and improve classification accuracy. However, SVM's performance depends on the choice of kernel and parameters, and it can be computationally intensive with large datasets [273]. It requires careful tuning of parameters and kernel selection to achieve optimal performance. Additionally, handling imbalanced datasets can be challenging and may require techniques such as oversampling or the use of different class weights [274].

A study by Antal and Hajdu [274] used SVM for DR classification by analysing features such as microaneurysms, hemorrhages, and exudates in retinal images. By mapping input data into high-dimensional space, SVMs create a hyperplane that best separates different classes, such as different stages of DR. This technique is highly effective for binary classification tasks, such as distinguishing between healthy and DR-affected eyes.

Sarni et al. [275] suggested a decision-support system for microaneurysms for DR screening. In order to classify the microaneurysms, Antel et al. used ensemble learning [276].

Using local binary patterns, Morales et al. performed classification to differentiate between Normal, DR, and AMD [277]. The diameter of blood vessels is another aspect of DR that changes. As a result, it is another feature used to categorize DR. Using a Gaussian filter, Nikita et al. segmented blood vessels, extracted texture and structural features, and then performed classification using SVM and ANN [95].

Bowd et al. [278] used SVMs for glaucoma classification by analysing features extracted from optic nerve head images, such as the cup-to-disc ratio and neuroretinal rim width. SVMs excel at distinguishing between glaucomatous and non-glaucomatous eyes, especially when dealing with high-dimensional feature spaces [278]. Floriano et al. [126] used SVM to classify AMD by identifying patterns in retinal images. SVMs are capable of handling high-dimensional data and are used to differentiate between healthy and AMD-affected eyes based on features such as drusen size, shape, and distribution [126]. SVMs maximize the margin between different classes, leading to robust classification results [279].

A study that showed potential in quickly and accurately detecting glaucoma has been put forward by some authors [111]. They acquired their fundus images privately.

The study showed that Bi-empirical mode decomposition (Bi-EMD) and the scalogram of continuous wavelet transform are used. Entropy features are a common feature extraction in fundus images because they are capable of accurately measuring pixel variation. To identify normal fundus images from abnormal ones, such as glaucoma, a retinal risk index has been established [111]. Both Bi-EMD and CWT have produced encouraging results with an accuracy of 88.6% using the SVM classifier and 92.48% using the RFF classifier. An accuracy of 96.2%, sensitivity of 95%, and specificity of 97.4% was achieved with a ten-fold cross-validation strategy using the KNN classifier. This novel algorithm has great potential in detecting glaucoma quickly and reliably.

DR is classified in a variety of ways using various databases [280], [281], [282]. Exudates were extracted by Du and Li [283], who then classified the samples into normal, NPDR, and PDR using SVM. Exudates were also extracted by Tjandrasa et al. [279], who also classified DR as mild, moderate, or severe using SVM as a classifier.

Gupta et al. [284] achieved an accuracy of 92% in detecting DR on the APTOS2019 and EyePacs datasets. They used the Life Choice-Based Optimizer (LCBO) algorithm, which selects the optimal features from the extracted set. These features are then fed into an optimised hybrid machine learning classifier, combining a Neural Network (NN) and a Deep Convolutional Neural Network (DCNN), where the Social Ski-Driver (SSD) algorithm is used to determine the best weight values for the hybrid classifier. This classifier categorises the severity of DR into mild, moderate, severe, proliferative DR, and normal.

K-Nearest Neighbors (KNN) is a simple yet effective ML technique that has been widely used for classifying DR, AMD, and glaucoma [261]. KNN is a non-parametric method that classifies a data point based on the majority class of its k-nearest neighbours [285]. KNN has been used to classify AMD by analysing features such as the presence and distribution of drusen, changes in retinal pigmentation, and other abnormalities [261]. Once these features are extracted, KNN can classify new retinal images by comparing them with previously labelled examples. The simplicity of KNN makes it a useful baseline model for AMD classification, providing a straightforward approach to identifying patterns in retinal images [261].

In a study by Kermany et al. [17], a multi-class comparison of different ophthalmic diseases using DL achieved an accuracy of 96.6%, with a sensitivity of 97.8% and a specificity of 97.4%. This study suggests that DL can be used to accurately classify eye diseases, which can have important implications for disease detection and monitoring.

Decision Trees (DT) classify glaucoma by sequentially splitting data based on features like cup-to-disc ratio, visual field test results, and intraocular pressure readings [286]. Pathan et al. [286] used decision trees (DT) to classify optic disc contours in fundus images, which is useful in detecting glaucoma. While DTs are easy to interpret, they

are prone to overfitting, and their performance improves significantly when used within an ensemble method like random forest (RF) [287]. DTs are used in classifying DR by evaluating features such as blood vessel abnormalities, microaneurysms, and exudates [288]. They are also used to classify AMD by recursively splitting the data based on the most significant features. DTs are straightforward to interpret and can effectively use features like drusen presence and retinal pigment epithelium abnormalities [289].

Random Forests (RF) are an ensemble learning method, that combines multiple decision trees to improve classification accuracy [290]. It can be seen from the literature that RFs are used to classify AMD by analysing various retinal features including texture, intensity, and colour information [17]. The ensemble approach of RFs reduces overfitting and improves generalisation, making it a reliable choice for AMD classification [17]. RFs have also been used to classify glaucoma by examining multiple features such as optic disc size, retinal nerve fibre layer thickness, and intraocular pressure. The robustness of RFs in handling feature variability and their resistance to overfitting make them suitable for glaucoma classification [291]. RFs are also extensively used for DR classification due to their ability to handle large datasets with many features. RFs analyse various retinal characteristics, including the number and severity of lesions, to classify the disease [53].

Naive Bayes (NB) is a simple yet powerful probabilistic classifier, that is advantageous in DR classification due to its simplicity and ability to handle both binary and multi-class problems [292]. It is particularly useful when dealing with missing data, as it can handle incomplete data without requiring imputation. However, the assumption of feature independence can be a limitation, especially when dealing with complex fundus images where features are often interrelated [274].

In glaucoma classification, NB can be applied to features extracted from optic nerve head images, such as the cup-to-disc ratio, neuroretinal rim width, and retinal nerve fibre layer thickness. The classifier calculates the probability of glaucoma given these features and assigns the diagnosis based on the highest probability [293]. However, the model's accuracy might be compromised due to the unrealistic assumption of feature independence, which can affect its performance in more complex cases [294].

El-Khalek et al. [295] achieved an accuracy of 96.85% in detecting AMD on a private dataset in 2024. Their proposed system extracted both local and global appearance markers from fundus images. These markers were obtained from the entire retina and iso-regions aligned with the optical disc. Their study used advanced classification schemes to locate and analyse the data. These algorithms include various methods, such as AdaBoost, RF, DT, logistic regression, SVM, KNN, and others. Their system not only achieved a high level of accuracy but also provided a detailed assessment of the severity of each retinal region.

ML techniques have shown great promise in classifying AMD, DR, and glaucoma. Each method has its strengths and limitations, but when applied appropriately, they provide valuable tools for the early detection of these eye diseases. The continued development and refinement of these techniques will enhance their accuracy and reliability, ultimately improving patient outcomes.

Recently, DL has been increasingly used by researchers to classify diseases from fundus images. DL models such as CNNs receive inputs in the form of pixels, sub-images, and entire images to perform classification [100]. Studies have shown that CNN models can match or even surpass expert ophthalmologists in detecting retinal diseases such as DR, AMD, and glaucoma [103]. There has been a lot of research on an automatic CNN-based system [213] for categorising retinal images into different severity levels. CNNs can combine the input images using an appropriate weight matrix and extract unique features of the input images while preserving the spatial arrangement information [213]. The scalability and generalisability of DL models make them suitable for widespread clinical use, as they perform well across diverse populations and varying image qualities by training on large datasets.

VGGNet is a CNN model known for its simplicity and use of small convolution filters, that has achieved high performance in image classification tasks [97]. ResNet is another DL model that introduces residual learning to address the vanishing gradient problem in deep networks, allowing for the training of very deep networks [296].

A study by Shyamalee and Meedeniya [297] compared the performance of three CNN architectures (Inception-v3, VGG19, ResNet50) using two datasets: RIM-ONE and ACRIMA. Pre-processing techniques such as dilation and Contrast Limited Adaptive Histogram Equalisation (CLAHE), enhanced image quality. The models were evaluated using 5-fold cross-validation on the RIM-ONE and ACRIMA datasets. The Inception-v3 model achieved the highest accuracy of 96.56% on the RIM-ONE dataset and 98.52% on the ACRIMA dataset [297].

Božić-Štulić and Stipaničev [298] used a DL algorithm to predict the presence of glaucoma from fundus images, with an accuracy of 97.3%. According to this study, DL may be able to identify glaucoma even in its early stages, which could significantly affect how the condition is managed and treated.

In a study by Ogundokun et al. [299], a DL method was contrasted with deep CNNs trained for automated evaluation of AMD. Automated identification was applied to a 2-class classification problem to distinguish between the AMD stage and the Normal stage and achieved an accuracy of 96.41%, a specificity of 94.82%, and an AUC of 0.9633. This study shows that DL models could perform a task in the current AMD management independent of skilled ophthalmologists.

Gulshan et al. [53] proposed a DL-based system for DR detection, which achieved high sensitivity and specificity comparable to human experts. Other studies obtained similar

results by utilising various CNN architectures. In [246], an automatic DL-based model for detecting DR severity is presented. The five modules that comprise the CNN-based automatic diabetic detection model for retinal images are pre-processing, exudates segmentation, blood vessel segmentation, texture feature extraction, and DR detection [300]. Adaptive histogram equalisation is used in the pre-processing stage to improve the quality of the input retinal images. In the second step, exudate and blood vessel segmentation are carried out by fuzzy c-means clustering and CNN. After extracting texture features from the exudates and blood vessels, an SVM implementation is used to identify DR.

Qomariah et al. [140] proposed a CNN and SVM-based automated system for the classification of DR and normal retinal images. Exudates, haemorrhage, and microaneurysms were characteristics. The proposed system was divided into two sections by the author: the first section included feature extraction based on neural networks, and the second section carried out classification using SVM. Researchers have proposed several methods for categorising DR, including pre-processing of the raw images, image enhancement, and post-processing, which are all fundamental aspects of image processing. After training, features are extracted, and classes are determined. Various features are extracted and used as training algorithm inputs. ANN was used to classify the disease stages by using features such as area, perimeter, and exudates count [301].

X. SEGMENTATION FOLLOWED BY CLASSIFICATION

Segmentation followed by classification methods, especially those based on fundus images, has seen significant advancements. In detecting ODs from fundus imaging, segmentation followed by classification involves two main steps: segmenting (identifying and isolating) specific regions or structures within a fundus image, and then classifying these segmented regions to detect a condition or categorise them into predefined classes. Segmentation ensures that the classification focuses on relevant regions, improving accuracy.

A study by Shyamalee et al. [302] performed segmentation using an attention U-Net with ResNet50 and classification using a modified InceptionV3. The attention U-Net with ResNet50 backbone achieved the highest segmentation accuracy for optic disc and optic cup on the RIM-ONE dataset. For classification, the modified Inception V3 model showed the highest performance. The final model predictions are based on the segmented images, and the cup-to-disc ratio is computed to support the classification results [302]. To make the DL model's decisions transparent, Grad-CAM and Grad-CAM++ generate heatmaps that highlight the regions of the fundus images influencing the predictions. These heatmaps help ophthalmologists understand the model's reasoning, increasing trust in the system.

Sangeetha and Maheswari [102] proposed a method for retinal image segmentation and blood vessel extraction using

TABLE 13. Summary of works done using classification and segmentation for automated detection of various eye classes.

Year	Author	Segmentation Method	Classification Method	Segmentation of:	Classification of:
2024	[302]	Attention U-Net with ResNet50	Modified Inception V3	Optic Disc, Optic Cup	Glaucoma
2022	[303]	Frangi filter	U-Net	Retinal Blood Vessel	DR
2022	[304]	Attention U-Net with Inceptionv3/VGG19/ResNet50	Inceptionv3/VGG19/ResNet50	Optic Disc, Optic Cup	Glaucoma
2022	[305]	U-Net	MFO-KELM	Optic Disc, Optic Cup	Glaucoma
2022	[306]	DNN	DNN	Optic Disc, Optic Cup	Glaucoma
2022	[307]	MSGANet-RAV	MSGANet-RAV	Optic Disc, Optic Cup	Glaucoma
2021	[308]	U-Net	U-Net	Retinal Blood Vessel	DR
2020	[272]	Multi-scale DL	U-Net	Drusen	AMD
2020	[309]	U-Net	CNN	Optic Disc, Optic Cup	Glaucoma
2019	[310]	Dual-Tree Complex Wavelet Transform	CNN, RNN	Retinal Blood Vessel	DR
2018	[262]	Generalized low rank approximation of matrices	SVM	Drusen	AMD
2018	[102]	Morphological operations, Binarisation thresholding	CNN	Retinal Blood Vessel	DR
2015	[215]	CNN	RF	Retinal Blood Vessel	DR

morphological processing, thresholding, edge detection, and adaptive histogram equalization. For the automatic diagnosis of DR from fundus images, they developed a CNN to accurately classify the severity of the disease. This network was trained on a high-end graphical processor unit (GPU) using publicly available datasets such as DRIVE, DIARETDB0, and DIARETDB1, as well as images collected from the Aravind Eye Hospital in Coimbatore, India. The proposed CNN achieved a sensitivity of 98%, a specificity of 93%, and an accuracy of 96.9% on a database of 854 images [102].

Yin et al. [303] developed a Deep Fusion Network, incorporating multiscale fusion, feature fusion, and classifier fusion for multi-source vessel image segmentation for DR detection. The multiscale fusion module enabled the network to detect blood vessels of various scales. The feature fusion module combines deep features with vessel responses extracted from a Frangi filter to create a compact and domain-invariant feature representation. The classifier fusion module enhances network supervision. DF-Net also predicts the Frangi filter's parameters, eliminating the need for manual parameter selection. The learned Frangi filter improves the feature map of the multiscale network and restores edge information lost during down-sampling operations. This proposed end-to-end network is easy to train, and the inference time for one image is 41ms on a GPU. The model outperforms state-of-the-art methods, achieving accuracies of 96.14%, 97.04%, and 98.02% on three publicly available fundus image datasets: DRIVE, STARE, and CHASEDB1, respectively [303].

In a study by Hervella et al. [306], a novel multi-task approach is proposed for the simultaneous classification of glaucoma and segmentation of the optic disc and cup. This approach aims to improve overall performance by leveraging both pixel-level and image-level labels during network training. Furthermore, the predicted segmentation maps, alongside the diagnosis, allow for the extraction of relevant biomarkers such as the cup-to-disc ratio. The proposed methodology introduces two significant technical innovations. First, a network architecture that enables simultaneous segmentation and classification by increasing the number of shared parameters between both tasks. Second, a multi-adaptive optimization strategy ensures that both tasks

contribute equally to the parameter updates during training, thus avoiding the need for loss-weighting hyperparameters. To validate this proposal, extensive experiments were conducted on the public REFUGE and DRISHTI-GS datasets. The results demonstrate that this approach outperforms comparable multi-task baselines and is highly competitive with existing state-of-the-art methods. Additionally, the provided ablation study indicates that both the network architecture and the optimization strategy independently contribute to the advantages of multi-task learning [306].

Another study by Shyamalee and Meedeniya [304] proposed a DL model to segment and classify retinal fundus images for glaucoma detection. Various data augmentation techniques were applied to prevent overfitting, along with several data pre-processing approaches to enhance image quality and achieve high accuracy. The segmentation models were based on an attention U-Net architecture, utilising three different convolutional neural network (CNN) backbones: Inception-v3, Visual Geometry Group 19 (VGG19), and Residual Neural Network 50 (ResNet50). The classification models also employ modified versions of these three CNN architectures. Using the RIM-ONE dataset, the attention U-Net with the ResNet50 model as the encoder backbone achieved the highest accuracy of 99.58% in segmenting the optic disc. Among the evaluated segmentation and classification architectures, the Inception-v3 model achieved the highest accuracy of 98.79% for glaucoma classification.

A study by Chowdhury et al. [307] proposed a multiscale guided attention network named MSGANet-RAV for pixel-wise retinal artery-vein classification. The proposed architecture integrates multiscale feature exploration with a sequence of GF and context-learnable SVA modules. As a joint task of pixel identification in ophthalmic images, the model incorporates a learnable joint-task loss method, balancing the weights of individual task losses to enhance artery-vein classification. Multiscale features of these images are refined through a two-stage GA module. In the first stage, the structural information of variant vessels is explored, while in the second stage, more refined feature representations are obtained by fusing contextual vessel information with the vessel skeleton (probability map). MSGANet-RAV achieved state-of-the-art performance on the LEI-CENTRAL dataset and

demonstrated comparable performance on the AV-DRIVE dataset, according to several benchmark metrics [307].

A study by Lim et al. [311] introduced the CNN-FE model, which enhances input features by highlighting disc pallor and vessel obstructions in fundus images. This model refines pixel-level probability maps by incorporating known retinal morphology, thereby improving segmentation validity and classification performance. Such integration of segmentation and classification processes leads to more accurate and reliable diagnostic outcomes by focusing on morphological features and improving confidence in the results [311].

Researchers have utilised transfer learning to adapt pre-trained models, such as Inception and ResNet, for retinal disease detection [312]. These pre-trained models have been fine-tuned to classify retinal images and demonstrated improved performance compared to models trained from scratch. The use of ensemble methods has been suggested in several studies. To improve performance overall, ensemble approaches aggregate the predictions of several machine learning models. These methods can help mitigate overfitting and improve model generalisation. Systems for detecting retinal diseases have been made more accurate and resilient through the use of ensemble approaches. For example, Sahlsten et al. used an ensemble of DL models to detect DR more effectively than they could with individual models [313].

Calleja and Medina [314] used a two-stage approach to detect DR that included LBP for feature extraction and ML, particularly SVM and RF, for classification. The results showed that RF outperformed SVM with an accuracy of 97.46%.

A study by Koh et al. [315] has conducted research on diagnosing retinal health in fundus eye images using a pyramid histogram of oriented gradients (PHOG) and speeded-up robust features of fundus images (SURF). Canonical correlation analysis was used to fuse the extracted correlated features. It achieved an accuracy of 96%, a sensitivity of 95%, and a specificity of 97% using the KNN classifier. The outcomes show that this method is useful for automatically classifying eye conditions like glaucoma.

Ren et al. [262] proposed a supervised feature learning method designed to create discriminative and compact descriptors for drusen segmentation in retinal images. This method integrates generalised low-rank approximation of matrices with supervised manifold regularization to derive new features from image patches sampled from retinal images. These learned features are specifically related to drusen and are potentially free from redundant information that could interfere with distinguishing drusen from the background. The features are then vectorised and used to train a support vector machine (SVM) classifier. Finally, the trained SVM classifier is utilised to classify the pixels in the test images as drusen or non-drusen. The proposed method's performance is validated on the STARE and DRIVE databases, achieving average sensitivity, specificity,

and accuracy of 90.03%, 97.06%, and 96.92%, respectively, on STARE, and 87.41%, 94.93%, and 94.81%, respectively, on DRIVE.

Overall, the literature suggests that by analysing fundus images and other imaging modalities, both ML and DL may be able to increase the accuracy of identifying eye conditions such as glaucoma, DR, and AMD. The performance of DL increases over ML as more images are added to the dataset. Further investigation is required to assess the algorithms' performance in larger and more varied datasets in order to validate their generalisability and determine the models' therapeutic potential. In addition, there is also ongoing research in developing new algorithms that can improve the performance of these models.

XI. DISCUSSION

From the literature we reviewed, various methods have been proposed for image-level classification, microaneurysm, exudate or blood vessel segmentation (at the pixel or object level), or segmentation of the optic disc and optic cup, which are important for estimating clinical parameters and facilitating the diagnostic process. Methods for image-level classification have been developed for DR, AMD, and glaucoma. These classifications are mainly binary, distinguishing between healthy and pathological conditions. However, more nuanced classifications have also been proposed, such as differentiating between no-glaucoma, suspicious glaucoma, and glaucoma, as well as up to six classes for DR, and AMD.

In our review, it has been shown that current ML and DL systems have achieved high accuracy in detecting retinal diseases. DL models, particularly CNNs, have achieved sensitivity and specificity rates often above 90% for detecting conditions like DR, AMD, and glaucoma [134]. Studies report AUC values frequently exceeding 0.90, indicating excellent diagnostic performance. However, the accuracy can vary depending on the dataset quality, model architecture, and the specific disease being detected.

We have found from our review that various features are extracted to detect retinal diseases. For glaucoma, features include the cup-to-disc ratio (CDR), optic nerve head morphology, and retinal nerve fibre layer thickness. Diabetic retinopathy features include microaneurysms, haemorrhages, exudates, and neovascularisation. AMD features include drusen size and distribution, retinal pigment epithelium changes, and geographic atrophy.

Pixel-level segmentation is a challenging but important task. Optic disc and optic cup segmentation are essential for a comprehensive and interpretable assessment of glaucoma. Segmenting retinal lesions such as drusen, exudates, hemorrhages, and microaneurysms allows for the estimation of their areas, locations, and changes over time, which is crucial for the precise diagnosis and monitoring of DR and AMD. For these two diseases, automatic segmentation has been used to provide more detailed image-level classification and disease grading.

However, the various retinal pathologies have generally been treated independently, with specific methods developed for each. This means that the development of algorithms for recognising one specific pathology often does not incorporate the knowledge gained from developing methods for detecting other pathologies.

The initial methods developed for this purpose relied on conventional image processing techniques such as thresholding, morphological operations, and model matching to recognize specific shapes like ellipses for the optic disc and small circles for drusen. These methods showed promising results on the datasets they were developed and tested on, but they failed to perform adequately on new, unseen images.

ML methods improved upon those image-processing techniques and achieved better results. Various supervised and unsupervised learning methods have been developed to assess different pathologies, enabling both image-level classification and segmentation of the optic disc, and optic cup. In this context, image pre-processing and feature selection play crucial roles. Pre-processing aims to reduce noise using techniques like moving average filters, median filters, and Gaussian filters, and to improve contrast, often using CLAHE. Feature selection involves identifying and extracting various features from the image and selecting the most significant ones, a process initially done manually and highly dependent on the scientist's expertise. To minimize subjectivity, a wide range of features was generally identified. Once all possible features were derived, Principal Component Analysis (PCA) was commonly used to reduce the feature space by selecting the most informative features.

Despite the promising results of ML, these methods were not robust against inter-subject anatomical variability (such as the appearance and shape of the OD), pathological changes (like the onset or variation of lesions), differences in acquisition systems (from different vendors), and limitations of the acquisition systems (such as noise and illumination drifts in the images).

In recent years, numerous deep learning (DL) techniques have been introduced, significantly enhancing retinal image classification and segmentation. DL methods automatically perform feature extraction and selection, allowing them to be applied directly to images without extensive pre-processing. Most proposed DL methods utilise standard and pre-trained CNNs, which, through transfer learning, achieve impressive performance even on limited datasets. The literature surveys indicate that ensemble learning can further enhance these results.

Key factors in developing a robust and high-performing DL model include the number of images, class imbalance, demographics, and clinical variables such as race, sex, and age. Another critical factor is the accuracy of the manual annotations made by clinicians, which are used to train and test the models. Reliable disease annotations require input from multiple clinical experts. Retrospective or prospective

clinical or laboratory exams are also used to confirm these annotations.

Despite the exceptional performance of DL methods in detecting retinal diseases, their clinical applicability is limited by their lack of interpretability and explainability, which makes them less trustworthy for automatic clinical decision-making. Recent research efforts are focused on developing interpretation techniques, such as class activation mapping, which highlight the parts of the image that most contribute to the model's prediction.

There have been some recent developments in increasing interpretability and explainability. A web application called GlaucoCare was developed, which provides a user-friendly interface for testing fundus images [302]. Users can upload images, and the system generates segmentation masks, heatmaps, and CDR values, along with the glaucoma prediction. The application aims to support clinicians by providing a second opinion and improving diagnostic accuracy [302].

Developing such models that provide interpretable and explainable results can help ophthalmologists understand the underlying reasoning behind AI-generated predictions. This can foster trust in AI systems and facilitate the integration of AI-generated insights into clinical decision-making.

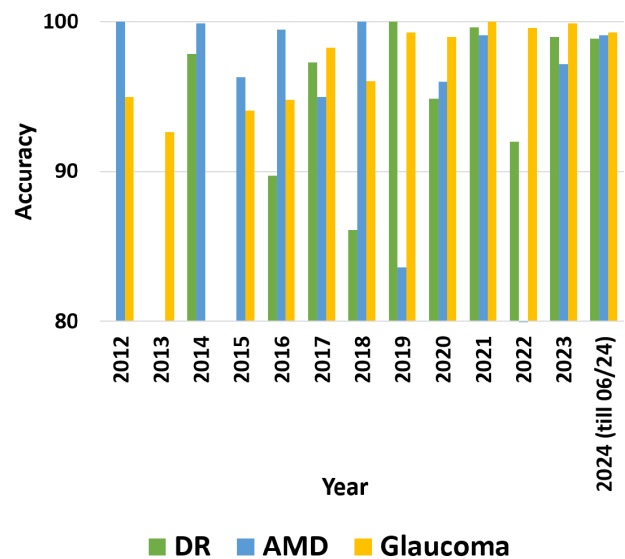


FIGURE 5. Bar graph displaying the maximum accuracy for each OD for each year between 2012 and 2024 (till June 2024).

Figure 5 shows the highest accuracy reported each year from 2012–2024 for studies detecting AMD, DR, and glaucoma. Accuracy levels have improved over time, starting in the 70-90% range in early years but reaching 95-100% by 2023. This demonstrates that ML/DL techniques have become more effective, likely due to larger datasets, better model architectures, and optimisation of imaging and pre-processing. Accuracy levels vary across diseases. This may indicate remaining challenges in detecting certain lesions or features associated with these diseases.

Out of all the studies we reviewed, Sivapriya et al. [316] achieved the highest accuracy of 98.88% in detecting DR on the MESSIDOR-2 dataset in 2024. They proposed a novel DL method, ResEAD2Net, for automatically segmenting the blood vessels and classifying DR [316]. The primary goal of this novel approach is to identify pathological changes in the retinal vascular structure indicative of DR. The proposed system includes three stages: pre-processing, vessel segmentation, and classification. Initially, the input images are processed to remove noise, followed by green channel extraction and enhancement using CLAHE and gamma correction. Segmenting the retinal vascular structure is crucial for detecting various stages of DR by identifying microaneurysms, hemorrhages, and exudates. The U-Net architecture is used to develop the segmentation model. The U-Net's contracting path features four consecutive downsampling and upsampling layers with skip connections [316]. However, this four-time downsampling may overlook information on small blood vessels. To address this, the study introduced ResEAD2Net, which reduces the number of downsampling and upsampling layers to two and incorporates two contracting and expansion paths in the network. This design retains detailed semantic information effectively.

Liu et al. [317] achieved the highest accuracy of 99.1% in classifying AMD on the Ichallenge dataset in 2024. A general self-supervised machine learning framework is proposed to handle diverse fundus diseases from unlabeled fundus images. This method achieved an AUC that surpasses existing supervised approaches by 15.7%. Additionally, the model adapts well to various datasets from different regions, races, and heterogeneous image sources or qualities from multiple cameras or devices.

Das et al. [113] achieved the highest accuracy of 99.3% in detecting glaucoma in 2024. They proposed a lightweight multi-scale CNN architecture, CDAM-Net, which was evaluated on a private dataset of 1426 fundus images, of which 837 were glaucoma and 589 were normal. Additionally, an attention module, channel shuffle dual attention (CSDA), was introduced, consisting of a channel attention block, a spatial attention block, and a channel shuffle unit. This module focuses on significant regions in the fundus images, thereby extracting class-specific features. The CDAM-Net primarily comprises multi-scale feature representation blocks, which enable the extraction of multi-scale features from fundus images. Each MFR block is followed by a CSDA module, further enriching the feature representation. The results indicate that CDAM-Net achieves promising classification performance compared to existing techniques [113].

In 2023, the highest accuracy of 99% in detecting DR was achieved by Abramovich et al. [318]. They proposed a DL model, FundusQ-Net, which obtained an accuracy of 99% on the DRIMDB database. FundusQ-Net utilises in-domain pre-training and semi-supervised learning to perform

the regression task of fundus image quality estimation. The model's high performance has been demonstrated on both local and external test sets.

Gu et al. [319] proposed the development of an intelligent model for classifying the severity of DR using fundus images. This model aimed to detect all five stages of DR, from no DR to proliferative DR, by integrating a Vision Transformer and residual attention mechanisms [319]. The proposed model consisted of two main components: the Feature Extraction Block (FEB) and the Grading Prediction Block (GPB). The FEB utilised a Vision Transformer to capture fine-grained attention on retinal haemorrhage and exudate areas, while the GPB employed residual attention to effectively identify spatial regions occupied by different classes of DR lesions. This combination allowed the model to classify the severity of DR with high accuracy [319].

The study conducted comprehensive experiments on the DDR dataset, demonstrating that the proposed model achieved superior performance compared to benchmark methods. The model was trained and tested on two public datasets: the DDR dataset, which included 13,673 fundus images from various hospitals in China, and the IDRiD dataset, which contained typical DR images representing the Indian population. It developed a Vision Transformer-based model for extracting fundus image features and integrated a residual attention module to enhance classification accuracy by focusing on spatial regions specific to each class [319]. It achieved state-of-the-art performance in DR classification tasks, particularly in distinguishing between different severity levels of DR. Despite its success, the study acknowledged limitations due to the imbalance and a limited number of labelled samples in the datasets [319].

Another study that used a transformer-based model to achieve high accuracy was a study by Xu et al. [320], which achieved an accuracy of 97.2% in detecting AMD in 2023. This study introduced DeepDrAMD, a hierarchical vision transformer-based deep learning model that incorporates data augmentation techniques and the SwinTransformer to detect AMD and distinguish between its subtypes using fundus images. DeepDrAMD excelled in distinguishing wet AMD subtypes, achieving an AUC of 0.9936. Comparative analysis demonstrated that DeepDrAMD outperformed conventional deep learning models and expert-level diagnosis.

