



# Artificial intelligence for early detection of diabetes mellitus complications via retinal imaging

Navid Sobhi<sup>1</sup> · Yasin Sadeghi-Bazargani<sup>2</sup> · Majid Mirzaei<sup>2</sup> · Mirsaeed Abdollahi<sup>1</sup> · Ali Jafarizadeh<sup>1</sup> · Siamak Pedrammehr<sup>3,4</sup> · Roohallah Alizadehsani<sup>3</sup> · Ru-San Tan<sup>5,6</sup> · Sheikh Mohammed Shariful Islam<sup>7,8,9</sup> · U. Rajendra Acharya<sup>10,11</sup>

Received: 3 October 2024 / Accepted: 23 February 2025 / Published online: 12 April 2025  
© The Author(s) 2025

## Abstract

**Background** Diabetes mellitus (DM) increases the risk of vascular complications, and retinal vasculature imaging serves as a valuable indicator of both microvascular and macrovascular health. Moreover, artificial intelligence (AI)-enabled systems developed for high-throughput detection of diabetic retinopathy (DR) using digitized retinal images have become clinically adopted. This study reviews AI applications using retinal images for DM-related complications, highlighting advancements beyond DR screening, diagnosis, and prognosis, and addresses implementation challenges, such as ethics, data privacy, equitable access, and explainability.

**Methods** We conducted a thorough literature search across several databases, including PubMed, Scopus, and Web of Science, focusing on studies involving diabetes, the retina, and artificial intelligence. We reviewed the original research based on their methodology, AI algorithms, data processing techniques, and validation procedures to ensure a detailed analysis of AI applications in diabetic retinal imaging.

**Results** Retinal images can be used to diagnose DM complications including DR, neuropathy, nephropathy, and atherosclerotic cardiovascular disease, as well as to predict the risk of cardiovascular events. Beyond DR screening, AI integration also offers significant potential to address the challenges in the comprehensive care of patients with DM.

**Conclusion** With the ability to evaluate the patient's health status in relation to DM complications as well as risk prognostication of future cardiovascular complications, AI-assisted retinal image analysis has the potential to become a central tool for modern personalized medicine in patients with DM.

**Keywords** Diabetes mellitus · Artificial intelligence · Diabetic retinopathy · Diabetes complications · Retina

Navid Sobhi, Yasin Sadeghi-Bazargani and Majid Mirzaei contributed equally to this work.

✉ Ali Jafarizadeh  
Jafarizadeha@tbzmed.ac.ir; Ali.jafarizadeh.md@gmail.com

✉ Siamak Pedrammehr  
s.pedrammehr@deakin.edu.au

<sup>1</sup> Nikookari Eye Center, Tabriz University of Medical Sciences, Tabriz, Iran

<sup>2</sup> Student Research Committee, Tabriz University of Medical Sciences, Tabriz, Iran

<sup>3</sup> Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, 75 Pigdons Rd, Waurin Ponds, VIC 3216, Australia

<sup>4</sup> Faculty of Design, Tabriz Islamic Art University, Tabriz, Iran

<sup>5</sup> National Heart Centre Singapore, Singapore, Singapore

<sup>6</sup> Duke-NUS Medical School, Singapore, Singapore

<sup>7</sup> Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Melbourne, VIC, Australia

<sup>8</sup> Cardiovascular Division, The George Institute for Global Health, Newtown, Australia

<sup>9</sup> Sydney Medical School, University of Sydney, Camperdown, Australia

<sup>10</sup> School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, QLD 4300, Australia

<sup>11</sup> Centre for Health Research, University of Southern Queensland, Springfield, Australia

## Introduction

Diabetes mellitus (DM) affects more than 422 million people globally and is among the top ten causes of mortality and morbidity [1, 2]. It significantly increases the risk of macrovascular complications, including atherosclerotic cardiovascular diseases such as coronary heart disease, cerebrovascular disease, and peripheral arterial disease (PAD), with mortality rates from cardiovascular diseases being two to six times higher in individuals with diabetes than in non-diabetic populations [1]. Additionally, it predisposes patients to microvascular complications, such as peripheral neuropathy [3], nephropathy, where the prevalence of end-stage kidney disease is ten times greater in diabetic populations [4], and retinopathy.

Diabetic retinopathy (DR) is the leading cause of preventable adult blindness [5] with a prevalence of 34.6% among diabetic patients [6]. DR can be categorized into non-proliferative DR (NPDR) and proliferative DR (PDR). NPDR is the early stage of DR, in which existing retinal blood vessels weaken, leading to microaneurysms and possible fluid leakage. It progresses from mild to severe stages without new blood vessel formation. However, PDR refers to the advanced stage of DR characterized by the growth of new abnormal blood vessels on the retina or optic disc, which can lead to serious vision complications [7]. The incidence of sight-threatening diabetic macular edema (DME), where fluid accumulates in the macula due to leaking retinal blood vessels, causing swelling and potential vision loss [8, 9], can occur at any stage and increases as DR progresses [5]. With the rising prevalence and incidence of DM, the burden of DR and the consequent vision loss has increased. Globally, 1.4 million patients with DM have severe NPDR and PDR and will require treatment to prevent or slow further vision deterioration [6]. Therefore, the term referable DR refers to cases in which DR has progressed to a level requiring specialist referral, typically encompassing moderate NPDR or worse, or the presence of DME [10]. Preemptive systematic and continual screening of asymptomatic diabetic patients, with early initiation of antiproliferative treatment where indicated, is obligatory for preventing clinical sequelae and reducing the human costs of DR [11].

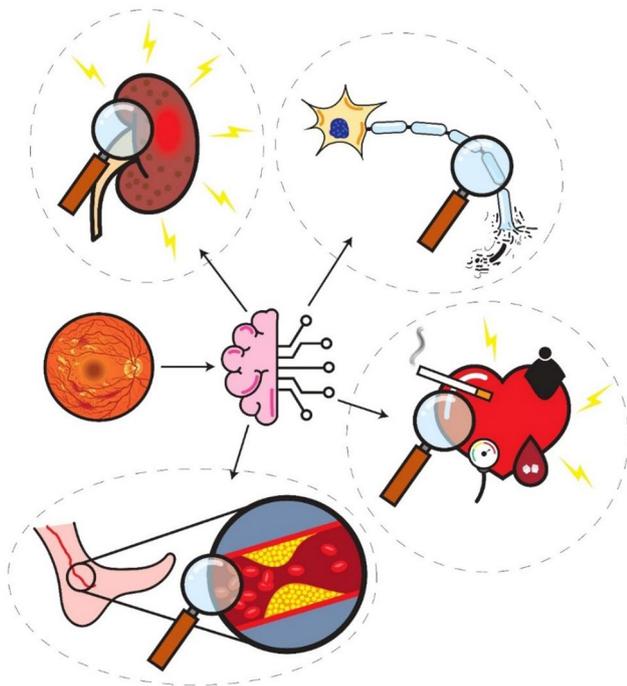
Retinal vasculature is the only part of the human body's microcirculation and can be visualized noninvasively [12]. The retina is a veritable "window" to the state of health of the microcirculation, providing insights into DR as well as other diabetic macrovascular and microvascular complications such as neuropathy and nephropathy [13]. Fundoscopic retinal examination is traditionally used in clinics to screen for DR. However, the introduction of digital fundus photography, which enables offline analysis of retinal images by experts, has contributed to the successful implementation of

national DR screening programs. With the advent of artificial intelligence (AI) applications in medicine, especially in the interpretation of digital medical images (which are by nature objective and accessible), many AI-based systems utilizing AI's ability to handle high-throughput, complex image data have been developed for the detection of DR using digitized retinal images and are now clinically adopted [14, 15].

Beyond DR screening, AI integration also holds immense potential for addressing the challenges associated with the comprehensive care of patients with DM complications. As mentioned above, retinal vasculature images contain diagnostic and prognostic clues to the microvascular and macrovascular health of the whole body. In contrast to existing review studies centered solely on DR within the context of DM, the retina, and AI, our study takes a broader approach [16–24]. Our review takes a comprehensive perspective, covering all DM-related complications detectable through retinal imaging rather than focusing solely on DR, as in most existing studies. We explored AI's potential to identify a broad range of complications, including DME as well as cardiovascular, neurological, and kidney diseases. Additionally, most studies emphasize diagnostic capabilities, with little attention given to the role of AI in predicting disease progression or guiding personalized treatments. Real-world implementation challenges, including adapting AI models to diverse healthcare settings, data privacy concerns, and regulatory hurdles, are underexplored. Ethical issues, such as algorithm bias and equitable access to AI tools, have received minimal discussion, especially in the context of low-resource settings where DM complications are prevalent. This narrative review aims to address these gaps by exploring broader applications of AI in diabetes care, synthesizing insights into multi-complication detection, and highlighting barriers to clinical adoption. Various DM-related complications are visualized in Fig. 1.

## Literature search method

We performed our search in PubMed, Scopus, and Web of Sciences databases for scholarly articles published in the English language until June 2024, using various combinations (by using Boolean operators: AND, OR, NOT) of the following main keywords: "diabetes mellitus," "retina," "artificial intelligence." Other keywords were "diabetic care," "diabetes complications," "diabetic foot," "diabetic nephropathy," "diabetic neuropathy," "diabetic retinopathy," "diabetic macular edema," "cardiovascular risk assessment," "peripheral arterial disease," "fundus image," "personalized medicine," "deep learning," "machine learning," and "machine vision".



**Fig. 1** Illustration of various diabetic complications diagnosed by retinal imaging using AI. (Graphical abstract)

In selecting articles for inclusion in our study, we employed an approach focused on identifying peer-reviewed original studies written in English with clear research goals, meticulous methodology, and extensive technical details. In our evaluation process, we defined the ‘technical details’ to include the technological aspects and analytical methods used in the studies. We also excluded non-peer-reviewed and low-quality articles with unclear research goals, duplicates, editorial articles, and conference abstracts. This entailed examining the types of AI algorithms deployed, data-processing techniques, specific retinal imaging methods, and any advanced computational strategies implemented. We also discuss software applications, statistical techniques, and validation procedures in these studies. Our focus on these technical details is to ensure a comprehensive assessment of the methodologies and technological advancements in AI applications in diabetic retinal image analysis.

## Diabetic retinopathy and diabetic macular edema

### DR screening

In a landmark study by Abramoff et al., an AI system based on supervised machine learning (ML) with logistic regression attained 96.8% sensitivity, 59.4% specificity, and 0.937 (95% CI, 0.916–0.959) AUC (a generic term

referring to the area under a curve, typically used to evaluate classifier performance) for the detection of referable DR. Despite its modest specificity, the system was projected to halve the screening burden compared with manual screening by experts [10]. The team developed a deep learning (DL) model, based on convolutional neural network (CNN) architectures inspired by AlexNet and VGGNet, to enhance the detection of referable DR. This model demonstrated a specificity of 87% (95% CI: 84.2–89.4%) and an AUC of 0.980 (95% CI: 0.968–0.992), achieving these improvements without compromising sensitivity [25]. EyeArt and Retmarker, two AI systems designed to detect referable DR, reported sensitivities of 93.8% (92.9–94.6%) and 85% (83.6–86.2%), respectively, and have been linked to reduced costs in DR screening [26]. Other economic models studies also have also shown that semi-automated and fully automated screening methods by human experts are more cost-effective than traditional manual screening [27, 28]. IDx-DR, the first medical device implementing AI for detecting moderate and severe DR, was approved by the FDA in 2018 [29]. A cost-effectiveness analysis of a nationwide DR screening program in China involving 251,535 participants revealed that AI models need a sensitivity of at least 88.2% and specificity of 80.4% to be cost-effective. The most cost-effective scenario showed a higher sensitivity (96.3%) but lower specificity (80.4%) than the most accurate model (93.3% sensitivity and 87.7% specificity). This study showed that urban regions and younger populations demand higher sensitivity for optimal cost-effectiveness [30]. Additionally, another cost-effectiveness analysis revealed that implementing AI-based DR screening in primary care for both non-indigenous and Indigenous Australians with DM is highly effective and cost-saving, with significant reductions in blindness cases and healthcare costs projected over 40 years [31].

Recent studies have highlighted the potential of AI to improve DR screening and diagnosis. In a study by Riazi-Esfahani et al. [32], 84 eyes of 57 individuals were analyzed to differentiate between normal eyes, NPDR, and PDR. They utilized optical coherence tomography (OCT) images, evaluated an ML-based segmentation system, and found significant differences in the retinal layer area and smoothness index (SI) across the groups. Specifically, the inner nuclear layer (INL) area was the most effective in distinguishing DR stages, with an accuracy of 87.6%, while the SI of the inner plexiform layer (IPL) in the nasal zone distinguished PDR from NPDR with 97.2% accuracy. The DL model developed by Saranya et al. [33] for detecting DR showed remarkable metrics. The model’s accuracy, sensitivity, specificity, and F1-scores (The harmonic mean of precision and recall used to evaluate the balance between false positives and false negatives in a binary classification) were

97.54%, 90.34%, 98.24%, and 93.28% on the MESSIDOR dataset, and 96.32%, 95.73%, 97.12%, and 96.74% on the E-ophtha Ex dataset. In another study, an AI screening model for referable DR in Rwanda demonstrated 92% (95% CI, 86.3–96.8%) sensitivity, 85% (95% CI, 75.1–88.2%) specificity, and 33.2% of patients were referred for follow-up [34]. Similarly, in Tanzania, a trial involving 2364 participants aims to improve the follow-up rates for referable DR cases using AI-supported screening pathways [35]. Another study comparing 21 AI algorithms for DR screening found varying performances. The mean agreement for referable DR was 79.4%, with sensitivity and specificity of 77.5% and 80.6%, respectively. These findings emphasize the need for real-world validation before clinical application [36]. Additionally, an AI model developed in Greenland for DR screening showed an AUC of 0.99 but requires further optimization for practical clinical application [37]. Lastly, AI analysis of retinal microvascular changes demonstrated the potential for the early diagnosis of diabetic complications by identifying significant retinal abnormalities [38].

In low-income countries with limited healthcare resources, the lack of established DM screening programs and insufficient blood glucose control results in a higher incidence of DM-related complications. In one study conducted in sub-Saharan Africa, the rate of progression from no DR to sight-threatening DR (STDR) was five times higher than in Europe [39, 40]. However, universal annual DR screening for all patients with DM remains elusive goal [40]. The solution may lie in using technological advancements, such as AI-based analysis of digital fundus photographs, which has the potential to reduce the costs of nationwide screening programs while delivering comparable or even superior diagnostic performance [6, 41, 42]. However, whether models developed and validated in countries with different ethnic compositions can be effectively applied to local populations is still a valid concern. In a study by Bellema et al. [43], an ensemble model combining adapted VGG-Net and ResNet architectures trained on a Singaporean database of color retinal images, demonstrated robust diagnostic performance in a real-world diabetic population in Zambia. Specifically, the model achieved an AUC of 0.973 (95% CI: 0.969–0.978) for detecting referable DR, with a sensitivity of 99.42% (95% CI: 99.15–99.68) for STDR and 97.19% (95% CI: 96.61–97.77) for DME. Separately, a real-world prospective interventional cohort study in Thailand utilized a DL-based system to detect STDR and reported an accuracy of 94.7% (95% CI 93.0–96.2), compared to 93.5% (91.7–95.0) achieved by human experts [44]. Moreover, among novel cost-effective approaches, retinal images can now be captured using smartphone-based retinal cameras. In one study, this method proved feasible when using the EyeArt software, an ensemble of deep artificial neural

networks, which achieved a sensitivity of 99.1% (95% CI: 95.1–99.9) and a specificity of 80.4% (95% CI: 73.9–85.9) for detecting STDR based on online-processed smartphone retinal images evaluated against human expert assessments [45].

