SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

An Investigation on Diabetes-Based Eye Disease Detection Methods using Deep Learning and GAN Techniques

Arwa Albelaihi[1], Dina M. Ibrahim[1,2,*]

[1] Department of Information Technology, College of Computer, Qassim University, Buraydah 51452, Saudi Arabia
[2] Department of Computers and Control Engineering, Faculty of Engineering, Tanta University, Tanta 31733, Egypt

Keywords:

Deep learning; Diseases; Diabetic eye disease; Retinal fundus images; Transfer learning, Generative adversarial networks; GAN techniques

ABSTRACT

Machine learning and deep learning play successful and influential roles in image detection and classification in many medical imaging diagnosis areas. Diabetes is becoming a significant health concern, and diabetic eye diseases (DEDs) will be the leading cause of vision loss worldwide. This paper presents a review on the state-of-the-art studies concerned with the detection, classification, segmentation, and grading of diabetic eye diseases, including the common four eye diseases: diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma, and cataracts. We classify the model techniques into three main categories: classification-based, localization-based, and generative adversarial network (GAN) models. We investigate the current and commonly used datasets available for fundus images of diabetic eye diseases. We also review research that employed different GAN techniques to improve the fundus image dataset or increase the size of the images in the datasets. Finally, we illustrate the performance measures used in the previous studies for evaluating various models of diabetic eye diseases.

1. Introduction

Deep learning (DL) is applied in various fields to find novel solutions to pressing problems, and is showing tremendously successful results in classification tasks. The medical sector is suitable for the application of artificial intelligence (AI) tools and techniques. AI is one of the most powerfully transformative technologies in the 21th century. This transformation occurred through the use of powerful machine learning (ML) tools and techniques such as deep convolutional networks, generative adversarial networks (GANs), deep reinforcement learning (DRL), convolutional neural networks (CNNs), and artificial neural networks (ANNs). Recently, DL has been succeeding in traditional AI in critical tasks such as recognizing speech, characterizing images, and generating natural and readable sentences. DL has been shown to be a successful and influential method in many medical imaging diagnosis areas in image detection and classification. DL can be used for detecting and classifying eye diseases, including diabetic eye disease by employing fundus images to

* *Corresponding author.*
E-mail address: d.hussein@qu.edu.sa

https://doi.org/10.37934/araset.56.2.246274

analyse and diagnose eye diseases. Diabetic eye disease contains a group of eye diseases, including diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma, and cataract. Diabetic eye disease can cause severe vision decline and blindness in patients aged 20 to 74 years. The advanced detection of diabetic eye disease is para-mount in preventing vision loss. Researches confirmed that 90% of patients with diabetes can avoid diabetic eye disease through early detection [1].

This paper reviews the most recent research published to enhance the DL detection models for diabetic eye diseases. In DL models, researchers have to collect both healthy and diabetic eye disease fundus images. Then, some image preprocessing techniques are applied to reduce noise in the images. The pre-processed images are input to the DL architecture to extract features and automatically learn the analysis rules. The feature weights are optimized recursively to guarantee the most reliable detection results. The unseen set of images are tested on optimized weights. The DL model requires a large number of datasets, and an insufficient dataset restricts the performance of the model. Multiple review studies have used artificial intelligence and machine learning techniques to detect and diagnose diabetic retinopathy and ophthalmology. Authors in previous work [2] published a review focusing on eye conditions such as diabetic retinopathy, glaucoma, and age-related macular diseases. They selected papers published between 2016 and 2018 and reviewed them in their study, summarizing those that used fundus and optical coherence tomography images, and TL methods. Their research did not include current (2019--2021) publications that incorporated TL methods into their approach, and they omitted the identification of eye cataract disease from their study scope.

Similarly, authors in [3] provided a review of current articles using AI in ophthalmology, but their focus did not cover comprehensive AI methodologies. The authors in [4] reviewed the computer-aided diagnosis of glaucoma (Gl) using fundus images. They addressed computer-aided methods focused on optical disc segmentation. A variety of studies that used deep learning and machine learning methods for Gl detection were not discussed in their review. It is therefore important to review studies that considered existing approaches to DED diagnostics. The work in [5] revised computer-aided DR detection studies, which are largely DR-lesion-based.

Authors in [6] published a review on the automatic detection of diabetic eye disease through deep learning using fundus images. They provided a comprehensive overview of the state of the art on diabetic eye disease (DED) detection methods using papers published between 2014 and 2020. They did not include the recent machine learning DED detection methods that apply the GAN technique and its subtypes like CycleGAN, CGAN, and DCGAN. In recent studies, a combination of neural networks and the adversarial idea has been used in medical image processing, obtaining good results with medical images. Some review studies focused on diabetic retinopathy (DR) detection in retinal fundus photographs through deep learning techniques [7,8]. The authors in [9,10] presented surveys on developments in the automatic detection of diabetic retinopathy using deep learning methods. However, others focused on glaucoma disease [11], in which the detection of glaucoma using retinal fundus images was comprehensive reviewed. Other authors conducted a systematic review of the applications of deep learning in the detection of glaucoma. Recently, [12] reviewed the applications of deep learning in fundus images and in glaucoma with optical coherence tomography.