Adak et al. [321] proposed a study in 2023, that focused on leveraging the capabilities of transformer networks to capture crucial features in retinal images to improve the performance of DR severity detection models. The study employed and fine-tuned various transformer-based models, including Vision Transformer, Class-attention in image Transformers (CaiT), Data-efficient image Transformer (DeiT), and Bidirectional Encoder representations for image Transformer (BEiT). These models were used individually and in ensembles to predict the severity stages of DR from fundus images. The researchers utilised the publicly

available APTOS-2019 limited vision detection dataset for their experiments [321].

The proposed solution architecture involved preprocessing raw fundus images, applying data augmentation techniques, and using transformer networks to extract features and classify the images into five severity stages: negative, mild, moderate, severe, and proliferative. The ensemble models showed promising results, achieving high accuracy and outperforming traditional ML and CNN-based methods [321]. Additionally, the study explored the impact of hyper-parameters, conducted ablation studies to assess the importance of individual transformers. In their study, ViT, DeiT, BEiT, CaiT achieved accuracies of 82.21%, 85.65%, 86.74%, and 86.91% respectively [321].

Haider et al. [322] achieved the highest accuracy of 99.91% in detecting glaucoma on the REFUGE dataset in 2023. Segmentation of optic disc and optic cup is commonly done for automated glaucoma screening. Their proposed model, FBSS-Net utilises both internal and external feature blending to enhance overall segmentation performance [322]. Internal feature blending empowers features at intervals, while external feature blending improves the network's learning capabilities, leading to better performance.

Pham et al. [323] achieved the highest accuracy of 58.2% for detecting AMD by using MuMO-GAN on a private dataset in 2022. In the study, generative adversarial networks (GANs) were utilised with additional drusen masks to preserve pathological information. The dataset comprised 8,196 fundus images from 1,263 AMD patients. The proposed GAN-based model, named Multi-Modal GAN (MuMo-GAN), was trained to generate synthetic predicted future fundus images. The DL model demonstrates that the inclusion of drusen masks aids in learning AMD progression. The model effectively generates future fundus images with accurate pathological features, accurately depicting drusen development over time. Both qualitative and quantitative experiments indicate that the model is more efficient in monitoring AMD progression compared to other studies [323].

Elangovan and Nath [324] achieved the highest accuracy of 99.6% for detecting glaucoma on the LAG-R dataset in 2022. The study developed a deep ensemble model using the stacking ensemble learning technique to achieve optimal performance in classifying glaucomatous and normal images. Thirteen pre-trained models, including Alexnet, Googlenet, VGG-16, VGG-19, Squeezenet, Resnet-18, Resnet-50, Resnet-101, Efficientnet-b0, Mobilenet-v2, Densenet-201, Inception-v3, and Xception, were implemented and compared in 65 different configurations, combining 13 CNN architectures with five different classification approaches. A two-stage ensemble selection technique was proposed to identify the optimal configurations, which were then pooled using a probability averaging technique. The final classification was performed using an SVM classifier.

Jabbar et al. [325] proposed a transfer learning-based model in 2022, based on a pre-trained VGGNet architecture,

modified to suit the needs of DR detection. The model comprises 16 layers with specific configurations designed for this task. Training the model involved fine-tuning hyperparameters, including the learning rate, batch size, and epochs, using Adam's optimisation function. The model was evaluated using the EyePACS dataset, split into training (80%) and testing (20%) sets.

The results demonstrated that the proposed VGGNet model achieved an accuracy of 96.6%, surpassing other models like ResNet, GoogLeNet, and AlexNet [325]. The model's robustness and high performance in detecting and classifying DR at various severity levels were evident. The authors conclude that their framework effectively enhances DR detection using transfer learning and data augmentation. They suggest that future work could involve integrating hand-engineered features with CNNs to further improve classification accuracy [325]. The study presents significant contributions to the field by developing a VGGNet-based model for DR detection, employing effective preprocessing and data augmentation techniques, and achieving high classification accuracy on a large dataset.

Another study by Chen et al. [326] used ViTs on fundus images to detect glaucoma in 2022. The study achieved a specificity and sensitivity of 91.2% and 92.3% on the ORIGA dataset and a specificity and sensitivity of 95.7%, and 94.1% on the RIM-ONEv3 dataset [326].

Shinde [105] achieved the highest accuracy of 100% in glaucoma detection in 2021. The system was developed utilising image processing, DL, and ML techniques. LeNet architecture is employed for input image validation, while the brightest spot algorithm is used for region of interest (ROI) detection. Optic disc and optic cup segmentation are performed using the U-Net architecture, followed by classification using SVM, Neural Network, and Adaboost classifiers.

Sun et al. [327] proposed a model in 2021 to address the challenges of DR grading and lesion discovery using a novel lesion-aware transformer (LAT) model. The authors proposed a unified deep model that jointly performed DR grading and lesion discovery using an encoder-decoder structure, incorporating a pixel relation-based encoder and a lesion filter-based decoder [327]. This model was the first to formulate lesion discovery as a weakly supervised lesion localization problem via a transformer decoder, learning lesion filters with only image-level labels. The study introduced two mechanisms for effective lesion filter learning: lesion region importance and lesion region diversity [327].

Extensive experiments on three challenging benchmarks, including Messidor-1, Messidor-2, and EyePACS, demonstrated that the proposed LAT model outperformed state-of-the-art methods in DR grading and lesion discovery [327]. The LAT model effectively captured the correlation between pixels for robust feature learning, evaluated the importance of different lesion regions, and ensured diversity in lesion-aware features to cover various lesion types [327]. The

study highlighted the effectiveness of the pixel relation-based encoder in adapting to pixel appearance variations and the lesion filter-based decoder in identifying diverse lesion regions. The proposed mechanisms for learning lesion region importance and diversity further improved the model's performance, making it a significant advancement in automated DR diagnosis. The study concluded that the LAT model, with its encoder-decoder structure and classification module, provided an effective solution for joint DR grading and lesion discovery, setting a new benchmark in the field [327].

In 2021, Wu et al. [56] proposed the application of transformers, specifically Vision Transformers, for DR grade recognition, contrasting it with the traditionally dominant CNNs. In their study, transformers utilised multi-head attention mechanisms to capture long-range contextual relations between image pixels, as opposed to the convolution layers used in CNNs [56]. The study proposed a method where fundus images were subdivided into non-overlapping patches, flattened into sequences, and processed through linear and positional embedding. These sequences were then input into multiple multi-head attention layers to generate the final representation, which was classified using a softmax layer [56].

The study aimed to demonstrate the suitability of the pure attention mechanism for DR grade recognition and to establish that transformers could replace traditional CNNs in this task. A Vision Transformer-based method was proposed. Fundus images were divided into patches, converted into sequences through flattening and embedding, and processed through multi-head attention layers [56]. The first token sequence was classified using a softmax layer. The method was tested on a dataset of fundus images with varying resolutions, achieving impressive performance: an accuracy of 91.4%, specificity of 97.7%, sensitivity of 92.6%, and an AUC of 0.986. Comparative experiments indicated that the proposed Vision Transformer model was competitive with current methods and highlighted its promise for DR grade recognition [56].

Balasubramanian [135] achieved the highest accuracy of 99% in detecting glaucoma using a dataset of 1155 fundus images. In 2020, a 25-layer convolutional neural network (CNN) was developed and trained to efficiently extract highly robust features from retinal fundus images. The proposed DL approach effectively detects and grades glaucoma from fundus images, demonstrating high accuracy and robustness.

Prabhu et al. [120] achieved the highest accuracy of 100% using RF and ANN to detect DR on a private dataset in 2019. This paper proposed an automatic DR detection system based on the identification of bright lesions on the retina, a key symptom of DR. The optic disc is removed from the fundus image because its brightness is similar to that of the bright lesions. Exudates, which are indicative of DR, are extracted from the image and various features are obtained. A feature-based hierarchical classification is performed to detect different stages of the disease. This method mirrors the

logical steps followed by ophthalmologists, ensuring more accurate classification results.

Floriano et al. [126] achieved the highest accuracy of 83.6% in AMD detection in 2019. This study introduces an approach that combines mathematical morphology, and an SVM. The method presented in this study offers a powerful tool for the non-invasive pre-diagnosis of AMD by detecting drusen in fundus images.

Rehman et al. [107] achieved the highest accuracy of 99.2% in glaucoma detection in 2019. The study used SVM on the DRIONS-DB dataset with 110 fundus images.

Lin et al. [328] achieved the highest accuracy of 86.1% in DR detection by using CNN on the EyePacs dataset in 2018. The study used entropy images, which quantify the amount of information in the fundus photographs, and significantly improved the detection accuracy, sensitivity, and specificity of referable DR in a deep learning-based system. Entropy imaging efficiently enhances the feature maps generated by the CNN, making it a valuable tool for increasing the performance of automated DR detection systems.

Singh et al. [112] achieved the highest accuracy of 94.8% in detecting glaucoma using KNN on the VERC dataset in 2016. This study presents a method for detecting glaucoma using wavelet feature extraction from segmented optic disc images, followed by optimized genetic feature selection and various learning algorithms. The focus on the segmented optic disc with blood vessels removed enhances the accuracy of glaucoma identification, achieving a high accuracy rate.

Kumari et al. [329] achieved the highest accuracy of 96.32% in detecting AMD in 2015. The study proposed an automated method for detecting and segmenting drusen using retinal fundus images. The method begins with gradient-based segmentation to accurately identify the true edges of drusen. Following this, connected component labelling is employed to remove suspicious pixels from the drusen region. The final step involves edge linking, which connects all labelled pixels into a coherent boundary, resulting in a meaningful segmentation of the drusen. In addition to detecting drusen, the method quantifies them to grade the severity of AMD. The detected drusen are categorized into small, intermediate, and large.

Akram et al. [330] achieved the highest accuracy of 97.89% in detecting DR on the STARE dataset in 2014. This paper proposes a system for detecting retinal lesions using a novel hybrid classifier. The system comprises several stages: pre-processing, extraction of candidate lesions, feature set formulation, and classification. During pre-processing, the system removes background pixels and extracts the blood vessels and optic discs from the digital retinal image. In the candidate lesion detection phase, filter banks are used to identify all regions that might contain lesions. A feature set is formulated for each potential candidate region using various descriptors such as shape, intensity, and statistical properties. These features assist in the classification process. This paper extends the m-Mediods-based modelling approach

by combining it with a Gaussian Mixture Model to form an ensemble, creating a hybrid classifier that enhances classification accuracy.

Noronha et al. [331] achieved the highest accuracy of 92.65% in detecting glaucoma on the KMC database in 2013. Their proposed system classifies images into three categories: normal, mild glaucoma, and moderate/severe glaucoma. The methodology involves extracting 3rd order HOS cumulant features from the transformed fundus images. These features are then subjected to linear discriminant analysis (LDA) to reduce their number while retaining clinically significant information. The reduced features are fed into SVM and NB classifiers to automate the detection process. The system was validated using a dataset of 272 fundus images, which included 100 normal images, 72 images with mild glaucoma, and 100 images with moderate/severe glaucoma. The validation employed a ten-fold cross-validation method to ensure robustness. For the three-class classification task, the system achieved an average accuracy of 92.65%, a sensitivity of 100%, and a specificity of 92% using the NB classifier.

Mookiah et al. [332] achieved the highest accuracy of 95% in the detection of glaucoma on the KMC dataset in 2012.

Hijazi et al. [333] achieved the highest accuracy of 100% in detecting AMD on the ARIA dataset in 2012. This paper proposes and compares two data mining techniques to support the automated screening for AMD. The first technique employs spatial histograms, which preserve both the colour and spatial information of the images for representation. A case-based reasoning (CBR) classification technique is then applied to these spatial histograms. The second technique is based on a hierarchical decomposition of the image set, generating a tree representation. A weighted frequent sub-graph mining technique is applied to this representation to identify sub-trees that frequently occur across the dataset. These identified sub-trees are encoded as feature vectors, to which standard classification techniques can be applied. By comparing these two methods, the paper aims to find effective automated screening approaches that reduce the need for manual inspection and improve the efficiency of early AMD detection.

Figure 6 displays the highest specificity reported each year for detecting the three diseases. Like accuracy, specificity has increased to high levels of over 90% for all diseases. AMD and DR detection tend to have slightly lower specificity than glaucoma detection. Since specificity relates to the ability to correctly identify negative cases, this suggests ML/DL models may have more difficulty excluding these diseases compared to glaucoma. Data imbalances and subtle imaging features could contribute to this discrepancy.

Among the studies we reviewed, El-Khalek et al. [295] achieved the highest specificity of 97.89% in detecting AMD on a private dataset in 2024. Das et al. [113] achieved the highest specificity of 100% in detecting glaucoma using MFR-Net and CDAM-Net on a private dataset of 1426 fundus images in 2024.

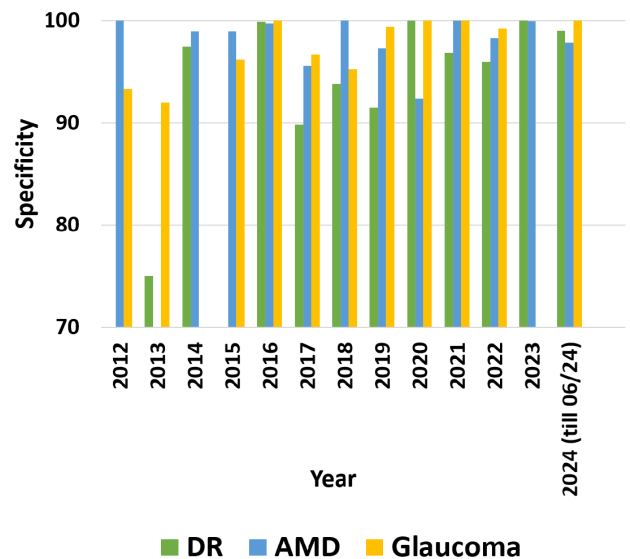


FIGURE 6. Bar graph displaying the maximum specificity attained annually for each of the ODs from 2012 to 2024 (till June 2024).

Sivapriya et al. [316] achieved the highest specificity of 99.01% in the detection of DR on the STARE dataset in 2024. The study proposed ResEAD2Net, for automatically segmenting the blood vessels and classifying DR [316]. The primary goal of this novel approach is to identify pathological changes in the retinal vascular structure indicative of DR.

In 2023, Song et al. [114] achieved the highest specificity of 94%. A generative adversarial network (GAN) model was trained using pairs of CF and FAF images to generate synthetic FAF images. The quality of these synthesized FAF images was assessed using standard generation metrics. The clinical effectiveness of the generated FAF images for AMD classification was evaluated by measuring the area under the curve (AUC) on the LabelMe dataset. When combined with CF images, the generated FAF images improved AMD specificity from 93.2% to 94%.

Abramovich et al. [318] achieved the highest specificity of 100% in detecting DR in 2023. They achieved this result using their proposed a DL model, FundusQ-Net, on the DRIMDB database.

Mahmoud et al. [121] achieved the highest specificity of 96.88% in detecting DR on the CHASE dataset in 2021. In this study, a hybrid inductive machine learning algorithm (HIMLA) is proposed as an automated diagnostic tool for DR detection. HIMLA processes and classifies coloured fundus images into healthy (no retinopathy) or unhealthy (presence of DR) categories by accurately identifying the appropriate medical cases of DR. The algorithm involves four main stages: pre-processing, segmentation, feature extraction, and classification. In the pre-processing stage, coloured fundus images are normalised to a specific brightness level to enhance their quality. During segmentation, the processed

images are encoded and decoded to isolate relevant regions, improving image clarity. Feature extraction and classification are performed using multiple instance learning (MIL), which aids in identifying and categorising the images based on the presence of DR. The proposed method was evaluated on the CHASE datasets, achieving an accuracy of 96.62%, sensitivity of 95.31%, and specificity of 96.88%.

Zapata et al. [334] achieved the highest specificity of 92.4% in detecting AMD in 2020. Their study developed five algorithms and evaluated them in the Optretina dataset.

Jiang et al. [335] achieved the highest specificity of 91.5% in detecting DR in 2019. This paper presents an automatic image-level DR detection system that leverages multiple well-trained DL models. To enhance the system's performance, several DL models are integrated using the Adaboost algorithm, which helps to reduce the bias inherent in individual models. To provide clear explanations for the DR detection results, the system generates weighted class activation maps (CAMs). These maps highlight the suspected positions of lesions, offering valuable insights into the detection process [335]. In the pre-processing stage, eight different image transformation techniques are applied to augment the diversity of fundus images. This augmentation step helps improve the robustness and performance of the detection models. Experimental results demonstrate that the proposed method exhibits stronger robustness and superior performance compared to individual DL models. By combining multiple models and employing advanced techniques like Adaboost and image augmentation, this system achieves more accurate and reliable DR detection.

Lin et al. [328] achieved the highest specificity of 93.81% in DR detection by using CNN on the EyePacs dataset in 2018.

Maheshwari et al. [123] achieved the highest specificity of 96.7% in detecting glaucoma in 2017. This paper presents a methodology for the automated detection of glaucoma, employing the empirical wavelet transform (EWT). The EWT is utilised to decompose the fundus images, and correntropy features are extracted from the decomposed EWT components. These extracted features are then ranked using the t-value feature selection algorithm, ensuring that the most significant features are chosen for classification. The classification of normal and glaucoma images is performed using the least-squares support vector machine (LS-SVM) classifier. The LS-SVM is tested with various kernels, including the radial basis function, Morlet wavelet, and Mexican-hat wavelet kernels, to determine the most effective approach. The proposed method achieves a classification accuracy of 98.33% with threefold cross-validation and 96.67% with tenfold cross-validation. These results highlight the effectiveness of the EWT-based feature extraction and the LS-SVM classifier in accurately detecting glaucoma from fundus images, offering a promising, low-cost alternative to traditional scanning methods [123].

Imani and Pourreza [336] achieved the highest specificity of 99.93% in detecting DR in 2016. This paper introduces an automatic method for the detection of retinal exudates, featuring an approach that utilises the Morphological Component Analysis (MCA) algorithm to distinguish lesions from normal retinal structures, thereby facilitating the detection process. In the initial stage, the MCA algorithm, equipped with appropriate dictionaries, separates blood vessels from lesions. Following this, the lesion segments of the retinal images are processed to detect exudate regions. Dynamic thresholding and mathematical morphology techniques are then applied to create the final exudate map. The performance of the proposed method was evaluated using three publicly available datasets: DiaretDB, HEI-MED, and e-optha. The method achieved Area Under the Curve (AUC) scores of 0.961, 0.948, and 0.937 on these datasets, respectively, surpassing most state-of-the-art methods. These results underscore the effectiveness of the MCA-based approach in accurately detecting retinal exudates, contributing to the early diagnosis and treatment of diabetic retinopathy.

Mittal et al. [161] achieved the highest specificity of 99% in detecting AMD in 2015. Their proposed method begins with gradient-based segmentation to identify the true edges of the drusen. This is followed by connected component labelling, which removes suspicious pixels from the drusen region, isolating relevant features. The final step involves edge linking to connect all labelled pixels into coherent boundaries, forming a meaningful segmentation of the drusen. The proposed method significantly outperforms existing techniques, achieving an accuracy of 96.17%, sensitivity of 89.81%, and specificity of 99% on two publicly available retinal image databases. Furthermore, to assess the severity of AMD, the detected drusen are quantified into three categories: small, intermediate, and large. The method achieves classification accuracies of 88.46% for small drusen, 98.55% for intermediate drusen, and 88.37% for large drusen. This automated approach enhances the accuracy and efficiency of drusen detection and provides a reliable means of grading the severity of AMD.

Akram et al. [330] achieved the highest specificity of 97.43% in detecting DR on the STARE dataset in 2014. This paper proposes a system for detecting retinal lesions using a novel hybrid classifier.

Noronha et al. [331] achieved the highest specificity of 92% in detecting glaucoma on the KMC database in 2013. Their proposed system classifies images into three categories: normal, mild glaucoma, and moderate/severe glaucoma.

Zheng et al. [9] achieved the highest specificity of 100% in detecting AMD in 2012. This study aimed to describe and evaluate an automated grading system for AMD using colour fundus photography. An automated "disease/no disease" grading system for AMD was developed using image-mining techniques. The process began with image pre-processing to normalise the colour and correct the nonuniform illumination of the fundus images. This step also defined a region of

interest and identified and removed pixels belonging to retinal vessels. To represent images for the prediction task, a graph-based image representation using quadrees was adopted. Following this, a graph-mining technique was applied to the generated graphs to extract relevant features, in the form of frequent subgraphs, from images of both AMD patients and healthy volunteers. Features from the training data were then used to train a classifier generator, which was subsequently employed to classify new, unseen images. The algorithm was evaluated using two publicly available fundus-image datasets comprising a total of 258 images (160 AMD and 98 normal). Ten-fold cross-validation was utilised to assess performance. The experiments yielded a best specificity of 100%, a best sensitivity of 99.4%, and an overall accuracy of 99.6%.

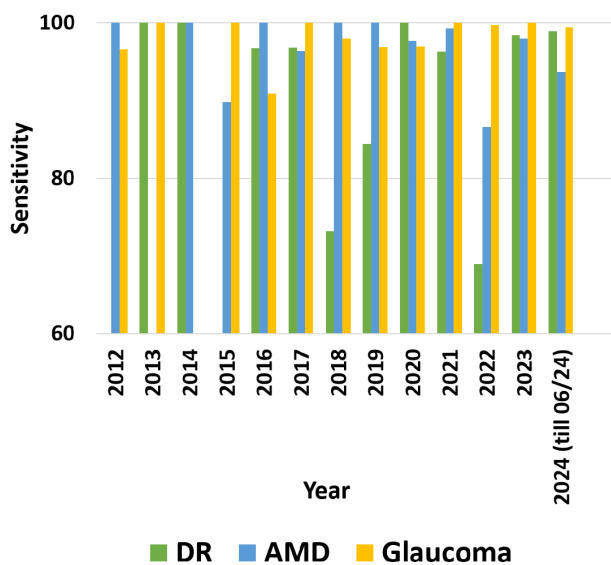


FIGURE 7. Bar graph displaying the maximum sensitivity attained annually for each of the different ODs from 2012 to 2024 (till June 2024).

Figure 7 shows sensitivity reported from 2012-2024, reflecting the ability to correctly detect positive disease cases. Sensitivity follows a similar trend to accuracy and specificity, rising from early years. Although the results were slightly lower for some studies in 2022. Overall, the results indicate AI methods have become very proficient at identifying true cases.

Among the reviewed studies, Sivapriya et al. [316] achieved the highest sensitivity of 98.91% in detecting DR using a novel DL method, ResEAD2Net, on the MESSIDOR-2 dataset in 2024.

Xu et al. [320] achieved the highest sensitivity of 96.75% in detecting AMD in 2023. This study introduced DeepDrAMD, a hierarchical vision transformer-based deep learning model that incorporates data augmentation techniques and the SwinTransformer to detect AMD and distinguish between its subtypes using fundus images.

Pham et al. [323] achieved the highest sensitivity of 56% for detecting AMD by using MuMo-GAN on a

private dataset in 2022. In the study, generative adversarial networks (GANs) were utilised with additional drusen masks to preserve pathological information. The dataset comprised 8,196 fundus images from 1,263 AMD patients. The proposed GAN-based model, named Multi-Modal GAN (MuMo-GAN), was trained to generate synthetic predicted future fundus images. The low number of sensitivity in this study indicates that there was a high rate of false positives [323].

Math et al. [138] achieved the highest sensitivity of 96.37% in detecting DR on the Kaggle and DIARET-DB1 database in 2021. This paper proposed a segment-based learning approach for detecting DR that jointly learns classifiers and features from the data, leading to significant improvements in recognising DR images and identifying lesions within them. Specifically, the approach involves adapting a pre-trained CNN to obtain segment-level diabetic retinopathy estimation (DRE). The segment-level results are then integrated to classify diabetic retinopathy images. This end-to-end segment-based learning approach effectively handles the irregular lesions characteristic of diabetic retinopathy. The proposed method was evaluated on the Kaggle dataset and achieved sensitivity and specificity rates of 96.37%. The segment-based learning approach proposed in this paper offers a robust solution for the detection of diabetic retinopathy, leveraging the strengths of pre-trained CNNs and integrated segment-level analysis.

Zapata et al. [334] achieved the highest specificity of 97.7% in detecting AMD on the Optretina dataset in 2020.

Rehman et al. [107] achieved the highest sensitivity of 96.9% in glaucoma detection on the DRIONS-DB dataset in 2019.

Soltani et al. [337] achieved the highest sensitivity of 97.8% in detecting glaucoma in 2018. This study introduces a new Fuzzy Expert System for the early diagnosis of glaucoma. The process begins with pre-treating original ONH images using appropriate filters to remove noise. The Canny detector algorithm is then employed to detect contours within the images. Key parameters are extracted after identifying the elliptical forms of both the optic disc and excavation, using the Randomized Hough Transform. The final stage involves a classification algorithm based on fuzzy logic approaches to determine the condition of the patients. This system is advantageous as it considers both instrumental parameters and risk factors such as age, race, and family history, which are crucial for accurately identifying cases suspected of having glaucoma. The proposed system was tested on a real dataset comprising ophthalmologic images of both normal and glaucomatous cases. By combining advanced image processing techniques with fuzzy logic and considering essential risk factors, the system offers a significant improvement in identifying glaucomatous conditions.

Yang et al. [139] achieved the highest sensitivity of 96.87% in detecting DR using DCNN on the EyePacs dataset in 2017.

Abramoff et al. [32] achieved the highest sensitivity of 96.8% in detecting DR in 2016 on the Messidor-2 database. Their proposed DL-enhanced algorithm demonstrated a sensitivity of 96.8% and a specificity of 87%. There were 6 false negatives out of 874 cases, resulting in a negative predictive value of 99%. Notably, no cases of severe NPDR, PDR, or ME were missed.

Mittal et al. [161] achieved the highest sensitivity of 89.81% in detecting AMD in 2015.

Hijazi et al. [338] achieved the highest sensitivity of 99.5% in detecting AMD in 2014. This paper investigates three alternative approaches to classifying retinal images, distinctively not relying on individual lesion segmentation for feature generation but instead using encodings focused on the entire image. The three different mechanisms for encoding retinal image data considered are time series, tabular, and tree-based representations. The evaluation utilised two publicly available retinal fundus image datasets, specifically in the context of screening for AMD. Statistical significance tests were conducted to assess the performance of these approaches. The results were impressive, with sensitivity, specificity, and accuracy rates all exceeding 99%. Notably, the tree-based approach demonstrated the best performance, achieving a sensitivity of 99.5%.

Tavakoliet al. [339] achieved the highest sensitivity of 100% in detecting DR in 2013. This study presents an algorithm using the Radon transform (RT) and multi-overlapping windows. This method focuses on detecting retinal landmarks and lesions to detect DR effectively. The proposed method begins by detecting and masking the optic nerve head (ONH). In the pre-processing stage, top-hat transformation and averaging filters are applied to remove the background. In the main processing section, the preprocessed image is divided into sub-images. Each sub-image is then segmented, and the vascular tree is masked by applying the RT. After detecting and masking the retinal vessels and ONH, MAs are identified and counted using RT and appropriate thresholding techniques. The algorithm was evaluated on three different retinal image databases: the Mashhad Database with 120 FA fundus images, the Second Local Database from Tehran with 50 FA retinal images, and a subset of the Retinopathy Online Challenge (ROC) database with 22 images. The performance of the automated DR detection method demonstrated a sensitivity and specificity of 94% and 75%, respectively, for the Mashhad database. For the Second Local Database, the method achieved a sensitivity and specificity of 100% and 70%, respectively.

Mookiah et al. [332] achieved the highest sensitivity of 96.6% in the detection of glaucoma on the KMC dataset in 2012. The system automates the identification of normal and glaucoma-affected eyes using features extracted from Higher Order Spectra (HOS) and Discrete Wavelet Transform (DWT). These features are input into an SVM classifier, which is tested with various kernel functions, including linear, polynomial (orders 1, 2, and 3), and Radial Basis

Function (RBF), to determine the best kernel for automated decision-making. In this study, the SVM classifier with a polynomial order 2 kernel function demonstrated the ability to distinguish between glaucoma and normal images with an accuracy of 95%. The system also achieved sensitivity and specificity rates of 93.33% and 96.67%, respectively. Additionally, the paper introduced a novel integrated index called the Glaucoma Risk Index (GRI), which combines HOS and DWT features to detect unknown cases using a single metric. This GRI aims to help clinicians make quicker glaucoma diagnoses during mass screenings of normal and glaucoma images [332].

The proposed automated system offers a cost-effective and efficient solution for glaucoma screening. It could potentially improve early detection and management of the disease while making the screening process accessible to a broader population.

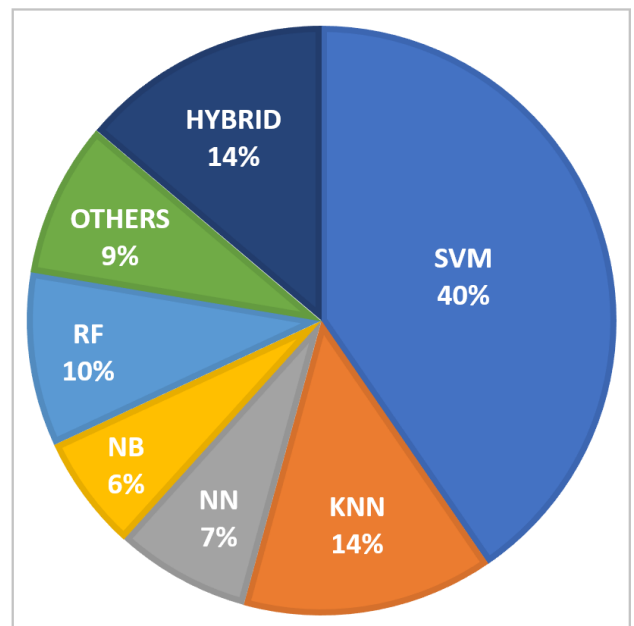


FIGURE 8. Pie chart diagram displaying the percentage of ML models used to automatically identify glaucoma.