### Advanced DR screening for DME

DME, an accumulation of fluid in the macula, can lead to visual loss at any stage of DR and mandates referral to specialists. Definitive diagnosis requires macular thickness measurement using OCT or visualization of edema on fluorescein angiography, which is the gold standard diagnostic tool. Hard exudates within one optic disc diameter on the color fundus photograph (CFP) are a surrogate for DME [46], with a significant false positive rate of up to 42% in the UK screening program [47]. The Danish guidelines recommend OCT as a second-line screening method for DR [48]. Diagnostic criteria for DME, as identified through OCT imaging, include central retinal thickness (CRT) > 250–300  $\mu\text{m}$ , presence of intraretinal or subretinal fluid, hyperreflective dots, and disruption of retinal layers [49]. Moreover, diagnostic criteria for DME, as identified through fluorescein angiography, include leakage of fluorescein dye, areas of capillary nonperfusion, and evidence of macular ischemia [50].

AI plays a significant role in diagnosing and classifying DME [23]. A DL model trained on fundus photographs attained an AUC of 0.89 (95% CI: 0.87–0.91) for discriminating center-involved DME, with superior specificity and positive predictive value compared to human experts [51]. More DL systems have been developed for the automatic detection of DME on OCT images, which can also discern grades of severity, ranging from diffuse retinal thickening cystoid macular edema to serous retinal detachment [52]. A study by Manikandan et al. reported 96% (95%CI: 0.94–0.98) and 94% (95% CI: 0.90–0.96) sensitivities for AI-assisted DME detection based on OCT and fundus images, respectively [53]. Similarly, another study by Lam et al. evaluated the performance of various DL models in detecting DME using fundus photography and OCT images. They reported a pooled area under the receiver operating characteristic curve (AUROCs) of 0.964 (95%CI: 0.964–0.964) for fundus photography-based algorithms and 0.985 (95%CI: 0.985–0.985) for OCT-based algorithms, with sensitivities of 92.6% and 95.9% and specificities of 91.1% and 97.9% [54]. AUROC measures the ability of a classifier to distinguish between classes across all thresholds, with values closer to 1 indicating better performance. Another investigation reported a pooled 96.0% (95% CI: 93.9–97.3%) sensitivity and 99.3% (95% CI: 98.2–99.7%) specificity for DME detection using DL with OCT images [55]. In a comparison with the other two DL models, the

Optic-Net model (98% accuracy, 100% specificity) outperformed Dense-Net (94% accuracy, 96% specificity) for OCT-based DME classification [56]. Liu et al. [57] trained a faster R-CNN model with ResNet101 backbone on more than 50,000 labeled fundus images and 20,000 OCT B-scans acquired from patients across multiple centers and reported a remarkable 97.78% sensitivity, 98.38% specificity, and 0.981 (95%CI: 0.966–0.990) AUC for detection of referable DR; and 91.30% sensitivity, 97.46% specificity, and 0.944 (95% CI 0.922–0.962) AUC for detection of DME. The incremental diagnostic utility of OCT-based AI analysis—combined CFP plus OCT screening—detected cases of DME that would have been missed on CFP analysis alone— supports the addition of OCT to standard CFP-based AI screening programs; however, cost-effectiveness should be further assessed.

AI evaluation of OCT angiography (OCTA), which generates angiographic images of retinal vessels from motion imaging of volumetric blood flow signals without contrast, has recently attracted attention. Ryu et al. [58] compared two models for detecting early DR using OCTA. The first model was a ML-based classifier that combined segmentation—via a U-Net capable of isolating blood vessels and the foveal avascular zone in OCTA images—followed by feature extraction and classification. The second model employed a CNN-based classifier that processed the OCTA images directly using a ResNet101 architecture. Notably, the CNN-based model achieved excellent performance, with an accuracy of 91–98%, sensitivity of 86–97%, specificity of 94–99%, and AUC values between 0.919 and 0.976, comparing favorably against ultra-widefield fluorescein angiography.

## Grading of DR severity

Classification of CFP images into distinct DR severity grades—no retinopathy, mild NPDR, moderate NPDR, severe NPDR, and PDR [24]—based on retinal vascular changes can provide valuable insights into disease progression and prognosis [59]. This is traditionally performed through expert manual fundus examinations for DR [60], a resource that is both limited and not easily accessible. Promising AI-based tools can categorize DR grades and reduce healthcare costs and burdens [27, 61]. Gulshan et al. [62] developed a CNN model with Inception-v3 architecture and transfer learning (TL) which outputs five independent binary classifiers for DR grading. The model attained good performance for grading DR, and achieved 84.0% (95% CI, 75.3–90.6%) sensitivity and 98.8% (95% CI, 98.5–99.0%) specificity for detecting severe or worse DR. Additionally, it attained 90.8% (95% CI, 86.1–94.3%) sensitivity and 98.7% (95% CI, 98.4–99.0%) specificity for detecting DME. Takahashi

et al. [63] trained a modified GoogLeNet model on 9,939 fundus photographs (four 45° color fundus photographs per eye) using a refined Davis grading protocol. When tested on 496 images (5% of the dataset), the model achieved 81% accuracy (correctly classifying 402 images) and a prevalence- and bias-adjusted Fleiss' kappa (PABAK) of 0.64 compared with the modified Davis grading. Against a real prognosis grading, the model attained a PABAK of 0.37 and 96% accuracy. Other researchers have also reported high sensitivity and specificity rates for DR grading using various models [64–66]. Figure 2 shows a schematic illustration of various DR severity grades.

In fundus examination, DR-induced vascular changes often manifest heterogeneously across the retina. Nevertheless, traditional manual analyses generally classify the entire fundus image according to the most severe lesion identified in any region of the image. With AI analysis of CFP images, the basis for severity grading is not apparent, which limits the explainability of AI decision-making. To address this, researchers have developed DL neural networks for lesion-based classification [62, 66, 67]. Wang et al. [68] trained a CNN-based DR lesion classifier, Lesion-Net, on 12,252 fundus images of patients with DM. The model contains two branches: the inferior branch uses a fully convolutional network (FCN-32s) for lesion segmentation predictions of eight types of common DR lesions to diagnose and stage DR, and the superior branch integrates convolutional layers from the Inception-v3 architecture to classify referable vs. non-referable DR classes. Lesion-Net attained a good AUC of 0.943, sensitivity of 90.6%, and specificity of 80.7% for the five-stage DR grading as well as an improved AUC of 0.936 vs. 0.928 for the internal test set and 0.977 vs. 0.964 for the external test set for binary referable vs. non-referable DR classification. According to a recent study, the Inception-v3 model, trained on a dataset of 35,126 retinal images, delivered the best performance for DR diagnosis among ML and DL models, yielding an accuracy rate of approximately 97% in feature extraction [69]. Concurrently, the MobileNetV3 model has been identified as optimal for data classification tasks in the same domain, achieving an accuracy of 98.56% [70].

Sandhu et al. combined OCT and OCTA images with primary clinical and demographic data collected from 111 patients to train an AI model for DR screening and staging [71]. They developed a novel computer-aided design system to grade NPDR into mild and moderate stages. They reported 98.7% accuracy, 100% sensitivity, 97.8% specificity, 99% differential scanning calorimetry (DSC), and 0.981 AUC (progressive improvements in almost all metrics were observed as OCTA, clinical, and demographic data were incrementally added to the model). Wang et al. [72] used ultra-widefield fluorescein angiographic images from 399

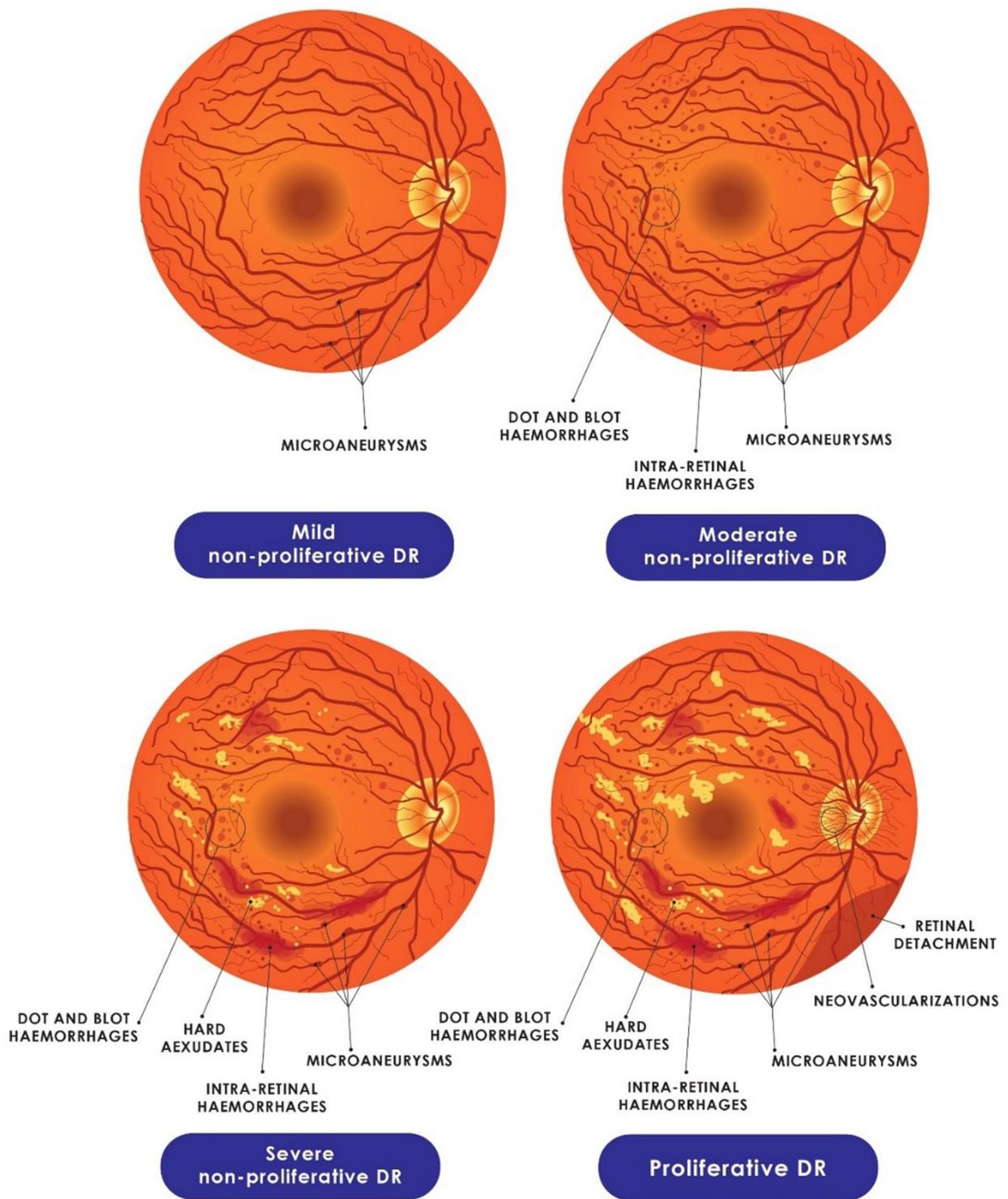


Fig. 2 Schematic representation of funduscopy showing the DR grading

patients to train an AI model for differentiating normal retina, NPDR, and PDR. The resulting developed model attained a classification accuracy of 88.50%.

### Individualized DR screening

The current guideline-based practice for DR screening involves annual fundoscopic examination. However, this practice fails to account for the heterogeneity of the clinical course of different patients. For example, only 11% of individuals with mild NPDR in both eyes progress to STDR each year, whereas all patients should undergo annual screening [73]. Given the increasing incidence of DM and projected increase in DM-related healthcare resource utilization, individualized screening at varying intervals and approaches (home visit, AI-assisted, or specialized screening) may be more cost-effective as an alternative to routine annual DR screening.

Mathematical models have been developed to predict the risk of progression from non-DR or mild DR to STDR. Aspelund et al. used six parameters—sex, DM type, duration of DM, glycated hemoglobin, blood pressure, and the existence and grading of DR—to statistically model optimal DR screening intervals [74–76]. Compared to routine annual DR screening, their model, along with those developed by other researchers, achieved a 40–60% reduction in the number of screening visits, which resulted in an 11–40% increase in the incidence of STDR detected at subsequent screening visits [74–76]. Moreover, AI can potentially improve the prediction accuracy of mathematical models using the least amount of data. Piri et al. analyzed demographic, laboratory, and comorbidity data of more than 1.4 million patients with DM using supervised ML models and artificial neural networks where only a routine blood test could predict DR with an accuracy of 92.76% [76]. These observations suggest that risk stratification without the need for retinal imaging could facilitate the early diagnosis of DR while reducing the need for mass screening, notably in regions with poor infrastructure and low compliance.

### Individualized DR follow-up

There is considerable variability in the progression of DR between patients with DM, even among those diagnosed with the same DR grade. A clear understanding of the anticipated progression of the disease greatly enhances the ability to develop the most effective screening and treatment plans for each patient. This enables customization of treatment options, scheduling of visit frequencies, and timing of interventions to optimize individual patient outcomes [77]. Skevofilakas et al. [78] developed a decision support system (DSS) for assessing DR progression using clinical data

collected from the electronic health records of 55 type I DM (T1DM) patients who were followed over five years. The DSS used strongly linked risk factors—age, DM duration, glycated hemoglobin, cholesterol, triglyceride, hypertension (HTN), incidence rate, and DM treatment duration—to predict DR progression. Receiver operating characteristic (ROC) analysis showed that the most accurate primary classifier had an accuracy of 97%, while the overall combined DSS had higher sensitivity, specificity, and accuracy.

Developing an AI model based on images, demographics, and clinical information to predict the risk of progression enables clinicians to make informed decisions. Arcadu et al. [79] developed an algorithm to predict the worsening of DR using 7-field CFPs acquired from patients with DM who were monitored over two years. The baseline CFPs were used to train the model to indicate the presence of 2-step or greater worsening of DR on the Early Treatment DR Scale during the follow-up visits. Based on an Inception-v3 architecture, separate deep CNNs (DCNN) were trained in parallel for each of the seven fields using TL initialized with ImageNet weights. Ultimately, the generated probabilities were consolidated using random forests (RF) to enhance the overall prediction accuracy. The trained DL model predicted worsening of DR at multiple intervals. It achieved an AUC of  $0.68 \pm 0.13$ , sensitivity of  $66\% \pm 23\%$ , and specificity of  $77\% \pm 12\%$  at 6 months. At 12 months, it achieved an AUC of  $0.79 \pm 0.05$ , sensitivity of  $91\% \pm 8\%$ , and specificity of  $65\% \pm 12\%$ . For the 24-month prediction, the model achieved an AUC of  $0.77 \pm 0.04$ , a sensitivity of  $79\% \pm 12\%$ , and a specificity of  $72\% \pm 14\%$ . Their results provided evidence for the use of AI in refining and personalizing the prognostication of DR. The ability of the AI model to accurately predict disease progression, even among individuals within the same DR grade as assessed by traditional manual methods, shows its potential to enhance clinical decision-making and patient management. Notably, the aggregate result was also superior to the output of each field-specific DCNN. Furthermore, they discovered that the peripheral parts of the retina contributed more to the predictions than the central regions, with the model performance dropping significantly when the peripheral zones were excluded. Additionally, the attribution map showed that the model could differentiate between microaneurysms, hemorrhages, and hard exudates, providing a foundation for potential lesion-specific analyses and predictions. Finally, the results of this study suggest that, in contrast to traditional manual grading assessments, which can only predict disease outcomes in a group of patients with similar signs and symptoms, DL algorithms can approximately predict the course of DR in a single patient.

Similarly, another study improved the automated ML models to predict DR progression from retinal images. The

models were trained and validated on images with mild or moderate NPDR over three years using images from a tertiary diabetes center. The models achieved an Area Under the Precision-Recall Curve (AUPRC), which summarizes the performance of a classifier in handling imbalanced datasets by focusing on precision and recall trade-offs across thresholds of 0.717 for mild NPDR and 0.863 for moderate NPDR, with sensitivity and specificity ranging from 0.63 to 0.80 [80]. Moreover, another trial evaluated the effectiveness of autonomous AI-based diabetic eye screening at the point-of-care in increasing examination completion rates among youth with type 1 and type 2 diabetes mellitus. This study compared AI-assisted examinations with traditional eye care provider referrals and found that the completion rate was significantly higher in the AI group (100%) than in the control group (22%). This study showed that AI could significantly improve DR follow-up adherence in a diverse youth population [81].