Therefore, to address the limitations of the above-mentioned studies, this article presents a review of the state-of-the-art advances in the detection, classification, segmentation, and grading of diabetic eye diseases, including the common four eye diseases: diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma, and cataract. We classify the model techniques into three main categories: classification-based, localization-based, and generative adversarial network (GAN) models. The three first categories, including transfer learning, deep learning, machine learning, and GAN techniques, are reviewed using papers published until the beginning of 2021. We review the

state of the art from those perspectives: DL models for diabetic eye disease detection, datasets available for diabetic eye disease, and performance measures for diabetic eye disease detection evaluation. Although surface disorders, vascular occlusions, and neuropathy are among diabetic eye diseases (DED), in this study we focus on diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma, and cataracts diabetic eye diseases.

2. Background

Diabetes is a major public health concern around the globe. Approximately 422 million adults have diabetes worldwide, and the number is projected to reach 629 million by 2045 [13]. As these trends continue, diabetes-related complications, such as diabetic retinopathy, glaucoma, diabetic macular edema, and cataract, will be the leading cause of vision loss around the globe. These diabetic eye diseases result from chronically high blood glucose levels that damage retinal capillaries. In the United States, diabetic retinopathy is a leading cause of blindness among patients aged between 20 and 64 years. Although uncontrolled type 1 and 2 diabetes and the associated microvascular complications may significantly affect one's vision, recent technologies have facilitated routine eye examinations and the management of vision-threatening diabetic retinal diseases. Digital photography, optical coherence tomography examination, and widefield imaging are promising for screening and detecting diabetic eye diseases [14].

The use of artificial intelligence has also proved useful in detecting serious eye diseases among people with diabetes [15]. Thus, because of the debilitating impact of DEDs, the gold standard in ophthalmology for diabetic patients should be to leverage modern technology to prevent vision loss. DR is the leading cause of vision loss among diabetic patients. This condition is marked by progressive damage to the retinal microvasculature. In 2015, about 2.6 million people across the globe were visually impaired or blind because of diabetic retinopathy [13]. In 2016, the World Health Organization (WHO) reported that about 146 million adults had DR [16]. This condition is mainly characterized by retinal ischemia, intraregional haemorrhages, and microvascular abnormalities [17]. Other symptoms include yellow fluid accumulation, which can cause retinal tear if left untreated.

Glaucoma is the progressive damage to the optic nerve, which causes initial loss of vision in the periphery. Glaucoma can develop into severe vision impairment if it progresses [16]. This eye condition affects about 80 million people across the globe [18]. The WHO estimated that about 76 million people across the globe would have glaucoma by the end of 2020. Because glaucoma damages the optic nerve, common symptoms include a hollowed-out optic nerve and retinal nerve fibre ganglion cell loss [17]. A lack of timely treatment can lead to permanent vision loss. DME is the thickening of the retina due to accumulated intraretinal fluid. The condition affects the central vision and causes a decline in vision, characterized by blurring and blindness. It is mainly triggered by the inability of the fluid to permeate the retinal vasculature [19]. The WHO estimated that half of the 15 million people with DME are undiagnosed and 50\% of the 8 million people affected do not have eye care. The WHO reported that the prevalence of DME will rise as the number of diabetic patients continue to increase [19]. These trends mean that only excellent metabolic control can reduce the incidence of DME.

A cataract occurs as a result of the opacification in the lens. The WHO defines this eye condition as a visual acuity of less than 3/60 in the better eye [20]. Cataracts affect approximately 18 million people around the globe every year, and about 90\% of the reported cases occur in developing nations. This disorder damages the optic refractive power, thereby reducing the amount of light that penetrates the retina. The main symptoms associated with this condition include blurring, glaring, and change in the colour spectrum [17]. It is one of the common causes of visual impairment in

diabetic patients. The next section describes the recent and various DL model techniques used to detect diabetic eye diseases through fundus images, which we divide into three main categories: classification-based, localization-based, and GAN models.

The DL model concerns training minimal data and datasets with a class imbalance between different diseases. If the training set is small, it may not provide satisfying results in terms of accuracy. There are two possible solutions:

i. using a range of enhancement methods, including rotating, shifting, cropping, and color setting (classical data augmentation)
ii. applying GAN to increase the dataset's size, increase the class balance between classes, and enhance the fundus image quality.

GAN has demonstrated robust generation abilities in the image generation of different types of medical images so that the distribution of the samples is smoother while the fundus data are expanding.