The pie chart in Figure 8 represents the distribution of machine learning (ML) algorithms used for glaucoma detection from fundus eye images, with support vector machines (SVM) being most common at 40%, followed by KNN at 14%. Neural networks (NN), random forests (RF), Naive Bayes (NB) make up 23% combined, whereas others contribute 9% and hybrid models are 14% of the total.

The SVM algorithm is used in 40% of the cases, making it the most frequently employed technique for detecting these eye conditions. It is particularly effective in high-dimensional spaces and is well-suited for cases where the number of dimensions exceeds the number of samples. Medical images, including fundus eye images, have high-dimensional data that SVM can handle efficiently [340].

SVM also uses regularisation parameters to control over-fitting, making it robust for small to medium-sized datasets [340]. This is crucial in medical imaging, where the number of labelled images can be limited.

KNN is utilised in 14% of the cases, showing its role as a common technique. KNN is one of the simplest machine learning algorithms. It is easy to understand and implement, which makes it a popular choice for initial exploratory analysis and in situations where interpretability is crucial. This may be the case as KNN can be effective with relatively small datasets, which is common in medical imaging where acquiring large amounts of labelled data can be challenging. Since KNN makes predictions based on the closest data points, it can perform well even with limited training data.

By considering the majority vote of its neighbours, KNN can be resilient to noise in the data. This can help in making robust predictions in medical images that may contain some level of noise or variability. Hybrid methods, which combine multiple algorithms, account for 14% of the usage, indicating a significant reliance on integrated approaches. Other ML approaches collectively make up 9% of the total.

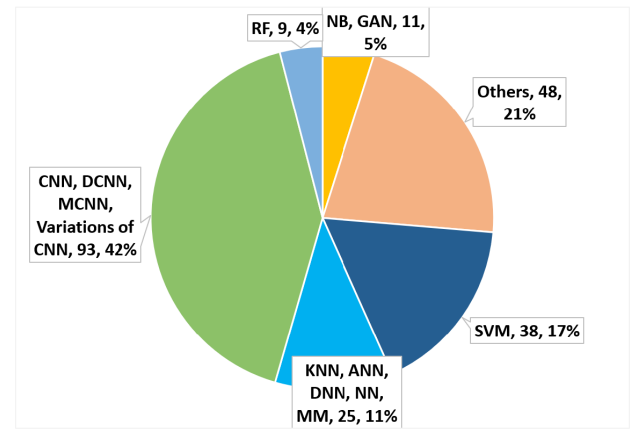


FIGURE 9. Pie chart diagram displaying the percentage contribution (%) of different ML and DL classifiers to the automated detection of glaucoma, DR, and AMD.

Figure 9 summarizes all models used for the three diseases, showing convolutional neural networks (CNN) and their variations are now the most widely used (42%), followed by other models and SVM. The chart clearly shows that CNN and its variations are the predominant choice for the automated detection of glaucoma, DR, and AMD from fundus eye images, followed by a mix of other methods that collectively contribute a significant portion.

This distribution reflects the effectiveness and versatility of CNNs in handling complex image data. CNNs and their variations, such as Deep Convolutional Neural Networks (DCNN) and Multi-Channel Convolutional Neural Networks (MCNN), constitute the largest segment, representing 39% of the classifiers used. This indicates a strong preference for CNN-based methods due to their effectiveness in image processing and feature extraction.

Support Vector Machines contribute 17% to automated detection, highlighting their role as a significant yet less dominant technique compared to CNNs. This segment includes K-Nearest Neighbours (KNN), Artificial Neural Networks (ANN), Deep Neural Networks (DNN), Neural Networks (NN), and other miscellaneous methods (MM), together making up 11% of the classifiers used.

Naive Bayes (NB) and Generative Adversarial Networks (GAN) together contribute 5% to the automated detection. NB is a probabilistic classifier that is simple and efficient. It assumes independence between features, which can be a limitation, but it performs well in certain diagnostic tasks where this assumption holds approximately true. GANs are used for generating synthetic data that can augment training datasets, thereby improving the robustness and accuracy of diagnostic models. They can also help in enhancing image quality and creating realistic variations of fundus images for training purposes.

Random Forest (RF) classifiers account for 4% of the usage, indicating their utility in eye disease detection. RF is an ensemble learning method that constructs multiple decision trees and combines their outputs. It is robust to overfitting and can handle a large number of features. RFs are useful in medical imaging for their ability to provide important scores for different features, helping in the interpretability of the diagnostic process. Other ML and DL (21%) classifiers include a variety of other techniques that are applied to detect eye diseases from fundus images. These methods include decision trees, logistic regression, ensemble methods, etc.

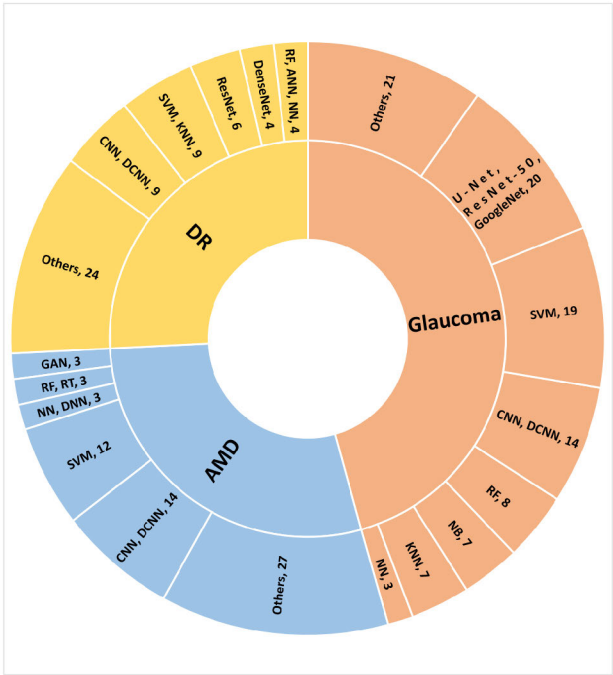


FIGURE 10. A sunburst graphic showing the prevalence of ML and DL procedures for AMD, DR, and glaucoma, along with their most common classifiers.

The sunburst chart in Figure 10 shows the distribution of various machine learning (ML) and deep learning (DL) algorithms used in the automated detection of three specific eye diseases: glaucoma, diabetic retinopathy (DR), and age-related macular degeneration (AMD). Each segment represents a different algorithm and its contribution to detecting these conditions.

Glaucoma:

SVM (19): Support Vector Machines are extensively used for glaucoma detection, accounting for a significant portion.

CNN, DCNN (14): Convolutional Neural Networks and Deep Convolutional Neural Networks are also widely used.

U-Net, ResNet-50, GoogleNet (20): Effectively, variations of CNN architectures, such as ResNet and GoogleNet, are employed.

RF (8): Random Forest is another method used for glaucoma detection.

NB (7): Naive Bayes classifiers contribute to the diagnostic process.

KNN (7): K-Nearest Neighbours are used as well.

NN (3): General Neural Networks are applied in some cases.

Others (21): This category includes various other methods.

Diabetic Retinopathy (DR):

CNN, DCNN (9): CNNs and their deep variations are predominantly used.

ResNet (6): ResNet, a type of CNN architecture, is also employed.

DenseNet (4): Another variation of CNNs used for DR detection.

RF, ANN, NN (4): Random Forest, Artificial Neural Networks, and general Neural Networks contribute to the process.

SVM, KNN (9): Support Vector Machines and K-Nearest Neighbors are used.

Others (24): Various other methods are included in this category.

Age-related Macular Degeneration (AMD):

CNN, DCNN (14): Convolutional Neural Networks and their deep variations are significantly employed.

SVM (12): Support Vector Machines are used.

NN, DNN (3): Neural Networks and Deep Neural Networks are also utilised.

RF, RT (3): Random Forest and Regression Trees are used.

GAN (3): Generative Adversarial Networks are employed in some cases.

Others (27): A large variety of other methods are used.

The reason why most studies that used ML used SVMs is that they are effective for classification in high-dimensional spaces, robust to overfitting, and suitable for small datasets [279]. Most DL studies used CNN as it is excellent at extracting spatial hierarchies of features from images, widely used in image-processing tasks [102]. DCNNs are deeper versions of CNNs capable of capturing more complex patterns. MCNN (Multi-Channel Convolutional Neural Networks): These process different

aspects of input images simultaneously, improving feature capture.

The chart highlights the diversity of ML and DL techniques applied to detect glaucoma, DR, and AMD from fundus images. CNN-based methods dominate due to their effectiveness in image processing. At the same time, other techniques like SVM, RF, and various neural network architectures also play significant roles in the automated detection process. It further indicates that CNNs account for the majority of deep learning approaches. This shift from conventional ML to deep learning reflects the power of CNNs for medical imaging tasks.

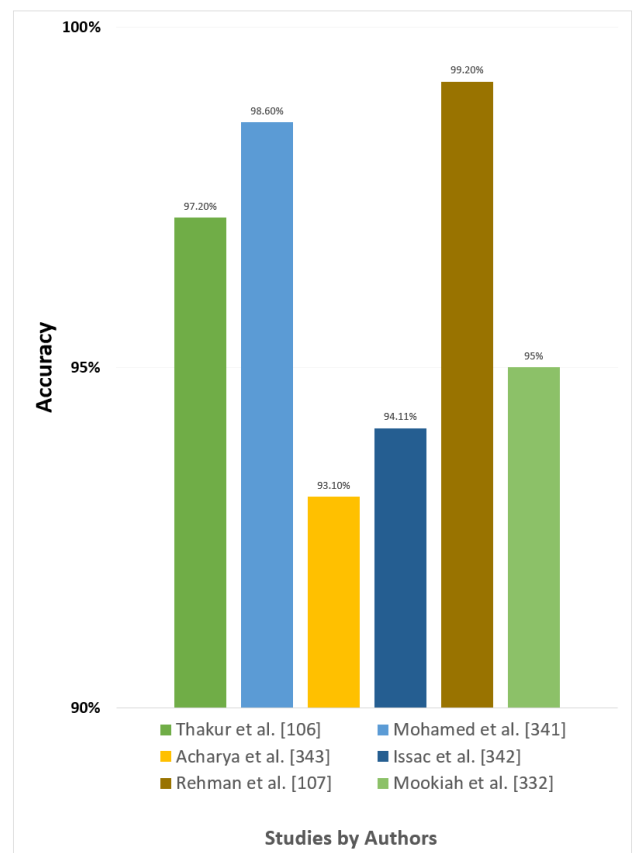


FIGURE 11. Bar graph illustrating the maximum accuracy in automated glaucoma detection using SVM classifiers that different authors have achieved.

The bar graph in Figure 11 illustrates the maximum accuracy achieved in automated glaucoma detection using SVM classifiers by different authors. Each bar represents a different study, with the height of the bar indicating the reported accuracy. The graph highlights the effectiveness of SVM classifiers in automated glaucoma detection, with reported accuracies ranging from 93.10% to 99.20%.

Out of these studies that used SVM as a classifier to detect glaucoma, Rehman et al. [107] achieved the highest accuracy of 99.20% on the DRIONS-DB database. They presented a multi-parametric optic disk detection and localisation method

for retinal fundus images. The method utilised region-based statistical and textural features to accurately identify the optic disc. Highly discriminative features are selected based on the mutual information criterion. The study then conducts a comparative analysis of four benchmark classifiers: SVM, RF, AdaBoost, and RusBoost. SVM achieved an accuracy of 99.20%, a specificity of 99.30%, and a sensitivity of 96.9% in their study.

Mohamed et al. [341] achieved the second-highest accuracy of 98.60% among the authors who used SVM to classify glaucoma. The proposed method was tested on the RIM-One database. This paper proposed a novel approach to developing an automatic glaucoma screening system based on superpixel classification using high-quality input images. Initially, input images undergo pre-processing to remove noise and correct illumination using an anisotropic diffusion filter and illumination correction methods. The processed images are then divided into superpixels using the Simple Linear Iterative Clustering (SLIC) approach. Features based on histogram data and textural information are extracted from each superpixel using the statistical pixel-level (SPL) method. These prominent features are then fed into a Support Vector Machine (SVM) classifier, which classifies each superpixel into categories such as optic disc, optic cup, blood vessel, and background regions. The SVM classifier also determines the boundaries of the optic disc and optic cup. The segmented optic disc and optic cup are subsequently used to determine the presence of glaucoma by measuring the cup-to-disc ratio (CDR). This method effectively combines preprocessing, feature extraction, and classification to provide a comprehensive analysis of the fundus images.

Thakur et al. [106] achieved the third highest accuracy of 97.20% using SVM to classify glaucoma on the DRISHTI-GS and RIM-ONE datasets. This paper introduces a new approach that derives reduced hybrid features from both structural and non-structural aspects to classify retinal fundus images. The structural features include the Disc Damage Likelihood Scale (DDLS) and the Cup-to-Disc Ratio (CDR), while the non-structural features encompass the Grey Level Run Length Matrix (GLRM), Grey Level Co-occurrence Matrix (GLCM), First Order Statistical (FoS) features, Higher Order Spectra (HOS), Higher Order Cumulant (HOC), and Wavelets. The methodology involved extracting these features and using them to train and evaluate various ML classifiers, including SVM, KNN, RF, NB, NN. SVM achieved an accuracy of 97.20%, a specificity of 96%, and a sensitivity of 97% in their study.

Mookiah et al. [332] achieved an accuracy of 95% using SVM to classify glaucoma on the KMC dataset with 60 images (30 normal and 30 glaucoma). The system identified normal and glaucoma classes through Higher Order Spectra (HOS) and Discrete Wavelet Transform (DWT) features, which are fed into a Support Vector Machine (SVM) classifier with various kernel functions (linear, polynomial

order 1, 2, 3, and Radial Basis Function). The SVM classifier with a polynomial order 2 kernel achieved an accuracy of 95%, with sensitivity and specificity of 93.33% and 96.67%, respectively

Issac et al. [342] achieved an accuracy of 94.11% using SVM to classify glaucoma. The fundus images used in this study were sourced from the Venu Eye Research Centre in New Delhi, India. The study involved 67 images from patients aged 18 to 75, comprising 35 normal images and 32 glaucoma images, all labelled by doctors. They employed an adaptive threshold using local features from the fundus image, making it resilient to image quality and noise, thus enhancing its applicability. Experimental results demonstrated that these features are more significant than the statistical or textural features used in previous studies. The proposed method achieves an accuracy of 94.11% and a sensitivity of 100%.

Acharya et al. [343] achieved an accuracy of 93.10% using SVM to classify glaucoma on the KMC dataset. They used 510 fundus images categorised into normal (266), mild (72), moderate (86), and severe (86) glaucoma classes.

They introduced an automated glaucoma detection method using various features extracted from the Gabor transform applied to fundus images. Features such as mean, variance, skewness, kurtosis, energy, and Shannon, Renyi, and Kapoor entropies were extracted from the Gabor transform coefficients. These features were then subjected to principal component analysis (PCA) to reduce dimensionality. Various ranking methods, including the Bhattacharyya space algorithm, t-test, Wilcoxon test, Receiver Operating Characteristic (ROC) curve, and entropy, were used to rank the features. The t-test ranking method achieved the highest performance, with an average accuracy of 93.10%, sensitivity of 89.75%, and specificity of 96.20% using 23 features with an SVM classifier.

SVM achieved high accuracy in glaucoma detection in these studies for several reasons. Glaucoma detection often involves analysing high-dimensional data, such as pixel intensity values and texture features from fundus images. SVMs are particularly effective in such high-dimensional spaces because they find the optimal hyperplane that separates the different classes (healthy vs. glaucomatous eyes) with maximum margin.

Some of the datasets used in the studies by these authors here are relatively small. SVMs are effective with smaller datasets because they focus on the support vectors (the most critical data points) rather than the entire dataset. SVMs use regularisation techniques to prevent overfitting, which is crucial when dealing with medical image data where the number of features can be very high relative to the number of samples. This robustness ensures that the SVM model generalises well to new, unseen data.

SVMs can effectively handle feature selection and dimensionality reduction, either inherently through the use of certain kernel functions or in combination with pre-processing techniques. This helps focus on the most relevant features for

glaucoma detection, improving accuracy. These properties make SVMs highly suitable for glaucoma detection from fundus images, resulting in high accuracy and reliable performance from different studies that employed them for glaucoma detection.

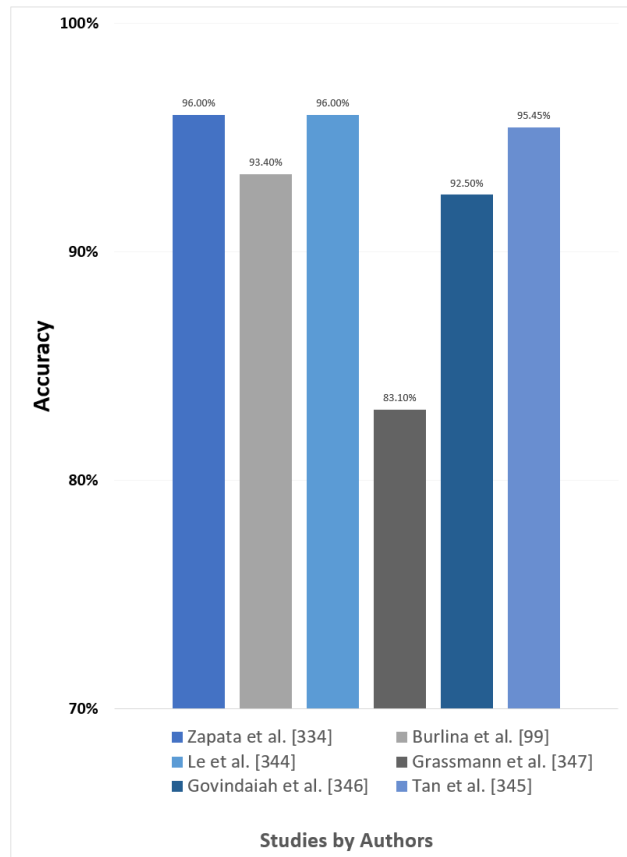


FIGURE 12. Bar graph illustrating the maximum accuracy attained by different authors for CNN classifier-based automated AMD identification.

The bar graph in Figure 12 shows the highest level of accuracy that various authors were able to achieve for CNN classifier-based automated age-related macular degeneration (AMD) identification. Each bar represents a different study, with the height of the bar indicating the reported accuracy.

The graph highlights the effectiveness of CNN classifiers in automated AMD identification, with reported accuracies ranging from 83.1% to 96.60%.

Out of the studies shown that used CNN for AMD detection, the highest accuracy was achieved by Zapata et al. [334] and Le et al. [344], both achieved an accuracy of 96% using CNN or one of its variants to classify AMD.

In their study, Zapata et al. [334] developed five algorithms and evaluated them in the Optretina dataset. Three different retinal specialists classified all the images. The dataset was split per patient into training (80%) and testing (20%) sets. Three different CNN architectures were employed, two of which were custom-designed to minimize the number of

parameters while maintaining accuracy. The main outcome measure was the area under the curve (AUC), along with accuracy, sensitivity, and specificity. The models were effectively used for data cleaning, quality assessment, eye orientation classification, and disease detection (AMD and GON). The custom-designed CNN architectures achieved these tasks with minimized parameters while maintaining high accuracy, demonstrating the potential for practical application in automated retinal image analysis and detection.

Le et al. [344] used fundus images from the Department of Ophthalmology at King Chulalongkorn Memorial Hospital in Thailand were collected for transfer learning, along with other publicly available datasets for testing. Seven models based on CNN—VGG19, Xception, DenseNet201, EfficientNetB7, InceptionV3, NASNetLarge, and ResNet152V2—were selected for training in 2-label (Normal vs. AMD) and 3-label (Normal vs. Dry AMD vs. Wet AMD) classifications. The experimental results indicated that the DenseNet201 model, with its Dense block structure, showed the best efficacy in both 2-label and 3-label AMD classifications, consistently ranking among the Top-3 models in terms of accuracy and generalisation performance, as measured by total accuracy and total F1-Score. It received an accuracy and sensitivity of 96%.

Tan et al. [345] achieved an accuracy of 95.50% in using CNN to detect AMD. They developed a fourteen-layer deep CNN model designed to automatically and accurately detect AMD at an early stage. The performance of the model was evaluated using blindfold and ten-fold cross-validation strategies, achieving accuracies of 91.17% and 95.45%, respectively

Burlina et al. [99] achieved an accuracy of 93.4% in using CNN to detect AMD. Using 5664 colour fundus images from the NIH AREDs dataset, this paper details the approach using deep learning for ARIA and AMD analysis. The researchers used transfer learning and universal features derived from deep convolutional neural networks (DCNN) to address clinically relevant 4-class, 3-class, and 2-class AMD severity classification problems.

Govindaiah et al. [346] achieved an accuracy of 92.50% in using CNN to detect AMD. They used the Age-Related Eye Disease Study (AREDS) dataset, which contained over 150,000 images along with qualitative grading information provided by expert graders and ophthalmologists. They employed a modified VGG16 neural network with batch normalisation in the last fully connected layers. The study involved two experiments. In the first experiment, the images were categorised into two classes based on clinical significance: No or early AMD, and Intermediate or Advanced AMD. In the second experiment, the images were categorized into four classes: No AMD, Early AMD, Intermediate AMD, and Advanced AMD. The modified VGG16 network achieved the best accuracy of 92.5% for the two-class problem with over 100,000 images. The results demonstrated that training a deep neural network with a sufficient

number of images yielded better performance than using a pre-trained network, especially for AMD detection and screening.

Grassmann et al. [347] achieved an accuracy of 83.10% in using CNN to detect AMD. Their study included 120,656 manually graded colour fundus images from 3,654 participants in the Age-Related Eye Disease Study (AREDS). Participants were over 55 years old, and those with non-AMD sight-threatening diseases were excluded from recruitment. Additionally, the algorithm's performance was evaluated using 5,555 fundus images from the population-based Kooperative Gesundheitsforschung in der Region Augsburg (KORA; Cooperative Health Research in the Region of Augsburg) study. The researchers defined 13 classes (9 AREDS steps, 3 late AMD stages, and 1 for ungradable images) and trained several convolutional deep learning architectures. An ensemble of network architectures was used to improve prediction accuracy. The performance of the algorithm was evaluated on an independent dataset. The primary measures were k statistics and accuracy to evaluate the concordance between the algorithm's predictions and expert human grader classifications.

Most of these studies achieved a high accuracy using CNN and its variants to detect AMD. CNNs excel at automatically extracting hierarchical features from images.

In the context of AMD detection, CNNs can identify intricate patterns, textures, and structures in fundus images that are indicative of the disease. This ability to learn and extract relevant features from raw images is crucial for accurate detection. CNNs use convolutional layers that apply filters across the input image, capturing spatial hierarchies and relationships within the image. This spatial invariance helps in detecting AMD features regardless of their location within the image, improving the robustness of the model. Overall, these studies demonstrate the high potential of CNN classifiers in accurately detecting AMD from fundus images.

The bar graph in Figure 13 illustrates the maximum accuracy achieved by different authors for CNN classifier-based automated diabetic retinopathy (DR) detection. Each bar represents a different study, with the height of the bar indicating the reported accuracy.

The graph highlights the effectiveness of CNN classifiers in automated DR detection, with reported accuracies ranging from 75.70% to 99.62%.

Gayathri et al. [137] achieved the highest accuracy of 99.62% for detecting DR using CNN. In their study, a multipath convolutional neural network (M-CNN) is employed for global and local feature extraction from fundus images. These features are then classified according to the severity of DR using various ML classifiers. The proposed model is evaluated using several publicly available databases: IDRiD, Kaggle (for DR detection), and MESSIDOR. Different ML classifiers, including SVM, RF, and J48, are used for categorisation. The experiments demonstrate that the M-CNN network combined with the J48 classifier produces

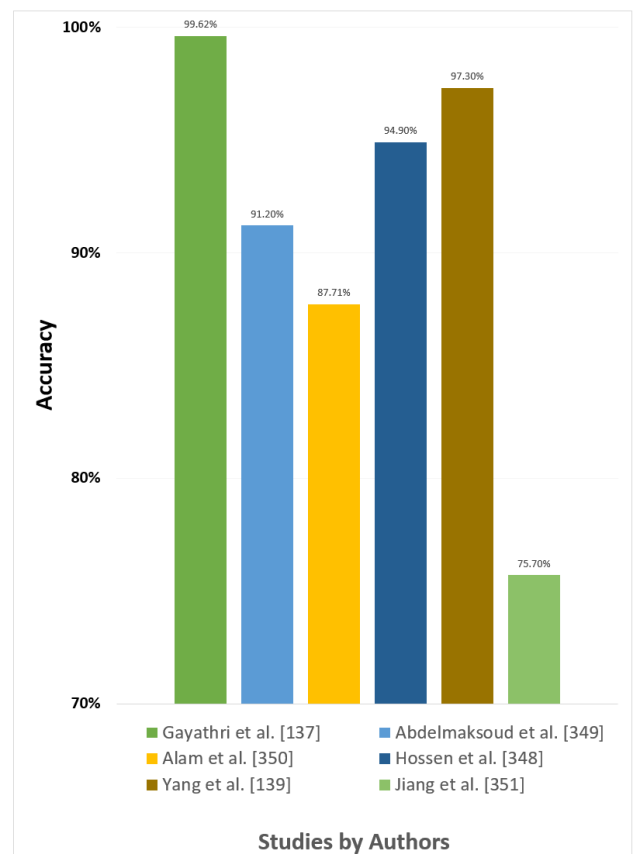


FIGURE 13. Bar graph displaying the maximum accuracy attained by different authors for CNN classifier-based automated DR detection.

the best results. The classifiers are assessed using features from pre-trained networks and existing DR grading methods.

Yang et al. [139] achieved the second-highest accuracy of 97.3% in detecting DR using DCNN on the EyePacs dataset. They proposed an automatic DR analysis algorithm based on a two-stage deep convolutional neural network. The algorithm can identify lesions in fundus colour images and provide DR severity grades. By introducing an imbalanced weighting scheme, the algorithm focuses more on lesion patches during DR grading, significantly improving grading performance under the same implementation setup.

Hossen et al. [348] achieved an accuracy of 94.9% in detecting DR with CNN. The study involved developing a DR classifier using a transfer learning technique with a DenseNet architecture-based pre-trained model. The classification of DR from retinal fundus images was based on its severity level. The identification of DR was achieved by detecting features such as micro-aneurysms, exudates, and hemorrhages in retinal images. Additionally, the preprocessing and augmentation of image data were conducted to enhance the model's ability to detect retinopathy. After the training and validation procedures, the developed classifier achieved a validation accuracy of 94.9%. The study demonstrates that using a DenseNet architecture can effectively detect Diabetic

Retinopathy by classifying retinal fundus images according to severity levels. The preprocessing and augmentation of image data significantly benefit the model, resulting in high training and validation accuracies.

Abdelmaksoud et al. [349] achieved an accuracy of 91.20% in detecting DR with CNN. This paper presents a novel hybrid deep learning technique named E-DenseNet, which integrates EyeNet and DenseNet models using transfer learning. The traditional EyeNet was customized by incorporating dense blocks and optimising the hyperparameters of the resulting hybrid E-DenseNet model. This approach aims to accurately detect healthy and various DR grades from both small and large ML colour fundus images. The model was trained and tested on four different datasets (EyePACS, APTOS, MESSIDOR, IDRiD). These datasets provided a comprehensive range of images necessary for robust training and validation of the proposed system. The E-DenseNet model demonstrates promising results compared to other systems, showcasing its effectiveness in accurately detecting various DR grades. By leveraging the strengths of both EyeNet and DenseNet through transfer learning and dense block integration, the proposed system provides a robust solution for the automated analysis of DR.

Alam et al. [350] achieved an accuracy of 87.71% in detecting DR with CNN. This study proposed a segmentation-assisted DR classification methodology that enhances current methods by using a fully convolutional network (FCN) to segment retinal neovascularisations (NV) in retinal images before classification. The study used the Kaggle EyePacs dataset, which contains fundus images from patients with varying degrees of DR (mild, moderate, severe NPDR, and PDR). The FCN was trained to locate neovascularisation in 669 retinal fundus photographs labelled with PDR status according to NV presence. The trained segmentation model was then used to locate probable NV in images from the classification dataset. Subsequently, a CNN was trained to classify the combined images and probability maps into categories of PDR. The segmentation-assisted classification achieved an accuracy of 87.71%. The study demonstrates that segmentation assistance improves the identification of the most severe stage of diabetic retinopathy.

Jiang et al. [351] achieved an accuracy of 75.7% in detecting DR with CNN. A total of 10,551 fundus images from the Kaggle fundus image dataset were collected for the experiment. The images were first pre-processed using histogram equalisation and image augmentation techniques. A CNN was then constructed and trained using the Caffe framework, with 8,626 images used for training the model. The performance of the trained CNN model was validated by classifying 1,925 fundus images into DR and non-DR categories. The results indicated that the CNN achieved an accuracy of 75.70% in classifying the 1,925 test fundus images.

Overall, these studies demonstrate the high potential of CNN classifiers in accurately detecting DR from fundus

images. CNNs achieve high accuracy in DR detection for several reasons. CNNs automatically extract features from fundus images, identifying patterns and structures associated with DR, such as microaneurysms, hemorrhages, and exudates [140]. They also capture spatial hierarchies and relationships within images, essential for detecting varying stages of DR across different regions of the retina [140]. The deep layers of CNNs allow them to learn both low-level and high-level features, crucial for accurate DR detection. Using pre-trained CNN models on large datasets and fine-tuning them for DR detection leverages learned features, enhancing accuracy even with smaller medical datasets [136].