### DR treatment: who to treat and how to treat?

Intravitreal injection of anti-vascular endothelial growth factor (VEGF) medications, such as ranibizumab, bevacizumab, and aflibercept, is an indicated treatment for STDR, especially DME [82–85]. OCT is often used to monitor therapeutic responses. Through the analysis of OCT images, AI models can predict individual patient responses to anti-VEGF therapy and potentially facilitate personalized treatment approaches for DME [86]. Advanced AI algorithms have been developed to evaluate OCT parameters—central macular fluid volume, integrity of the ellipsoid zone, intraretinal fluid, subretinal fluid, hyperreflective retina foci, and external limiting membrane [87]—to predict visual acuity trajectories in DME, thereby providing clinicians with objective parameters for DME diagnosis and follow-up. Liu et al. [88] developed an ensemble ML system that consisted of DL models and five classical ML (CML). They trained AlexNet, VGG16, ResNet18, and an ensemble of three DL architectures on a dataset of 304 pretreatment OCT images of patients with DME. Fifteen OCT features generated by the DL ensemble model were then used to train conventional ML algorithms—Least Absolute Shrinkage and Selection Operator (LASSO), support vector machine (SVM), decision tree (DT), and RF—to predict post-treatment central foveal thickness (CFT) and best-corrected visual acuity (BCVA) values one month after three months of anti-VEGF injections [88]. However, the model failed to accurately predict post-treatment CFT and BCVA values, suggesting that OCT images alone were insufficient as the sole model inputs. However, adding further clinical information associated with treatment outcomes in CML models improves prediction accuracy.

An AI system's prediction of a poor treatment response may justify the adoption of alternative therapeutic strategies, thereby minimizing unnecessary risks and costs, while ensuring that patients who could benefit from appropriate treatments are not excluded. In a retrospective study, Gallardo et al. [89] developed an ML system to assess the burden of anti-VEGF treatment—defined as low, moderate, and high based on the interval of injections—in a treat-and-extend (T&E) regimen for DME and retinal vein occlusion using demographic data and OCT images obtained from patients at two consecutive clinic visits [89]. The proposed RF-based supervised ML model predicts the 1-year treatment requirement with a reasonable AUC so that the decision-making process is rendered interpretable. In particular, all features related to intraretinal fluid are important for predicting low and high treatment demands.

In addition to predicting treatment response, AI can be applied to the direct planning of therapeutic procedures. Focal or grid laser photocoagulation, in which a series of controlled photocoagulations are delivered to the pathological areas of the retina [90, 91], is another indicated treatment for DME and PDR. This treatment induces regression of neovascularization by normalizing oxygen partial pressures in the peripheral avascular regions of the retina. As a result, the rates of vitreous hemorrhage and membrane formation were reduced.

Treatment efficacy is highly dependent on the siting and dosing of the administered photocoagulants [92]. Standard predetermined patterns for photocoagulation cannot account for individual differences in the shapes and patterns of macular edema and anatomical variations of the retinal vasculature [90, 93]. Furthermore, manual mapping of the coagulation pattern requires surgical expertise and considerable time [93]. AI can be used to automate retinal segmentation, such that only personalized, predetermined areas of the retina are coagulated, thereby increasing the precision of laser photocoagulation and minimizing unwanted side effects. Novel AI software has generated personalized, high-quality coagulation maps by processing patient information. The system enhances precision in localizing the exact burn points and controlling the power delivered compared to manual methods, resulting in a nine-fold reduction in laser burns beyond the edema borders, shortened procedural preparation time, and fewer postoperative complications [93].

## Beyond the eye: systemic microvascular and macrovascular complications

Table 1 summarizes studies on AI models developed using retinal images to detect DM complications or sequelae outside the eye. The explanations are detailed in the following sections.

## Predicting cardiovascular risk using retinal photographs

DM increases the risk of cardiovascular events by two to four folds [1, 94], and cardiovascular diseases remain the leading cause of death in individuals with DM. Evaluating future cardiovascular risks may help guide treatment decisions and enable timely preventive measures. In response to this need, risk calculators have been developed using traditional statistical models that incorporate both clinical factors (e.g., age, sex, body mass index, and blood pressure) and laboratory measures (e.g., eGFR, HbA1c, HDL,

**Table 1** Summary of performance metrics of models and architectures for diabetes complications diagnosis

Study	AI modeling algorithms	Architecture	clinical predictors	Metric	Results
<b>Diabetic Cardiovascular disease</b>					
Poplin et al. [97]	DL	Inception-v3 (in the UKBiobank dataset)	Age	R2 (95% CI)	0.74 (0.73,0.75)
			Gender	AUC (95% CI)	0.97 (0.966,0.971)
			Smoking status	AUC (95% CI)	0.71 (0.70,0.73)
			BMI	R2 (95% CI)	0.13 (0.11,0.14)
			SBP	R2 (95% CI)	0.36 (0.35,0.37)
			DBP	R2 (95% CI)	0.32 (0.30,0.33)
			HbA1c	R2 (95% CI)	0.09 (0.03,0.16)
<b>Diabetic Neuropathy</b>					
Benson et al. [104]	DL, in conjunction with TL	VGG16 CNN	retinal changes as a predictor of diabetic neuropathy	Accuracy Sensitivity specificity	89% 78% 95%
Cervera et al. [105]	DL	SqueezeNet, Inception, and DenseNet	retinal changes as a predictor of diabetic neuropathy	AUC (95% CI)	0.8013 (±0.0257)
<b>Diabetic Nephropathy</b>					
Sabanayagam et al. [115]	DL	CondenseNet	Image only showing retinal changes as a predictor of diabetic nephropathy	AUC (95% CI) Sensitivity Specificity	0.911 (0.886–0.936) 0.83 0.83
			RF only includes age, sex, ethnicity, diabetes, and hypertension	AUC (95% CI) Sensitivity Specificity	0.916 (0.891–0.941) 0.82 0.84
			hybrid DLA combining image and RF	AUC (95% CI) Sensitivity Specificity	0.938 (0.917–0.959) 0.84 0.85
			Image only showing retinal changes as a predictor of diabetic nephropathy	AUC	0.829–0.918
			RF only (age, sex, height, weight, body-mass index, and blood pressure)	AUC	0.787–0.861
Zhang et al. [116]	DL (CNN)	ResNet-50	hybrid DLA combining image and RF	AUC	0.845–0.930
			<b>Diabetic peripheral arterial disease</b>		
Muel-ler et al. [123]	DL (CNN)	Multiple Instance Learning (MIL)	retinal changes as a predictor of diabetic PAD	Accuracy AUROC	0.674–0.837 0.653–0.890

LDL, and triglycerides) to generate predictions. Notably, such models are not specific to the DM population. Based on the same information, an AI model modestly improved risk prediction performance, increasing the AUC from 0.69 to 0.75 [95].

Retinal parameters have been shown to correlate with cardiovascular outcomes in DM patients, independent of underlying risk factors [96]. The abundance of retinal image databases from patients with DR has enabled researchers to extract new associations using AI. For instance, Poplin et al. trained an Inception-v3 model on 284,335 retinal images and reported 0.70 AUC for image-based prediction of major cardiovascular events [97]. Moreover, the accuracy of cardiovascular risk prediction can be enhanced using combined retinal images and accessible clinical and demographic data [98]. Finally, cardiovascular risk assessment could be part of DR screening in the path toward holistic diabetes care.

### Beyond the optic nerve: predicting diabetic neuropathy

Similar to DR, diabetic neuropathy is a common microvascular complication of DM. It causes skin ulcers, increases the risk of limb amputation, and reduces quality of life, accounting for 27% of annual DM-related healthcare costs [99–101]. The gold standard for diagnosing diabetic neuropathy is a combination of clinical examinations such as the monofilament test and electrophysiological tests, such as nerve conduction studies (NCS), which evaluate the function of the peripheral nerves [102]. However, routine annual screening using clinical neurological examination is insensitive and often fails to detect early changes before irreversible damage [103]. The retinal vasculature, which reflects systemic microcirculation, may also play a role in the development of diabetic neuropathy. In a study, Benson et al. used a pre-trained VGG16 CNN and a SVM classifier to analyze retinal images and then compared AI-based retinal analysis with physician-diagnosed diabetic peripheral neuropathy. They attained 89% accuracy, 78% sensitivity, and 95% specificity for detecting diabetic peripheral neuropathy [104]. Among the DL methods evaluated for diagnosing diabetic neuropathy using CFP, SqueezeNet, Inception, and DenseNet emerged as the most effective. These models achieved AUCs of 0.8013 ( $\pm 0.0257$ ) during validation and 0.7097 ( $\pm 0.0031$ ) during testing, based on analyses performed on various datasets with and without pre-trained weights [105].

### Eyes on the kidneys: screening for chronic kidney disease

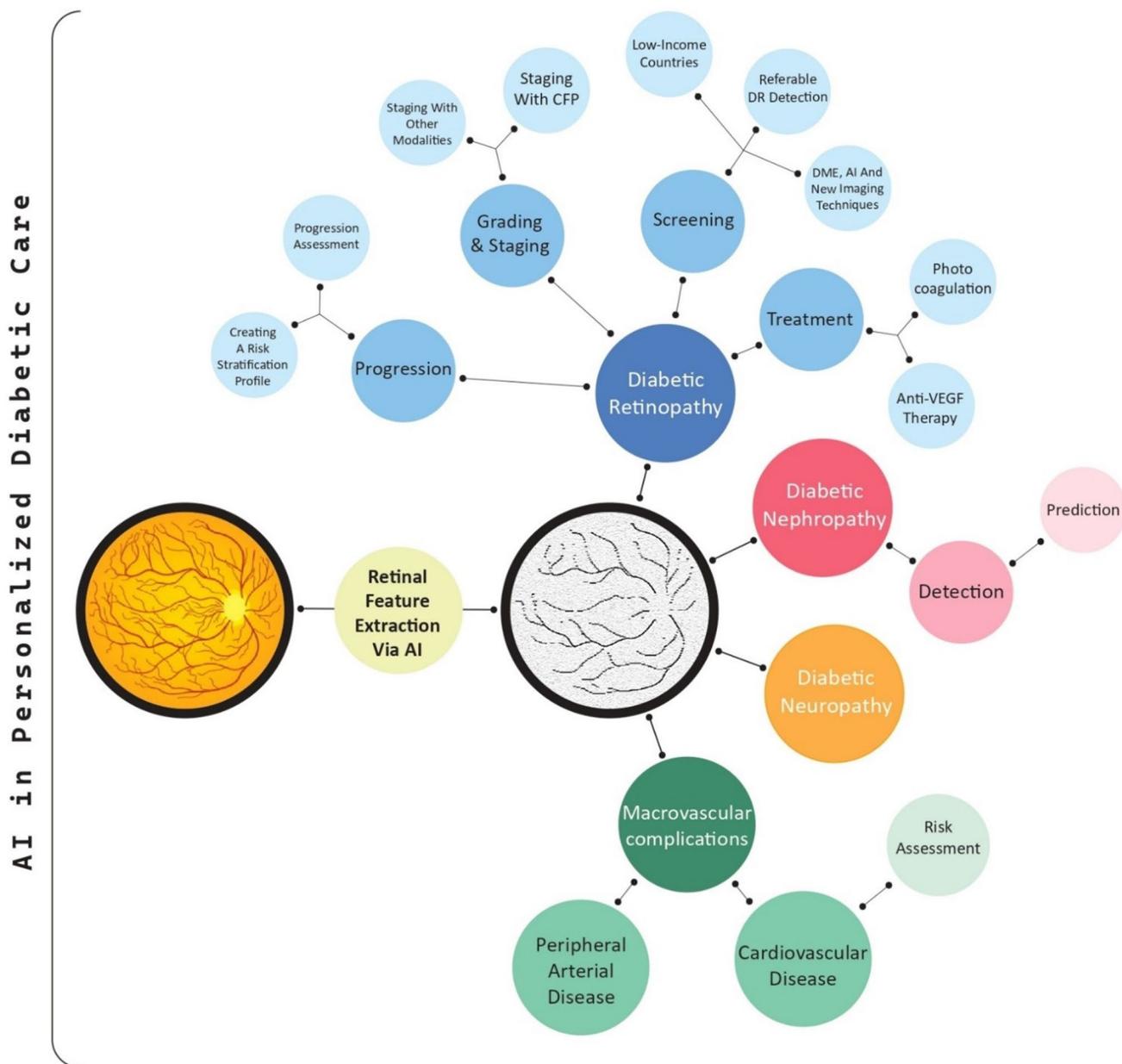
Diabetic nephropathy, a microvascular complication of DM, is a common cause of chronic kidney disease (CKD) [106–109]. The global prevalence of CKD is about 13.4% [110]. According to the Global Burden of Disease (GBD) study, CKD-related deaths increased by 41.5% between 1990 and 2017 [111]. In 2020, the World Health Organization ranked CKD as the tenth leading cause of death and projects that it will become the fifth leading cause by 2040 [112]. Early diagnosis by blood and urine screening assays, including urinary albumin-to-creatinine ratio (ACR), is recommended to identify at-risk patients and initiate interventions that slow the progression of kidney dysfunction. However, persistent albuminuria, which means  $ACR > 30$  mg/g, and  $eGFR < 60$  ml/min/1.73 m<sup>2</sup>, in the context of DM and in conjunction with other complications such as DR, are sufficient to establish the diagnosis of diabetic nephropathy [113, 114]. Retinal images provide clues to systemic microcirculatory health and may be used to screen for diabetic nephropathy. In a study by Sabanayagam et al. [115], clinical parameters and retinal images were utilized both independently and in combination within a CondenseNet DL model to predict diabetic nephropathy. The model achieved AUC values of 0.916 for clinical risk factors alone, 0.911 for retinal images alone, and 0.938 when both input types were combined. In a similar study, a ResNet50 DL model attained maximum AUCs of 0.861, 0.918, and 0.930 for clinical risk factors, retinal images, and hybrid inputs, respectively [116]. These findings corroborate the superior performance of AI-based diabetic nephropathy detection using combined clinical and retinal image inputs [117].

### Retinal images for diagnosing peripheral arterial disease

Peripheral arterial disease (PAD) is a macrovascular complication of DM, with an incidence rate of 25–30% [118]. The elevated risk of PAD in type 2 DM (T2DM) can be attributed to a combination of traditional cardiovascular risk factors—age, sex, race, smoking, pulse pressure, glycated hemoglobin, albuminuria, and hyperlipidemia [119]—and diabetes-specific factors, including postprandial hyperglycemia, advanced glycation end-products, lipoproteins, and hypercoagulability [120]. A low ankle-brachial pressure index is commonly used to screen for lower limb PAD non-invasively, but may be falsely high in elderly people with inelastic arteries. Alternative noninvasive diagnostic techniques are ultrasound Doppler waveform analysis and toe-brachial index [121]. However, the gold standard technique for diagnosing PAD is Computed tomography angiography

(CTA) [122]. Deep neural network architecture has recently been demonstrated to be a promising method for detecting PAD using CFP. The most successful model in this inventive methodology achieved an AUROC maximum score of 0.890. Moreover, visualizing the attention weights used by the network, provides valuable insights into its decision-making process, particularly the significance of ocular features in PAD. Statistical analysis of the model’s performance confirmed that the optic disc and temporal arcades were assigned significantly higher importance ( $p < 0.001$ ) than the retinal background in the detection process. These

results robustly support the feasibility and effectiveness of utilizing modern DL methodologies to detect PAD [123]. However, a major challenge in employing AI to diagnose PAD from retinal images is the difficulty of integrating high-quality images and assembling comprehensive training datasets with labeled PAD cases into the AI frameworks [124, 125]. Figure 3 shows a systematic overview of the use of AI in diabetes care.