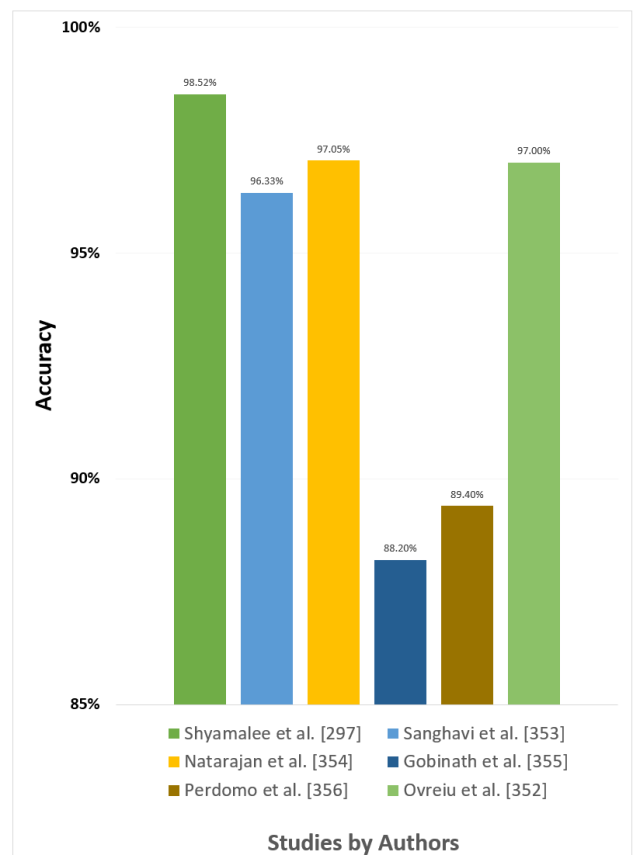


FIGURE 14. Maximum accuracy values attained by several authors for CNN classifier-based automated glaucoma detection.

Modern CNN architectures like ResNet and DenseNet include innovations that enhance feature extraction and model performance [352]. These factors contribute to the high accuracy of CNNs in DR detection, making them highly effective tools for automated detection from fundus images.

The bar graph in Figure 14 illustrates the maximum accuracy achieved by different authors for CNN classifier-based automated glaucoma detection. Each bar represents a different study, with the height of the bar indicating the reported accuracy. The graph highlights the effectiveness of CNN classifiers in automated glaucoma detection, with reported accuracies ranging from 88.20% to 98.52%.

Shyamalee et al. [297] achieved the highest accuracy of 98.52% in classifying glaucoma using CNN architecture Inception-v3 on the ACRIMA dataset. The study proposed developing an automated system for classifying glaucoma using DL, specifically through three various CNN architectures (Inception-v3, VGG19, ResNet50) on the publicly available RIM-ONE and ACRIMA databases. RIM-ONE includes three versions with a total of 942 fundus images (399 glaucomatous and 543 healthy) and ACRIMA consists of 705 images (396 glaucomatous and 309 healthy).

To enhance the image quality, they applied pre-processing techniques like dilation and Contrast Limited Adaptive Histogram Equalisation (CLAHE), which improve brightness and contrast. They also used data augmentation techniques such as rotation, shearing, zooming, flipping, and shifting to address data imbalance and increase the training dataset size. The models were trained using the Adam and SGD optimizers with binary cross-entropy as the loss function, over 150 epochs with a 70:15:15 split for training, testing, and validation sets [297].

The results of the study showed that the Inception-v3 model achieved the highest accuracy of 98.52% on the ACRIMA dataset and 96.56% on the RIM-ONE dataset [297]. VGG19 and ResNet50 also demonstrated high accuracy but slightly less than Inception-v3. The researchers evaluated model performance using metrics such as accuracy, precision, recall, F1-score, sensitivity, specificity, and AUC. Confusion matrices and ROC curves were used to illustrate the models' ability to correctly classify fundus images and their diagnostic performance.

The study provided a comparative analysis of the proposed CNN architectures with existing studies, highlighting the superior performance of Inception-v3 in classifying glaucoma. By addressing the class imbalance issue with augmentation techniques, the researchers improved the model's robustness and reduced overfitting [297].

Sanghavi and Kurhekar [353] achieved an accuracy of 96.33% in classifying glaucoma using CNN. This study utilised six widely used datasets of fundus images: DRISHTI-GS, ORIGA, ACRIMA, PAPILA, g1020, and RIM-ONE. This research investigates various segmentation and classification techniques for optic disk segmentation and the classification of normal and glaucomatous eyes. The approach begins with histogram processing to determine the type of image, which informs whether segmentation is necessary. Some datasets contained complete retinal images, while others included segmented optic disks. Segmented images are directly used for classification with the proposed convolutional neural network (CNN).

For complete retinal images, segmentation is performed using Simple Linear Iterative Clustering (SLIC) and normalised graph cut algorithms [353]. The performance of the proposed framework is compared with that of pre-trained neural networks, including VGG19, InceptionV3, and ResNet50V2, using major metrics. The study trained

and tested these architectures with 3115 images from six standard datasets. The proposed framework achieved superior performance, with an accuracy of 96.33%, outperforming all compared models.

Ovrieu et al. [352] achieved an accuracy of 97% in classifying glaucoma using CNN architecture DenseNet on the ACRIMA dataset. This paper proposed a method utilising densely connected neural networks (DenseNet) with 201 layers, initially pre-trained on the ImageNet dataset, and applied to the ACRIMA dataset. The method achieved an accuracy of approximately 97% and an F1-score of 0.969

Natarajan et al. [354] achieved an accuracy of 97.05% in classifying glaucoma using CNN architecture U-Net on the DRISHTI-GS1 dataset. This paper presents a two-stage deep learning framework called UNet-SNet for glaucoma detection. Initially, each fundus image is segmented into Gaussian Mixture Model (GMM) superpixels, and the Region of Interest (RoI) is separated using Cuckoo Search Optimisation (CSO). In the first stage, a regularised U-Net is trained with RoIs for optic disc (OD) segmentation. In the second stage, SqueezeNet is fine-tuned with deep features extracted from the ODs to classify the fundus images as either glaucomatous or normal. The U-Net was trained and tested with the RIGA and RIM-ONEv2 datasets, achieving accuracies of 97.84% and 99.85%, respectively. The classifier was trained with ODs segmented from the RIM-ONEv2 dataset and achieved an accuracy of 97.05% on the DRISHTI-GS1 dataset.

Gobinath [355] achieved an accuracy of 88.2% in classifying glaucoma using CNN. This study highlights the potential of using semi-supervised deep learning models over supervised methods. By utilising both labelled and unlabeled data on fundus images, the proposed semi-supervised GAN model comprises a SegNet, a real data generator, and a classifier to enhance segmentation performance. They demonstrate an accuracy of 88.2%, specificity of 90.8%, and sensitivity of 85%.

Perdomo et al. [356] achieved an accuracy of 89.04% in classifying glaucoma using CNN. This study introduced a multi-stage deep learning model for glaucoma detection, utilising a curriculum learning strategy. In curriculum learning, the model is trained sequentially to handle increasingly difficult tasks. The proposed model consists of several stages: segmentation of the optic disc and physiological cup, prediction of morphometric features from these segmentations, and prediction of the disease level (categorised as healthy, suspicious, or glaucomatous). The experimental evaluation demonstrates that the proposed method outperforms conventional convolutional deep learning models. Specifically, the method achieved an accuracy of 89.4% and an Area Under the Curve (AUC) of 0.82 on the RIM-ONE-v1 and DRISHTI-GS1 datasets, respectively. These results highlight the effectiveness of the multi-stage deep learning approach for glaucoma detection.

The figures highlight the top accuracy results achieved with SVM and CNN classifiers. SVMs still achieve high accuracy for glaucoma, up to 99.2% (Figure 11). Meanwhile, CNNs now consistently surpass 95% accuracy for AMD (Figure 12) and Glaucoma (Figure 14). Continued algorithm improvements and larger datasets for DL will be key to further boosting performance.

A review of the studies listed in Appendix A reveals that each study used different datasets and involved various subjects. The difference in datasets, the number of fundus images and the quality of the images can also make it difficult to gauge performance. The diversity in data and sample sizes makes it difficult to compare the performance of the proposed methods for detecting eye diseases accurately.

XII. CURRENT LIMITATIONS

Although the studies included in our comprehensive review demonstrate the potential of ML and DL in the detection of eye diseases, several limitations remain.

From the reviewed literature, we can see that the simultaneous occurrence of multiple pathologies has rarely been considered and evaluated, which could aid in the recognition and segmentation of retinal structures and lesions. For instance, in glaucoma detection, retinal lesions caused by DR and AMD are often ignored and not detected by the developed algorithms. When these lesions are close to the optic disc, they complicate the detection of its boundary, making it more error-prone. An algorithm that simultaneously identifies both the optic disc boundary and retinal lesions could be more effective.

Another limitation is the limited availability of high-quality, well-annotated eye image datasets, which are crucial for training robust ML and DL models. The scarcity of such datasets holds back the development of these algorithms. There is no dataset with images from the same subjects acquired at different time points, which hinders the validation of specific methods for disease monitoring. Such monitoring is crucial in clinical practice and should be considered in the development of automatic methods to support diagnosis.

Furthermore, comparing the performance of various studies is challenging due to the use of different datasets, which vary in terms of the number of subjects, data collection methods, and image quality. These factors, along with variations in image resolution and differences between imaging devices, can significantly impact the performance of the algorithms.

Some studies did not use enough fundus images; a larger dataset should allow the model to train itself with more data, leading to a more accurate diagnosis. A larger dataset must be used to clarify the performance disparities between ophthalmologists' diagnoses and AI models. Imbalanced datasets can also hinder model performance, as they may lead to biased predictions.

CNNs and Vision Transformers excel when they have access to large datasets. However, retinal images are rarely

available in substantial quantities and typically lack annotations. DL models often overfit when trained on limited data. Additionally, DL models are inherently complex and computationally intensive, which can hinder their seamless integration into clinical practice. Most of the included studies relied on a common reference standard for image classification decisions made by ophthalmologist graders. This implies that the algorithm may not perform well for images with subtle findings that many ophthalmologists might overlook.

Another limitation of the studies is using different performance metrics for evaluating ML and DL models in detecting ODs from fundus images. While some studies used only accuracy, sensitivity and specificity, others used precision, mean error, AUC, correlation coefficient, IoU, and Dice. The lack of standardisation in performance metrics has several disadvantages. Without standard metrics, comparing the performance of different models becomes challenging, as each study may choose different metrics that highlight its strengths but may not provide a complete picture of the model's performance.

Certain metrics might favour models that are good at handling specific types of data imbalances or particular aspects of the data. This could lead to the selection of models that are not necessarily the best overall but perform well on the chosen metric.

Focusing on a narrow set of metrics might lead to overlooking important aspects of model performance, such as how well a model can generalise to unseen data, its robustness to noise in the data, or its performance across different subgroups within the data. Different studies might interpret the same metric differently, especially in the absence of context or an understanding of what each metric truly measures. For example, a high accuracy might seem impressive, but it may not be as relevant in a dataset where there are only a handful of fundus images. It is crucial to consider all metrics to assess the success of classification fully. While various metrics offer valuable insights into different aspects of model performance, the lack of standardisation in their use across studies can complicate the evaluation and make it difficult to compare performance.

Addressing these limitations is essential for the successful implementation of ML and DL in ophthalmology and for realising their potential to revolutionize eye disease detection and management.

XIII. FUTURE DIRECTIONS

Our study identified several challenges in developing and deploying AI-based diagnostic tools for retinal diseases. One major challenge is the limited size and lack of diversity in current datasets, which hampers the model's ability to generalise across different populations and conditions, making it less reliable. Additionally, in real-world scenarios, images often contain noise due to varying quality, lighting conditions, and patient movement, which can obscure important diagnostic

features and reduce the accuracy of AI models. Another challenge is the variation in imaging devices used by different hospitals and clinics, noise present in images, resulting in inconsistencies in image quality and characteristics that pose a challenge for AI models trained on data from a single type of device. Moreover, gaining the trust of clinicians for the widespread adoption of AI models in clinical settings requires transparent, interpretable models and consistent performance across diverse clinical environments.

To address these challenges, several AI solutions are proposed, structured in a clear roadmap for future work. Firstly, collaboration with multiple hospitals and clinics globally to collect a large, diverse set of annotated images should be done, ensuring the inclusion of varied demographics (age, sex, race) and multiple pathologies. It would be beneficial for future research to associate each image with information not only on the presence or grading of a specific pathology but also on any additional pathologies present. This approach would facilitate the development of algorithms capable of screening and analysing multiple pathologies simultaneously, effectively managing signs related to other conditions that currently represent noise in images. Establishing standardised protocols for image annotation and data collection will ensure consistency and reliability across different sources. Secondly, to handle real-world image noise, the development and integration of advanced preprocessing algorithms to clean and enhance images by reducing noise and correcting lighting variations should be done. Training models using techniques such as data augmentation, adversarial training, and noise-robust algorithms like uncertainty quantification will improve their resilience to real-world noise [357].

Managing variation in imaging devices will involve creating calibration procedures to standardise images from different devices, ensuring that the AI model can handle variations in image quality and settings. Training AI models on datasets collected from various types of imaging devices will enhance their generalisability and robustness across different clinical environments. To build clinician trust through explainable AI (XAI), the development of interpretable models that provide clear, understandable insights into their decision-making processes should be done, integrating XAI techniques to highlight which parts of an image contributed to the diagnosis [357], [358]. Additionally, designing user-friendly interfaces that present AI findings in an easy-to-interpret manner should include visual aids, confidence scores, and clear explanations of the AI's reasoning [359]. Elsharkawy et al. proposed an automated, explainable artificial intelligence (XAI) system for diagnosing age-related macular degeneration [360]. This system mimicked physician perceptions by deriving clinically meaningful features from optical coherence tomography (OCT) B-scan images, enabling differentiation between a normal retina, various AMD grades (early, intermediate, geographic atrophy, inactive wet, active neovascular disease), and non-AMD diseases [360]. The XAI system extracted retinal

OCT-based clinical markers related to AMD progression, including subretinal tissue, sub-retinal pigment epithelial tissue, intraretinal fluid, subretinal fluid, and choroidal hyper transmission using a DeepLabV3+ network; merged retina layers using a novel convolutional neural network model; drusen detection via 2D curvature analysis; and retinal layer thickness and reflectivity features. These clinical features were utilised in a hierarchical decision tree process to grade the OCT images. Severe cases indicating advanced AMD were further analysed to diagnose specific conditions, while less severe cases were assessed for intermediate or early-stage AMD. The system was evaluated on 1285 OCT images and achieved 90.82% accuracy, demonstrating its capability to automatically distinguish between normal eyes, various AMD grades, and non-AMD diseases [360].

Model explainability is a continuing area of research, particularly with CNNs. Vision Transformers (ViTs) can be computationally demanding and require substantial training data, which can restrict their use in certain contexts [361]. Despite this, ViTs have inherent interpretability features, such as self-attention mechanisms, that enable the model to concentrate on important aspects of the input image. This characteristic makes ViTs potentially more suitable for developing explainable models compared to traditional CNNs [361]. Future work should be done to comprehend the balance between interpretability and computational complexity in ViTs and to discover the best methods for creating lightweight, explainable models using ViTs.

Combining modalities for comprehensive analysis is another key solution. For example, fundus imaging offers a colour photograph of the retina, providing a view of the retinal surface, while OCT offers a cross-sectional image, giving insight into the layers of the retina. Using both will allow for a more complete analysis of the retina's health. Each imaging modality may capture different aspects of retinal diseases. For instance, OCT can show subretinal fluid or macular edema not visible on fundus photography. AI models can learn to identify disease markers from both types of images, potentially improving diagnostic accuracy. Some changes may be more apparent or only visible in one type of imaging. By analysing both fundus and OCT images, AI can help in the early detection of diseases by picking up on these subtle changes, which is crucial for conditions like glaucoma and AMD, where early intervention can prevent progression. AI systems can cross-verify findings across both types of images to reduce false positives and negatives. For example, what appears to be an abnormality in a fundus image may be clarified as a normal variation in the OCT, leading to more confident diagnoses. Monitoring these diseases over time can be more effective when both types of images are available. AI models can track changes in both the retinal surface and sub-retinal structures, giving a clearer picture of how a disease is responding to treatment. The development of models that simultaneously analyse fundus and OCT images to provide a more comprehensive assessment of

TABLE 14. List of research papers that have used machine learning to classify Glaucoma cases. Note: AEH = Aravind eye hospital, Madurai, India; NOI = NIO hospital, Pune; OHT = Ophthalmology Hospital, Tunisia; KMC = Kasturba Medical College, India; BEVC = Biobank Eye Vision Consortium, UK; S1 - DRISHTI-GS, AEH (101); S2 - RIM-One; S3 - NOI (118); S4 - DRIONS-DB (110); S5 JSIEC (124); S6 - DRIVE (44); HCECE = Homogeneity, Contrast, Entropy, Correlation and Energy; PHOVW = Pyramid histogram of visual words; KRYE = Kapoor/Rényi/Yager entropy; SRKE = Shannon/Rényi/Kapoor entropy; LCP = Local configuration pattern; BV = Blood Vessels, SFEM = Superpixel feature extraction module, RHT = Randomized Hough transform, SFTMH = Statistical features/texton map histogram, HOCF = Higher order cumulant features, DCGGFHHW = DDLS/CDR/GLRM/GLCM/FoS/HOS/HOC/Wavelet; CHRf = corneal hysteresis + corneal resistance factor, VERC = Venu Eye Research Centre, India; IOP = Intraocular Pressure; CCT = Central Corneal Thickness; ACRIMA = A1.

Year	Authors	Classifier	Dataset	Feature Extracted	Accuracy %	Specificity %	Sensitivity%
2024	[365]	SVM	RIMONE	Various	96.21	97.16	94.18
2024	[366]	SVM	REFUGE, A1, ORIGA, S2, S1	Various	81	84	
2024	[366]	RF	REFUGE, A1, ORIGA, S2, S1	Various	88	98	
2024	[366]	Logistic Regression	REFUGE, A1, ORIGA, S2, S1	Various	95	98	
2024	[366]	NB Gaussian	REFUGE, A1, ORIGA, S2, S1	Various	70	84	
2024	[366]	NB Multinomial	REFUGE, A1, ORIGA, S2, S1	Various	70	70	
2024	[367]	Random Forest	SOURCE	IOP, CCT	79.2	86.5	43.9
2024	[367]	Decision Tree	REFUGE, A1, ORIGA, S2, S1	Various	85	77	
2024	[367]	Logistic Regression (L2,L1)	SOURCE	IOP, CCT	76.2, 77.5	80.8, 83.8	53.6,47.1
2024	[367]	Elastic Net Regression	SOURCE	IOP, CCT	76.6	81.8	50.9
2024	[367]	Gradient Boosted Trees	SOURCE	IOP, CCT	78.7	84.7	49.4
2024	[367]	XGBoost	SOURCE	IOP, CCT	68.5	82.2	55.1
2023	[368]	SVM, KNN	ORIGA	Various deep features	89.8, 81.9		89.8, 81.9
2023	[368]	Decision Tree, NB	ORIGA, HRF	Various deep features	98.8, 53.3		98.8, 53.3
2023	[368]	NB	ORIGA	Various deep features	69.4		69.4
2023	[368]	Logistic Regression	ORIGA	Various deep features	99.4		99.4
2023	[368]	KNN, DT, SVM, RF, LR	HRF	Various deep features	50		50
2023	[368]	KNN	DRISHTI-GS	Various deep features	79.2		100
2023	[368]	Decision Tree, SVM	DRISHTI-GS	Various deep features	96, 73.3		98.6, 100
2023	[368]	RF	DRISHTI-GS	Various deep features	95		98.6
2023	[368]	NB	DRISHTI-GS	Various deep features	65.3		54.3
2023	[368]	Logistic Regression	DRISHTI-GS	Various deep features	99		100
2022	[369]	ANFIS	Private	Various	97.1		97.3
2022	[104]	LDA,RF	S1, ACRIMA-DB	HCECE	73.33, 82.56		88.88, 84.68
2022	[104]	SVM, NB	S1, ACRIMA-DB	HCECE	82.64, 88.86		86.48, 90.25
2022	[324]	SVM	S1-R, ORIGA-R, S2-R	Various	93.4, 79.6, 91.3		
2022	[324]	SVM	A1-R, LAG-R	Various	99.5, 99.6		
2021	[116]	Densenet, InceptionResnetV4	BEVC, UK (2476 images)	IOP, CHRf			81
2021	[105]	SVM, NN, Adaboost	i. S1, ii. S2, iii. S3, iv. S4, v. S5, vi. S6	CDR+NRR +BV ISNT ratio	100	100	100
2020	[106]	KNN	i. DRISHTI-GS, AEH (101 images), ii. RIM-One (159 images)	DCGGFHHW	91.6	90	91
2020	[106]	NN	i. DRISHTI-GS, AEH (101 images), ii. RIM-One (159 images)	DCGGFHHW	94.4	93	94
2020	[106]	SVM	i. DRISHTI-GS, AEH (101 images), ii. RIM-One (159 images)	DCGGFHHW	97.2	96	97
2020	[106]	RF	i. DRISHTI-GS, AEH (101 images), ii. RIM-One (159 images)	DCGGFHHW	94.4	93	94
2020	[106]	NB	i. DRISHTI-GS, AEH (101 images), ii. RIM-One (159 images)	DCGGFHHW	89.6	89	88
2020	[84]	Inception-v3, Adam Optimizer	Private	Various	92.65		
2020	[370]	SVM	Private (298 images)	CDR	95.5		
2020	[370]	KNN	Private (298 images)	CDR	93.3		
2020	[370]	NB	Private (298 images)	CDR	94.35		
2019	[341]	SVM	RIM-One, Spain (166 images)	SFEM	98.6	97.6	92.3
2019	[371]	CNN, RF	Private (208 images), Helsinki	RNFL/GCC thickness map			
2019	[107]	SVM	DRIONS-DB (110 images)	SFTMH	99.2	99.3	96.9
2018	[117]	RF, CNN, SVM	Private (304 images)	RNFL/GCC thickness map			96
2018	[109]	RF	KMC (2220 images)	PHOVW, Fisher vector	96.79	96.96	96.73
2018	[337]	Fuzzy logic	OHT (104 images)	RHT	90.2	94.8	97.8
2017	[110]	KNN	KMC (702 images)	LCP	95.7	93.7	96.2
2017	[122]	LS_SVM	KMC (488 images)	KRYE, fractal dimensions	94.8	93.6	95.9
2017	[123]	LS_SVM	KMC (60 images)	EWT, Correntropy	98.3	96.7	100
2016	[112]	KNN	VERC(63 images)	Wavelet feature extraction	94.8	100	90.9
2015	[372]	SVM	RIM-One, Spain (158 images)	HWT	90.1	85.7	94.4
2015	[342]	SVM	VERC (67 images)	CDR, NRR, BV ISNT	94.1	90	100
2015	[343]	SVM	KMC (510 images)	Statistics/energy, SRKE	93.1	96.2	89.7
2013	[331]	SVM, NB	KMC (272 images)	HOCF	92.65	92	100
2012	[373]	SMO	KMC (60 images)	Wavelet energy	93.33		
2012	[332]	SVM	KMC (60 images)	Wavelet energy, HOS	95	93.3	96.6

TABLE 15. List of papers that used DL to classify Glaucoma. Note: LALES = Los Angeles Latino Eye Study. S = SVM, k = KNN, C/R-18/G = ResNet-18/GlaucomaNet; R = ResNet, G2 = GoogleNet, A = AlexNet, V = VGGNet.

Year	Authors	Classifier	Dataset	Accuracy %	Specificity %	Sensitivity%	AUC
2024	[304]	Inception-v3	RIM-ONE	98.79	95.58	99.22	0.99
2024	[374]	DNN	Private		93.51	92.05	0.96
2024	[353]	CNN	Private (3115 images)	96.33			
2024	[113]	MFR-Net	Private (1426 images)	98.36	100	97.6	0.98
2024	[113]	CDAM-Net	Private (1426 images)	99.3	100	98.8	0.99
2024	[367]	Deep Learning Embedding Model	SOURCE	77.7	83.4	50.1	
2024	[367]	Deep Learning FCN	SOURCE	72.5	75.3	58.8	
2024	[375]	GDA	Private (66,742 images)		90	96	0.98
2024	[375]	GDA	Private (66,742 images)		87	94	0.96
2024	[375]	GDA	Private (66,742 images)		85	91	0.93
2024	[375]	GDA	ACRIMA		65	85	0.87
2024	[375]	GDA	RIM-ONE		75	89	0.92
2023	[317]	LSVT-Net	REFUGE	99.6			0.99
2023	[322]	ESS-Net	REFUGE	99.88	99.91	97.98	
2023	[322]	FBSS-Net	REFUGE	99.91	99.95	97.41	
2023	[322]	ESS-Net	Rim-One-r3	99.69	99.74	97.48	
2023	[322]	FBSS-Net	Rim-One-r3	99.74	99.86	97.04	
2023	[322]	ESS-Net	Drions-DB	99.76	99.95	94.45	
2023	[322]	FBSS-Net	Drions-DB	99.82	99.91	97.41	
2023	[322]	ESS-Net	Drishiti-GS	99.85	99.97	96.45	
2023	[322]	FBSS-Net	Drishiti-GS	99.87	99.93	97.87	
2022	[154]	U-Net with Inception-v3 (on ODS)	RIM-ONE, ACRIMA	98.45, 97.91	97.14, 96.57	99.38, 98.69	
2022	[154]	U-Net with VGG19 (on ODS)	RIM-ONE, ACRIMA	99.02, 98.12	98.38, 96.03	99.47, 99.34	
2022	[154]	U-Net with ResNet50 (on ODS)	RIM-ONE, ACRIMA	99.53, 99.23	99.25, 98.45	99.73, 99.67	
2022	[154]	U-Net with Inception-v3 (on OCS)	RIM-ONE, ACRIMA	96.34, 96.46	97.46, 96.64	92.58, 96.16	
2022	[154]	U-Net with VGG19 (on OCS)	RIM-ONE, ACRIMA	96.55, 97.22	89.82, 97.61	97.95, 95.08	
2022	[154]	U-Net with ResNet50 (on OCS)	RIM-ONE, ACRIMA	97.37, 97.85	97.97, 98.05	96.53, 96.61	
2022	[297]	Inception-v3	RIM-ONE	96.56	90.9	98.31	
2022	[297]	VGG19	RIM-ONE	94.95	86.13	97.98	
2022	[297]	ResNet50	RIM-ONE	95.49	88.88	97.59	
2022	[297]	Inception-v3	ACRIMA	98.52	90.9	99.5	
2022	[297]	VGG19	ACRIMA	92.64	78.78	97.08	
2022	[297]	ResNet50	ACRIMA	95.58	83.33	98.21	
2022	[304]	U-Net with Inception-v3 (on ODS)	RIM-ONE	98.65	96.81	99.01	
2022	[304]	U-Net with Inception-v3 (on ODS)	ACRIMA	97.91	90	98.63	
2022	[304]	U-Net with VGG19 (on ODS)	RIM-ONE	99.07	97.46	99.38	
2022	[304]	U-Net with VGG19 (on ODS)	ACRIMA	98.18	91.66	98.78	
2022	[304]	U-Net with ResNet50 (on ODS)	RIM-ONE	99.58	98.69	99.75	
2022	[304]	U-Net with ResNet50 (on ODS)	ACRIMA	99.26	94.73	99.54	
2022	[304]	U-Net with Inception-v3 (on OCS)	RIM-ONE	96.39	91.27	97.49	
2022	[304]	U-Net with Inception-v3 (on OCS)	ACRIMA	96.65	82.25	98.01	
2022	[304]	U-Net with VGG19 (on OCS)	RIM-ONE	96.7	91.81	97.74	
2022	[304]	U-Net with VGG19 (on OCS)	ACRIMA	97.35	86.66	98.32	
2022	[304]	U-Net with ResNet50 (on OCS)	RIM-ONE	97.42	92.98	98.37	
2022	[304]	U-Net with ResNet50 (on OCS)	ACRIMA	98.05	89.83	98.78	
2022	[354]	U-Net	ACRIMA	99.86			
2022	[354]	U-Net	DRISHTIGS1	97.05			
2022	[326]	ViT	ORIGA, RIM-ONEv3		91.2, 95.7	92.3, 94.1	
2022	[361]	ViT	ORIGA				0.96
2021	[141]	InceptionResNetV2	Private (22972 images)		94.3-98.4	97.5-98.2	
2021	[355]	CNN	Private (1426 images)	88.2	90.8	85	
2021	[364]	TCNN	RIM-One, Spain/RIGA	91.5	92	90.5	
2021	[364]	SSCNN	RIM-One, Spain/RIGA	92.4	93.3	91.7	
2021	[364]	SSCNN-DAE	RIM-One, Spain/RIGA	93.8	90.5	98.9	
2021	[376]	ViT	ORIGA	73.7			0.964
2021	[377]	ViT	RIGA	90.2			0.975
2020	[352]	DenseNet-201	ACRIMA	97	100	94.1	0.971
2020	[298]	DCNNs	DRIVE, DIARETDB1	97.3			
2020	[135]	CNN	Private (1155 images)	99			
2020	[378]	ResNet-50	Private	97			
2019	[136]	AG-CNN	LAG	96.2	96.7	95.4	0.983
2019	[379]	GoogleNet, ResNet-50	RIM-ONE	91, 90	99, 94	17, 42	0.91, 0.84
2019	[380]	CNN, RF	Private (357 images)				0.94
2019	[381]	S, k, C/R-18/G	LALES (93 images)				0.91 to 0.92
2019	[382]	Multi-task CNN	REFUGE (400 images)				0.95
2019	[379]	GoogleNet ResNet	RIM-ONE	85, 86	91, 93	29, 21	0.75, 0.74
2018	[356]	DCNN	RIM-ONE	89.4			
2018	[22]	DCNN	LabelMe		92	95.6	0.986
2018	[383]	MB-NN	Private	91.5	90.9	92.3	
2018	[384]	CNN	Private (3239 images)		79.4	90	0.926
2018	[385]	DL	Private (50000 images)		92	95.6	0.986
2015	[118]	CNN	ORIGA SCES				0.831, 0.887

retinal health, capturing both surface and sub-surface retinal features, should be done. Future work should focus on creating algorithms capable of detecting and grading multiple

retinal pathologies simultaneously will involve associating each image with information about the presence and grading of various conditions, facilitating more holistic patient care.

TABLE 16. List of research papers which have used machine learning to classify AMD.

Year	Authors	Classifier	Dataset	Feature Extracted	Accuracy %	Specificity %	Sensitivity%
2024	[386]	SVM	Private	Various	96		
2024	[386]	SVM	Private	Various	92		
2024	[295]	SVM	Private	Various	96.85	97.89	93.72
2021	[387]	NN	Private			88.8	91.8
2020	[388]	ResNet34	Private		87.2	80-100	50
2019	[126]	MM, SVM	STARE	Various	83.58		
2017	[389]	RF, Wavelet transform	Private	Energy and entropy features	92.48	95.58	89.37
2017	[255]	SVM, PHOG	KMC, Manipal, India	Nonlinear features	83.3	84.8	82.6
2017	[18]	DNN	Private	Various	87.63		
2015	[161]	NN	Private	Edges of drusen	96.17	99	89.81
2015	[329]	Pixel-wise feature extraction	Private	Various	96.32	97	
2014	[338]	Tree-based approach	STARE		99.9	99	99.5
2014	[390]	Gabor filter	STARE	Various	96		
2014	[119]	SVM, Energy, entropies, DWT	KMC, Manipal, India	First-order statistics	93.7	96.3	91.11
2012	[333]	Hierarchical decomposition	ARIA		100	100	100
2012	[9]	WFSGM	ARIA + STARE	Various	99.6	100	99.4

TABLE 17. List of research papers that used deep learning to classify AMD.

Year	Authors	Classifier	Dataset	Accuracy %	Specificity %	Sensitivity%	AUC
2024	[317]	LSVT-Net	Ichallenge	99.1			0.989
2023	[114]	GAN	LabelMe	88.5	94	88.5	0.967
2023	[114]	GAN	LabelMe	83.7	93.2	83.7	0.931
2023	[320]	DeepDrAMD	Private	95.43		96.75	0.9876
2023	[320]	DeepDrAMD	Private (dryAMD vs wetAMD)	90.57		92.54	0.9346
2023	[320]	DeepDrAMD	Private (type-1 vs type 2 MNV)	97.2		96.08	0.9936
2023	[344]	VGG19	Private	83		91	
2023	[344]	Xception	Private	94		91	
2023	[344]	DenseNet201	Private	96		96	
2023	[344]	EfficientNetB7	Private	95		92	
2023	[344]	InceptionV3	Private	94		92	
2023	[344]	NASNetLarge	Private	95		93	
2023	[344]	ResNet152V2	Private	93		90	
2022	[323]	MuMo-GAN	Private	58.2	78.7	56	
2022	[323]	Pix2pix	Private	46.1	75.4	50.6	
2022	[323]	PAN	Private	50.4	75.4	51.2	
2021	[391]	CNN	Private		90	88	0.983
2021	[392]	AMDOCT-Net	Private	99.1, 95.7	100, 92	98.2, 99.3	
2021	[393]	ViT	MESSIDOR, APTOS	91.3			0.963
2020	[334]	CNN	Optretina	96	92.4	97.7	0.979
2019	[394]	CNN	Private		97	94	0.9925
2019	[394]	DCNN	THDJ (364 images)		97.31	100	
2019	[394]	DCNN	Private				0.998
2018	[346]	VGG16	AREDS	92.5			
2018	[345]	CNN	Private	95.45	93.5	96.43	
2018	[347]	Ensemble	AREDS		84.2	94.3	
2018	[17]	DCNN	Private	96.6	97.4	97.8	0.999
2018	[134]	DCNN	Private	100, 99.6, 99.8	100, 99.2, 99.6	100, 100, 100	
2018	[347]	CNN	Private	83.1			
2018	[385]	DL	Private (8000 images)		94	98	0.995
2017	[99]	DCNN	NIH AREDS	93.4			
2017	[395]	DCNN + SVM	AREDS	95	95.6	96.4	
2016	[396]	RT, DWT, LSDA	STARE	99.49	99.75	99.21	
2016	[396]	RT, DWT, LSDA	ARIA	96.89	91.67	100	
2016	[397]	DCNN	AREDS	95	95.6	96.4	

Uncertainty quantification in retinal health screening refers to the process of identifying, characterising, and managing the uncertainties inherent in predictive models [362]. Uncertainty quantification is crucial for providing reliable and robust diagnostic outcomes from ML and DL models [363]. Aleatoric uncertainty or data uncertainty, arises from inherent variability in the data due to noise or insufficient data. In retinal imaging, it might result from variations in image quality, patient demographics, or differences in imaging devices. Epistemic uncertainty or model uncertainty arises from the model's lack of knowledge, often due to insufficient training data or limitations in the model architecture. Uncertainty quantification provides a measure of confidence in the model's predictions and can help clinicians make more

informed decisions [363]. For instance, if a model predicts the presence of DR with high uncertainty, a clinician might decide to perform additional tests before confirming the diagnosis. Incorporating uncertainty quantification methods will provide measures of confidence in the model's predictions, helping clinicians make more informed decisions by highlighting cases that may require additional review or testing. Developing training programs for clinicians to understand and effectively use AI tools, including how to interpret uncertainty measures and integrate AI findings into their diagnostic workflow, will be crucial.

Our study found that ML and DL are the two emerging tools that are used for screening retinal diseases. ML algorithms, however, may not be as potent as DL for automatic

TABLE 18. List of research papers that have used machine learning to classify DR. Note: THDJ = Tsukazaki Hospital Database, Japan; MDKEH = Mashhad Database, Khatam-Al-Anbia Eye Hospital.

Year	Authors	Classifier	Dataset	Feature Extracted	Accuracy %	Specificity %	Sensitivity%
2024	[398]	RF	EyePACS1	Microaneurysms	93		88
2024	[398]	Linear SVM	EyePACS1	Microaneurysms	92		89
2024	[398]	Decision Tree	EyePACS1	Microaneurysms	89		87
2024	[398]	NB	EyePACS1	Microaneurysms	89		77
2024	[398]	Radial SVM	EyePACS1	Microaneurysms	94		89
2023	[399]	SVM	APTOS2019	Various	87.04		
2023	[245]	SVM, KNN, MLP, NB	E-optha	Microaneurysms	94	93.8	87
2022	[284]	NN	APTOS2019, EyePacs	Various	92		
2021	[400]	SVM, NB		Various	99		
2021	[401]	NN	Private	Various			95
2021	[121]	HIMLA	CHASE (28 images)	Various	96.62	96.88	95.31
2019	[120]	RF	Private	Exudates	100		
2019	[120]	ANN	Private	Exudates	100		
2019	[402]	NB	IDRiD	Microaneurysms	77.85		
2016	[233]	MRF	DIARETDB1	Microaneurysms		82	
2016	[336]	MCA	DIARETDB0	Microaneurysms		99.93	89.01
2016	[336]	MCA	HEI-MED	Microaneurysms		99.81	81.26
2016	[336]	MCA	e-Opha	Microaneurysms		99.83	80.32
2014	[330]	GMM	MESSIDOR	Various	97.59	97.03	97.61
2014	[330]	GMM	DRIVE	Various	94.03	94.74	94.26
2014	[330]	GMM	STARE	Various	97.89	97.43	98.79
2014	[330]	GMM	DIARETDB0	Various	92.96	92.76	93.08
2014	[98]	SVM, KNN, GMM, AdaBoost	Various		53.16	100	
2013	[339]	Radon transform	MDKEH (120 images)	Various		75	94
2013	[339]	Radon transform	Private (72 images)	Various		70	100

detection, as the user has to define each of the features to detect the disease. Future work should focus on developing a novel DL model to detect multi-retinal classes such as AMD, DR, glaucoma, cataract, comorbid, and high blood pressure subjects, from a combination of data from multiple imaging modalities such as fundus and OCT. However, DL models demand substantial computational resources for both training and testing, which can impede their scalability and practicality in clinical settings. Future work should focus on developing an efficient, lightweight DL model that can be trained and deployed on devices with limited resources.

The roadmap for these solutions includes short-term, medium-term, and long-term phases. In the short term (1-2 years), the aim should be to collect and standardise a large, diverse dataset from multiple sources, develop and implement advanced image preprocessing techniques, train and validate models on multi-device datasets, and begin integrating explainable AI techniques into model development. In the medium term (2-3 years), the plan should be to enhance model robustness against real-world noise through extensive testing and refinement, develop calibration procedures for cross-device standardization, launch user-friendly interfaces with integrated XAI for clinical trials, and start combining fundus and OCT images for dual-modality analysis. In the long term (3-4 years), the focus will be on establishing standardised evaluation metrics and protocols for clinical adoption, fully implementing multi-pathology detection algorithms, integrating uncertainty quantification into AI models and interfaces, and conducting large-scale clinical trials to validate model performance and gain regulatory approval.

A proposed model, like the one shown in Figure 15, can be created in the future, which takes different modalities such as fundus and OCT eye images as input, and with

the help of xAI makes the results more understandable for clinicians. Combining different methods, analysing images and data acquired through various modalities, and conducting simultaneous analyses of multiple pathologies or retinal lesions could lead to improved performance and predictions that are as intuitive as those of expert clinicians. Future work should also focus on establishing a standardised set of performance metrics that reflect the needs and priorities of both the medical and patient communities. The lack of standardised evaluation metrics can be a barrier to the clinical adoption of ML and DL models, as clinicians and regulatory bodies may find it difficult to assess the efficacy and safety of these models without a common framework for evaluation. Future work should focus on developing user-friendly interfaces that present uncertainty information clearly to clinicians, as that will be crucial for the integration of AI models in clinical settings. Models incorporating uncertainty quantification may find it easier to gain regulatory approval, as they can provide evidence of their reliability and limitations, aligning with the requirements for safety and efficacy.

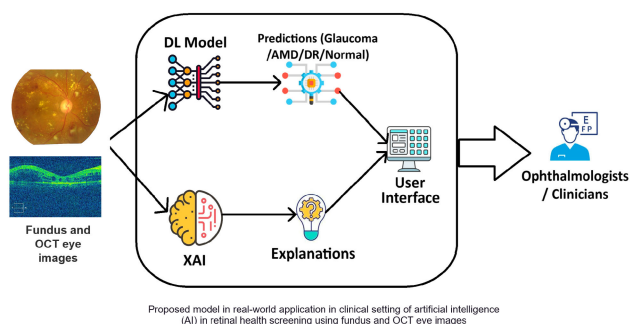
By addressing these challenges through a structured and phased approach, the aim should be to develop AI solutions that are reliable, interpretable, and broadly applicable in clinical settings, ultimately improving the diagnosis and management of retinal diseases.

XIV. CONCLUSION

Over the last decade, AI has significantly transformed retinal health detection by automating and improving the accuracy of diagnoses. In recent years, numerous automatic diagnostic support methods have been proposed with the goal of facilitating widespread screening and providing quantitative,

TABLE 19. List of papers that used DL to classify DR. E = EyePACS, A = APTOS, M = MESSIDOR, I = IDRiD, R = ResNet, G2 = GoogleNet, A2 = AlexNet, V = VGGNet.

Year	Authors	Classifier Type	Dataset	Accuracy %	Specificity %	Sensitivity%	AUC
2024	[403]	VGG16	APTOS2019	80.4		70	
2024	[403]	VGG19	APTOS2019	78.75		97.6	
2024	[403]	InceptionV3	APTOS2019	96.88		85.4	
2024	[403]	DenseNet169	APTOS2019	88.75		72.2	
2024	[316]	ResEAD2Net	STARE	98.07	99.01	90.24	0.975
2024	[316]	ResEAD2Net	DRIVE	97.55	98.01	90.07	0.976
2024	[316]	ResEAD2Net	MESSIDOR2	98.88	98.29	98.91	
2023	[317]	LSVT-Net	MESSIDOR	88.8			0.961
2023	[317]	LSVT-Net	MESSIDOR2	92.7			0.966
2023	[317]	LSVT-Net	APTOS2019	94			0.981
2023	[318]	DL (FundusQNet)	DRIMDB	99	100	98.4	0.999
2023	[142]	HemNet	DIARETDB1	97.12	94.76	90.98	0.988
2023	[404]	DL	EyePACS	75.7	79.5	75.7	0.939
2023	[404]	DL	MESSIDOR2	87	91.6	87	0.972
2023	[319]	Res-ViT	MESSIDOR-2, APTOS	89.3			0.981
2023	[321]	ensemble-ViT	MESSIDOR-2, APTOS	91.2			0.977
2023	[350]	FCN	EyePACS	87.71			
2022	[349]	CNN, E-DenseNet	E, A, M, I	91.2	96	69	
2022	[405]	VGG16	Kaggle	77.06			0.917
2022	[405]	VGG19	Kaggle	78.74			0.914
2022	[405]	InceptionV3	Kaggle	78.14			0.894
2022	[405]	ResNet50	Kaggle	79.06			0.909
2022	[405]	ResNet50V2	Kaggle	77.82			0.899
2022	[405]	ResNet152	Kaggle	79.14			0.909
2022	[405]	Xception	Kaggle	79.06			0.898
2022	[405]	InceptionResNetV2	Kaggle	79.62			0.901
2022	[405]	MobileNetV2	Kaggle	76.08			0.888
2022	[405]	DenseNet201	Kaggle	79.86			0.913
2022	[405]	EfficientNetB0	Kaggle	77.13			0.909
2022	[115]	VSUL-Net	DRIVE	96.95	98.21	83.8	0.985
2022	[115]	VSUL-Net	CHASE-DB1	97.17	98.35	81.73	0.987
2022	[115]	VSUL-Net	STARE	97.27	98.13	86.64	0.99
2022	[325]	R, G2, A2, V	Kaggle EyePACS	92.4 to 96.6		85.4 to 95.3	0.924 to 0.971
2021	[56]	ViT-DR	MESSIDOR-2				0.956
2021	[56]	ViT-DR	e-optha				0.975
2021	[56]	ViT-DR	APTOS				0.946
2021	[56]	ViT-DR	IDRiD				0.924
2021	[137]	M-CNN	IDRiD, Kaggle, MESSIDOR	99.62			
2021	[138]	CNN	Kaggle, DIARET-DB1		96.37	96.37	
2021	[327]	ViT	MESSIDOR-2, e-optha, A				0.956, 0.977, 0.947
2020	[348]	DenseNet-121	APTOS	94.9	97.1	92.6	0.88 (Kappa)
2020	[406]	AttenNet (DenseNet 169)	Z109 (public)		100	100	
2020	[406]	AttenNet (DenseNet 169)	B28K (private)		91.5	92.4	
2020	[407]	VGG-16, VGG-19	EyePACS	82	82	80	
2020	[175]	EfficientNet	APTOS				0.935 (Kappa)
2019	[335]	Inception-V3	Private	87.91	91.5	84.4	0.935
2019	[13]	Ensemble CNN	EyePACS	80.8	86.7	51.5	
2018	[328]	CNN	EyePACS	86.1	93.81	73.24	0.92
2017	[247]	ConvNet	EyePACS, e-optha, DiaretDB1				0.954, 0.949, 0.955
2017	[139]	DCNN	EyePACS	97.3, 95.9	86.3, 89.8	96.87, 96.87	
2017	[351]	DCNN	EyePACS	75.7			
2016	[408]	CKML, VNXX	Messidor, EyePACS	89.7, 89.3	90, 89.2	89.3, 90	0.891, 0.887
2016	[32]	AlexNet	Messidor-2		87	96.8	
2016	[409]	CNN	EyePACS		66.6	96.2	0.946

**FIGURE 15.** Proposed AI model which takes fundus and OCT eye images of the same patient as input and with the help of xAI makes the results more understandable for clinicians.

objective, and reproducible information on various retinal diseases such as DR, AMD, and glaucoma.

This paper presents an overview of traditional, ML, and modern DL techniques for detecting ophthalmic diseases using retinal fundus images. Traditional computer-aided diagnosis (CAD) systems have evolved into sophisticated ML and DL models, reducing the need for manual feature extraction and enabling real-time, large-scale screenings.

Our review details the main features and clinical parameters of each disease and describes various publicly available image datasets used for algorithm development. The paper also provides important critical insights and discusses research trends. Additionally, it reviews methods based on traditional image processing techniques, highlighting their crucial role in implementing pre-processing steps that are still necessary to enhance the performance of ML and DL models. We have seen from the literature, the recent trend of using DL

TABLE 20. List of acronyms used in this paper.

ADASYN	Adaptive Synthetic Sampling
AEH	Aravind Eye Hospital
AMD	Age-related Macular Degeneration
AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under Curve
ISNT	Inferior, Superior, Nasal, and Temporal
CAD	Computer-aided design
CDR	Cup-to-Disc Ratio
CFP	Color Fundus Photography
CLAHE	Contrast limited adaptive histogram equalisation
CNN	Convolutional Neural Networks
CWT	Continuous Wavelet Transform
DCN	Deep Convolutional Network
DDLS	Disc Damage Likelihood Scale
DL	Deep Learning
DT	Decision Tree
FA	Fluorescein Angiography
FoS	First-order Statistics
GAN	Generative Adversarial Networks
GDA	Glaucoma Domain Adaptation
GLCM	Grey Level Co-occurrence Matrix
GLRM	Grey Level Run Length Matrix
GMM	Gaussian Mixture Model
HMM	Hidden Markov model
HOC	Higher-order Cumulants
HOS	Higher-order Statistics
IOP	Intraocular Pressure
KMC	Kasturba Medical College
KNN	K-Nearest Neighbor
LGF	Laplacian-Gaussian Filter
LBP	Local Binary Patterns
LIME	Local Interpretable Model-Agnostic Explanations
LRP	Layer-Wise Relevance Propagation
MCA	Morphological Component Analysis
MFO-KELM	Mayfly optimization with kernel extreme learning machine
ML	Machine Learning
MSGANet-RAV	Multiscale Guided Attention Network
NN	Neural Network
NPDR	Non-proliferative diabetic retinopathy
NRR	Neuroretinal Rim
OCS	Optic Cup Segmentation
OCT	Optical Coherence Tomography
OD	Ophthalmic diseases
ODS	Optic Disc Segmentation
PCA	Principal Component Analysis
PDR	Proliferative diabetic retinopathy
PHOG	Pyramid histogram of oriented gradients
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSO	Particle Swarm Optimisation
RF	Random Forest
RH	Retinal Health
ROC	Retinopathy Online Challenge
SLO	Scanning laser ophthalmoscopy
SURF	Speeded-up robust features
SVM	Support Vector machine
UBM	Ultrasound Biomicroscopy
VEGF	Vascular endothelial growth factor
xAI	Explainable Artificial Intelligence

over ML due to its robustness and other advantages, such as no need for manual feature extractions.

Ophthalmologists have historically performed retinal screening using a labour-intensive, time-consuming manual procedure that can lead to subjective bias in the diagnosis [86]. By utilising an automated system with DL, analysis time will be shortened. Additionally, it will reduce the subjective variations in how observers interpret images [364].

Clinicians would have a better chance of diagnosing and treating these disorders if ML and DL algorithms were utilised to identify them in their early stages. Since AI models and approaches would be used, there would not be any physician subjectivity that could reduce diagnosis accuracy.

The performance of DL models in detecting glaucoma, DR, and AMD can be translated into real-world clinical settings, though challenges remain. Successful implementation

depends on integrating these models into clinical workflows, ensuring they are trained on diverse and representative datasets, and addressing regulatory and ethical considerations. However, continuous evaluation and adaptation are needed to ensure these models perform reliably across different populations and healthcare environments.

This review is useful for identifying the current main challenges and findings related to the automatic detection of each specific disease, as well as common aspects and discrepancies between various solutions developed for different diseases. While many review papers focus on the automatic detection of a single ophthalmic disease from fundus images, this comprehensive overview of the literature on all pathologies can facilitate the migration of the best solutions across different conditions, potentially leading to the development of more precise and clinically useful automatic analysis tools for all retinal diseases.

The continued development and integration of AI-based diagnostic tools in ophthalmology hold the potential to significantly improve patient outcomes and revolutionise the field of retinal health diagnostics. An automated retinal health screening system in clinical settings can be utilised to distinguish healthy eyes from ODs, hence cutting down on the amount of time needed for retinal screening sessions. Additionally, there would be less human error and no bias on the part of the clinicians. When effectively implemented, these methods would result in a faster and more consistent OD diagnosis than a human process. By addressing current limitations, such as the need for diverse datasets, explainability, and uncertainty quantification integration, researchers can continue to advance the capabilities of ML and DL systems in detecting and monitoring retinal diseases.

By overcoming existing challenges and capitalising on the promising results, researchers and clinicians can collaborate to develop more accurate, efficient, and accessible diagnostic tools for retinal diseases, explore the potential of personalised medicine in retinal health diagnostics, and ultimately improve patient outcomes and quality of life for millions of people worldwide.

APPENDIX A

The complete set of findings from the systematic review is displayed in Tables 14 to 19. In addition, the type of ML or DL in the study, the authors, the year that the article was published, and the performance are all displayed. The results presented here are based on 1601 significant research articles found in the Google Scholar, IEEE, PubMed, and Science Direct databases between January 2012 and June 2024.

ACKNOWLEDGMENT

The authors appreciate the editor's and reviewers' thoughtful comments on the article's original draft.

REFERENCES

- [1] *Blindness and Vision Impairment: 2023 Report*, WHO, Geneva, Switzerland, 2023.
- [2] *Ageing and Health. World Health Organization: 2016 Report*, WHO, Geneva, Switzerland, 2016.
- [3] M. Novak, *Issues Aging*, 3rd ed., Boston, MA, USA: Pearson, 2012.
- [4] J. K. Bambara, V. Wadley, C. Owsley, R. C. Martin, C. Porter, and L. E. Dreer, "Family functioning and low vision: A systematic review," *J. Vis. Impairment Blindness*, vol. 103, no. 3, pp. 137–149, Mar. 2009.
- [5] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," *Brit. J. Ophthalmology*, vol. 96, no. 5, pp. 614–618, May 2012.
- [6] M. Rosman, Y. Zheng, E. Lamoureux, S. M. Saw, T. Aung, J. J. Wang, P. Mitchell, and T. Y. Wong, "Review of key findings from the Singapore Malay eye study (Simes-1)," *Singap. Med. J.*, vol. 53, pp. 82–87, Feb. 2012.
- [7] *Who Releases New Global Estimates on Visual Impairment: 2012 Report*, WHO, Geneva, Switzerland, 2012.
- [8] M. Katibeh, M. Pakravan, M. Yaseri, M. Pakbin, and R. Soleimanizad, "Knowledge and awareness of age related eye diseases: A population-based survey," *Ophthalmic Epidemiology*, vol. 21, no. 5, pp. 338–345, 2014.
- [9] Y. Zheng, M. H. A. Hijazi, and F. Coenen, "Automated 'disease/no disease' grading of age-related macular degeneration by an image mining approach," *Investigative Ophthalmology Vis. Sci.*, vol. 53, no. 13, p. 8310, Dec. 2012.
- [10] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The pathophysiology and treatment of glaucoma: A review," *Jama*, vol. 311, no. 18, pp. 1901–1911, 2014.
- [11] D. S. Ting, L. Peng, A. V. Varadarajan, P. A. Keane, P. M. Burlina, M. F. Chiang, L. Schmetterer, L. R. Pasquale, N. M. Bressler, and D. R. Webster, "Deep learning in ophthalmology: The technical and clinical considerations," *Prog. Retinal Eye Res.*, vol. 72, Sep. 2019, Art. no. 100759.
- [12] N. Gour and P. Khanna, "Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network," *Biomed. Signal Process. Control*, vol. 66, Apr. 2021, Art. no. 102329.
- [13] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019.
- [14] P. Mehta, A. Lee, C. Lee, M. Balazinska, and A. Rokem, "Multilabel multiclass classification of OCT images augmented with age, gender and visual acuity data," *BioRxiv*, 2018, Art. no. 316349.
- [15] J. Y. Choi, T. K. Yoo, J. G. Seo, J. Kwak, T. T. Um, and T. H. Rim, "Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database," *PLoS ONE*, vol. 12, no. 11, Nov. 2017, Art. no. e0187336.
- [16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017.
- [17] D. S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.
- [18] C. S. Lee, D. M. Baughman, and A. Y. Lee, "Deep learning is effective for classifying normal versus age-related macular degeneration OCT images," *Ophthalmology Retina*, vol. 1, no. 4, pp. 322–327, Jul. 2017.
- [19] A. J. Tatham, R. N. Weinreb, and F. A. Medeiros, "Strategies for improving early detection of glaucoma: The combined structure-function index," *Clin. Ophthalmol.*, vol. 8, pp. 611–621, Mar. 2014.
- [20] J. U. Ahmed, S. Ahmed, A. Ahmed, and M. A. Majid, "Caring vision: Fighting against blindness in Bangladesh," *Int. Eye Sci.*, vol. 8, no. 1, pp. 17–24, 2019.
- [21] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 169–208, 2010.
- [22] Y. Li, X. Xia, and Y. M. Paulus, "Advances in retinal optical imaging," *Photonics*, vol. 5, no. 2, p. 9, Apr. 2018.
- [23] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher, "Automated detection of diabetic retinopathy on digital fundus images," *Diabetic Med.*, vol. 19, no. 2, pp. 105–112, Feb. 2002.
- [24] A. A. Sleman, A. Soliman, M. Elsharkawy, G. Giridharan, M. Ghazal, H. Sandhu, S. Schaal, R. Keynton, A. Elmaghraby, and A. El-Baz, "A novel 3D segmentation approach for extracting retinal layers from optical coherence tomography images," *Med. Phys.*, vol. 48, no. 4, pp. 1584–1595, Apr. 2021.

- [25] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," *Science*, vol. 254, no. 5035, pp. 1178–1181, 1991.
- [26] R. Brancato, F. Bandello, and R. Lattanzio, "Iris fluorescein angiography in clinical practice," *Surv. Ophthalmology*, vol. 42, no. 1, pp. 41–70, Jul. 1997.
- [27] N. N. El-Den, A. Naglah, M. Elsharkawy, M. Ghazal, N. S. Alghamdi, H. Sandhu, H. Mahdi, and A. El-Baz, "Scale-adaptive model for detection and grading of age-related macular degeneration from color retinal fundus images," *Sci. Rep.*, vol. 13, no. 1, p. 9590, Jun. 2023.
- [28] D. X. Hammer, R. Daniel Ferguson, J. C. Magill, M. A. White, A. E. Elsner, and R. H. Webb, "Image stabilization for scanning laser ophthalmoscopy," *Opt. Exp.*, vol. 10, no. 26, p. 1542, 2002.
- [29] C. J. Pavlin and F. S. Foster, "Ultrasound biomicroscopy: High-frequency ultrasound imaging of the eye at microscopic resolution," *Radiologic Clinics North Amer.*, vol. 36, no. 6, pp. 1047–1058, 1998.
- [30] M. F. Mafee, A. Karimi, J. D. Shah, M. Rapoport, and S. A. Ansari, "Anatomy and pathology of the eye: Role of MR imaging and CT," *Magn. Reson. Imag. Clinics North Amer.*, vol. 14, no. 2, pp. 249–270, May 2006.
- [31] R. Simó and C. Hernández, "Neurodegeneration in the diabetic eye: New insights and therapeutic perspectives," *Trends Endocrinology Metabolism*, vol. 25, no. 1, pp. 23–33, Jan. 2014.
- [32] M. D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, and M. Niemeijer, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investigative Ophthalmology Vis. Sci.*, vol. 57, no. 13, p. 5200, Oct. 2016.
- [33] American Diabetes Association, "Diagnosis and classification of diabetes mellitus," *Diabetes Care*, vol. 37, pp. S81–S90, Dec. 2013.
- [34] J. M. Forbes and M. E. Cooper, "Mechanisms of diabetic complications," *Physiological Rev.*, vol. 93, no. 1, pp. 137–188, 2013.
- [35] D. A. Antonetti, R. Klein, and T. W. Gardner, "Diabetic retinopathy," *New England J. Med.*, vol. 366, no. 13, pp. 1227–1239, 2012.
- [36] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kalviainen, and J. Pietila, "The DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Proc. Brit. Mach. Vis. Conf.*, 2007, pp. 15.1–15.10.
- [37] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, p. 3704, May 2021.
- [38] S. Haggag, A. Elnakib, A. Sharafeldeen, M. Elsharkawy, F. Khalifa, R. K. Farag, M. A. Mohamed, H. S. Sandhu, W. Mansoor, A. Sewelam, and A. El-Baz, "A computer-aided diagnostic system for diabetic retinopathy based on local and global extracted features," *Appl. Sci.*, vol. 12, no. 16, p. 8326, Aug. 2022.
- [39] M. Elsharkawy, A. Sharafeldeen, A. Soliman, F. Khalifa, M. Ghazal, E. El-Daydamony, A. Atwan, H. S. Sandhu, and A. El-Baz, "A novel computer-aided diagnostic system for early detection of diabetic retinopathy using 3D-OCT higher-order spatial appearance model," *Diagnostics*, vol. 12, no. 2, p. 461, Feb. 2022.
- [40] L. A. Levin and P. Peeples, "History of neuroprotection and rationale as a therapy for glaucoma," *Amer. J. Managed Care*, vol. 14, no. 1, pp. S11–S14, 2008.
- [41] R. J. Casson, G. Chidlow, J. P. Wood, J. G. Crowston, and I. Goldberg, "Definition of glaucoma: Clinical and experimental concepts," *Clin. Experim. Ophthalmology*, vol. 40, no. 4, pp. 341–349, May 2012.
- [42] J. B. Jonas, T. Aung, R. R. Bourne, A. M. Bron, R. Ritch, and S. Panda-Jonas, "Glaucoma," *Lancet*, vol. 390, no. 10108, pp. 2183–2193, 2017.
- [43] N. Tamim, M. Elsharkawy, and H. Nassar, "Accurate diagnosis of diabetic retinopathy and normal retinal images using hybrid features and genetic algorithm," *Appl. Sci.*, vol. 11, p. 6178, 2021.
- [44] J. Ambati, J. P. Atkinson, and B. D. Gelfand, "Mechanisms of age-related macular degeneration," *Nature Rev. Immunology*, vol. 13, no. 6, pp. 438–451, 2013.
- [45] H. Kolb, *Simple Anatomy of the Retina*, 1995.
- [46] R. van Leeuwen, C. C. Klaver, J. R. Vingerling, A. Hofman, and P. T. de Jong, "Epidemiology of age-related maculopathy: A review," *Eur. J. Epidemiology*, vol. 18, no. 9, pp. 845–854, 2003. [Online]. Available: <https://link.springer.com/article/10.1023/A:1025643303914>
- [47] J. Ambati, J. P. Atkinson, and B. D. Gelfand, "Immunology of age-related macular degeneration," *Nature Rev. Immunology*, vol. 13, no. 6, pp. 438–451, 2013.
- [48] M. Elsharkawy, M. Elrazzaz, M. Ghazal, M. Alhalabi, A. Soliman, A. Mahmoud, E. El-Daydamony, A. Atwan, A. Thanos, H. S. Sandhu, G. Giridharan, and A. El-Baz, "Role of optical coherence tomography imaging in predicting progression of age-related macular disease: A survey," *Diagnostics*, vol. 11, no. 12, p. 2313, Dec. 2021.
- [49] U. Schmidt-Erfurth and S. M. Waldstein, "A paradigm shift in imaging biomarkers in neovascular age-related macular degeneration," *Prog. Retinal Eye Res.*, vol. 50, pp. 1–24, Jan. 2016.
- [50] S. Chelaramani, M. Gupta, V. Agarwal, P. Gupta, and R. Habash, "Multi-task learning for fine-grained eye disease prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 734–749.
- [51] T. Walter, J. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236–1243, Oct. 2002.
- [52] G. G. Gardner, D. Keating, T. H. Williamson, and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool," *Brit. J. Ophthalmology*, vol. 80, no. 11, pp. 940–944, Nov. 1996.
- [53] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016.
- [54] M. Wassel, A. M. Hamdi, N. Adly, and M. Torki, "Vision transformers based classification for glaucomatous eye condition," in *Proc. 26th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2022, pp. 5082–5088.
- [55] Z. Jiang, L. Wang, Q. Wu, Y. Shao, M. Shen, W. Jiang, and C. Dai, "Computer-aided diagnosis of retinopathy based on vision transformer," *J. Innov. Opt. Health Sci.*, vol. 15, no. 2, Mar. 2022, Art. no. 2250009.
- [56] J. Wu, R. Hu, Z. Xiao, J. Chen, and J. Liu, "Vision transformer-based recognition of diabetic retinopathy grade," *Med. Phys.*, vol. 48, no. 12, pp. 7850–7863, Dec. 2021.
- [57] S. Alayón, J. Hernández, F. J. Fumero, J. F. Sigut, and T. Díaz-Alemán, "Comparison of the performance of convolutional neural networks and vision transformer-based systems for automated glaucoma detection with eye fundus images," *Appl. Sci.*, vol. 13, no. 23, p. 12722, Nov. 2023.
- [58] A. Parvaiz, M. A. Khalid, R. Zafar, H. Ameer, M. Ali, and M. M. Fraz, "Vision transformers in medical computer vision—A contemplative retrospection," *Eng. Appl. Artif. Intell.*, vol. 122, Jun. 2023, Art. no. 106126.
- [59] N. Vannadil and P. Kokil, "Automated age-related macular degeneration diagnosis in retinal fundus images via VIT," in *Proc. Int. Conf. Mach. Learn., Deep Learn. Comput. Intell. Wireless Commun.* Switzerland: Springer, 2012, pp. 271–280. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-47942-7_24
- [60] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereology*, vol. 33, no. 3, p. 231, Aug. 2014.
- [61] D. Marin, M. E. Gegundez-Arias, A. Suero, and J. M. Bravo, "Obtaining optic disc center and pixel region by automatic thresholding methods on morphologically processed fundus images," *Comput. Methods Programs Biomed.*, vol. 118, no. 2, pp. 173–185, Feb. 2015.
- [62] F. Fumero, S. Alayon, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," in *Proc. 24th Int. Symp. Computer-Based Med. Syst. (CBMS)*, Jun. 2011, pp. 1–6.
- [63] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *Int. J. Biomed. Imag.*, vol. 2013, pp. 1–11, Jun. 2013.
- [64] D. Marin, A. Aquino, M. E. Gegundez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Trans. Med. Imag.*, vol. 30, no. 1, pp. 146–158, Jan. 2011.
- [65] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
- [66] N. Sikder, M. S. Chowdhury, A. Shamim Mohammad Arif, and A.-A. Nahid, "Early blindness detection based on retinal images using ensemble learning," in *Proc. 22nd Int. Conf. Comput. Inf. Technol. (ICCIIT)*, Dec. 2019, pp. 1–6.