**Fig. 3** The diagram provides a structured overview of the integration of AI in diabetic care, showing the pathways in which AI has significant applications, including DR, nephropathy, neuropathy, cerebrovas-

cular disease, peripheral arterial disease, and cardiovascular assessment. Each branch highlights the potential of AI to enhance diagnosis, monitoring, and management within specialized areas of diabetic care

## AI techniques used in retinal images analysis

AI researchers have used various models to analyze retinal images; the two most common neural networks are the CNN-based architectures ResNet [125–127] and VGGNet [128, 129]. Figure 4 illustrates the architecture of a CNN specifically designed for image analysis. Deep neural networks often fail in practice due to difficulties in optimizing the networks caused by the diminishing gradient problem. ResNet overcomes this issue by residual learning, which uses skip or shortcut connections to skip one or more layers during forward and backward passes [125, 130]. This ensures that the deeper layers produce no higher training errors than their shallower counterparts, which makes it easier for the network to learn identity functions. ResNet's modular architecture also facilitates up or down scaling of the model by adjusting the number of residual blocks. This makes it an ideal choice for developers because they can tailor the architecture to specific needs and computational constraints. Belying its depth, ResNet consumes comparatively fewer computational resources, which underscores its efficiency. VGGNet is a family of deep CNNs distinguished by its simple design, ability to capture complex image characteristics, and excellent image classification performance. VGG16, a notable variant comprising 16 layers [128, 129], begins with layers containing 64 channels, with a  $3 \times 3$  filter size and consistent padding, which are succeeded by a max-pooling layer with a stride of (2, 2), and subsequent convolution layers that progressively increase in channels, to for example 128, while maintaining a uniform  $3 \times 3$  filter size. VGGNet is predominantly engineered for image classification on extensive datasets, notably the ImageNet database, which encompasses over 14 million images categorized based on the WordNet structure. Beyond image classification, advancements in VGG architectures have been pursued to cater to diverse computer vision applications and to augment its efficacy in classification problems [131, 132].

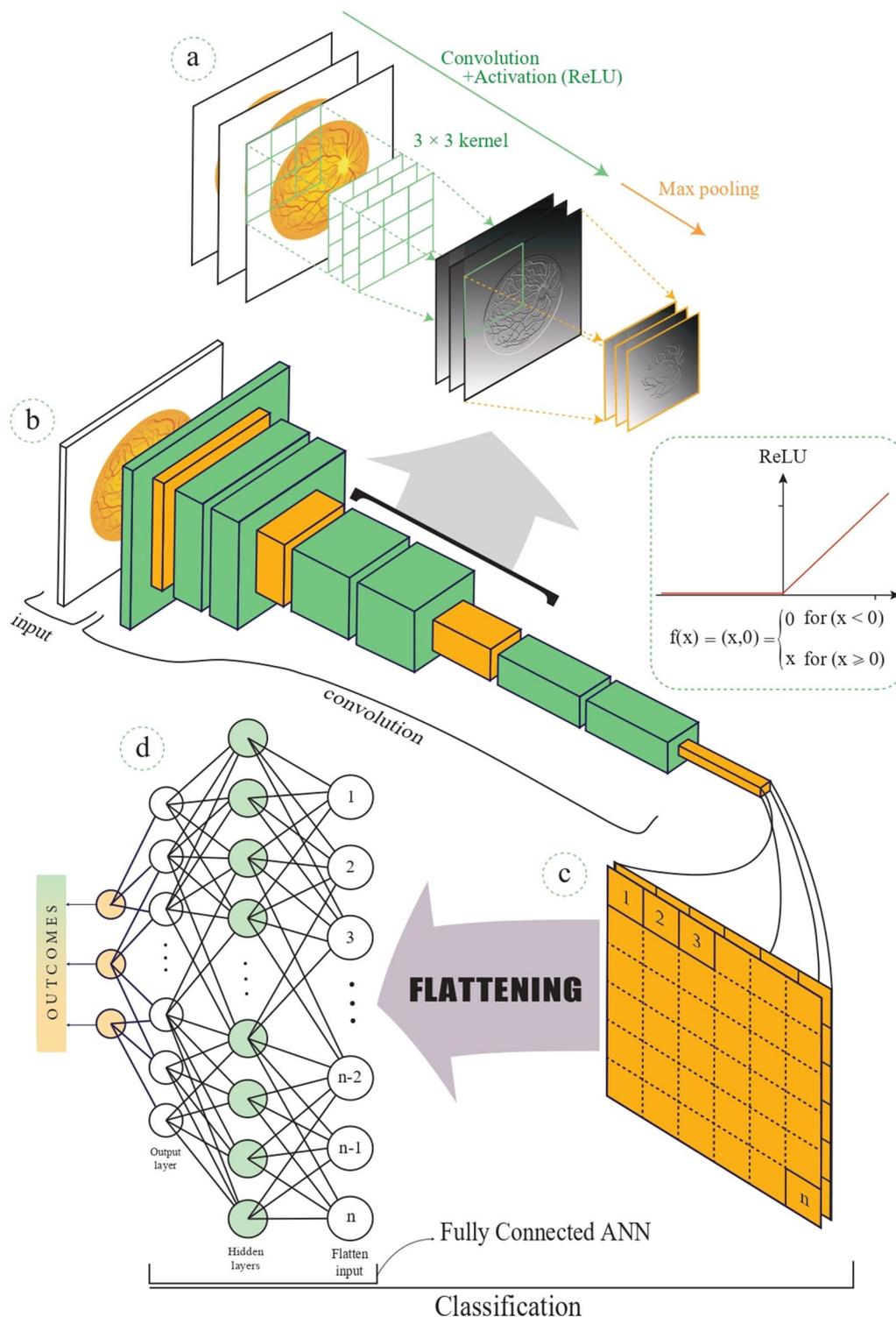
Inception-v3, DenseNet, AlexNet, ResNet, and U-Net are neural network architectures with distinct structural variations that have been applied to retinal image analysis [133]. Inception-v3, from the GoogLeNet family, was designed for multi-level feature extraction tailored to large-scale image recognition tasks, incorporating factorization and other approaches, such as batch normalization, to minimize parameters and maximize performance [134, 135]. DenseNet is well-suited for image classification and segmentation due to dense connections between layers, which improve gradient flow and encourage feature reuse [136]. AlexNet, an early CNN, is adept at image classification due to its deep architecture and use of the rectified linear unit (RELU) activation but may be outperformed by later models in terms of efficiency [137]. By expanding on residuals

and using split-attention methods, ResNet is able to train the model to focus on the most relevant characteristics for image classification [138]. Unlike Inception-v3 and AlexNet, which are more general-purpose models, U-Net was developed specifically for segmenting biological images [139]. Its “U-shaped” design allows for more accurate localization, which is important in diagnostic imaging [140].

Supervised ML learns by mapping the input signals or images to their respective labels. Supervised ML techniques have been extensively employed in retinal image analysis for the detection and classification of various retinal diseases, including DR. With labeled retinal images, the algorithms can effectively learn the patterns associated with DR, thereby facilitating early diagnosis and timely interventions to avert irreversible vision loss [25, 141]. TL uses pretrained datasets, such as ImageNet [142], to overcome the need for large retina-specific training datasets [143]. For retinal image analysis, the final layers of the pre-trained model are typically adjusted to cater to the specificities of the retinal diseases or features. After this adjustment, the model is fine-tuned on available retinal datasets. TL accelerates the training process, often yielding models with enhanced accuracy and robustness that can be applied to diverse tasks including DR detection, retinal lesion identification, and retinal vessel segmentation [25, 62]. A recent study analyzing the morphological characteristics of retinal vessels in DR patients demonstrated that TL can accurately quantify vascular changes, revealing significant differences between DR patients and healthy individuals, and showing that DME does not alter the overall retinal vascular pattern [144].

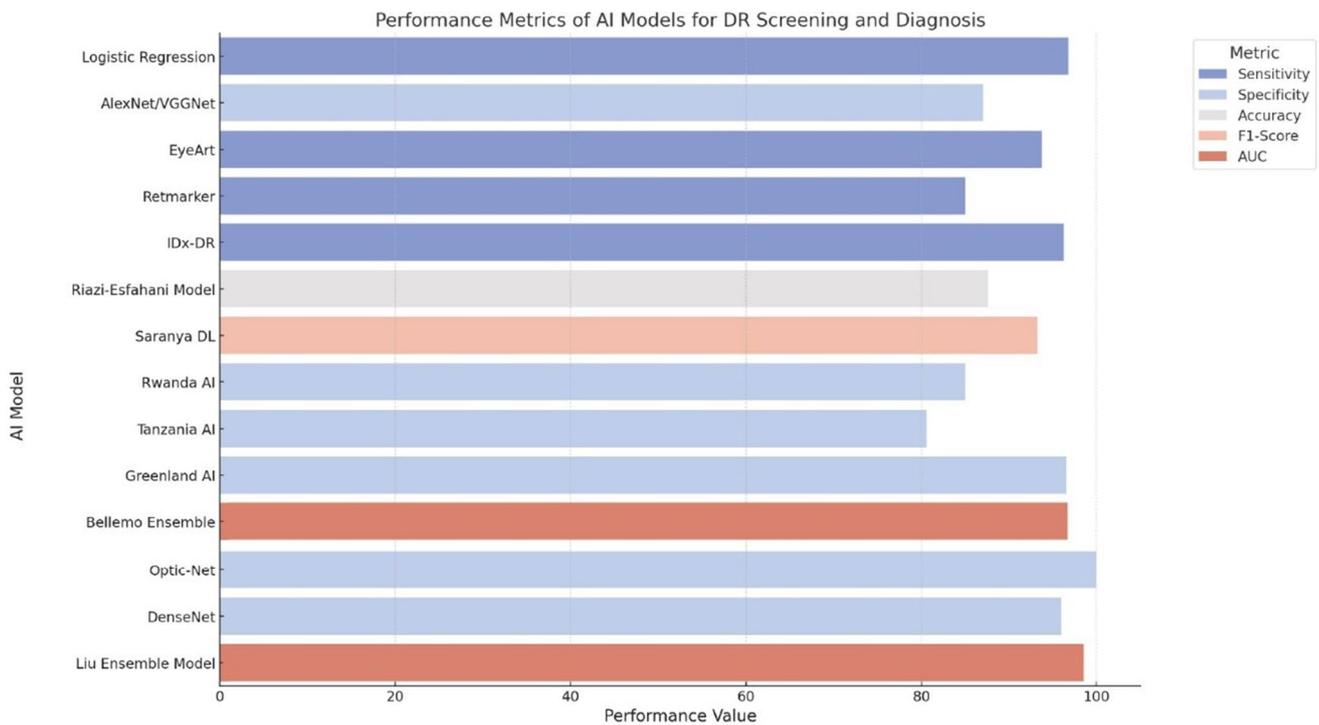
Various AI models are compared in terms of sensitivity, specificity, F1-score, and AUC for diagnosing and screening DR complications, as shown in Fig. 5. This highlights the quality and weakness of each model. For example, Logistic Regression and Saranya DL achieved high sensitivity (more than 96%), whereas the EyeArt and Bellemo Ensemble models did not exceed specificity and showed a balanced result between sensitivity and specificity. Optic-Net is a model that has maximum specificity (100%) and high accuracy (98%), making it the most dependable for DR screening with minimum false positives. Conversely, specificity, one of the important characteristics of the Rwanda AI model, is only moderate, which means there is a wide scope for further improvement, specifically when resources are limited.

Such diversity in the performance of AI represents AI's ability to fit in with a range of clinical needs. For instance, if the sensitivity of a test is extremely high, then many cases can be detected. Conversely, if the specificity is the main goal of a test, not so many referrals will be made. These indicators show that Optic-Net was the best model in terms of precision, whereas Saranya DL and Bellemo Ensemble



**Fig. 4** A detailed look at the architecture of a convolutional neural network (CNN) for image analysis. **a:** The convolutional layers (green) use kernels that slide over the three-channel RGB image to recognize key features from the input image. Following the convolution process, the Rectified Linear Unit (ReLU) activation function (green) is applied to introduce non-linearity, enhancing the network’s capability to learn intricate patterns. Subsequently, the max pooling process (orange) is applied, reducing the spatial dimensions by selecting the maximum

value within specified regions. **b:** this part is the multiple repetitions of convolution, ReLU activation function, and max pooling processes, creating the final feature map. **c:** shows flattening of the last max pooling layer, which converts the 2D feature maps into a 1D vector to prepare the data for the upcoming fully connected layers. **d:** the architecture transitions to fully connected layers, leading to a classification process where the features are used to provide definitive conclusions



**Fig. 5** Performance comparison of AI models across the best metrics for DR screening and diagnosis, as obtained from the included studies. All AUC values are multiplied by 100

adhered to the criteria of sensitivity and specificity, thus allowing broader screening. The Liu Ensemble Model also performs well, with an impressive AUC of 0.981, showing strong overall diagnostic reliability. The best choice depends on the clinical priorities and whether precision or balance is more critical.

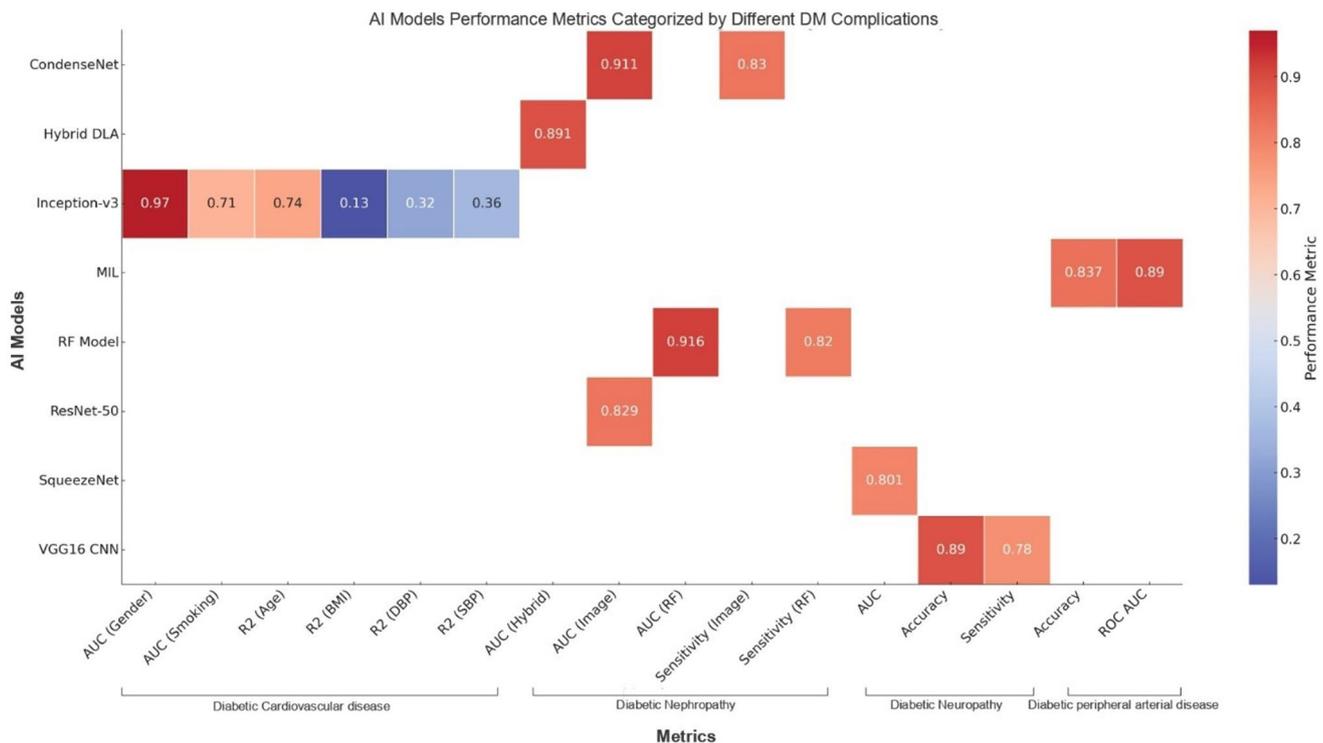
The nuanced interplay between the AI models and their performance metrics in diagnosing diabetic complications is shown in Fig. 6. Models such as VGG16 CNN exhibit strong specificity (0.95), making them reliable for reducing false positives in clinical screening. High AUC values, as seen with Hybrid DLA and Inception-v3, reflect their robust discriminatory power for differentiating cases, ensuring accurate diagnostic outcomes. Sensitivity-focused models, such as CondenseNet, effectively minimize missed diagnoses, which are critical for early detection in high-risk populations. Accuracy, exemplified by MIL and VGG16 CNN, highlights a model's overall reliability across diverse settings. This heatmap also reveals key patterns, such as the consistent link between models excelling in sensitivity and those designed for nuanced clinical predictions and the role of metrics such as AUC in indicating broad diagnostic capabilities. However, models optimized for specificity, like RF, may exhibit slightly lower sensitivity, suggesting trade-offs between reducing false positives and ensuring inclusivity in detection. While these performance metrics provide a holistic view of AI model effectiveness, contextual factors like

patient demographics, imaging protocols, and dataset variability must be considered to ensure the findings' clinical applicability and reliability.

## Challenges of AI in diabetes care

Only two Asian countries have DR screening programs that conform to International Council of Ophthalmology standards [145]. There is not only an imperative to standardize national protocols to mitigate gaps in screening and referral timelines [146] but also a need to establish comprehensive guidelines for AI implementation in DR screening to ensure standardized and effective practice [147]. AI-based DR screening can reduce economic constraints and enhance accessibility to healthcare services [148], but several obstacles remain to surmount [149]. Overcoming these challenges will require multidisciplinary cooperation, data standardization, resource sharing, real-world verification, and productization [150]. In particular, DL algorithms require large datasets with thousands or millions of images for training, which are costly to label and curate.

AI developers often resort to using available but limited training datasets, which may not be generalizable to the real world, where image quality may be affected by deficiencies in the clinical setting and are not of consistently high quality [151]. Suboptimal image quality and low target pathology



**Fig. 6** Comparison of AI model performance across various metrics for diabetic complication diagnosis, based on data extracted from the included studies

incidence may drive higher false-positive rates [152]. Even if image issues are resolved, AI algorithms that only focus on retinal image assessment will not be able to analyze the clinical and psychosocial aspects that can modulate the diagnosis, perception, coping mechanisms, and holistic management of individual patients [153].

The lack of compatibility among electronic health record vendors impedes such clinical data integration and may delay the execution of computer-interpretable guidelines for clinical decision-making in DR [154]. To garner wider adoption, there is a need for more research on the assessment and testing of AI diagnostic tools in clinical settings [155], even as validating AI algorithms in different populations and camera systems remains challenging [156]. Additionally, conducting rigorous clinical trials of AI models, particularly randomized controlled trials, requires meticulous methodological adjustments and considerations of clinical equipoise, informed consent, and fairness [157]. Practical challenges to building and deploying AI at scale include regulatory pressures, conflicting business goals, and data quality issues [158]. Even as AI promises solutions to problems through technological advances, AI in healthcare raises ethical and legal issues [159]. As AI systems become more autonomous, the need to incorporate ethical considerations and moral reasoning into model decision-making processes has become more pressing [160]. Equally important is the

obligation to identify and resolve ethical issues—inclusivity, bias, and social acceptability—to ensure fair user access to promising healthcare AI technologies [161]. Issues such as protecting sensitive medical information during image analysis or remote sharing via tele-retinal screening [162] also raise patient privacy concerns, requiring careful system security during AI solution development and deployment.

Current approaches to AI model design are delinked mainly from the complex healthcare environments they are intended for; as a result, the development of AI models has vastly outpaced their adoption into existing clinical workflows [163]. This separation contributes to models that lack clear use cases and are neither tested nor scaled in clinical settings. A mixed-methods approach that integrates design thinking and quality improvement methodologies—aiming to understand variations in healthcare processes and incorporating user-centered design to ensure model functionality in practice [164]—can potentially compensate this gap to smoothen AI integration within the healthcare domain [165], and gather wider clinical adoption [166]. Moreover, ensuring that AI models are interpretable is crucial for building trust among clinicians [167]. Techniques like Gradient-weighted Class Activation Mapping (Grad-CAM) play pivotal roles in this regard by generating visual explanations that highlight regions in retinal images significantly influencing the model's decisions, thereby allowing

clinicians to verify that AI systems focus on medically relevant areas [168]; moreover, integrating Grad-CAM with local interpretable model-agnostic explanations (LIME) further enhances explainability by enabling visualization of the decision-making process, which in turn improves trust in AI-generated outcomes [169]. For instance, Mercaldo et al. utilized Grad-CAM with CNN to differentiate between healthy eyes and those with DR, achieving an accuracy of 98%, and further distinguishing between stages of DR with 91% accuracy. Their study also introduced a similarity index to evaluate the robustness of heatmaps generated by class activation mapping algorithms, ensuring consistent localization of symptomatic areas in angiography [170].

Real-time AI processing can analyze retinal images within seconds, facilitating quicker diagnoses and timely interventions, which is crucial for preventing disease progression in diabetic patients [171]. Ruamviboonsuk et al. conducted a prospective study in Thailand to evaluate a DL system for real-time DR detection in a community-based screening program. The system demonstrated 94.7% accuracy, 91.4% sensitivity, and 95.4% specificity in detecting vision-threatening DR, comparable to retina specialists. This study emphasizes the importance of integrating socioenvironmental factors and workflows into large-scale AI-driven screening programs in low- and middle-income countries [44]. Natarajan et al. explored the use of a real-time AI system, Medios AI, for smartphone-based retinal imaging by minimally trained health workers in Mumbai, India. The system achieved a sensitivity of 100% and specificity of 88.4% for referable DR detection. These results highlight the potential of AI systems in community-level DR screening, particularly in remote areas lacking access to ophthalmologists [171].

AI screening systems have shown promising results in detecting DR from CFP and OCT images. Despite the advancements, there remains a critical need for validation, regulatory frameworks, safe implementation, and demonstration of clinical impact before these innovations can be widely adopted on a large scale. Testing AI models in clinical settings is crucial for identifying and addressing system issues before full deployment [172, 173]. Additionally, AI tools must be validated and calibrated for local populations and clinical contexts, as results from one setting may not be universally applicable [174]. Burlina et al. [175] demonstrated that AI tools trained on data from specific population groups can exhibit disparities in performance when applied to other groups, which introduce biases in the diagnosis of DR. Similarly, Rogers et al. [176] found that AI models trained on fundus images captured using standard desktop fundus cameras failed to provide accurate diagnoses when evaluated on images taken with handheld portable cameras. This highlights how variations in imaging

device configurations can negatively impact model performance. These findings show the importance of training AI models on diverse datasets that encompass various demographics and imaging conditions to enhance their generalizability and ensure equitable healthcare outcomes. To further address these challenges, tailored solutions, such as localized AI systems designed for specific regions or populations and collaborative initiatives sensitive to cultural and socioeconomic contexts, could significantly improve model performance across varied patient groups [177].

Interpretability is essential for gaining the trust of clinicians and patients alike, in complex AI models of DR diagnosis or treatment decision-making. There is a need for transparent and interpretable AI in the reasoning processes behind the generated model outputs. Explainable AI (XAI) refers to systems that provide understandable explanations for their decisions and actions. While explainable AI is an active area of research, achieving reliable and interpretable explanations remains a significant challenge [178].

Finally, regulatory compliance is also a crucial function that allows for the application and operation of AI models in the context of DR diagnosis while also ensuring their safety, efficiency, and universality. After passing through the demanding FDA process, which involves extreme validation, the like of IDx-DR can demonstrate the highest sensitivity and specificity in clinical settings. This example serves to highlight the transparency of the algorithm development process, the careful reporting of validation results, and the continuous monitoring of the post-market period [15]. Aligning with regulatory standards by promoting collaborations within the international setting creates AI research space from the challenges it faces and thus allows for the safe practice of AI in clinical care [177].

## Potential enhancements and future directions of AI in DR management

At the intersection of telemedicine and AI, tele-retinal image analysis promises to democratize access to screening and downstream healthcare services, transform the management of DR, and improve long-term patient outcomes while reducing financial and time costs for both patients and payers. AI-based DR screening has demonstrated encouraging outcomes, with DL algorithms yielding high levels of sensitivity and specificity. The operational efficiency of community-based tele-retinal image analysis may be enhanced: the gradability of retinal images can be assessed at source, expediting identification of poor-quality images for either manual or additional AI-based grading (the ultimate choice will depend on the associated labor cost vs. intrinsic diagnostic value-add of AI) [179–181]. Future technological

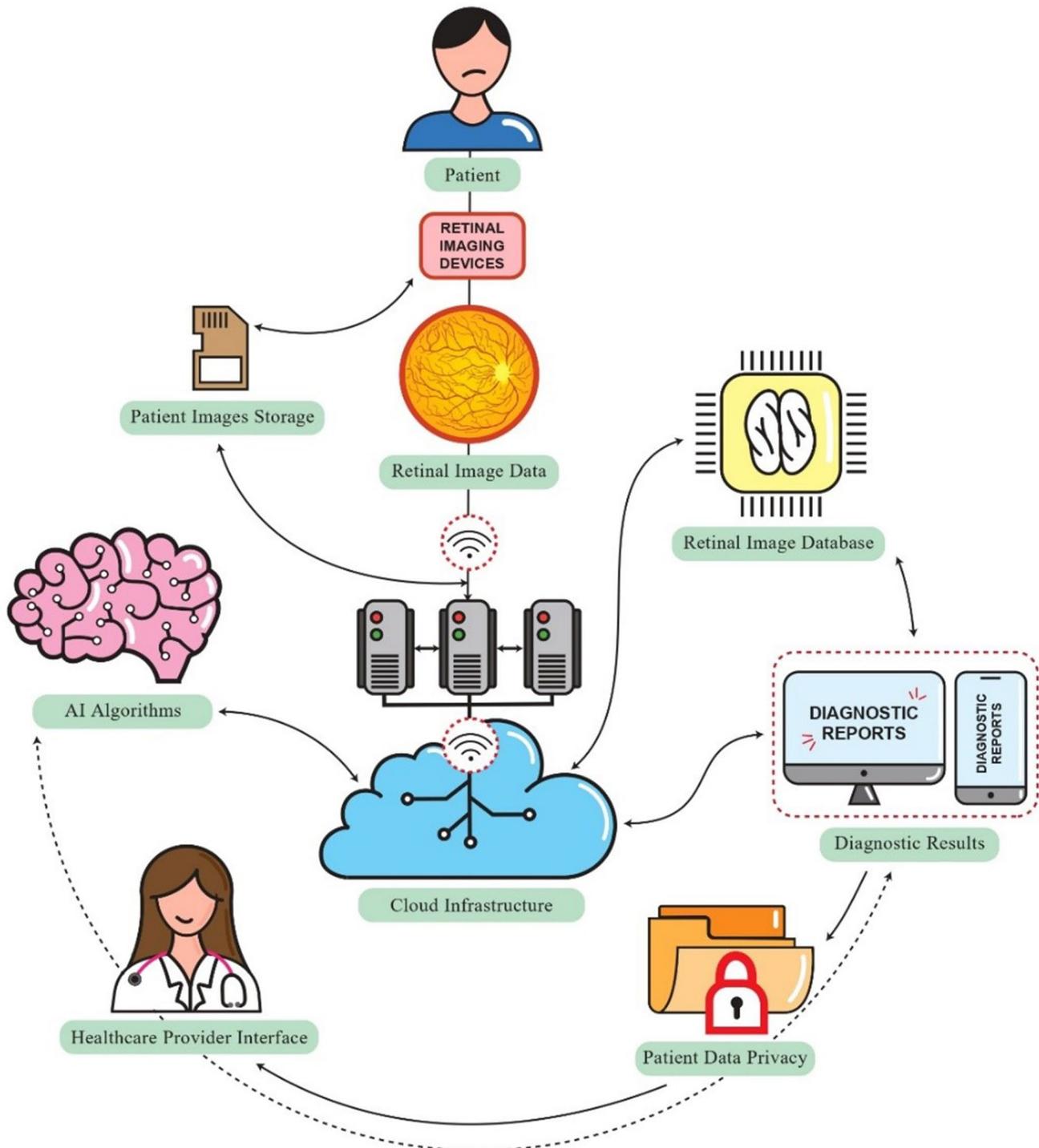
developments in AI can introduce significant opportunities for technical refinements to optimize DR diagnosis and downstream management. Faster AI software processing speeds will enhance system accuracy and efficiency, facilitate seamless and responsive navigation of the AI interface by clinician users, and enable more effective preventive therapeutic interventions. Independent of these developments, real-world implementation of AI technology presents its practical challenges, e.g., workflow integration, technical adaptability, ethical implications, cost-effectiveness considerations, etc. To ensure the ethical and balanced integration of AI into DR screening programs, the governance model for AI implementation must focus on honesty, equality, reliability, and responsibility [172–174].

Ensemble learning is a machine learning approach that does accurate diagnostics in medical applications by using multiple algorithms to make predictions that are more reliable and robust. Methods like bagging, boosting, and stacking are usually employed in this method. Bagging distributes data through different subsets and trains multiple models. Boosting which is the sequential improvement of the previous models, was demonstrated in modeling liver and diabetes. Stacking, which involves using another model (called a meta-learner) to combine predictions from the multiple models, has been proven to be the most effective among other methods in various diseases [182]. However, they do not only decrease the risk of overfitting but also increase a model's capability of generalizing over different datasets which is one of the reasons they are widely used in clinical settings. Moreover, ensemble learning is good for coping with problem areas like imbalanced datasets, which are usually seen in medical diagnostics, by strengthening model stability and reducing false outcomes [183]. Ensemble learning techniques, which combine multiple models, have shown to significantly improve accuracy in DR detection and classification compared to single-model approaches. Mondal et al. proposed an automated ensemble of DenseNet101 and ResNeXt models, combining DenseNet's efficient feature utilization with ResNeXt's advanced split-transform-merge strategy. Their approach, applied to preprocessed datasets (APTOS19 and DIARETDB1), achieved high accuracy for both two-class (96.98%) and five-class (86.08%) classifications, with robust precision and recall, particularly aided by GAN-based data augmentation to address class imbalance [184]. Similarly, Lukashovich et al. explored ensemble learning with a focus on hyperparameter optimization, combining grid search and random search techniques. Their gradient boosting model achieved strong performance, with a binary classification accuracy of 94% and a multi-class staging accuracy of 75.31%. Both studies highlight ensemble learning's capacity to improve DR screening and staging, emphasizing the importance of integrating advanced

algorithms and robust data preprocessing to achieve clinically relevant outcomes [185].

Explainable AI make the complex decision-making processes of AI models transparent and interpretable, which is essential, given the clinical impact of healthcare decisions, for gaining the trust and acceptance of AI model outputs by doctors and patients. The applications of explainable AI in the medical domain are vast and transformative [186, 187], encompassing diverse tasks, such as decision-making, risk management, predictions, and medical image analysis (the sensitivity of AI for detecting abnormalities often surpasses that of the human eye). Explainable AI aids in explaining AI-driven insights. For example, while a traditional AI model may provide a diagnosis with a confidence score ("DR detected with 95% confidence"), ophthalmologists may hesitate to Google Net trust the result or administer recommended treatment without knowing the basis for the diagnosis. With explainable AI, the system can highlight regions in the retina or specific lesions, like hemorrhages or microaneurysms, critical to the model decision via heat maps. These allow clinicians to relate to the results like traditional fundoscopic examination, and the maps may additionally serve as useful guides to the planning of photocoagulation therapy, where applicable [188, 189].

Cloud-based systems revolutionize diabetes management and prevention by enhancing data processing, enabling real-time interventions, and optimizing resources. They facilitate early detection, personalized therapy, and swift glucose fluctuation response. Migration to cloud structures reduces costs and administrative burdens, while user-friendly digital platforms support self-monitoring and community engagement. Integrating AI with cloud platforms promises sharper insights for combating diseases like Type II diabetes, leading to societal benefits. Moreover, some researchers have investigated the potential of cloud-based systems for diabetes management and prevention [190, 191]. A study by Salari et al. [192] has shown promising results in using mobile and cloud systems to improve self-care for chronic conditions, offering hope for more effective and user-friendly solutions. Similarly, Nasser et al. [193] have proposed innovative methods using advanced technology, such as AI and cloud computing, to predict glucose levels and integrate them with wearable devices. These advancements point toward a future where cloud-based systems could revolutionize how diabetes is managed, offering personalized and timely interventions to enhance health outcomes for individuals with the condition. Figure 7 illustrates the interconnected components of a cloud-based system for DR screening.