- [67] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotequi, G. Quéllec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ, and A. Chabouis, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *IRBM*, vol. 34, no. 2, pp. 196–203, Apr. 2013.
- [68] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informat. Med. Unlocked*, vol. 20, Aug. 2020, Art. no. 100377.
- [69] L. Li, M. Xu, X. Wang, L. Jiang, and H. Liu, "Attention based glaucoma detection: A large-scale database and CNN model," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10563–10572.
- [70] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.
- [71] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabudhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRID): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, Jul. 2018.
- [72] K. C. Jordan, M. Menolotto, N. M. Bolster, I. A. Livingstone, and M. E. Giardini, "A review of feature-based retinal image analysis," *Expert Rev. Ophthalmology*, vol. 12, no. 3, pp. 207–220, May 2017.
- [73] A. Diaz-Pinto, A. Colomer, V. Naranjo, S. Morales, Y. Xu, and A. F. Frangi, "Retinal image synthesis for glaucoma assessment using DCGAN and VAE models," in *Proc. 19th Int. Conf.*, Madrid, Spain, Switzerland: Springer, Nov. 2018, pp. 224–232. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-03493-1_24
- [74] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Tae, H. Pratt, and Y. Zheng, "Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis," *Symmetry*, vol. 10, no. 4, p. 87, Mar. 2018.
- [75] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB0: Evaluation database and methodology for diabetic retinopathy algorithms," *Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta Univ. Technol., Finland*, vol. 73, pp. 1–17, May 2006. [Online]. Available: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=bd7d2380e76fb9dfd367d669e311d4913f67f7d2>
- [76] X. Luo, J. Li, M. Chen, X. Yang, and X. Li, "Ophthalmic disease detection via deep learning with a novel mixture loss function," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 9, pp. 3332–3339, Sep. 2021.
- [77] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [78] A. K. Feeny, M. Tadarati, D. E. Freund, N. M. Bressler, and P. Burlina, "Automated segmentation of geographic atrophy of the retinal epithelium via random forests in AREDS color fundus images," *Comput. Biol. Med.*, vol. 65, pp. 124–136, Oct. 2015.
- [79] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. W. Pluim, R. Duits, and B. M. T. H. Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE Trans. Med. Imag.*, vol. 35, no. 12, pp. 2631–2644, Dec. 2016.
- [80] D. Bhulakshmi and D. S. Rajput, "A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images," *PeerJ Comput. Sci.*, vol. 10, p. e1947, Apr. 2024.
- [81] A. A. Soofi and Fazal-e-Amin, "Exploring deep learning techniques for glaucoma detection: A comprehensive review," 2023, *arXiv:2311.01425*.
- [82] X. Leng, R. Shi, Y. Wu, S. Zhu, X. Cai, X. Lu, and R. Liu, "Deep learning for detection of age-related macular degeneration: A systematic review and meta-analysis of diagnostic test accuracy studies," *PLoS ONE*, vol. 18, no. 4, Apr. 2023, Art. no. e0284060.
- [83] M. Zedan, M. Zulkifley, A. Ibrahim, A. Moubark, N. Kamari, and S. Abdani, "Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review," *Diagnostics*, vol. 13, no. 13, p. 2180, Jun. 2023.
- [84] D. M. S. Barros, J. C. C. Moura, C. R. Freire, A. C. Taleb, R. A. M. Valentim, and P. S. G. Morais, "Machine learning applied to retinal image processing for glaucoma detection: Review and perspective," *Biomed. Eng. OnLine*, vol. 19, no. 1, p. 20, Dec. 2020.
- [85] M. M. Islam, H.-C. Yang, T. N. Poly, W.-S. Jian, and Y.-C. Jack Li, "Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis," *Comput. Methods Programs Biomed.*, vol. 191, Jul. 2020, Art. no. 105320.
- [86] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, and V. Lakshminarayanan, "Ophthalmic diagnosis using deep learning with fundus images—A critical review," *Artif. Intell. Med.*, vol. 102, Jan. 2020, Art. no. 101758.
- [87] E. Pead, R. Megaw, J. Cameron, A. Fleming, B. Dhillon, E. Trucco, and T. MacGillivray, "Automated detection of age-related macular degeneration in color fundus photography: A systematic review," *Surv. Ophthalmology*, vol. 64, no. 4, pp. 498–511, Jul. 2019.
- [88] T. Saba, S. T. F. Bokhari, M. Sharif, M. Yasmin, and M. Raza, "Fundus image classification methods for the detection of glaucoma: A review," *Microsc. Res. Technique*, vol. 81, no. 10, pp. 1105–1121, Oct. 2018.
- [89] T. A. Soomro, J. Gao, T. Khan, A. F. M. Hani, M. A. U. Khan, and M. Paul, "Computerised approaches for the detection of diabetic retinopathy using retinal fundus images: A survey," *Pattern Anal. Appl.*, vol. 20, no. 4, pp. 927–961, Nov. 2017.
- [90] M. C. V. Stella Mary, E. B. Rajsingh, and G. R. Naik, "Retinal fundus image analysis for diagnosis of glaucoma: A comprehensive survey," *IEEE Access*, vol. 4, pp. 4327–4354, 2016.
- [91] A. Bhuiyan, "A review of disease grading and remote diagnosis for sight threatening eye condition: Age related macular degeneration," *J. Comput. Sci. Syst. Biol.*, vol. 7, no. 2, pp. 62–71, 2014.
- [92] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2136–2155, Dec. 2013.
- [93] Z. Zhang, R. Srivastava, H. Liu, X. Chen, L. Duan, D. W. Kee Wong, C. K. Kwok, T. Y. Wong, and J. Liu, "A survey on computer aided diagnosis for ocular diseases," *BMC Med. Informat. Decis. Making*, vol. 14, no. 1, p. 80, Dec. 2014.
- [94] D. Bizios, A. Heijl, J. L. Hougaard, and B. Bengtsson, "Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by stratus OCT," *Acta Ophthalmologica*, vol. 88, no. 1, pp. 44–52, Feb. 2010.
- [95] G. Nikita, C. Mehmet, and H. Bryan, "Machine learning identification of diabetic retinopathy from fundus images," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, 2014, pp. 1–7.
- [96] C. Zhu, B. Zou, R. Zhao, J. Cui, X. Duan, Z. Chen, and Y. Liang, "Retinal vessel segmentation in colour fundus images using extreme learning machine," *Computerized Med. Imag. Graph.*, vol. 55, pp. 68–77, Jan. 2017.
- [97] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, "Fundus image classification using VGG-19 architecture with PCA and SVD," *Symmetry*, vol. 11, no. 1, p. 1, Dec. 2018.
- [98] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "DREAM: Diabetic retinopathy analysis using machine learning," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 5, pp. 1717–1728, Sep. 2014.
- [99] P. Burlina, K. D. Pacheco, N. Joshi, D. E. Freund, and N. M. Bressler, "Comparing humans and deep learning performance for grading AMD: A study in using universal deep features and transfer learning for automated AMD analysis," *Comput. Biol. Med.*, vol. 82, pp. 80–86, Mar. 2017.
- [100] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Proc. Comput. Sci.*, vol. 90, pp. 200–205, Jan. 2016.
- [101] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, "Deep learning in medical imaging: General overview," *Korean J. Radiol.*, vol. 18, no. 4, p. 570, 2017.
- [102] S. N. Sangeetha and P. Uma Maheswari, "An intelligent model for blood vessel segmentation in diagnosing DR using CNN," *J. Med. Syst.*, vol. 42, no. 10, p. 175, Oct. 2018.
- [103] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: An extensive validation," *Biomed. Eng. OnLine*, vol. 18, no. 1, p. 29, Dec. 2019.
- [104] S. Nawaldgi and Y. S. Lalitha, "Automated glaucoma assessment from color fundus images using structural and texture features," *Biomed. Signal Process. Control*, vol. 77, Aug. 2022, Art. no. 103875.

- [105] R. Shinde, "Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms," *Intelligence-Based Med.*, vol. 5, Aug. 2021, Art. no. 100038.
- [106] N. Thakur and M. Juneja, "Classification of glaucoma using hybrid features with machine learning approaches," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102137.
- [107] Z. U. Rehman, S. S. Naqvi, T. M. Khan, M. Arsalan, M. A. Khan, and M. A. Khalil, "Multi-parametric optic disc segmentation using superpixel based feature classification," *Exp. Syst. Appl.*, vol. 120, pp. 461–473, Apr. 2019.
- [108] J. Koh, E. Ng, S. Bhandary, A. Laude, and U. Acharya, "Automated detection of retinal health using PHOG and SURF features extracted from fundus images," *Appl. Intell.*, vol. 48, no. 5, pp. 1379–1393, 2018.
- [109] J. E. W. Koh, E. Y. K. Ng, S. V. Bhandary, Y. Hagiwara, A. Laude, and U. R. Acharya, "Automated retinal health diagnosis using pyramid histogram of visual words and Fisher vector techniques," *Comput. Biol. Med.*, vol. 92, pp. 204–209, Jan. 2018.
- [110] U. R. Acharya, S. Bhat, J. E. W. Koh, S. V. Bhandary, and H. Adeli, "A novel algorithm to detect glaucoma risk using texture and local configuration pattern features extracted from fundus images," *Comput. Biol. Med.*, vol. 88, pp. 72–83, Sep. 2017.
- [111] U. R. Acharya, M. R. K. Mookiah, J. E. W. Koh, J. H. Tan, S. V. Bhandary, A. K. Rao, H. Fujita, Y. Hagiwara, C. K. Chua, and A. Laude, "Automated screening system for retinal health using bi-dimensional empirical mode decomposition and integrated index," *Comput. Biol. Med.*, vol. 75, pp. 54–62, Aug. 2016.
- [112] A. Singh, M. K. Dutta, M. ParthaSarathi, V. Uher, and R. Burget, "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image," *Comput. Methods Programs Biomed.*, vol. 124, pp. 108–120, Feb. 2016.
- [113] D. Das, D. R. Nayak, S. V. Bhandary, and U. R. Acharya, "CDAM-net: Channel shuffle dual attention based multi-scale CNN for efficient glaucoma detection using fundus images," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108454.
- [114] F. Song, W. Zhang, Y. Zheng, D. Shi, and M. He, "A deep learning model for generating fundus autofluorescence images from color fundus photography," *Adv. Ophthalmology Pract. Res.*, vol. 3, no. 4, pp. 192–198, Nov. 2023.
- [115] M. Arsalan, A. Haider, J. H. Koo, and K. R. Park, "Segmenting retinal vessels using a shallow segmentation network to aid ophthalmic analysis," *Mathematics*, vol. 10, no. 9, p. 1536, May 2022.
- [116] P. Mehta, C. A. Petersen, J. C. Wen, M. R. Banitt, P. P. Chen, K. D. Bojikian, C. Egan, S.-I. Lee, M. Balazinska, A. Y. Lee, and A. Rokem, "Automated detection of glaucoma with interpretable machine learning using clinical data and multimodal retinal images," *Amer. J. Ophthalmology*, vol. 231, pp. 154–169, Nov. 2021.
- [117] M. Abe, "Assisting glaucoma diagnosis with optical coherence tomography and color fundus images using machine learning approach," *Investigative Ophthalmology Vis. Sci.*, vol. 59, p. 2079, Jul. 2018.
- [118] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 715–718.
- [119] M. R. K. Mookiah, U. R. Acharya, J. E. W. Koh, C. K. Chua, J. H. Tan, V. Chandran, C. M. Lim, K. Noronha, A. Laude, and L. Tong, "Decision support system for age-related macular degeneration using discrete wavelet transform," *Med. Biol. Eng. Comput.*, vol. 52, no. 9, pp. 781–796, Sep. 2014.
- [120] N. Prabhu, D. Bhoir, and N. Shanbhag, "Diabetic retinopathy screening using machine learning for hierarchical classification," *Int. J. Innov. Technol. Exploring Eng.*, vol. 8, no. 10, pp. 1943–1948, 2019.
- [121] M. H. Mahmoud, S. Alameery, H. Fouad, A. Altinawi, and A. E. Youssef, "An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm," *Pers. Ubiquitous Comput.*, vol. 27, no. 3, pp. 751–765, Jun. 2023.
- [122] S. Maheshwari, R. B. Pachori, V. Kanhangad, S. V. Bhandary, and U. R. Acharya, "Iterative variational mode decomposition based automated detection of glaucoma using fundus images," *Comput. Biol. Med.*, vol. 88, pp. 142–149, Sep. 2017.
- [123] S. Maheshwari, R. B. Pachori, and U. R. Acharya, "Automated diagnosis of glaucoma using empirical wavelet transform and coreentropy features extracted from fundus images," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 3, pp. 803–813, May 2017.
- [124] H. Wang and D. Hu, "Comparison of SVM and LS-SVM for regression," in *Proc. Int. Conf. Neural New. Brain*, vol. 1, 2005, pp. 279–283.
- [125] R. Bock, J. Meier, L. G. Nyúl, J. Hornegger, and G. Michelson, "Glaucoma risk index: Automated glaucoma detection from color fundus images," *Med. Image Anal.*, vol. 14, no. 3, pp. 471–481, Jun. 2010.
- [126] A. García-Florián, Á. Ferreira-Santiago, O. Camacho-Nieto, and C. Yáñez-Márquez, "A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images," *Comput. Electr. Eng.*, vol. 75, pp. 218–229, May 2019.
- [127] K. Argade, K. Deshmukh, M. Narkhede, N. Sonawane, and S. Jore, "Automatic detection of diabetic retinopathy using image processing and data mining techniques," in *Proc. Int. Conf. Green Comput. Internet Things (ICGCIoT)*, 2015, pp. 517–521.
- [128] J. Nandy, W. Hsu, and M. Lee, "An incremental feature extraction framework for referable diabetic retinopathy detection," in *Proc. IEEE 28th Int. Conf. Tools Artif. Intell. (ICTAI)*, May 2016, pp. 908–912.
- [129] L. Jain, H. V. S. Murthy, C. Patel, and D. Bansal, "Retinal eye disease detection using deep learning," in *Proc. 14th Int. Conf. Inf. Process. (ICINPRO)*, Dec. 2018, pp. 1–6.
- [130] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [131] J. H. Tan, U. R. Acharya, S. V. Bhandary, K. C. Chua, and S. Sivaprasad, "Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network," *J. Comput. Sci.*, vol. 20, pp. 70–79, May 2017.
- [132] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1153–1159, May 2016.
- [133] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
- [134] S. Kaymak and A. Serener, "Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning," in *Proc. IEEE 14th Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Cluj-Napoca, Romania, Sep. 2018, pp. 265–269.
- [135] N. B. A. K. Balasubramanian, "Deep learning approach to predict and grade glaucoma from fundus images through constitutional neural networks," in *Artificial Intelligence Trends for Data Analytics Using Machine Learning and Deep Learning Approaches*, vol. 1, 2020, pp. 71–86.
- [136] L. Li, M. Xu, H. Liu, Y. Li, X. Wang, L. Jiang, Z. Wang, X. Fan, and N. Wang, "A large-scale database and a CNN model for attention-based glaucoma detection," *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 413–424, Feb. 2020.
- [137] S. Gayathri, V. P. Gopi, and P. Palanisamy, "Diabetic retinopathy classification based on multipath CNN and machine learning classifiers," *Phys. Eng. Sci. Med.*, vol. 44, no. 3, pp. 639–653, Sep. 2021.
- [138] L. Math and R. Fatima, "Adaptive machine learning classification for diabetic retinopathy," *Multimedia Tools Appl.*, vol. 80, no. 4, pp. 5173–5186, Feb. 2021.
- [139] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *Medical Image Computing and Computer Assisted Intervention*. Switzerland: Springer, Sep. 2017, pp. 533–540. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-66179-7_61
- [140] D. U. N. Qomariah, H. Tjandrasa, and C. Fatichah, "Classification of diabetic retinopathy and normal retinal images using CNN and SVM," in *Proc. 12th Int. Conf. Inf. Commun. Technol. Syst. (ICTS)*, Jul. 2019, pp. 152–157.
- [141] Z. Li, C. Guo, D. Lin, D. Nie, Y. Zhu, C. Chen, L. Zhao, J. Wang, X. Zhang, M. Dongye, D. Wang, F. Xu, C. Jin, P. Zhang, Y. Han, P. Yan, Y. Han, and H. Lin, "Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images," *Brit. J. Ophthalmology*, vol. 105, no. 11, pp. 1548–1554, Nov. 2021.
- [142] T. Aziz, C. Charoenlarnopparut, and S. Mahapakulchai, "Deep learning-based hemorrhage detection for diabetic retinopathy screening," *Sci. Rep.*, vol. 13, no. 1, p. 1479, Jan. 2023.

- [143] F. Girard, T. Hurtut, C. Kavalec, and F. Cheriet, "Atlas-based score for automatic glaucoma risk stratification," *Computerized Med. Imag. Graph.*, vol. 87, Jan. 2021, Art. no. 101797.
- [144] N. Thakur and M. Juneja, "Optic disc and optic cup segmentation from retinal images using hybrid approach," *Expert Syst. Appl.*, vol. 127, pp. 308–322, Aug. 2019.
- [145] M. S. Haleem, L. Han, J. V. Hemert, B. Li, A. Fleming, L. R. Pasquale, and B. J. Song, "A novel adaptive deformable model for automated optic disc and cup segmentation to aid glaucoma diagnosis," *J. Med. Syst.*, vol. 42, no. 1, pp. 1–18, Jan. 2018.
- [146] A. Chakravarty and J. Sivaswamy, "Joint optic disc and cup boundary extraction from monocular fundus images," *Comput. Methods Programs Biomed.*, vol. 147, pp. 51–61, Aug. 2017.
- [147] M. P. N. Maldhure and P. V. V. Dixit, "Glaucoma detection using optic cup and optic disc segmentation," *Int. J. Eng. Trends Technol.*, vol. 20, no. 2, pp. 52–55, Feb. 2015.
- [148] S. Chandrika and K. Nirmala, "Analysis of CDR detection for glaucoma diagnosis," *Int. J. Eng. Res. Appl.*, vol. 2, no. 4, pp. 23–27, 2013.
- [149] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1019–1032, Jun. 2013.
- [150] R. Ingle and P. Mishra, "Cup segmentation by gradient method for the assessment of glaucoma from retinal image," *Int. J. Eng. Trends Technol.*, vol. 4, no. 6, pp. 2540–2543, 2013.
- [151] K. Narasimhan, K. Vijayarekha, K. JogiNarayana, P. SivaPrasad, and V. SatishKumar, "Glaucoma detection from fundus image using opencv," *Res. J. Appl. Sci., Eng. Technol.*, vol. 4, no. 24, pp. 5459–5463, 2012.
- [152] S. Yan, X. Pan, and Y. Wang, "MRSNet: Joint consistent optic disc and cup segmentation based on large kernel residual convolutional attention and self-attention," *Digit. Signal Process.*, vol. 145, Feb. 2024, Art. no. 104308.
- [153] Y. Yi, Y. Jiang, B. Zhou, N. Zhang, J. Dai, X. Huang, Q. Zeng, and W. Zhou, "C2FTFNet: Coarse-to-fine transformer network for joint optic disc and cup segmentation," *Comput. Biol. Med.*, vol. 164, Sep. 2023, Art. no. 107215.
- [154] T. Shyamalee and D. Meedeniya, "Attention U-Net for glaucoma identification using fundus image segmentation," in *Proc. Int. Conf. Decis. Aid Sci. Appl. (DASA)*, Mar. 2022, pp. 6–10.
- [155] L. Liu, Z. Zhang, S. Li, K. Ma, and Y. Zheng, "S-CUDA: self-cleansing unsupervised domain adaptation for medical image segmentation," *Med. Image Anal.*, vol. 74, Dec. 2021, Art. no. 102214.
- [156] Y. Zhang, X. Cai, Y. Zhang, H. Kang, X. Ji, and X. Yuan, "TAU: Transferable attention U-Net for optic disc and cup segmentation," *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106668.
- [157] S. Wang, L. Yu, X. Yang, C.-W. Fu, and P.-A. Heng, "Patch-based output space adversarial learning for joint optic disc and cup segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2485–2495, Nov. 2019.
- [158] V. G. Edupuganti, A. Chawla, and A. Kale, "Automatic optic disk and cup segmentation of fundus images using deep learning," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 2227–2231.
- [159] A. Sevastopolsky, "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network," *Pattern Recognit. Image Anal.*, vol. 27, no. 3, pp. 618–624, Jul. 2017.
- [160] J. Zilly, J. M. Buhmann, and D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," *Computerized Med. Imag. Graph.*, vol. 55, pp. 28–41, Jan. 2017.
- [161] D. Mittal and K. Kumari, "Automated detection and segmentation of drusen in retinal fundus images," *Comput. Electr. Eng.*, vol. 47, pp. 82–95, Oct. 2015.
- [162] H. X. Y. Fu, S. Lin, D. W. K. Wong, and L. J. Deepvessel, "Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 132–139.
- [163] H. Fu, J. Cheng, Y. Xu, and J. Liu, "Glaucoma detection based on deep learning network in fundus image," in *Proc. Deep Learn. Convolutional Neural Netw. Med. Imag. Clin. Inform.*, 2019, pp. 119–137.
- [164] R. Kapoor, B. T. Whigham, and L. A. Al-Aswad, "The role of artificial intelligence in the diagnosis and management of glaucoma," *Current Ophthalmology Rep.*, vol. 7, pp. 136–142, Mar. 2019.
- [165] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, and M. He, "Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs," *Ophthalmology*, vol. 125, no. 8, pp. 1199–1206, 2018.
- [166] T. Walter and J.-C. Klein, "Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques," in *Proc. Int. Symp. Med. Data Anal. Germany: Springer*, 2001, pp. 282–287. [Online]. Available: https://link.springer.com/chapter/10.1007/3-540-45497-7_43
- [167] D. Welfer, J. Scharcanski, C. M. Kitamura, M. M. Dal Pizzol, L. W. B. Ludwig, and D. R. Marinho, "Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach," *Comput. Biol. Med.*, vol. 40, no. 2, pp. 124–137, Feb. 2010.
- [168] A. Dehghani, H. A. Moghaddam, and M.-S. Moin, "Optic disc localization in retinal images using histogram matching," *EURASIP J. Image Video Process.*, vol. 2012, no. 1, pp. 1–11, Dec. 2012.
- [169] S. Sekhar, W. Al-Nuaimy, and A. K. Nandi, "Automated localisation of retinal optic disk using Hough transform," in *Proc. 5th IEEE Int. Symp. Biomed. Imaging: From Nano to Macro*, May 2008, pp. 1577–1580.
- [170] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1189–1195, Oct. 2004.
- [171] M. Mishra, M. K. Nath, and S. Dandapat, "Glaucoma detection from color fundus images," *Int. J. Comput. Commun. Technol. (IJCCT)*, vol. 2, no. 6, pp. 7–10, 2011.
- [172] I. Mehdi, D. E. C. Belkhiat, and D. Jabri, "A high accuracy segmentation method for retinal blood vessel detection based on hybrid filters and an adaptive thresholding," *J. Ambient Intell. Humanized Comput.*, vol. 15, no. 1, pp. 323–335, Jan. 2024.
- [173] N. A. Alkhaldi and H. T. Halawani, "Intelligent machine learning enabled retinal blood vessel segmentation and classification," *Comput., Mater. Continua*, vol. 74, no. 1, pp. 399–414, 2023.
- [174] P. Kuppusamy, M. M. Basha, and C.-L. Hung, "Retinal blood vessel segmentation using random forest with Gabor and Canny edge features," in *Proc. Int. Conf. Smart Technol. Syst. Next Gener. Comput. (ICSTSN)*, Mar. 2022, pp. 1–4.
- [175] Z. Zhang, "Deep-learning-based early detection of diabetic retinopathy on fundus photography using EfficientNet," in *Proc. the 4th Int. Conf. Innov. Artif. Intell.*, May 2020, pp. 70–74.
- [176] S. K. Saroj, R. Kumar, and N. P. Singh, "Fréchet PDF based matched filter approach for retinal blood vessels segmentation," *Comput. Methods Programs Biomed.*, vol. 194, Oct. 2020, Art. no. 105490.
- [177] C. Zhou, X. Zhang, and H. Chen, "A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model," *Comput. Methods Programs Biomed.*, vol. 187, Apr. 2020, Art. no. 105231.
- [178] X. Wang, X. Jiang, and J. Ren, "Blood vessel segmentation from fundus image by a cascade classification framework," *Pattern Recognit.*, vol. 88, pp. 331–341, Apr. 2019.
- [179] A. Khawaja, T. M. Khan, K. Naveed, S. S. Naqvi, N. U. Rehman, and S. J. Nawaz, "An improved retinal vessel segmentation framework using frangi filter coupled with the probabilistic patch based denoiser," *IEEE Access*, vol. 7, pp. 164344–164361, 2019.
- [180] H. Xia, F. Jiang, S. Deng, J. Xin, and R. Doss, "Mapping functions driven robust retinal vessel segmentation via training patches," *IEEE Access*, vol. 6, pp. 61973–61982, 2018.
- [181] M. Javidi, H.-R. Pourreza, and A. Harati, "Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation," *Comput. Methods Programs Biomed.*, vol. 139, pp. 93–108, Feb. 2017.
- [182] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 16–27, Jan. 2017.
- [183] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Iterative vessel segmentation of fundus images," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1738–1749, Jul. 2015.
- [184] E. Imani, M. Javidi, and H.-R. Pourreza, "Improvement of retinal blood vessel detection using morphological component analysis," *Comput. Methods Programs Biomed.*, vol. 118, no. 3, pp. 263–279, Mar. 2015.
- [185] Y. Wang, G. Ji, P. Lin, and E. Trucco, "Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition," *Pattern Recognit.*, vol. 46, no. 8, pp. 2117–2133, Aug. 2013.