**Fig. 7** This schematic illustrates the interconnected components of a cloud-based system for DR screening. The cloud infrastructure, depicted as remote servers and databases, processes retinal image data captured by a retinal imaging device, such as a fundus camera or OCT scanner. AI algorithms analyze the images within the cloud, generating diagnostic results like risk scores for DR. These results

are accessible to healthcare providers through interfaces on computers or mobile devices, ensuring prompt patient care. Patient data privacy measures safeguard sensitive information, including encryption and secure transmission protocols. Additionally, a feedback loop may exist, where diagnostic results contribute to the continuous improvement of AI algorithms over time

## Conclusion

Given the increasing prevalence and incidence of DM, developing cost-effective, population-based DR screening and management strategies is becoming increasingly important nationally. AI has shown significant promise in accurately diagnosing DR, with potential applicability extending beyond the diagnosis and grading of DR. This includes the diagnosis of diabetic neuropathy, diabetic nephropathy, and cardiovascular diseases. These applications may involve using CFP independently or in conjunction with other advanced diagnostic techniques, such as OCT and various clinical parameters. AI's capacity to assess a patient's health about DM complications and to forecast the risk of future complications positions AI-assisted retinal image analysis as a potentially key component in personalized medicine for individuals with DM.

**Acknowledgements** We appreciate the editing and paraphrasing services provided by Grammarly and Ref-N-Write.

**Author contributions** AJ conceptualized the manuscript. YSB, MM, MA, AJ, and NS conducted a literature review. YSB, MM, MA, and NS drafted the manuscript. SP, URA, RA, RST, SMSI, and AJ provided comments and revised the manuscript. NS and AJ created and designed all the figures. All authors read and approved the final version of the manuscript. It should be noted that YSB, MM, and NS contributed equally to this study. Every author adhered to the COPE guidelines as universally acknowledged.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions  
None.

**Data availability** N/A.

## Declarations

**Ethics approval and consent to participate** This study was conducted under the principles outlined in the Declaration of Helsinki.

**Consent for publication** N/A.

**Competing interests** The authors declare that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Glovaci D, Fan W, Wong ND. Epidemiology of diabetes mellitus and cardiovascular disease. *Curr Cardiol Rep.* 2019;21(4):21. <https://doi.org/10.1007/s11886-019-1107-y>.
- Arokiasamy P, Salvi S, Selvamani Y. Global burden of diabetes mellitus. *Handbook of global health.* Springer; 2021. pp. 1–44.
- Jeffcoate WJ, Harding KG. Diabetic foot ulcers. *Lancet.* 2003;361(9368):1545–51. [https://doi.org/10.1016/S0140-6736\(03\)13169-8](https://doi.org/10.1016/S0140-6736(03)13169-8).
- Koye DN, Magliano DJ, Nelson RG, Pavkov ME. The global epidemiology of diabetes and kidney disease. *Adv Chronic Kidney Dis.* 2018;25(2):121–32. <https://doi.org/10.1053/j.ackd.2017.10.011>.
- Cheung N, Mitchell P, Wong TY. Diabetic retinopathy. *Lancet.* 2010;376(9735):124–36. [https://doi.org/10.1016/s0140-6736\(09\)6124-3](https://doi.org/10.1016/s0140-6736(09)6124-3).
- Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care.* 2012;35(3):556–64. <https://doi.org/10.2337/dc11-1909>.
- Singh R, Ramasamy K, Abraham C, Gupta V, Gupta A. Diabetic retinopathy: an update. *Indian J Ophthalmol.* 2008;56(3):179–88.
- Otani T, Kishi S, Maruyama Y. Patterns of diabetic macular edema with optical coherence tomography. *Am J Ophthalmol.* 1999;127(6):688–93.
- Yanoff M, Fine BS, Brucker AJ, Eagle RC Jr. Pathology of human cystoid macular edema. *Surv Ophthalmol.* 1984;28:505–11.
- Abramoff MD, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol.* 2013;131(3):351–7. <https://doi.org/10.1001/jamaophthalmol.2013.1743>
- Kollias AN, Ulbig MW. Diabetic retinopathy: early diagnosis and effective treatment. *Dtsch Arztebl Int.* 2010;107(5):75–83. <https://doi.org/10.3238/arztebl.2010.0075>. quiz 4.
- McMillan DE. The microcirculation in diabetes. *Microcirc Endothelium Lymphatics.* 1984;1(1):3–24.
- Abbas K, Lu Y, Bavishi S, Mishra N, TomThundyil S, Sawant SA, et al. A simple review of small vessel disease manifestation in the brain, retina, and kidneys. *J Clin Med.* 2022. <https://doi.org/10.3390/jcm11195546>.
- Li S, Zhao R, Zou H. Artificial intelligence for diabetic retinopathy. *Chin Med J (Engl).* 2021;135(3):253–60. <https://doi.org/10.1097/cm9.0000000000001816>.
- Rajesh AE, Davidson OQ, Lee CS, Lee AY. Artificial intelligence and diabetic retinopathy: AI framework, prospective studies, Head-to-head validation, and Cost-effectiveness. *Diabetes Care.* 2023;46(10):1728–39. <https://doi.org/10.2337/dci23-0032>.
- Uppamma P, Bhattacharya S. Deep learning and medical image processing techniques for diabetic retinopathy: A survey of applications, challenges, and future trends. *J Healthc Eng.* 2023;2023:2728719. <https://doi.org/10.1155/2023/2728719>.
- Gonçalves MB, Nakayama LF, Ferraz D, Faber H, Korot E, Malerbi FK, et al. Image quality assessment of retinal fundus photographs for diabetic retinopathy in the machine learning era: a review. *Eye (Lond).* 2024;38(3):426–33. <https://doi.org/10.1038/s41433-023-02717-3>.
- Balakrishnan U, Venkatachalapathy K, Marimuthu GS. A hybrid PSO-DEFS based feature selection for the identification of diabetic retinopathy. *Curr Diabetes Rev.* 2015;11(3):182–90. <https://doi.org/10.2174/1573399811666150330150038>.
- Mudie LI, Wang X, Friedman DS, Brady CJ. Crowdsourcing and automated retinal image analysis for diabetic retinopathy. *Curr*

- Diab Rep. 2017;17(11):106. <https://doi.org/10.1007/s11892-017-0940-x>.
20. Raman R, Srinivasan S, Virmani S, Sivaprasad S, Rao C, Rajalakshmi R. Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. *Eye (Lond)*. 2019;33(1):97–109. <https://doi.org/10.1038/s41433-018-0269-y>.
  21. Asiri N, Hussain M, Al Adel F, Alzaidi N. Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey. *Artif Intell Med*. 2019;99:101701. <https://doi.org/10.1016/j.artmed.2019.07.009>.
  22. Rajalakshmi R, Prathiba V, Arulmalar S, Usha M. Review of retinal cameras for global coverage of diabetic retinopathy screening. *Eye (Lond)*. 2021;35(1):162–72. <https://doi.org/10.1038/s41433-020-01262-7>.
  23. Shahriari MH, Sabbaghi H, Asadi F, Hosseini A, Khorrami Z. Artificial intelligence in screening, diagnosis, and classification of diabetic macular edema: A systematic review. *Surv Ophthalmol*. 2023;68(1):42–53. <https://doi.org/10.1016/j.survophthal.2022.08.004>.
  24. Yang Z, Tan TE, Shao Y, Wong TY, Li X. Classification of diabetic retinopathy: past, present and future. *Front Endocrinol (Lausanne)*. 2022;13:1079217. <https://doi.org/10.3389/fendo.2022.1079217>.
  25. Abramoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. *Invest Ophthalmol Vis Sci*. 2016;57(13):5200–06. <https://doi.org/10.1167/iovs.16-19964>
  26. Tufail A, Rudisill C, Egan C, Kapetanakis VV, Salas-Vega S, Owen CG, et al. Automated diabetic retinopathy image assessment software: diagnostic accuracy and Cost-Effectiveness compared with human graders. *Ophthalmology*. 2017;124(3):343–51. <https://doi.org/10.1016/j.ophtha.2016.11.014>.
  27. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekaran DV, et al. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a National programme: an economic analysis modelling study. *Lancet Digit Health*. 2020;2(5):e240–9. [https://doi.org/10.1016/S2589-7500\(20\)30060-1](https://doi.org/10.1016/S2589-7500(20)30060-1).
  28. Huang X-M, Yang B-F, Zheng W-L, Liu Q, Xiao F, Ouyang P-W, et al. Cost-effectiveness of artificial intelligence screening for diabetic retinopathy in rural China. *BMC Health Serv Res*. 2022;22(1):260. <https://doi.org/10.1186/s12913-022-07655-6>.
  29. Taylor NP, FDA APPROVES FIRST AI DEVICE TO DETECT DIABETIC RETINOPATHY. <https://www.fiercebiotech.com/medtech/fda-approves-diabetic-retinopathy-detecting-ai-algorithm> (2018). Accessed.
  30. Wang Y, Liu C, Hu W, Luo L, Shi D, Zhang J, et al. Economic evaluation for medical artificial intelligence: accuracy vs. cost-effectiveness in a diabetic retinopathy screening case. *Npj Digit Med*. 2024;7(1):43. <https://doi.org/10.1038/s41746-024-01032-9>.
  31. Hu W, Joseph S, Li R, Woods E, Sun J, Shen M, et al. Population impact and cost-effectiveness of artificial intelligence-based diabetic retinopathy screening in people living with diabetes in Australia: a cost effectiveness analysis. *eClinicalMedicine*. 2024;67:102387. <https://doi.org/10.1016/j.eclinm.2023.102387>.
  32. Riazi-Esfahani H, Jafari B, Azimi H, Rahimi M, Saedian J, Pouya P, et al. Assessment of area and structural irregularity of retinal layers in diabetic retinopathy using machine learning and image processing techniques. *Sci Rep*. 2024;14(1):4013. <https://doi.org/10.1038/s41598-024-54535-6>.
  33. Saranya P, Umamaheswari KM. Detection of exudates from retinal images for non-proliferative diabetic retinopathy detection using deep learning model. *Multimedia Tools Appl*. 2023;83(17):52253–73. <https://doi.org/10.1007/s11042-023-17462-8>.
  34. Whitestone N, Nkurikiye J, Patnaik JL, Jaccard N, Lanouette G, Cherwek DH, et al. Feasibility and acceptance of artificial intelligence-based diabetic retinopathy screening in Rwanda. *Br J Ophthalmol*. 2024;108(6):840–5. <https://doi.org/10.1136/bjo-2022-322683>.
  35. Cleland CR, Bascaran C, Makupa W, Shilio B, Sandi FA, Philipin H, et al. Artificial intelligence-supported diabetic retinopathy screening in Tanzania: rationale and design of a randomised controlled trial. *BMJ Open*. 2024;14(1):e075055–e. <https://doi.org/10.1136/bmjopen-2023-075055>.
  36. Kubin A-M, Huhtinen P, Ohtonen P, Keskitalo A, Wirkkala J, Hautala N. Comparison of 21 artificial intelligence algorithms in automated diabetic retinopathy screening using handheld fundus camera. *Ann Med*. 2024;56(1). <https://doi.org/10.1080/07853890.2024.2352018>.
  37. Larsen TJ, Pettersen MB, Nygaard Jensen H, Lynge Pedersen M, Lund-Andersen H, Jørgensen ME, Byberg S. The use of artificial intelligence to assess diabetic eye disease among the Greenlandic population. *Int J Circumpolar Health*. 2024;83(1). <https://doi.org/10.1080/22423982.2024.2314802>.
  38. Rui Z, Ying Z, Wenji NI, Ya H, Dandan LI, Tao JIN, Yong Z. Application value of artificial intelligence-based retinal microvascular analysis in diagnosis of diabetes complications. *J Practical Med*. 2024;40(8):1142.
  39. Burgess PI, Harding SP, García-Fiñana M, Beare NAV, Glover S, Cohen DB, et al. Incidence and progression of diabetic retinopathy in Sub-Saharan Africa: A five year cohort study. *PLoS ONE*. 2017;12(8):e0181359. <https://doi.org/10.1371/journal.pone.0181359>.
  40. Zheng Y, He M, Congdon N. The worldwide epidemic of diabetic retinopathy. *Indian J Ophthalmol*. 2012;60(5):428–31. <https://doi.org/10.4103/0301-4738.100542>.
  41. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF diabetes atlas: global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract*. 2017;128:40–50. <https://doi.org/10.1016/j.diabres.2017.03.024>.
  42. Huemer J, Wagner SK, Sim DA. The evolution of diabetic retinopathy screening programmes: A chronology of retinal photography from 35 mm slides to artificial intelligence. *Clin Ophthalmol*. 2020;14:2021–35. <https://doi.org/10.2147/ophth.S261629>.
  43. Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT, et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digit Health*. 2019;1(1):e35–44. [https://doi.org/10.1016/s2589-7500\(19\)30004-4](https://doi.org/10.1016/s2589-7500(19)30004-4).
  44. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, et al. Real-time diabetic retinopathy screening by deep learning in a multisite National screening programme: a prospective interventional cohort study. *Lancet Digit Health*. 2022;4(4):e235–44. [https://doi.org/10.1016/S2589-7500\(22\)00017-6](https://doi.org/10.1016/S2589-7500(22)00017-6).
  45. Rajalakshmi R, Subashini R, Anjana RM, Mohan V. Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*. 2018;32(6):1138–44. <https://doi.org/10.1038/s41433-018-0064-9>.
  46. Wong RL, Tsang CW, Wong DS, McGhee S, Lam CH, Lian J, et al. Are we making good use of our public resources? The false-positive rate of screening by fundus photography for diabetic macular oedema. *Hong Kong Med J*. 2017;23(4):356–64. <https://doi.org/10.12809/hkmj166078>.
  47. Mackenzie S, Schmermer C, Charnley A, Sim D, Vikas T, Dumskyj M, et al. SDOCT imaging to identify macular pathology in patients diagnosed with diabetic maculopathy by a