- [186] M. U. Akram and S. A. Khan, "Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy," *Eng. Comput.*, vol. 29, no. 2, pp. 165–173, Apr. 2013.
- [187] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, Sep. 2012.
- [188] L. K. Singh, M. Khanna, S. Thawkar, and R. Singh, "Deep-learning based system for effective and automatic blood vessel segmentation from retinal fundus images," *Multimedia Tools Appl.*, vol. 83, no. 2, pp. 6005–6049, Jan. 2024.
- [189] C. P. Darmo, L. I. Kesuma, and D. Geovani, "The combination of black hat transform and U-Net in image enhancement and blood vessel segmentation in retinal images," *Comput. Eng. Appl. J.*, vol. 12, no. 3, pp. 129–146, Oct. 2023.
- [190] S. V. Deshmukh and A. Roy, *Retinal Blood Vessel Segmentation Based on Modified CNN and Analyze the Perceptual Quality of Segmented Images*. Springer, 2022, pp. 609–625.
- [191] E. Abdelmaksoud, S. El-Sappagh, S. Barakat, T. Abuhmed, and M. Elmogy, "Automatic diabetic retinopathy grading system based on detecting multiple retinal lesions," *IEEE Access*, vol. 9, pp. 15939–15960, 2021.
- [192] S. Feng, Z. Zhuo, D. Pan, and Q. Tian, "CcNet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features," *Neurocomputing*, vol. 392, pp. 268–276, Jun. 2020.
- [193] Z. Zhuo, J. Huang, K. Lu, D. Pan, and S. Feng, "A size-invariant convolutional network with dense connectivity applied to retinal vessel segmentation measured by a unique index," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105508.
- [194] X. Li, Y. Jiang, M. Li, and S. Yin, "Lightweight attention convolutional neural network for retinal vessel image segmentation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1958–1967, Mar. 2021.
- [195] D. Chen, Y. Ao, and S. Liu, "Semi-supervised learning method of U-Net deep learning network for blood vessel segmentation in retinal images," *Symmetry*, vol. 12, no. 7, p. 1067, Jun. 2020.
- [196] D. Wang, A. Haytham, J. Pottenburgh, O. Saeedi, and Y. Tao, "Hard attention net for automatic retinal vessel segmentation," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 12, pp. 3384–3396, Dec. 2020.
- [197] D. Yang, G. Liu, M. Ren, B. Xu, and J. Wang, "A multi-scale feature fusion method based on U-Net for retinal vessel segmentation," *Entropy*, vol. 22, no. 8, p. 811, Jul. 2020.
- [198] P. Yin, R. Yuan, Y. Cheng, and Q. Wu, "Deep guidance network for biomedical image segmentation," *IEEE Access*, vol. 8, pp. 116106–116116, 2020.
- [199] G. Alfonso Francia, C. Pedraza, M. Aceves, and S. Tovar-Arriaga, "Chaining a U-Net with a residual U-Net for retinal blood vessels segmentation," *IEEE Access*, vol. 8, pp. 38493–38500, 2020.
- [200] A. Farahani and H. Mohseni, "Medical image segmentation using customized U-Net with adaptive activation functions," *Neural Comput. Appl.*, vol. 33, no. 11, pp. 6307–6323, Jun. 2021.
- [201] X. Guo, C. Chen, Y. Lu, K. Meng, H. Chen, K. Zhou, Z. Wang, and R. Xiao, "Retinal vessel segmentation combined with generative adversarial networks and dense U-Net," *IEEE Access*, vol. 8, pp. 194551–194560, 2020.
- [202] Y. Lv, H. Ma, J. Li, and S. Liu, "Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation," *IEEE Access*, vol. 8, pp. 32826–32839, 2020.
- [203] K. J. Noh, S. J. Park, and S. Lee, "Scale-space approximated convolutional neural networks for retinal vessel segmentation," *Comput. Methods Programs Biomed.*, vol. 178, pp. 237–246, Sep. 2019.
- [204] M. Z. Alom, C. Yakopcic, M. Hasan, T. M. Taha, and V. K. Asari, "Recurrent residual U-net for medical image segmentation," *J. Med. Imag.*, vol. 6, no. 1, 2019, Art. no. 014006.
- [205] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "CE-net: Context encoder network for 2D medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 38, no. 10, pp. 2281–2292, Oct. 2019.
- [206] S. Lian, L. Li, G. Lian, X. Xiao, Z. Luo, and S. Li, "A global and local enhanced residual U-Net for accurate retinal vessel segmentation," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 852–862, May 2021.
- [207] C. Wang, Z. Zhao, Q. Ren, Y. Xu, and Y. Yu, "Dense U-net based on patch-based learning for retinal vessel segmentation," *Entropy*, vol. 21, no. 2, p. 168, Feb. 2019.
- [208] T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, M. Paul, and L. Zheng, "Strided U-net model: Retinal vessels segmentation using dice loss," *Digit. Image Comput., Techn. Appl. (DICTA)*, vols. 10–13, pp. 1–8, Dec. 2018.
- [209] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on fully convolutional neural networks," *Expert Syst. Appl.*, vol. 112, pp. 229–242, Dec. 2018.
- [210] Z. Jiang, H. Zhang, Y. Wang, and S.-B. Ko, "Retinal blood vessel segmentation using fully convolutional network with transfer learning," *Computerized Med. Imag. Graph.*, vol. 68, pp. 1–15, Sep. 2018.
- [211] Y. Guo, Ü. Budak, and A. Şengür, "A novel retinal vessel detection approach based on multiple deep convolution neural networks," *Comput. Methods Programs Biomed.*, vol. 167, pp. 43–48, Dec. 2018.
- [212] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
- [213] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "DeepVessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Switzerland: Springer, 2016, pp. 132–139. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-46723-8_16
- [214] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 109–118, Jan. 2016.
- [215] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomputing*, vol. 149, pp. 708–717, Feb. 2015.
- [216] S. S. Lachure, R. V. Mante, and P. N. Chatur, "A simple review on supervised blood vessel segmentation methodologies in retinal images along with databases," *Int. J. Managment, IT Eng.*, vol. 3, no. 2, pp. 456–476, 2013.
- [217] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos, "Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics," *Comput. Methods Programs Biomed.*, vol. 158, pp. 71–91, May 2018.
- [218] B. S. Y. Lam, Y. Gao, and A. W. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," *IEEE Trans. Med. Imag.*, vol. 29, no. 7, pp. 1369–1381, Jul. 2010.
- [219] M. S. Miri and A. Mahloojifar, "Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 5, pp. 1183–1192, May 2011.
- [220] T. A. Soomro, A. J. Afifi, L. Zheng, S. Soomro, J. Gao, O. Hellwich, and M. Paul, "Deep learning models for retinal blood vessels segmentation: A review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019.
- [221] T. A. Soomro, T. M. Khan, M. A. U. Khan, J. Gao, M. Paul, and L. Zheng, "Impact of ICA-based image enhancement technique on retinal blood vessels segmentation," *IEEE Access*, vol. 6, pp. 3524–3538, 2018.
- [222] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Blood vessel segmentation of fundus images by major vessel extraction and subimage classification," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 3, pp. 1118–1128, May 2015.
- [223] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351, Munich, Germany. Cham, Switzerland: Springer, 2015, pp. 234–241.
- [224] H. Xiong, S. Liu, R. V. Sharan, E. Coiera, and S. Berkovsky, "Weak label based Bayesian U-Net for optic disc segmentation in fundus images," *Artif. Intell. Med.*, vol. 126, Apr. 2022, Art. no. 102261.
- [225] J. Du, B. Zou, C. Chen, Z. Xu, and Q. Liu, "Automatic microaneurysm detection in fundus image based on local cross-section transformation and multi-feature fusion," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105687.
- [226] D. J. Derwin, S. T. Selvi, O. J. Singh, and B. P. Shan, "A novel automated system of discriminating microaneurysms in fundus images," *Biomed. Signal Process. Control*, vol. 58, Apr. 2020, Art. no. 101839.
- [227] N. Mazlan, H. Yazid, H. Arof, and H. Mohd Isa, "Automated microaneurysms detection and classification using multilevel thresholding and multilayer perceptron," *J. Med. Biol. Eng.*, vol. 40, no. 2, pp. 292–306, Apr. 2020.

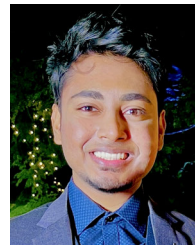
- [228] D. J. Derwin, S. T. Selvi, and O. J. Singh, "Secondary observer system for detection of microaneurysms in fundus images using texture descriptors," *J. Digit. Imag.*, vol. 33, no. 1, pp. 159–167, Feb. 2020.
- [229] S. Bakar, H. Yazid, S. Rahim, and S. Basah, "Microaneurysms detection using blob analysis for diabetic retinopathy," *Int. J. Integr. Eng.*, vol. 11, no. 6, pp. 196–203, 2019.
- [230] W. Cao, J. Shan, N. Czarnek, and L. Li, "Microaneurysm detection in fundus images using small image patches and machine learning methods," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2017, pp. 325–331.
- [231] R. T. Mamilla, V. K. R. Ede, and P. R. Bhima, "Extraction of microaneurysms and hemorrhages from digital retinal images," *J. Med. Biol. Eng.*, vol. 37, no. 3, pp. 395–408, Jun. 2017.
- [232] P. Navarro, D. Alonso, and K. Stathis, "Automatic detection of microaneurysms in diabetic retinopathy fundus images using the lab colour space," *J. Opt. Soc. Amer. A Opt. Image Sci. Vis.*, vol. 33, no. 1, pp. 74–83, 2016.
- [233] R. Ganjee, R. Azmi, and M. Ebrahimi Moghadam, "A novel microaneurysms detection method based on local applying of Markov random field," *J. Med. Syst.*, vol. 40, no. 3, p. 74, Jan. 2016.
- [234] A. Sopharak, B. Uyyanonvara, and S. Barman, "Simple hybrid method for fine microaneurysm detection from non-dilated diabetic retinopathy retinal images," *Computerized Med. Imag. Graph.*, vol. 37, nos. 5–6, pp. 394–402, Jul. 2013.
- [235] I. Lazar and A. Hajdu, "Retinal microaneurysm detection through local rotating cross-section profile analysis," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 400–407, Feb. 2013.
- [236] X. Zhang, J. Wu, M. Meng, Y. Sun, and W. Sun, "Feature-transfer network and local background suppression for microaneurysm detection," *Mach. Vis. Appl.*, vol. 32, no. 1, p. 1, Jan. 2021.
- [237] J. Xue, S. Yan, J. Qu, F. Qi, C. Qiu, H. Zhang, M. Chen, T. Liu, D. Li, and X. Liu, "Deep membrane systems for multitask segmentation in diabetic retinopathy," *Knowledge-Based Syst.*, vol. 183, Nov. 2019, Art. no. 104887.
- [238] Z. Yan, X. Han, C. Wang, Y. Qiu, Z. Xiong, and S. Cui, "Learning mutually local-global U-Nets for high-resolution retinal lesion segmentation in fundus images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 597–600.
- [239] S. Guo, T. Li, H. Kang, N. Li, Y. Zhang, and K. Wang, "L-seg: An end-to-end unified framework for multi-lesion segmentation of fundus images," *Neurocomputing*, vol. 349, pp. 52–63, Jul. 2019.
- [240] R. R. Akut, "FILM: Finding the location of microaneurysms on the retina," *Biomed. Eng. Lett.*, vol. 9, no. 4, pp. 497–506, Nov. 2019.
- [241] N. Eftekhari, H.-R. Pourreza, M. Masoudi, K. Ghiasi-Shirazi, and E. Saedi, "Microaneurysm detection in fundus images using a two-step convolutional neural network," *Biomed. Eng. OnLine*, vol. 18, no. 1, p. 67, Dec. 2019.
- [242] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 158, pp. 185–192, May 2018.
- [243] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, "Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network," *Inf. Sci.*, vol. 420, pp. 66–76, Dec. 2017.
- [244] J. Shan and L. Li, "A deep learning method for microaneurysm detection in fundus images," in *Proc. IEEE 1st Int. Conf. Connected Health: Appl., Syst. Eng. Technol. (CHASE)*, Jun. 2016, pp. 357–358.
- [245] X. Zhang, Y. Ma, Q. Gong, and J. Yao, "Automatic detection of microaneurysms in fundus images based on multiple preprocessing fusion to extract features," *Biomed. Signal Process. Control*, vol. 85, Aug. 2023, Art. no. 104879.
- [246] C. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," *AMIA Joint Summits Transl. Sci. Proc.*, vol. 2017, pp. 147–155, Jan. 2018.
- [247] G. Quellec, K. Charrrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Med. Image Anal.*, vol. 39, pp. 178–193, Jul. 2017.
- [248] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284–302, Aug. 1996.
- [249] M. U. Akram, S. Khalid, and S. A. Khan, "Identification and classification of microaneurysms for early detection of diabetic retinopathy," *Pattern Recognit.*, vol. 46, no. 1, pp. 107–116, Jan. 2013.
- [250] S. Sreng, N. Maneerat, and K. Hamamoto, "Automated microaneurysms detection in fundus images using image segmentation," in *Proc. Int. Conf. Digit. Arts, Media Technol. (ICDAMT)*, Mar. 2017, pp. 19–23.
- [251] M. Haloi, "Improved microaneurysm detection using deep neural networks," 2015, *arXiv:1505.04424*.
- [252] C. Kou, W. Li, W. Liang, Z. Yu, and J. Hao, "Microaneurysms segmentation with a U-Net based on recurrent residual convolutional neural network," *J. Med. Imag.*, vol. 6, no. 2, p. 1, Jun. 2019.
- [253] R. Phillips, J. Forrester, and P. Sharp, "Automated detection and quantification of retinal exudates," *Graefes Arch. Clin. Experim. Ophthalmology*, vol. 231, no. 2, pp. 90–94, Feb. 1993.
- [254] O. Perdomo, J. Arevalo, and F. A. González, "Convolutional network to detect exudates in eye fundus images of diabetic subjects," *Proc. SPIE*, vol. 10160, Jan. 2017, pp. 235–240.
- [255] U. R. Acharya, Y. Hagiwara, J. E. W. Koh, J. H. Tan, S. V. Bhandary, A. K. Rao, and U. Raghavendra, "Automated screening tool for dry and wet age-related macular degeneration (ARMD) using pyramid of histogram of oriented gradients (PHOG) and nonlinear features," *J. Comput. Sci.*, vol. 20, pp. 41–51, May 2017.
- [256] S. Farsiu, S. J. Chiu, R. V. O'Connell, F. A. Folgar, E. Yuan, J. A. Izatt, and C. A. Toth, "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, pp. 162–172, Jan. 2014.
- [257] C. Kou, W. Li, Z. Yu, and L. Yuan, "An enhanced residual U-Net for microaneurysms and exudates segmentation in fundus images," *IEEE Access*, vol. 8, pp. 185514–185525, 2020.
- [258] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017.
- [259] G. B. Kande, T. S. Savithri, and P. V. Subbaiah, "Automatic detection of microaneurysms and hemorrhages in digital fundus images," *J. Digit. Imag.*, vol. 23, no. 4, pp. 430–437, Aug. 2010.
- [260] M. Badar, M. Shahzad, and M. Fraz, "Simultaneous segmentation of multiple retinal pathologies using fully convolutional deep neural network," in *Proc. 22nd Conf. Med. Image Understand. Anal. (MIUA)*, Southampton, U.K. Switzerland: Springer, Jul. 2018, pp. 313–324. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-95921-4_29
- [261] P. P. Srinivasan, L. A. Kim, P. S. Mettu, S. W. Cousins, G. M. Comer, J. A. Izatt, and S. Farsiu, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomed. Opt. Exp.*, vol. 5, no. 10, p. 3568, 2014.
- [262] X. Ren, Y. Zheng, Y. Zhao, C. Luo, H. Wang, J. Lian, and Y. He, "Drusen segmentation from retinal images via supervised feature learning," *IEEE Access*, vol. 6, pp. 2952–2961, 2018.
- [263] Y. J. Kim and K. G. Kim, "Automated segmentation methods of drusen to diagnose age-related macular degeneration screening in retinal images," *Comput. Math. Methods Med.*, vol. 2018, pp. 1–8, May 2018.
- [264] A. D. Mora, P. M. Vieira, A. Manivannan, and J. M. Fonseca, "Automated drusen detection in retinal images using analytical modelling algorithms," *Biomed. Eng. OnLine*, vol. 10, no. 1, p. 59, 2011.
- [265] S. M. Mohaimin, S. K. Saha, A. M. Khan, A. S. Mohammad Arif, and Y. Kanagasigam, "Automated method for the detection and segmentation of drusen in colour fundus image for the diagnosis of age-related macular degeneration," *IET Image Process.*, vol. 12, no. 6, pp. 919–927, Jun. 2018.
- [266] Z. B. Sheh, L. D. Cohen, G. Mimoun, G. Coscas, and G. Soubrane, "An adaptive contrast method for segmentation of drusen," in *Proc. Int. Conf. Image Process.*, vol. 1, 1997, pp. 255–258.
- [267] K. Rapantzikos, M. Zervakis, and K. Balas, "Detection and segmentation of drusen deposits on human retina: Potential in the diagnosis of age-related macular degeneration," *Med. Image Anal.*, vol. 7, no. 1, pp. 95–108, Mar. 2003.
- [268] Z. Liang, D. W. K. Wong, J. Liu, K. L. Chan, and T. Y. Wong, "Towards automatic detection of age-related macular degeneration in retinal fundus images," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 4100–4103.
- [269] G. Quellec, S. R. Russell, and M. D. Abramoff, "Optimal filter framework for automated, instantaneous detection of lesions in retinal images," *IEEE Trans. Med. Imag.*, vol. 30, no. 2, pp. 523–533, Feb. 2011.
- [270] S. S. Parvathi and N. Devi, "Automatic drusen detection from colour retinal images," in *Proc. Int. Conf. Comput. Intell. Multimedia Appl. (ICCIMA)*, Dec. 2007, pp. 377–381.

- [271] L. Brandon and A. Hoover, "Drusen detection in a retinal image using multi-level analysis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2003, pp. 618–625.
- [272] Q. T. M. Pham, S. Ahn, S. J. Song, and J. Shin, "Automatic drusen segmentation for age-related macular degeneration in fundus images using deep learning," *Electronics*, vol. 9, no. 10, p. 1617, Oct. 2020.
- [273] H. A. Hassan, N. M. Tahir, A. Zabidi, I. M. Yassin, and M. Karbasi, "Classification of visualization exudates fundus images results using support vector machine," *J. Fundam. Appl. Sci.*, vol. 9, no. 4S, p. 19, Jan. 2018.
- [274] B. Antal and A. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowledge-Based Syst.*, vol. 60, pp. 20–27, Apr. 2014.
- [275] S. Sarni, J. Chisina, P. Vasile, and S. James, "Automatic detection of microaneurysms in fundus images for diabetic retinopathy screening," *Computerized Med. Imag. Graph.*, vol. 49, pp. 1–10, 2015.
- [276] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1720–1726, Jun. 2012.
- [277] S. Morales, K. Engan, V. Naranjo, and A. Colomer, "Retinal disease screening through local binary patterns," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 1, pp. 184–192, Jan. 2017.
- [278] C. Bowd, L. M. Zangwill, C. C. Berry, E. Z. Blumenthal, C. Vasile, C. Sanchez-Galeana, C. F. Bosworth, P. A. Sample, and R. N. Weinreb, "Detecting early glaucoma by assessment of retinal nerve fiber layer thickness and visual function," *Investigative Ophthalmology Vis. Sci.*, vol. 42, no. 9, pp. 1993–2003, 2001.
- [279] H. Tjandrasa, R. E. Putra, A. Y. Wijaya, and I. Ashanti, "Classification of non-proliferative diabetic retinopathy based on hard exudates using soft margin SVM," in *Proc. IEEE Int. Conf. Control Syst., Comput. Eng.*, 2013, pp. 376–380.
- [280] C. Sinthanayothin, V. Kongbunkiat, S. Phoojaruenchanachai, and A. Singalavaniya, "Automated screening system for diabetic retinopathy," in *Proc. Conf. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 1, Jun. 2003, pp. 915–920.
- [281] P. Prasanna, S. Jain, N. Bhagat, and A. Madabhushi, "Decision support system for detection of diabetic retinopathy using smartphones," in *Proc. 7th Int. Conf. Pervasive Comput. Technol. for Healthcare Workshops*, May 2013, pp. 176–179.
- [282] M. Jahiruzzaman and A. B. M. A. Hossain, "Detection and classification of diabetic retinopathy using K-means clustering and fuzzy logic," in *Proc. 18th Int. Conf. Comput. Inf. Technol. (ICCIT)*, Dec. 2015, pp. 534–538.
- [283] N. Du and Y. Li, "Automated identification of diabetic retinopathy stages using support vector machine," in *Proc. 32nd Chin. Control Conf.*, Jul. 2013, pp. 3882–3886.
- [284] S. Gupta, S. Thakur, and A. Gupta, "Optimized hybrid machine learning approach for smartphone based diabetic retinopathy detection," *Multimedia Tools Appl.*, vol. 81, no. 10, pp. 14475–14501, Apr. 2022.
- [285] H. Zhang, A. C. Berg, M. Maire, and J. Malik, "SVM-KNN: Discriminative nearest neighbor classification for visual category recognition," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, 2006, pp. 2126–2136.
- [286] S. Pathan, P. Kumar, R. Pai, and S. V. Bhandary, "Automated detection of optic disc contours in fundus images using decision tree classifier," *Biocybernetics Biomed. Eng.*, vol. 40, no. 1, pp. 52–64, Jan. 2020.
- [287] A. R. Chowdhury, T. Chatterjee, and S. Banerjee, "A random forest classifier-based approach in the detection of abnormalities in the retina," *Med. Biol. Eng. Comput.*, vol. 57, no. 1, pp. 193–203, Jan. 2019.
- [288] S. M. Hosseini, M. Tazhibi, M. Amini, H. Jahani Hashemi, A. Zaree, and S. Roosta, "Using classification tree for prediction of diabetic retinopathy on type II diabetes," *J. Isfahan Med. School*, vol. 28, no. 104, pp. 17–27, 2010.
- [289] P. Fraccaro, M. Nicolo, M. Bonetto, M. Giacomini, P. Weller, C. E. Traverso, M. Prosperi, and D. OSullivan, "Combining macula clinical signs and patient characteristics for age-related macular degeneration diagnosis: A machine learning approach," *BMC Ophthalmology*, vol. 15, no. 1, pp. 1–9, Dec. 2015.
- [290] R. Casanova, S. Saldana, E. Y. Chew, R. P. Danis, C. M. Greven, and W. T. Ambrosius, "Application of random forests methods to diabetic retinopathy classification analyses," *PLoS ONE*, vol. 9, no. 6, Jun. 2014, Art. no. e98587.
- [291] P. Shanmugam, J. Raja, and R. Pitchai, "An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier," *Appl. Soft Comput.*, vol. 109, Sep. 2021, Art. no. 107512.
- [292] K. Nirmala, N. Venkateswaran, and C. V. Kumar, "HoG based naive Bayes classifier for glaucoma detection," in *Proc. IEEE Region 10 Conf.*, Nov. 2017, pp. 2331–2336.
- [293] S. Dua, U. R. Acharya, and P. Dua, *Machine Learning in Healthcare Informatics*, vol. 56. Germany: Springer, 2014. [Online]. Available: <https://link.springer.com/book/10.1007/978-3-642-40017-9>
- [294] N. Gharaibeh, O. M. Al-hazimeh, A. Abu-Ein, and K. Nahar, "A hybrid SVM Naïve-bayes classifier for bright lesions recognition in eye fundus images," *Int. J. Electr. Eng. Inf.*, vol. 13, no. 3, pp. 530–545, 2021.
- [295] A. A. Abd El-Khalek, H. M. Balaha, N. S. Alghamdi, M. Ghazal, A. T. Khalil, M. E. A. Abo-Elsoud, and A. El-Baz, "A concentrated machine learning-based classification system for age-related macular degeneration (AMD) diagnosis using fundus images," *Sci. Rep.*, vol. 14, no. 1, p. 2434, Jan. 2024.
- [296] D. Zhang, W. Bu, and X. Wu, "Diabetic retinopathy classification using deeply supervised ResNet," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2017, pp. 1–6.
- [297] T. Shyamalee and D. Meedeniya, "CNN based fundus images classification for glaucoma identification," in *Proc. 2nd Int. Conf. Adv. Res. Comput. (ICARC)*, Feb. 2022, pp. 200–205.
- [298] D. Božić-Štulić, M. Braović, and D. Stipaničević, "Deep learning based approach for optic disc and optic cup semantic segmentation for glaucoma analysis in retinal fundus images," *Int. J. Electr. Comput. Eng. Syst.*, vol. 11, no. 2, pp. 111–120, Jul. 2020.
- [299] R. O. Ogundokun, A. T. Abdulahi, A. R. Adenike, A. N. Babatunde, and R. S. Babatunde, "Inception V3 based approach for the recognition of age-related macular degeneration disease," in *Proc. Int. Conf. Sci., Eng. Bus. Sustain. Develop. Goals (SEB-SDG)*, vol. 1, Apr. 2023, pp. 1–7.
- [300] K. V. Maya and K. S. Adarsh, "Detection of retinal lesions based on deep learning for diabetic retinopathy," in *Proc. 5th Int. Conf. Electr. Energy Syst. (ICEES)*, Feb. 2019, pp. 1–5.
- [301] M. P. Paing, S. Choomchuay, and M. D. Rapeeporn Yodprom, "Detection of lesions and classification of diabetic retinopathy using fundus images," in *Proc. 9th Biomed. Eng. Int. Conf. (BMEiCON)*, Dec. 2016, pp. 1–5.
- [302] T. Shyamalee, D. Meedeniya, G. Lim, and M. Karunarathne, "Automated tool support for glaucoma identification with explainability using fundus images," *IEEE Access*, vol. 12, pp. 17290–17307, 2024.
- [303] P. Yin, H. Cai, and Q. Wu, "DF-net: Deep fusion network for multi-source vessel segmentation," *Inf. Fusion*, vol. 78, pp. 199–208, Feb. 2022.
- [304] T. Shyamalee and D. Meedeniya, "Glaucoma detection with retinal fundus images using segmentation and classification," *Mach. Intell. Res.*, vol. 19, no. 6, pp. 563–580, Dec. 2022.
- [305] I. K. Gupta, A. Choubey, and S. Choubey, "Mayfly optimization with deep learning enabled retinal fundus image classification model," *Comput. Electr. Eng.*, vol. 102, Sep. 2022, Art. no. 108176.
- [306] Á. S. Hervella, J. Rouco, J. Novo, and M. Ortega, "End-to-end multi-task learning for simultaneous optic disc and cup segmentation and glaucoma classification in eye fundus images," *Appl. Soft Comput.*, vol. 116, Feb. 2022, Art. no. 108347.
- [307] A. Z. M. E. Chowdhury, G. Mann, W. H. Morgan, A. Vukmirovic, A. Mehnert, and F. Soheli, "MSGANet-RAV: A multiscale guided attention network for artery-vein segmentation and classification from optic disc and retinal images," *J. Optometry*, vol. 15, pp. S58–S69, 2022.
- [308] J. Morano, Á. S. Hervella, J. Novo, and J. Rouco, "Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images," *Artif. Intell. Med.*, vol. 118, Aug. 2021, Art. no. 102116.
- [309] J. Civit-Masot, M. J. Domínguez-Morales, S. Vicente-Díaz, and A. Civit, "Dual machine-learning system to aid glaucoma diagnosis using disc and cup feature extraction," *IEEE Access*, vol. 8, pp. 127519–127529, 2020.
- [310] R. P. Muthusamy, S. Vinod, and M. Tholkapiyan, "Automatic detection of abnormalities in retinal blood vessels using DTCWT, GLCM feature extractor and CNN-RNN classifier," *Int. J. Recent Technol. Eng.*, vol. 7, no. 6, pp. 3–1890, 2019.
- [311] G. Lim, Y. Cheng, W. Hsu, and M. L. Lee, "Integrated optic disc and cup segmentation with deep learning," in *Proc. IEEE 27th Int. Conf. Tools with Artif. Intell. (ICTAI)*, Nov. 2015, pp. 162–169.