- digital photographic retinal screening programme. *PLoS ONE*. 2011;6(5):e14811. <https://doi.org/10.1371/journal.pone.0014811>.
48. Grauslund J, Andersen N, Andresen J, Flesner P, Haamann P, Heegaard S, et al. Evidence-based Danish guidelines for screening of diabetic retinopathy. *Acta Ophthalmol*. 2018;96(8):763–9. <https://doi.org/10.1111/aos.13936>.
  49. Browning DJ, Glassman AR, Aiello LP, Bressler NM, Bressler SB, Danis RP, et al. Optical coherence tomography measurements and analysis methods in optical coherence tomography studies of diabetic macular edema. *Ophthalmology*. 2008;115(8):1366–71. e1.
  50. Keane PA, Ruiz-Garcia H, Sadda SR. Clinical applications of long-wavelength (1,000-nm) optical coherence tomography. *Ophthalmic Surg Lasers Imaging Retina*. 2011;42(4):S67–74.
  51. Varadarajan AV, Bavishi P, Ruamviboonsuk P, Chotcomwongse P, Venugopalan S, Narayanaswamy A, et al. Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning. *Nat Commun*. 2020;11(1):130. <https://doi.org/10.1038/s41467-019-13922-8>.
  52. Wu Q, Zhang B, Hu Y, Liu B, Cao D, Yang D, DETECTION OF MORPHOLOGIC PATTERNS OF DIABETIC MACULAR EDEMA USING A DEEP LEARNING APPROACH BASED ON OPTICAL COHERENCE TOMOGRAPHY IMAGES, et al. *Retina*. 2021;41(5):1110–7. <https://doi.org/10.1097/IAE.0000000000002992>.
  53. Manikandan S, Raman R, Rajalakshmi R, Tamilselvi S, Surya RJ. Deep learning–based detection of diabetic macular edema using optical coherence tomography and fundus images: A meta-analysis. *Indian J Ophthalmol*. 2023;71(5):1783–96. [https://doi.org/10.4103/IJO.IJO\\_2614\\_22](https://doi.org/10.4103/IJO.IJO_2614_22).
  54. Lam C, Wong YL, Tang Z, Hu X, Nguyen TX, Yang D, et al. Performance of artificial intelligence in detecting diabetic macular edema from fundus photography and optical coherence tomography images: A systematic review and Meta-analysis. *Diabetes Care*. 2024;47(2):304–19. <https://doi.org/10.2337/dc23-0993>.
  55. Li H-Y, Wang D-X, Dong L, Wei W-B. Deep learning algorithms for detection of diabetic macular edema in OCT images: A systematic review and meta-analysis. *Eur J Ophthalmol*. 2023;33(1):278–90. <https://doi.org/10.1177/11206721221094786>.
  56. Pavithra K, Kumar P, Geetha M, Bhandary S. Statistical Analysis of Deep Learning Models for Diabetic Macular Edema Classification using OCT Images. 2022 International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER): IEEE; 2022. pp. 305–10.
  57. Liu R, Li Q, Xu F, Wang S, He J, Cao Y, et al. Application of artificial intelligence-based dual-modality analysis combining fundus photography and optical coherence tomography in diabetic retinopathy screening in a community hospital. *Biomed Eng Online*. 2022;21(1):47. <https://doi.org/10.1186/s12938-022-01018-2>.
  58. Ryu G, Lee K, Park D, Park SH, Sagong M. A deep learning model for identifying diabetic retinopathy using optical coherence tomography angiography. *Sci Rep*. 2021;11(1):23024. <https://doi.org/10.1038/s41598-021-02479-6>.
  59. Ferris FL. How effective are treatments for diabetic retinopathy? *JAMA*. 1993;269(10):1290–1.
  60. Harding SP, Broadbent DM, Neoh C, White MC, Vora J. Sensitivity and specificity of photography and direct ophthalmoscopy in screening for sight threatening eye disease: the Liverpool diabetic eye study. *BMJ*. 1995;311(7013):1131–5. <https://doi.org/10.1136/bmj.311.7013.1131>.
  61. Ye X, Wang J, Chen Y, Lv Z, He S, Mao J, et al. Automatic screening and identifying myopic maculopathy on optical coherence tomography images using deep learning. *Translational Vis Sci Technol*. 2021;10(13):10.
  62. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402–10.
  63. Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H. Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. *PLoS ONE*. 2017;12(6):e0179790.
  64. Ruamviboonsuk P, Krause J, Chotcomwongse P, Sayres R, Raman R, Widner K, et al. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. *NPJ Digit Med*. 2019;2(1):25.
  65. Abbas Q, Fondon I, Sarmiento A, Jiménez S, Alemany P. Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. *Med Biol Eng Comput*. 2017;55:1959–74.
  66. Ting DSW, Cheung CY-L, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA*. 2017;318(22):2211–23.
  67. Li Z, Keel S, Liu C, He Y, Meng W, Scheetz J, et al. An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. *Diabetes Care*. 2018;41(12):2509–16.
  68. Wang Y, Yu M, Hu B, Jin X, Li Y, Zhang X, et al. Deep learning-based detection and stage grading for optimising diagnosis of diabetic retinopathy. *Diab/Metab Res Rev*. 2021;37(4):e3445.
  69. Aishwarya BP, Chandan MS, Preethu KS, Swathi KA. Diabetic retinopathy using inception V3 model. *Int J Res Appl Sci Eng Technol*. 2023;11(4):4612–6. <https://doi.org/10.22214/ijraset.2023.51304>
  70. Prajapati M, Baliarsingh SK, Hota J. MobileNetV3 based Classification Model for Diabetic Retinopathy. 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN): IEEE; 2023. pp. 213–7.
  71. Sandhu HS, Elmogy M, Sharafeldeen AT, Elsharkawy M, El-Adawy N, Eltanboly A, et al. Automated diagnosis of diabetic retinopathy using clinical biomarkers, optical coherence tomography, and optical coherence tomography angiography. *Am J Ophthalmol*. 2020;216:201–6.
  72. Wang X, Ji Z, Ma X, Zhang Z, Yi Z, Zheng H, et al. Automated grading of diabetic retinopathy with Ultra-Widefield fluorescein angiography and deep learning. *J Diabetes Res*. 2021;2021:2611250. <https://doi.org/10.1155/2021/2611250>.
  73. García-Fiñana M, Hughes DM, Cheyne CP, Broadbent DM, Wang A, Komárek A, et al. Personalized risk-based screening for diabetic retinopathy: A multivariate approach versus the use of stratification rules. *Diabetes Obes Metab*. 2019;21(3):560–8. <https://doi.org/10.1111/dom.13552>.
  74. Aspelund T, Þórisdóttir Ó, Ólafsdóttir E, Gudmundsdóttir A, Einarsdóttir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. *Diabetologia*. 2011;54(10):2525. <https://doi.org/10.1007/s00125-011-2257-7>.
  75. Soto-Pedre E, Pinies JA, Hernaez-Ortega MC. External validation of a risk assessment model to adjust the frequency of eye-screening visits in patients with diabetes mellitus. *J Diabetes Complications*. 2015;29(4):508–11. <https://doi.org/10.1016/j.jdiacom.2014.12.020>.
  76. Piri S, Delen D, Liu T, Zolbanin HM. A data analytics approach to Building a clinical decision support system for diabetic retinopathy: developing and deploying a model ensemble. *Decis Support Syst*. 2017;101:12–27. <https://doi.org/10.1016/j.dss.2017.05.012>.
  77. Group ETDRSR. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified airline

- house classification: ETDRS report number 10. *Ophthalmology*. 1991;98(5):786–806.
78. Skevofilakas M, Zarkogianni K, Karamanos BG, Nikita KS. A hybrid decision support system for the risk assessment of retinopathy development as a long term complication of type 1 diabetes mellitus. 2010 Annual Int Conf IEEE Eng Med Biology. 2010;IEEE:6713–6.
  79. Arcadu F, Benmansour F, Maunz A, Willis J, Haskova Z, Prunotto M. Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *NPJ Digit Med*. 2019;2(1):1–9.
  80. Silva PS, Zhang D, Jacoba CMP, Fickweiler W, Lewis D, Leitmeyer J, et al. Automated machine learning for predicting diabetic retinopathy progression from Ultra-Widefield retinal images. *JAMA Ophthalmol*. 2024;142(3):171. <https://doi.org/10.1001/jamaophthalmol.2023.6318>.
  81. Wolf RM, Channa R, Liu TYA, Zehra A, Bromberger L, Patel D, et al. Autonomous artificial intelligence increases screening and follow-up for diabetic retinopathy in youth: the ACCESS randomized control trial. *Nat Commun*. 2024;15(1):421. <https://doi.org/10.1038/s41467-023-44676-z>.
  82. Rosenfeld PJ, Fung AE, Puliafito CA. Optical coherence tomography findings after an intravitreal injection of bevacizumab (avastin) for macular edema from central retinal vein occlusion. *Ophthalmic Surg Lasers Imaging*. 2005;36(4):336–9.
  83. Ng EW, Shima DT, Calias P, Cunningham ET Jr., Guyer DR, Adamis AP. Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov*. 2006;5(2):123–32. <https://doi.org/10.1038/nrd1955>.
  84. Chun DW, Heier JS, Topping TM, Duker JS, Bankert JM. A pilot study of multiple intravitreal injections of Ranibizumab in patients with center-involving clinically significant diabetic macular edema. *Ophthalmology*. 2006;113(10):1706–12. <https://doi.org/10.1016/j.ophtha.2006.04.033>.
  85. Nguyen QD, Oatlipinar S, Shah SM, Haller JA, Quinlan E, Sung J, et al. Vascular endothelial growth factor is a critical stimulus for diabetic macular edema. *Am J Ophthalmol*. 2006;142(6):961–9. <https://doi.org/10.1016/j.ajo.2006.06.068>.
  86. Rasti R, Allingham MJ, Mettu PS, Kavusi S, Govind K, Cousins SW, Farsi S. Deep learning-based single-shot prediction of differential effects of anti-VEGF treatment in patients with diabetic macular edema. *Biomedical Opt Express*. 2020;11(2):1139. <https://doi.org/10.1364/BOE.379150>.
  87. Ye X, Gao K, He S, Zhong X, Shen Y, Wang Y, et al. Artificial Intelligence-Based quantification of central macular fluid volume and VA prediction for diabetic macular edema using OCT images. *Ophthalmol Therapy*. 2023;12(5):2441–52. <https://doi.org/10.1007/s40123-023-00746-5>.
  88. Liu B, Zhang B, Hu Y, Cao D, Yang D, Wu Q et al. Automatic prediction of treatment outcomes in patients with diabetic macular edema using ensemble machine learning. *Annals Translational Med*. 2021;9(1).
  89. Gallardo M, Munk MR, Kurmann T, De Zanet S, Mosinska A, Karagoz IK, et al. Machine learning can predict Anti-VEGF treatment demand in a Treat-and-Extend regimen for patients with neovascular AMD, DME, and RVO associated macular edema. *Ophthalmol Retina*. 2021;5(7):604–24. <https://doi.org/10.1016/j.oret.2021.05.002>.
  90. Zamytskiy EAZA, Karlova E, Zamytskiy P. Analysis of the coagulates intensity in laser treatment of diabetic macular edema in a Navilas robotic laser system. 2017.
  91. Kozak I, Luttrull JK. Modern retinal laser therapy. *Saudi J Ophthalmol*. 2015;29(2):137–46. <https://doi.org/10.1016/j.sjopt.2014.09.001>.
  92. Moutray T, Evans JR, Lois N, Armstrong DJ, Peto T, Azuara-Blanco A. Different lasers and techniques for proliferative diabetic retinopathy. *Cochrane Database Syst Reviews*. 2018(3).
  93. Ilyasova N, Demin N, Shirokanev A, Andriyanov N. Automated system for the personalization of retinal laser treatment in diabetic retinopathy based on the intelligent analysis of OCT data and fundus images. *Intelligent decision technologies*. Springer; 2022. pp. 171–81.
  94. Larsson SC, Wallin A, Håkansson N, Stackelberg O, Bäck M, Wolk A. Type 1 and type 2 diabetes mellitus and incidence of seven cardiovascular diseases. *Int J Cardiol*. 2018;262:66–70. <https://doi.org/10.1016/j.ijcard.2018.03.099>.
  95. Wang M, Francis F, Kunz H, Zhang X, Wan C, Liu Y, et al. Artificial intelligence models for predicting cardiovascular diseases in people with type 2 diabetes: A systematic review. *Intelligence-Based Med*. 2022;6:100072. <https://doi.org/10.1016/j.ibmed.2022.100072>.
  96. Guo VY, Chan JC, Chung H, Ozaki R, So W, Luk A, et al. Retinal information is independently associated with cardiovascular disease in patients with type 2 diabetes. *Sci Rep*. 2016;6:19053. <https://doi.org/10.1038/srep19053>.
  97. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomedical Eng*. 2018;2(3):158–64. <https://doi.org/10.1038/s41551-018-0195-0>.
  98. Mordi IR, Trucco E, Syed MG, MacGillivray T, Nar A, Huang Y, et al. Prediction of major adverse cardiovascular events from retinal, clinical, and genomic data in individuals with type 2 diabetes: A population cohort study. *Diabetes Care*. 2022;45(3):710–6. <https://doi.org/10.2337/dc21-1124>.
  99. Juster-Switlyk K, Smith AG. Updates in diabetic peripheral neuropathy. *F1000Res*. 2016;5. <https://doi.org/10.12688/f1000research.7898.1>.
  100. Gordois A, Scuffham P, Shearer A, Oglesby A, Tobian JA. The health care costs of diabetic peripheral neuropathy in the US. *Diabetes Care*. 2003;26(6):1790–5. <https://doi.org/10.2337/diacare.26.6.1790>.
  101. Quattrini C, Tesfaye S. Understanding the impact of painful diabetic neuropathy. *Diabetes Metab Res Rev*. 2003;19(Suppl 1):S2–8. <https://doi.org/10.1002/dmrr.360>.
  102. Pop-Busui R, Boulton AJ, Feldman EL, Bril V, Freeman R, Malik RA, et al. Diabetic neuropathy: a position statement by the American diabetes association. *Diabetes Care*. 2017;40(1):136.
  103. Petropoulos IN, Ponirakis G, Khan A, Almuhammad H, Gad H, Malik RA. Diagnosing diabetic neuropathy: something old, something new. *Diabetes Metab J*. 2018;42(4):255–69. <https://doi.org/10.4093/dmj.2018.0056>.
  104. Benson J, Estrada T, Burge M, Soliz P. Diabetic peripheral neuropathy risk assessment using digital fundus photographs and machine learning. *Annu Int Conf IEEE Eng Med Biol Soc*. 2020;2020:1988–91. <https://doi.org/10.1109/embc44109.2020.9175982>.
  105. Cervera DR, Smith L, Diaz-Santana L, Kumar M, Raman R, Sivaprasad S. Identifying peripheral neuropathy in colour fundus photographs based on deep learning. *Diagnostics*. 2021;11(11):1943. <https://doi.org/10.3390/diagnostics11111943>.
  106. Lee WJ, Sobrin L, Lee MJ, Kang MH, Seong M, Cho H. The relationship between diabetic retinopathy and diabetic nephropathy in a Population-Based study in Korea (KNHANES V-2, 3). *Investig Ophthalmol Vis Sci*. 2014;55(10):6547–53. <https://doi.org/10.1167/iovs.14-15001>.
  107. Ahmed MH, Elwali ES, Awadalla H, Almobarak AO. The relationship between diabetic retinopathy and nephropathy in Sudanese adult with diabetes: population based study. *Diabetes Metabolic Syndrome: Clin Res Reviews*. 2017;11:S333–6. <https://doi.org/10.1016/j.dsx.2017.03.011>.
  108. Pedro R-A, Ramon S-A, Marc B-B, Juan F-B, Isabel M-M. Prevalence and relationship between diabetic retinopathy and nephropathy, and its risk factors in the North-East of Spain, a