- [312] J. I. Orlando, "REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Med. Image Anal.*, vol. 59, Jan. 2020, Art. no. 101570.
- [313] J. Sahlsten, J. Jaskari, J. Kivinen, L. Turunen, E. Jaanio, K. Hietala, and K. Kaski, "Deep learning fundus image analysis for diabetic retinopathy and macular edema grading," *Sci. Rep.*, vol. 9, no. 1, p. 10750, Jul. 2019.
- [314] L. De Calleja and M. Medina, "LBP and machine learning for diabetic retinopathy detection," in *Intelligent Data Engineering and Automated Learning*, vol. 8669. Switzerland: Springer, 2014, pp. 110–117. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-10840-7_14
- [315] J. E. W. Koh, E. Y. K. Ng, S. V. Bhandary, A. Laude, and U. R. Acharya, "Automated detection of retinal health using PHOG and SURF features extracted from fundus images," *Int. J. Speech Technol.*, vol. 48, pp. 1379–1393, Aug. 2017.
- [316] G. Sivapriya, R. Manjula Devi, P. Keerthika, and V. Praveen, "Automated diagnosis classification of diabetic retinopathy with microvascular structure of fundus images using deep learning method," *Biomed. Signal Process. Control*, vol. 88, Feb. 2024, Art. no. 105616.
- [317] Y. Liu, M. Kang, S. Gao, C. Zhang, Y. Liu, S. Li, Y. Qi, A. Nathan, W. Xu, C. Tang, E. Occhipinti, M. Yusufu, N. Wang, W. Bai, and L. Occhipinti, "Diagnosis of multiple fundus disorders amidst a scarcity of medical experts via self-supervised machine learning," 2024, *arXiv:2404.13388*.
- [318] O. Abramovich, H. Pizem, J. Van Eijgen, I. Oren, J. Melamed, I. Stalmans, E. Z. Blumenthal, and J. A. Behar, "FundusQ-net: A regression quality assessment deep learning algorithm for fundus images quality grading," *Comput. Methods Programs Biomed.*, vol. 239, Sep. 2023, Art. no. 107522.
- [319] Z. Gu, Y. Li, Z. Wang, J. Kan, J. Shu, and Q. Wang, "Classification of diabetic retinopathy severity in fundus images using the vision transformer and residual attention," *Comput. Intell. Neurosci.*, vol. 2023, no. 1, Jan. 2023, Art. no. 1305583.
- [320] K. Xu, S. Huang, Z. Yang, Y. Zhang, Y. Fang, G. Zheng, B. Lin, M. Zhou, and J. Sun, "Automatic detection and differential diagnosis of age-related macular degeneration from color fundus photographs using deep learning with hierarchical vision transformer," *Comput. Biol. Med.*, vol. 167, Dec. 2023, Art. no. 107616.
- [321] C. Adak, T. Karkera, S. Chattopadhyay, and M. Saqib, "Detecting severity of diabetic retinopathy from fundus images using ensemble transformers," *Comput. Vis. Pattern Recognit.*, vol. 1, pp. 1–9, 2023.
- [322] A. Haider, M. Arsalan, C. Park, H. Sultan, and K. R. Park, "Exploring deep feature-blending capabilities to assist glaucoma screening," *Appl. Soft Comput.*, vol. 133, Jan. 2023, Art. no. 109918.
- [323] Q. T. M. Pham, S. Ahn, J. Shin, and S. J. Song, "Generating future fundus images for early age-related macular degeneration based on generative adversarial networks," *Comput. Methods Programs Biomed.*, vol. 216, Apr. 2022, Art. no. 106648.
- [324] P. Elangovan and M. K. Nath, "En-ConvNet: A novel approach for glaucoma detection from color fundus images using ensemble of deep convolutional neural networks," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 6, pp. 2034–2048, Nov. 2022.
- [325] M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, "Transfer learning-based model for diabetic retinopathy diagnosis using retinal images," *Brain Sci.*, vol. 12, no. 5, p. 535, Apr. 2022.
- [326] Y. Chen, J. Yan, X. Zhang, Y. Yang, G. Xie, and Y. Huang, "Glaucoma diagnosis using vision transformer on retinal fundus images," *J. Healthcare Eng.*, vol. 2022, pp. 1–11, 2022.
- [327] R. Sun, Y. Li, T. Zhang, Z. Mao, F. Wu, and Y. Zhang, "Lesion-aware transformers for diabetic retinopathy grading," in *Proc. CVPR*, Nashville, TN, USA, Jun. 2021, pp. 10938–10947.
- [328] G.-M. Lin, M.-J. Chen, C.-H. Yeh, Y.-Y. Lin, H.-Y. Kuo, M.-H. Lin, M.-C. Chen, S. D. Lin, Y. Gao, A. Ran, and C. Y. Cheung, "Transforming retinal photographs to entropy images in deep learning to improve automated detection for diabetic retinopathy," *J. Ophthalmology*, vol. 2018, pp. 1–6, Sep. 2018.
- [329] K. Kumari and D. Mittal, "Automated drusen detection technique for age-related macular degeneration," *Brit. J. Healthcare Med. Res.*, vol. 2, p. 18, Mar. 2015.
- [330] M. Usman Akram, S. Khalid, A. Tariq, S. A. Khan, and F. Azam, "Detection and classification of retinal lesions for grading of diabetic retinopathy," *Comput. Biol. Med.*, vol. 45, pp. 161–171, Feb. 2014.
- [331] K. P. Noronha, U. R. Acharya, K. P. Nayak, R. J. Martis, and S. V. Bhandary, "Automated classification of glaucoma stages using higher order cumulant features," *Biomed. Signal Process. Control*, vol. 10, pp. 174–183, Mar. 2014.
- [332] M. R. K. Mookiah, U. R. Acharya, C. M. Lim, A. Petznick, and J. S. Suri, "Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features," *Knowl.-Based Syst.*, vol. 33, pp. 73–82, Sep. 2012.
- [333] M. H. A. Hijazi, F. Coenen, and Y. Zheng, "Data mining techniques for the screening of age-related macular degeneration," *Knowl.-Based Syst.*, vol. 29, pp. 83–92, May 2012.
- [334] M. A. Zapata, D. Royo-Fibla, O. Font, J. I. Vela, I. Marcantonio, E. U. Moya-Sanchez, A. Sanchez-Perez, D. Garcia-Gasulla, U. Cortes, E. Ayguade, and J. Labarta, "Artificial intelligence to identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma," *Clin. Ophthalmology*, vol. 14, pp. 419–429, Feb. 2020.
- [335] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, and W. Qian, "An interpretable ensemble deep learning model for diabetic retinopathy disease classification," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 2045–2048.
- [336] E. Imani and H.-R. Pourreza, "A novel method for retinal exudate segmentation using signal separation algorithm," *Comput. Methods Programs Biomed.*, vol. 133, pp. 195–205, Sep. 2016.
- [337] A. Soltani, T. Battikh, I. Jabri, and N. Lakhoua, "A new expert system based on fuzzy logic and image processing algorithms for early glaucoma diagnosis," *Biomed. Signal Process. Control*, vol. 40, pp. 366–377, Feb. 2018.
- [338] M. Hijazi, F. Coenen, and Y. Zheng, "Data mining for AMD screening: A classification based approach," *Comput. J.*, vol. 15, no. 2, pp. 57–69, 2014.
- [339] M. Tavakoli, R. P. Shahri, H. Pourreza, A. Mehdizadeh, T. Banaee, and M. H. Bahreini Toosi, "A complementary method for automated detection of microaneurysms in fluorescein angiography fundus images to assess diabetic retinopathy," *Pattern Recognit.*, vol. 46, no. 10, pp. 2740–2753, Oct. 2013.
- [340] T. Balasubramanian, S. Krishnan, M. Mohanakrishnan, K. R. Rao, C. V. Kumar, and K. Nirmala, "HOG feature based SVM classification of glaucomatous fundus image with extraction of blood vessels," in *Proc. IEEE Annu. India Conf. (INDICON)*, Dec. 2016, pp. 1–4.
- [341] N. A. Mohamed, M. A. Zulkifley, W. M. D. W. Zaki, and A. Hussain, "An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101454.
- [342] A. Issac, M. Partha Sarathi, and M. K. Dutta, "An adaptive threshold based image processing technique for improved glaucoma detection and classification," *Comput. Methods Programs Biomed.*, vol. 122, no. 2, pp. 229–244, Nov. 2015.
- [343] U. R. Acharya, E. Y. K. Ng, L. W. J. Eugene, K. P. Noronha, L. C. Min, K. P. Nayak, and S. V. Bhandary, "Decision support system for the glaucoma using Gabor transformation," *Biomed. Signal Process. Control*, vol. 15, pp. 18–26, Jan. 2015.
- [344] N. T. Le, T. Le Truong, P. F. Pongsachareonnont, D. Suwajanakorn, A. Mavichak, R. Itthipanichpong, W. Asdornwised, S. Chaitusaney, and W. Benjapolakul, "Deep learning approach for age-related macular degeneration detection using retinal images: Efficacy evaluation of different deep learning models," *Egyptian Informat. J.*, vol. 24, no. 4, Dec. 2023, Art. no. 100402.
- [345] J. H. Tan, S. V. Bhandary, S. Sivaprasad, Y. Hagiwara, A. Bagchi, U. Raghavendra, A. Krishna Rao, B. Raju, N. S. Shetty, A. Gertych, K. C. Chua, and U. R. Acharya, "Age-related macular degeneration detection using deep convolutional neural network," *Future Gener. Comput. Syst.*, vol. 87, pp. 127–135, Oct. 2018.
- [346] A. Govindaiah, M. A. Hussain, R. T. Smith, and A. Bhuiyan, "Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Washington, DC, USA, Apr. 2018, pp. 1525–1528.
- [347] F. Grassmann, J. Mengelkamp, C. Brandl, S. Harsch, M. E. Zimmermann, B. Linkohr, A. Peters, I. M. Heid, C. Palm, and B. H. F. Weber, "A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography," *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, Sep. 2018.

- [348] M. S. Hossen, A. A. Reza, and M. C. Mishu, "An automated model using deep convolutional neural network for retinal image classification to detect diabetic retinopathy," in *Proc. Int. Conf. Comput. Advancements*, Dhaka, Bangladesh, Jan. 2020, pp. 1–8.
- [349] E. Abdelmaksoud, S. Barakat, and M. Elmogy, "A computer-aided diagnosis system for detecting various diabetic retinopathy grades based on a hybrid deep learning technique," *Med. Biol. Eng. Comput.*, vol. 60, no. 7, pp. 2015–2038, Jul. 2022.
- [350] M. Alam, E. Zhao, C. Lam, and D. Rubin, "Segmentation-assisted fully convolutional neural network enhances deep learning performance to identify proliferative diabetic retinopathy," *J. Clin. Med.*, vol. 12, no. 1, p. 385, Jan. 2023.
- [351] Y. Jiang, H. Wu, and J. Dong, "Automatic screening of diabetic retinopathy images with convolution neural network based on caffe framework," in *Proc. 1st Int. Conf. Med. Health Informat.*, Taichung, Taiwan, May 2017, pp. 90–94.
- [352] S. Ovreiu, E. A. Paraschiv, and E. Ovreiu, "Deep learning & digital fundus images: Glaucoma detection using densenet," in *Proc. ECAI*, 2021, pp. 1–4.
- [353] J. Sanghavi and M. Kurhekar, "An efficient framework for optic disk segmentation and classification of glaucoma on fundus images," *Biomed. Signal Process. Control*, vol. 89, Mar. 2024, Art. no. 105770.
- [354] D. Natarajan, E. Sankaralingam, K. Balraj, and S. Karuppusamy, "A deep learning framework for glaucoma detection based on robust optic disc segmentation and transfer learning," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 1, pp. 230–250, Jan. 2022.
- [355] M. Gobinath, "A survey on deep learning models in glaucoma detection using fundus images(feasibility study of semi-supervised learning)," in *Proc. Ann. Romanian Soc. Cell Biol.*, 2021, pp. 9519–9530.
- [356] O. Perdomo, V. Andrearczyk, F. Meriaudeau, H. Müller, and F. A. González, "Glaucoma diagnosis from eye fundus images based on deep morphometric feature estimation," in *Computational Pathology and Ophthalmic Medical Image Analysis*. Switzerland: Springer, 2018, pp. 319–327. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-00949-6_38
- [357] S. Seoni, V. Jahmunah, M. Salvi, P. Barua, F. Molinari, and U. Acharya, "Application of uncertainty quantification to artificial intelligence in healthcare: A review of last decade (2013–2023)," *Comput. Biol. Med.*, vol. 165, Oct. 2023, Art. no. 107441.
- [358] A. Holzinger, "The next frontier: Ai we can really trust," in *Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. Switzerland: Springer, 2021, pp. 427–440. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-93736-2_33
- [359] O. Asan, A. E. Bayrak, and A. Choudhury, "Artificial intelligence and human trust in healthcare: Focus on clinicians," *J. Med. Internet Res.*, vol. 22, no. 6, Jun. 2020, Art. no. e15154.
- [360] M. Elsharkawy, A. Sharafeldien, F. Khalifa, A. Soliman, A. Elnakib, M. Ghazal, A. Sewelam, A. Thanos, H. S. Sandhu, and A. El-Baz, "A clinically explainable AI-based grading system for age-related macular degeneration using optical coherence tomography," *IEEE J. Biomed. Health Informat.*, vol. 28, no. 4, pp. 2079–2090, Apr. 2024.
- [361] S. Muchuchuti and S. Viriri, "Retinal disease detection using deep learning techniques: A comprehensive review," *J. Imag.*, vol. 9, no. 4, p. 84, Apr. 2023.
- [362] A. Kurz, K. Hauser, H. A. Mehrtens, E. Kriehoff-Henning, A. Hekler, J. N. Kather, S. Fröhling, C. von Kalle, and T. J. Brinker, "Uncertainty estimation in medical image classification: Systematic review," *JMIR Med. Informat.*, vol. 10, no. 8, Aug. 2022, Art. no. e36427.
- [363] M. A. Ahsan, A. Qayyum, A. Razi, and J. Qadir, "An active learning method for diabetic retinopathy classification with uncertainty quantification," *Med. Biol. Eng. Comput.*, vol. 60, no. 10, pp. 2797–2811, Oct. 2022.
- [364] M. Alghamdi and M. Abdel-Mottaleb, "A comparative study of deep learning models for diagnosing glaucoma from fundus images," *IEEE Access*, vol. 9, pp. 23894–23906, 2021.
- [365] D. Gautam, "Improved machine learning-based glaucoma detection from fundus images using texture features in FAWT and LS-SVM classifier," *Multimedia Tools Appl.*, vol. 83, no. 29, pp. 72845–72860, Feb. 2024.
- [366] Y. Subbarayudu, G. V. Reddy, M. R. Kumar, M. Aravind Naik, R. V. S. Nagini, and B. Singh, "An efficient novel approach for glaucoma classification on retinal fundus images through machine learning paradigm," in *Proc. MATEC Web Conf.*, vol. 392, 2024, p. 01108.
- [367] S. Y. Wang, "Prediction models for glaucoma in a multicenter electronic health records consortium: The sight outcomes research collaborative," *Ophthalmology Sci.*, vol. 4, no. 3, May 2024, Art. no. 100445.
- [368] R. Thanki, "A deep neural network and machine learning approach for retinal fundus image classification," *Healthcare Analytics*, vol. 3, Nov. 2023, Art. no. 100140.
- [369] G. Latha and P. Aruna Priya, "Glaucoma retinal image detection and classification using machine learning algorithms," *J. Phys., Conf. Ser.*, vol. 2335, no. 1, Sep. 2022, Art. no. 012025.
- [370] L. K. Singh, H. Garg, N. A. Pooja, and M. Khanna, "Performance analysis of machine learning techniques for glaucoma detection based on textural and intensity features," *Int. J. Innov. Comput. Appl.*, vol. 11, no. 4, p. 216, 2020.
- [371] G. An, K. Omodaka, K. Hashimoto, S. Tsuda, Y. Shiga, N. Takada, T. Kikawa, H. Yokota, M. Akiba, and T. Nakazawa, "Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images," *J. Healthcare Eng.*, vol. 2019, pp. 1–9, Feb. 2019.
- [372] C. Raja and N. Gangatharan, "A hybrid swarm algorithm for optimizing glaucoma diagnosis," *Comput. Biol. Med.*, vol. 63, pp. 196–207, Aug. 2015.
- [373] S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, "Wavelet-based energy features for glaucomatous image classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 80–87, Jan. 2012.
- [374] M. Jang, J. Shin, Y. H. Kim, T.-Y. Jeong, S. Jo, S.-J. Kim, V. Devaraj, J. Kang, E.-J. Choi, J. E. Lee, and J.-W. Oh, "3D superstructure based metabolite profiling for glaucoma diagnosis," *Biosensors Bioelectron.*, vol. 244, Jan. 2024, Art. no. 115780.
- [375] Y. Madadi, H. Abu-Serhan, and S. Yousefi, "Domain adaptation-based deep learning model for forecasting and diagnosis of glaucoma disease," *Biomed. Signal Process. Control*, vol. 92, Jun. 2024, Art. no. 106061.
- [376] Y. Chen, J. Yan, X. Zhang, Y. Yang, and Y. Huang, "Detecting glaucoma from fundus images with vision transformers," *J. Healthcare Eng.*, vol. 2021, pp. 1–9, 2021.
- [377] W. Lee, D. Lee, J. Lim, S. Lee, and E. Jung, "Glaucoma diagnosis using vision transformer-based classification," 2021, *arXiv:2110.01721*.
- [378] S. Ovreiu, I. Cristescu, F. Balta, A. Sultana, and E. Ovreiu, "Early detection of glaucoma using residual networks," in *Proc. 13th Int. Conf. Commun. (COMM)*, 2020, pp. 161–164.
- [379] A. Serener and S. Serte, "Transfer learning for early and advanced glaucoma detection with convolutional neural networks," in *Proc. Med. Technol. Congr. (TIPTeKNO)*, 2019, pp. 1–4.
- [380] M. Akiba, Y. Hayashi, M. Salman, Y. Miwa, Y. Awatsuji, R. Kazahaya, A. Kanemaru, E. Nitta, H. Aikawa, and A. Ishibazawa, "Evaluation of glaucoma diagnosis machine learning models based on color optical coherence tomography and color fundus images," *Investigative Ophthalmology Vis. Sci.*, vol. 60, no. 9, p. 1298, 2019.
- [381] P. Wang, J. Shen, R. Chang, M. Moloney, M. Torres, B. Burkemper, X. Jiang, D. Rodger, R. Varma, and G. M. Richter, "Machine learning models for diagnosing glaucoma from retinal nerve fiber layer thickness maps," *Ophthalmology Glaucoma*, vol. 2, no. 6, pp. 422–428, Nov. 2019.
- [382] A. Chakravarty and J. Sivswamy, "A deep learning based joint segmentation and classification framework for glaucoma assesment in retinal color fundus images," 2018, *arXiv:1808.01355*.
- [383] Y. Chai, H. Liu, and J. Xu, "Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models," *Knowl.-Based Syst.*, vol. 161, pp. 147–156, Dec. 2018.
- [384] D. Wu, N. Hammel, R. C. Dunn, L. Peng, and J. E. Webster, "Machine learning for identifying glaucoma related features in fundus images," *Investigative Ophthalmology Vis. Sci.*, vol. 59, no. 9, p. 1720, 2018.
- [385] M. He, Z. Li, S. Keel, and R. T. Chang, "A deep learning system for detecting glaucomatous optic neuropathy and age-related macular degeneration based on color fundus photographs," *Investigative Ophthalmology Vis. Sci.*, vol. 59, no. 9, p. 2086, 2018.
- [386] G. Borkhade and S. Akashe, "Machine learning classifier to analyze prognostic value for current imaging biomarker in age macular degeneration," in *Proc. IEEE Int. Conf. Interdiscipl. Approaches Technol. Manage. Social Innov. (IATMSI)*, Mar. 2024, pp. 1–6.
- [387] R. Cheung, J. Chun, T. Sheidow, M. Motolko, and M. S. Malvankar-Mehta, "Diagnostic accuracy of current machine learning classifiers for age-related macular degeneration: A systematic review and meta-analysis," *Eye*, vol. 36, no. 5, pp. 994–1004, May 2022.

- [388] T. C. Lin, Y. C. Jheng, S. J. Chen, and S. H. Chiou, "Artificial intelligence machine learning of optical coherence tomography angiography for the diagnosis of age-related macular degeneration," *Investigative Ophthalmology Vis. Sci.*, vol. 61, no. 7, p. 2031, 2020.
- [389] J. E. W. Koh, U. R. Acharya, Y. Hagiwara, U. Raghavendra, J. H. Tan, S. V. Sree, S. V. Bhandary, A. K. Rao, S. Sivaprasad, K. C. Chua, A. Laude, and L. Tong, "Diagnosis of retinal health in digital fundus images using continuous wavelet transform (CWT) and entropies," *Comput. Biol. Med.*, vol. 84, pp. 89–97, May 2017.
- [390] S. Waseem, M. U. Akram, and B. A. Ahmed, "Drusen detection from colored fundus images for diagnosis of age related macular degeneration," in *Proc. 7th Int. Conf. Inf. Autom. Sustainability*, Dec. 2014, pp. 1–5.
- [391] L. Dong, Q. Yang, R. H. Zhang, and W. B. Wei, "Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: A systematic review and meta-analysis," *EClinicalMedicine*, vol. 35, May 2021, Art. no. 100875.
- [392] H. Y. You, "An AMD-OCT-NET for automated AMD detection under evaluations of different image size, denoising and cropping," in *Proc. IEEE 3rd Eurasia Conf. Biomed. Eng., Healthcare Sustainability (ECBIOS)*, Taiwan, May 2021, pp. 138–142.
- [393] S. Vasudevan, B. Rajendran, A. Yadav, K. Ravi, N. Nair, and A. Ramakrishnan, "ViT-DR: Vision transformers in diabetic retinopathy grading using fundus images," in *Proc. IEEE 10th Region Humanitarian Technol. Conf.*, 2021, pp. 167–172.
- [394] S. Matsuba, H. Tabuchi, H. Ohsugi, H. Enno, N. Ishitobi, H. Masumoto, and Y. Kiuchi, "Accuracy of ultra-wide-field fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration," *Int. Ophthalmology*, vol. 39, no. 6, pp. 1269–1275, Jun. 2019.
- [395] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, "Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks," *JAMA Ophthalmology*, vol. 135, no. 11, p. 1170, Nov. 2017.
- [396] U. R. Acharya, M. R. K. Mookiah, J. E. W. Koh, J. H. Tan, K. Noronha, S. V. Bhandary, A. K. Rao, Y. Hagiwara, C. K. Chua, and A. Laude, "Novel risk index for the identification of age-related macular degeneration using radon transform and DWT features," *Comput. Biol. Med.*, vol. 73, pp. 131–140, Jun. 2016.
- [397] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, "Detection of age-related macular degeneration via deep learning," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Prague, Czech Republic, Apr. 2016, pp. 184–188.
- [398] A. Jabbar, H. B. Liaqat, A. Akram, M. U. Sana, I. D. Azpíroz, I. D. L. T. Diez, and I. Ashraf, "A lesion-based diabetic retinopathy detection through hybrid deep learning model," *IEEE Access*, vol. 12, pp. 40019–40036, 2024.
- [399] M. Nahiduzzaman, M. Robiul Islam, M. Omaer Faruq Goni, M. Shamim Anower, M. Ahsan, J. Haider, and M. Kowalski, "Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier," *Expert Syst. Appl.*, vol. 217, May 2023, Art. no. 119557.
- [400] N. Gharaibeh, O. M. Al-hazaimeh, A. Abu-Ein, and K. M. O. Nahar, "A hybrid SVM Naïve-bayes classifier for bright lesions recognition in eye fundus images," *Int. J. Electr. Eng. Informat.*, vol. 13, no. 3, pp. 530–545, Sep. 2021.
- [401] J.-H. Wu, T. Y. A. Liu, W.-T. Hsu, J. H.-C. Ho, and C.-C. Lee, "Performance and limitation of machine learning algorithms for diabetic retinopathy screening: meta-analysis," *J. Med. Internet Res.*, vol. 23, no. 7, Jul. 2021, Art. no. e23863.
- [402] D. K. Kirange, J. P. Chaudhari, K. P. Rane, K. S. Bhagat, and N. Chaudhri, "Diabetic retinopathy detection and grading using machine learning," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 6, pp. 3570–3576, 2019.
- [403] F. Kallel and A. Echtioui, "Retinal fundus image classification for diabetic retinopathy using transfer learning technique," *Signal, Image Video Process.*, vol. 18, no. 2, pp. 1143–1153, Mar. 2024.
- [404] D. Shi, W. Zhang, S. He, Y. Chen, F. Song, S. Liu, R. Wang, Y. Zheng, and M. He, "Translation of color fundus photography into fluorescein angiography using deep learning for enhanced diabetic retinopathy screening," *Ophthalmology Sci.*, vol. 3, no. 4, Dec. 2023, Art. no. 100401.
- [405] M. Saini and S. Susan, "Diabetic retinopathy screening using deep learning for multi-class imbalanced datasets," *Comput. Biol. Med.*, vol. 149, Oct. 2022, Art. no. 105989.
- [406] J. Wu, "AttenNet: Deep attention based retinal disease classification in oct images," in *MultiMedia Modeling*. Switzerland: Springer, Jan. 2020, pp. 565–576. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-37734-2_46
- [407] Q. H. Nguyen, R. Muthuraman, L. Singh, G. Sen, A. C. Tran, B. P. Nguyen, and M. Chua, "Diabetic retinopathy detection using deep learning," in *Proc. 4th Int. Conf. Mach. Learn. Soft Comput. (ICMLSC)*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 103–107. [Online]. Available: <https://doi.org/10.1145/3380688.3380709>
- [408] H. H. Vo and A. Verma, "New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space," in *Proc. ISM*, San Jose, CA, USA, Dec. 2016, pp. 209–215.
- [409] E. Colas, A. Besse, A. Orgogozo, B. Schmauch, N. Meric, and E. Besse, "Deep learning approach for diabetic retinopathy screening," *Acta Ophthalmologica*, vol. 94, no. S256, p. 97, Oct. 2016.



SAAD ISLAM received the Bachelor of Science degree (Hons.) in computer science and engineering from North South University, Dhaka, Bangladesh, in 2019, the master's degree in business management and organizations, and marketing from Macquarie University, Sydney, Australia, in 2021, the Diploma and the Advanced Diploma degrees in information technology from Sydney Business and Technology College, Australia, in 2022. His Ph.D. thesis focuses on "Automated and Accurate Detection of Retinal Health Using Artificial Intelligence," aiming to devise an effective AI-driven diagnostic system designed to significantly enhance capabilities in the medical industry and reduce costs for patients worldwide. He was featured as their Youngest Distinguished Alumnus at 25 by NSU OEA, in 2020. He was a Software Engineer with one of the leading pharmaceutical companies in Bangladesh. He is currently a Ph.D. Researcher with the School of Mathematics, Physics and Computing, University of Southern Queensland, Australia. He is also a member with the University of Southern Queensland's Advanced Data Analytics Research Laboratory. He has actively engaged in various research initiatives, collaborating with highly cited academics to co-author several journal articles. His commitment to sharing knowledge extends beyond academia, as evidenced by his success as a renowned writer and best-selling author of four books, written in both English and Bengali.



RAVINESH C. DEO (Senior Member, IEEE) received the B.Sc. degree from The University of the South Pacific, Fiji, the M.Sc. degree from the University of Canterbury, New Zealand, and the Ph.D. degree in signal processing from The University of Adelaide, Australia. He leads the UniSQ's Advanced Data Analytics Research Laboratory, University of Southern Queensland, Australia. His field of research is machine learning, artificial intelligence, and applied computing.

His research interests include artificial intelligence and applications to health, energy, wireless communication, national intelligence and defense, and engineering domains. He mentored more than 40 higher degree students, many of them receiving a best thesis award. He received the Excellence in Research Award and the Ronel Erwee Memorial Award for Excellence in Postgraduate Supervision, the Faculty Research Excellence, the University and School Publication, the Faculty Teaching Excellence, and Elsevier Highly Cited Papers for influential contributions. He has secured significant research grants and has published more than 300 articles that have cumulative citations of more than 15,400 and an H-index of more than 71. He is a Highly Cited Researcher ranked in Top 1% by Clarivate Analytics and Web of Science, in 2021, and Top 2% of World Scientists by Stanford University, USA, from 2021 to 2023. He was educated and received the Senior School National Dux Award at Labasa College, Fiji, before completing the Ph.D. degree.



PRABAL DATTA BARUA received the bachelor's degree (Hons.) in mass communication and journalism from the University of Dhaka, the master's degree in information systems from the University of Ballarat, Australia, the master's degree in mass communication and journalism from the University of Dhaka, the Master of Education degree in management and leadership from the University of Sydney, and the Ph.D. degree in information systems from the University of Southern Queensland, Australia. He is currently pursuing the second Ph.D. degree in the application of AI in health care with Kumamoto University, Japan. He accredited Higher Degree by Research (HDR) Supervisor with UniSQ. He is the Academic Dean of the Australian International Institute of Higher Education; an Adjunct Professor with the University of Southern Queensland (UniSQ); a member with the ICT Industry Advisory Board, UniSQ; a member with the ICT Programs Leadership and Governance Committee, UniSQ; and a member of Business in AI Steering Committee, UniSQ. He is an Adjunct Professor with Taylor's University Malaysia; an Adjunct Professor in New England (UNE); an accredited HDR Supervisor in UNE; and a Visiting Professor with the SRM University of Science and Technology, India. He is also the CEO of Cogninet Australia, which conducts research in medical and educational informatics using artificial intelligence (AI). He is also an Honorary Professional/Industry Fellow with the University of Technology Sydney (UTS). He is a prolific author with more than 130 publications in top-tier academic journals and has contributed as a guest editor in many international journals. He has secured over A\$3M in government and industry research grants for conducting cutting-edge research in applying artificial intelligence in health informatics, education analytics, and ICT for business transformation. He is supervising more than 15 research students at both master's and doctoral levels. He received the Graduate Certificate in Data Analytics from the Melbourne Institute of Technology and multiple diplomas and advanced diploma courses in management and leadership, education and training, project management, programming, and software development from the TAFE NSW.



JEFFREY SOAR is currently a Professor and the Personal Chair in human-centered technology of the School of Business, University of Southern Queensland. He came to academic research from a long and distinguished career in industry, including as Chief Information Officer in government agencies in Australia and New Zealand. His research interests include AI, e-business, e-health, technology and development, and social and organizational change.



PING YU (Senior Member, IEEE) is currently a Professor with the School of Computing and Information Technology, University of Wollongong, Australia, and an Elected Fellow with Australasian Institute of Digital Health. She is the International Leader in collaborative, industry-based research that advocates and supports the digital transformative agenda of health and aged care sectors in Australia and overseas. She has delivered more than 30 industry-based research projects, supervised 15 Ph.D. graduates, and published 200 peer-refereed papers. Her projects focus on detecting symptom changes, automatic clinical assessment, risk prediction, and dementia care. Her contributions continue to shape the future of healthcare globally. Her current research interests include natural language processing, generative AI, large language models, ontology, and machine learning in health, and aged care. Notably, she received the 2022 Brilliant Women in Digital Health Award from Telstra.



U. RAJENDRA ACHARYA received the Ph.D. degree from the National Institute of Technology Karnataka, Surathkal, India, and the D.Eng. degree from Chiba University, Japan. He is a Professor in Artificial Intelligence in Health at the University of Southern Queensland, Australia; a Distinguished Professor with the International Research Organization for Advanced Science and Technology, Kumamoto University, Japan; an Adjunct Professor with the University of Malaya, Malaysia; and an Adjunct Professor with Asia University, Taiwan. His funded research has accrued cumulative grants exceeding six million Singapore dollars. He has authored more than 600 publications, including 550 in refereed international journals, 42 in international conference proceedings, and 17 books. He has received over 67,500 citations on Google Scholar (with an H-index of 130). His research interests include biomedical imaging and signal processing, data mining, and visualization, and applications of biophysics for better healthcare design and delivery. He has been ranked in the top 1% of the highly cited researchers for the last seven consecutive years in computer science, according to the Essential Science Indicators of Thomson, 2016 to 2022. He is on the editorial boards of many journals and has served as a guest editor on several AI-related issues.

...