- Population-Based study. *Ophthalmic Epidemiol.* 2010;17(4):251–65. <https://doi.org/10.3109/09286586.2010.498661>.
109. Gupta M, Rao IR, Nagaraju SP, Bhandary SV, Gupta J, Babu GTC. Diabetic retinopathy is a predictor of progression of diabetic kidney disease: A systematic review and Meta-Analysis. *Int J Nephrol.* 2022;2022:3922398. <https://doi.org/10.1155/2022/3922398>.
  110. Ulasi II, Awobusuyi O, Nayak S, Ramachandran R, Musso CG, Depine SA, et al. Chronic kidney disease burden in low-resource settings: regional perspectives. *Seminars in Nephrology*: Elsevier; 2022. p. 151336.
  111. Bikbov B, Purcell CA, Levey AS, Smith M, Abdoli A, Abebe M, et al. Global, regional, and National burden of chronic kidney disease, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet.* 2020;395(10225):709–33.
  112. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. *Lancet.* 2018;392(10159):2052–90.
  113. Eknoyan G, Lameire N, Eckardt K, Kasiske B, Wheeler D, Levin A, et al. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int.* 2013;3(1):5–14.
  114. Caramori ML, Fioretto P, Mauer M. Low glomerular filtration rate in normoalbuminuric type 1 diabetic patients: an indicator of more advanced glomerular lesions. *Diabetes.* 2003;52(4):1036–40.
  115. Sabanayagam C, Xu D, Ting DSW, Nusinovi S, Banu R, Hamzah H, et al. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *Lancet Digit Health.* 2020;2(6):e295–302. [https://doi.org/10.1016/s2589-7500\(20\)30063-7](https://doi.org/10.1016/s2589-7500(20)30063-7).
  116. Zhang K, Liu X, Xu J, Yuan J, Cai W, Chen T, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomedical Eng.* 2021;5(6):533–45. <https://doi.org/10.1038/s41551-021-00745-6>.
  117. Zhao L, Ren H, Zhang J, Cao Y, Wang Y, Meng D, CLASSIFIED USING THE LESION-AWARE DEEP LEARNING SYSTEM, PREDICTS DIABETIC END-STAGE RENAL DISEASE IN CHINESE PATIENTS. *DIABETIC RETINOPATHY.* *Endocr Pract.* 2020;26(4):429–43. <https://doi.org/10.4158/ep-2019-0512>
  118. Diehm C, Lawall H, Diabetes. *Herzchirurgie und periphere gefäße.* *Clin Res Cardiol.* 2006;95(S1):i63–9. <https://doi.org/10.1007/s00392-006-1123-y>.
  119. Althouse AD, Abbott JD, Forker AD, Bertolet M, Barinas-Mitchell E, Thurston RC, et al. Risk factors for incident peripheral arterial disease in type 2 diabetes: results from the bypass angioplasty revascularization investigation in type 2 diabetes (BARI 2D) trial. *Diabetes Care.* 2014;37(5):1346–52. <https://doi.org/10.2337/dcl3-2303>.
  120. Yamagishi SI, Nakamura K, Matsui T, Ueda SI, Imaizumi T. Role of postprandial hyperglycaemia in cardiovascular disease in diabetes. *Int J Clin Pract.* 2007;61(1):83–7. <https://doi.org/10.1111/j.1742-1241.2006.01168.x>.
  121. Achim A, Stanek A, Homorodean C, Spinu M, Onea HL, Lazăr L, et al. Approaches to peripheral artery disease in diabetes: are there any differences?? *Int J Environ Res Public Health.* 2022;19(16):9801. <https://doi.org/10.3390/ijerph19169801>.
  122. Hirsch AT, Kalsi H, Rooke TW. *Peripheral arterial diseases.* *Cardiovascular Medicine.* Springer; 2007. pp. 1681–703.
  123. Mueller S, Wintergerst MWM, Falahat P, Holz FG, Schaefer C, Schahab N, et al. Multiple instance learning detects peripheral arterial disease from high-resolution color fundus photography. *Sci Rep.* 2022;12(1):1389. <https://doi.org/10.1038/s41598-022-05169-z>.
  124. Panda BB, Thakur S, Mohapatra S, Parida S. Artificial intelligence in ophthalmology: A new era is beginning. *Artif Intell Med Imaging.* 2021;2(1):5–12. <https://doi.org/10.35711/aimi.v2.i1.5>.
  125. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition2016.* pp. 770–8.
  126. Kaiming H, Xiangyu Z, Shaoqing R, Jian S. Identity mappings in deep residual networks. *European conference on computer vision*: Springer; 2016. pp. 630–45.
  127. Kurama V. A Review of Popular Deep Learning Architectures: ResNet, InceptionV3, and SqueezeNet. <https://blog.paperspace.com/popular-deep-learning-architectures-resnet-inceptionv3-squeeze-net/> (2020). Accessed.
  128. Anwar A. Difference between Alexnet, Vggnet, Resnet and inception. *Medium-Towards Data Sci.* 2019.
  129. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556.* 2014. <https://doi.org/10.48550/arXiv.1409.1556>
  130. Jafarizadeh A, Maleki SF, Pouya P, Sobhi N, Abdollahi M, Pedrammehr S, et al. Current and future roles of artificial intelligence in retinopathy of prematurity. *arXiv preprint arXiv:2402.09975.* 2024. <https://doi.org/10.48550/arXiv.2402.09975>
  131. Rosebrock A. *Imagenet: Vggnet, Resnet, inception, and Xception with Keras.* Mars. 2017.
  132. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: A review. *Neurocomputing.* 2016;187:27–48.
  133. Farabi Maleki S, Yousefi M, Afshar S, Pedrammehr S, Lim CP, Jafarizadeh A, Asadi H. Artificial intelligence for multiple sclerosis management using retinal images: Pearl, peaks, and pitfalls. *Semin Ophthalmol.* 2023;1–18. <https://doi.org/10.1080/08820538.2023.2293030>.
  134. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition2016.* pp. 2818–26.
  135. Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International conference on machine learning: pmlr;* 2015. pp. 448–56.
  136. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition2017.* pp. 4700–8.
  137. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst.* 2012;25.
  138. Zhang H, Wu C, Zhang Z, Zhu Y, Lin H, Zhang Z et al. Resnet: Split-attention networks. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition2022.* pp. 2736–46.
  139. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60–88. <https://doi.org/10.1016/j.media.2017.07.005>.
  140. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18:* Springer; 2015. pp. 234–41.
  141. Jeong Y, Hong YJ, Han JH. Review of machine learning applications using retinal fundus images. *Diagnostics (Basel).* 2022;12(1). <https://doi.org/10.3390/diagnostics12010134>.
  142. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans Knowl Data Eng.* 2010;22(10):1345–59. <https://doi.org/10.1109/TKDE.2009.191>.

143. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. *Artificial Neural Networks and Machine Learning–ICANN*. 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4–7, 2018, Proceedings, Part III 27: Springer; 2018. pp. 270–9.
144. Deng X-Y, Zhang Z-X, Li H-X, Wang J, Chen Y-Q, Mao J-B, et al. Retinal vascular morphological characteristics in diabetic retinopathy: an artificial intelligence study using a transfer learning system to analyze ultra-wide field images. *Int J Ophthalmol*. 2024;17(6):1001–6. <https://doi.org/10.18240/ijjo.2024.06.03>.
145. Kanagasingam Y, Xiao D, Vignarajan J, Preetham A, Tay-Kearney M-L, Mehrotra A. Evaluation of artificial Intelligence–Based grading of diabetic retinopathy in primary care. *JAMA Netw Open*. 2018;1(5):e182665–e. <https://doi.org/10.1001/jamanetworkopen.2018.2665>.
146. Das T, Takkar B, Sivaprasad S, Thanksphon T, Taylor H, Wiedemann P, et al. Recently updated global diabetic retinopathy screening guidelines: commonalities, differences, and future possibilities. *Eye*. 2021;35(10):2685–98. <https://doi.org/10.1038/s41433-021-01572-4>.
147. Wang LZ, Cheung CY, Tapp RJ, Hamzah H, Tan G, Ting D, et al. Availability and variability in guidelines on diabetic retinopathy screening in Asian countries. *Br J Ophthalmol*. 2017;101(10):1352–60. <https://doi.org/10.1136/bjophthalmol-2016-310002>.
148. Martins TGS. Comment on: how is the risk of being diagnosed with referable diabetic retinopathy affected by failure to attend diabetes eye screening appointments? *Eye*. 2021;35(7):2045. <https://doi.org/10.1038/s41433-020-1040-8>.
149. Rahim SS, Palade V, Holzinger A. Image Processing and Machine Learning Techniques for Diabetic Retinopathy Detection: A Review. 2020. pp. 136–54.
150. Shuai M. Research status and prospect of deep learning algorithm-based artificial intelligence in assisted diagnosis of diabetic retinopathy. *Chin J Experimental Ophthalmol*. 2019:684–8.
151. Chu A, Squirell D, Phillips AM, Vaghefi E. Essentials of a robust deep learning system for diabetic retinopathy screening: A systematic literature review. *J Ophthalmol*. 2020;2020:1–11. <https://doi.org/10.1155/2020/8841927>.
152. Morris A. AI can diagnose diabetic retinopathy. *Nat Reviews Endocrinol*. 2018;14(2):65. <https://doi.org/10.1038/nrendo.2017.177>.
153. Pieczynski J, Kuklo P, Grzybowski A. The role of telemedicine, In-Home testing and artificial intelligence to alleviate an increasingly burdened healthcare system: diabetic retinopathy. *Ophthalmol Therapy*. 2021;10(3):445–64. <https://doi.org/10.1007/s40123-021-00353-2>.
154. Wang L, Shang Y, Weng X, Yang X, Sang A, Dong J et al. The Development of Standardized and Sharable Computer Interpretable Guidelines for Diabetic Retinopathy. Proceedings of the 2019 International Conference on Intelligent Medicine and Health. New York, NY, USA: ACM; 2019. pp. 1–7.
155. Nsoesie EO. Evaluating artificial intelligence applications in clinical settings. *JAMA Netw Open*. 2018;1(5):e182658–e. <https://doi.org/10.1001/jamanetworkopen.2018.2658>.
156. Xie Y, Gunasekaran DV, Balaskas K, Keane PA, Sim DA, Bachmann LM, et al. Health economic and safety considerations for artificial intelligence applications in diabetic retinopathy screening. *Translational Vis Sci Technol*. 2020;9(2):22. <https://doi.org/10.1167/tvst.9.2.22>.
157. Grote T. Randomised controlled trials in medical AI: ethical considerations. *J Med Ethics*. 2022;48(11):899–906. <https://doi.org/10.1136/medethics-2020-107166>.
158. Chen J, Storch V, Kurshan E. Beyond Fairness Metrics: Roadblocks and Challenges for Ethical AI in Practice. 2021.
159. Amer S. Ethical Issues of the Use of AI in Healthcare. 2021. pp. 843–53.
160. Casas-Roma J, Conesa J, Caballé S, Education. Ethical Dilemmas and AI: From Ethical Design to Artificial Morality. 2021. pp. 167–82.
161. Morris MR. AI and Accessibility: A Discussion of Ethical Considerations. 2019. <https://doi.org/10.1145/3356727>
162. Ogunyemi OI, Gandhi M, Tayek C. Predictive Models for Diabetic Retinopathy from Non-Image Teleretinal Screening Data. AMIA Joint Summits on Translational Science proceedings AMIA Joint Summits on Translational Science. 2019;2019:472–7.
163. Verma H, Mlynar J, Schaer R, Reichenbach J, Jreige M, Prior J et al. Rethinking the Role of AI with Physicians in Oncology: Revealing Perspectives from Clinical and Research Workflows. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM; 2023. pp. 1–19.
164. Juluru K, Shih H-H, Keshava Murthy KN, Elnajjar P, El-Rowmeim A, Roth C, et al. Integrating ai algorithms into the clinical workflow. *Radiology: Artif Intell*. 2021;3(6). <https://doi.org/10.1148/ryai.2021210013>.
165. Hua D, Petrina N, Young N, Cho J-G, Sacks A, Poon S. Using AI-Driven triaging to optimise clinical workflows in Non-Emergency outpatient settings. 2023 Australasian computer science week. New York, NY, USA: ACM; 2023. pp. 240–3.
166. Smith M, Sattler A, Hong G, Lin S. From code to bedside: implementing artificial intelligence using quality improvement methods. *J Gen Intern Med*. 2021;36(4):1061–6. <https://doi.org/10.1007/s11606-020-06394-w>.
167. Heisler M, Karst S, Lo J, Mammo Z, Yu T, Warner S, et al. Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography. *Translational Vis Sci Technol*. 2020;9(2):20.
168. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vision*. 2020;128:336–59.
169. Garreau D, Luxburg U. Explaining the explainer: A first theoretical analysis of LIME. International conference on artificial intelligence and statistics: PMLR; 2020. pp. 1287–96.
170. Mercaldo F, Di Giammarco M, Apicella A, Di Iadarola G, Cesarelli M, Martinelli F, Santone A. Diabetic retinopathy detection and diagnosis by means of robust and explainable convolutional neural networks. *Neural Comput Appl*. 2023;35(23):17429–41.
171. Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S. Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol*. 2019;137(10):1182–8.
172. Ahmed Adel S. An overview of artificial intelligence use in diabetic retinopathy treatment: a narrative review. *Futurity Med*. 2023;4–13. <https://doi.org/10.57125/FEM.2023.03.30.01>.
173. Daich Varela M, Sen S, De Guimaraes TAC, Kabiri N, Pontikos N, Balaskas K, Michaelides M. Artificial intelligence in retinal disease: clinical application, challenges, and future directions. *Graefes Archive Clin Experimental Ophthalmol*. 2023. <https://doi.org/10.1007/s00417-023-06052-x>.
174. Happe M, Gill H, Salem DH, Janga SC, Hajrasouliha A. Clinical features distinguishing diabetic retinopathy severity using artificial intelligence. *Proc IMPRS*. 2023;5(1). <https://doi.org/10.1806/026676>.
175. Burlina P, Joshi N, Paul W, Pacheco KD, Bressler NM. Addressing artificial intelligence bias in retinal diagnostics. *Translational Vis Sci Technol*. 2021;10(2):13.
176. Rogers TW, Gonzalez-Bueno J, Garcia Franco R, Lopez Star E, Méndez Marín D, Vassallo J, et al. Evaluation of an AI system for the detection of diabetic retinopathy from images captured with

- a handheld portable fundus camera: the MAILOR AI study. *Eye*. 2021;35(2):632–8.
177. Abdollahi M, Jafarizadeh A, Ghafouri-Asbagh A, Sobhi N, Pourmoghtader K, Pedrammehr S, et al. Artificial intelligence in assessing cardiovascular diseases and risk factors via retinal fundus images: A review of the last decade. *WIREs Data Min Knowl Discov*. 2024;14(6):e1560. <https://doi.org/10.1002/widm.1560>.
178. Nolan BM, DayBranch ER, Barton K, Korsen N. Patient and provider experience with artificial intelligence screening technology for diabetic retinopathy in a rural primary care setting. *J Maine Med Cent*. 2023;5(2). <https://doi.org/10.46804/2641-2225.1144>.
179. Land MR, Patel PA, Bui T, Jiao C, Ali A, Ibnamasud S, et al. Examining the role of telemedicine in diabetic retinopathy. *J Clin Med*. 2023;12(10):3537. <https://doi.org/10.3390/jcm12103537>.
180. Bryan JM, Bryar PJ, Mirza RG. Convolutional neural networks accurately identify ungradable images in a diabetic retinopathy telemedicine screening program. *Telemedicine e-Health*. 2023;29(9):1349–55. <https://doi.org/10.1089/tmj.2022.0357>.
181. Lin S, Ma Y, Xu Y, Lu L, He J, Zhu J, et al. Artificial intelligence in Community-Based diabetic retinopathy telemedicine screening in urban China: Cost-effectiveness and Cost-Utility analyses with Real-world data. *JMIR Public Health Surveillanc*. 2023;9:e41624–e. <https://doi.org/10.2196/41624>.
182. Mahajan P, Uddin S, Hajati F, Moni MA. Ensemble learning for disease prediction: A review. *Healthcare: MDPI*; 2023. p. 1808.
183. Shree R, Madagaonkar S, Prateek LA, Tony A, Rathnamma M, Venkata Ramana V, Chandrasekaran K. Application of Ensemble Methods in Medical Diagnosis. *International Conference on Innovations in Data Analytics*: Springer; 2022. pp. 355–67.
184. Mondal SS, Mandal N, Singh KK, Singh A, Izonin I. Eddr: an ensemble deep learning technique for detection and classification of diabetic retinopathy. *Diagnostics*. 2022;13(1):124.
185. Lukashevich MM. Diabetic retinopathy fundus image classification using ensemble methods. *Pattern Recognit Image Anal*. 2024;34(2):331–9.
186. Loh HW, Ooi CP, Seoni S, Barua PD, Molinari F, Acharya UR. Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022). *Computer Methods and Programs in Biomedicine*. 2022:107161.
187. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. 2017. <https://doi.org/10.48550/arXiv.1702.08608>
188. Ribeiro MT, Singh S, Guestrin C. Why should i trust you? Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*2016. pp. 1135–44.
189. Chen JH, Asch SM. Machine learning and prediction in medicine—beyond the peak of inflated expectations. *N Engl J Med*. 2017;376(26):2507.
190. Shakeel PM, Baskar S, Dhulipala VRS, Jaber MM. Cloud based framework for diagnosis of diabetes mellitus using K-means clustering. *Health Inform Sci Syst*. 2018;6(1):16. <https://doi.org/10.1007/s13755-018-0054-0>.
191. Sharma SK, Zamani AT, Abdelsalam A, Muduli D, Alabrah AA, Parveen N, Alanazi SM. A diabetes monitoring system and Health-Medical service composition model in cloud environment. *IEEE Access*. 2023;11:32804–19. <https://doi.org/10.1109/ACCESS.2023.3258549>.
192. Salari R, Niakan Kalhori R, GhaziSaeedi S, Jeddi M, Nazari M, Fatehi M. Mobile-Based and Cloud-Based system for Self-management of people with type 2 diabetes: development and usability evaluation. *J Med Internet Res*. 2021;23(6):e18167–e. <https://doi.org/10.2196/18167>.
193. Nasser AR, Hasan AM, Humaidi AJ, Alkhayyat A, Alzubaidi L, Fadhel MA, et al. IoT and cloud computing in Health-Care: A new wearable device and cloud-Based deep learning algorithm for monitoring of diabetes. *Electronics*. 2021;10(21):2719. <https://doi.org/10.3390/electronics10212719>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.