3. Methods for Detecting Diabetic Eye Diseases Through Deep Learning and GAN Techniques

3.1 Diabetic Eye Diseases Classification Methods using Deep Learning

This section reviews the state-of-the-art models by categorizing them into three subsections depending on their techniques, starting with a section on modelling based on transfer learning techniques, followed by a section on DL and ML hybrid models, and ending with a tailored DL classifier model. The TL principle focuses on reusing the features learned from DL models' primary task and adapting to the secondary task. The idea is to minimize the complexity of computing when training a (resource intensive) neural network architecture. With TL, instead of random generation, the parameters are initialized from previous learning. The first layers intuitively learn to extract basic features such as edges, textures, etc. The top layers, such as blood vessels and exudates, are more specific to the task. In cases where the data are insufficient to train a neural network from scratch (high volume of data required), TL can be successfully employed. The TL approach has been commonly adopted in diabetic eye disease detection [21-25]

Pan *et al.,* compared three CNNs models, DenseNet, ResNet50, and VGG16 [21], on four types of lesions of DR, including non-perfusion regions (NP), leakages, microaneurysms, and laser scars. They created datasets of 4067 fundus fluorescein angiography images from 435 eyes (218 left eyes and 217 right eyes). DR lesion detection and the process were found to be efficient in terms of computing. Experimental findings showed that DenseNet is an effective model to identify and differentiate retinal lesions automatically in multi-label classified FFA images. However, the process does not accurately identify microaneurysms because they are easily misclassified in the pervasive presence of fluorescein. Samanta *et al.,* introduced transfer-learning-based CNN architecture for colour fundus photography, which performs relatively well in identifying DR from hard exudates, blood vessels, and texture on a much smaller dataset [22]. The dataset was trained using their model on four classes (No DR, Mild DR, Moderate DR, and Proliferative DR) and reached a Cohen's kappa score on the validation set of 0.8836, with 0.9809 on the training set. Their model uses several architectures such as Inceptionv1, Inceptionv2, Inceptionv3, Xception, VGG16, ResNet-50, DenseNet, and AlexNet. Zhang *et al.,* presented a DeepDR framework for DR detection [23]. DeepDR actively detects the existence and severity of DR from fundus images by ensemble learning and transfer learning, using several transfer and ensemble learning methods including ResNet50, InceptionV3, DenseNets, Xception, and InceptionResNetV2. Additionally, [23] introduced a new dataset for DR images called

macula-centred retinal fundus images (13,767 images of 1872 patients). The proposed network achieved 97.5% sensitivity, with a specificity of 97.7%. However, their model needs to be assessed with a more complex and larger dataset.

The CNN model and the Lookahead optimizer were used for image classification of cataract disease in [24] to enhance accuracy and shorten the processing time. The fundus image datasets were derived from the Kaggle dataset, consisting of normal images of the fundus and images of the cataract fundus. The model effectively identified the label of images through CNN AlexNet architecture with the Lookahead optimizer on a stochastic gradient descent with Adam. As a result, the accuracy of Adam was increased by about 20\% and the optimizer Stochastic gradient descent was improved by about 2.5%. Sarki *et al.,* introduced a DL architecture with a pre-trained CNN merged with image processing techniques for the early detection of diabetic eye diseases (DR, DME, and glaucoma) [26]. They recognized specific work limitations in the early classification of diabetic eye disease. Later, the same authors developed an automated classification system examining mild multi-class diabetic eye disease [25]. Their model was applied to datasets from various sources: Messidor, Messidor-2, DRISHTI-GS, and Retina. The research was conducted applying the VGG16 and InceptionV3, and different performance enhancement methods were used, i.e., fine tuning, optimization, and contrast enhancement. The highest accuracy of 88.3% was achieved on the VGG16 model for multi-class classification, and 85.95% for mild multi-class classification.

Many transfers learning implementations, including GoogLeNet, AlexNet, and VGG16, for object recognition are mostly available in DL to retrain a new image, such as a medical image set. However, these architectures are less suitable for medical images in terms of classification efficacy. For example, Pan *et al.,* [21] used VGG16 for DR diagnosis employing eye fundus images and attained nearly 79.6% specificity because these TL frameworks were created for objectives such as animals, flowers, etc. So, TL techniques may be inappropriate for real-time medical images. Some authors combined DL with ML classifiers for diabetic eye diseases detection [27-29], such as random forest (RF), support vector machine (SVM), naïve Bayes, and decision tree. Grassmann *et al.,* [27] designed a DL algorithm prediction method for the severity of age-related macular degeneration from colour fundus photography based on a huge dataset (120,656 manually graded colour fundus images). In their study, multiple CNNs were trained independently, and based on the results of the single CNNs, the RF algorithm was trained to build a model ensemble. The model correctly classified 94.3% of healthy fundus photographs.

Malik *et al.,* developed a framework with a multiple ML algorithm including decision tree, RF, naïve Bayes, and artificial neural network algorithms [28]. The authors found that tree-based methods performed better than the artificial neural network. The dataset used in their research analysis consisted of real-time data to which data mining techniques and classification algorithms were applied. Although the RF algorithm seemed to perform better than the artificial neural network, it had a marginally longer execution time than the decision tree algorithm. Several studies have combined DL with ML via SVM, as SVM is a fast and dependable algorithm for classification and performs well with a limited amount of data. Theera-Umpon *et al.,* [29] developed a model to detect hard exudates with a limited dataset named DIARETDB1 containing 89 fundus images. They used supervised learning techniques including SVM, hierarchical adaptive neurofuzzy inference system (hierarchical ANFIS), multilayer perceptron (MLP) network, and CNNs. When the MLP network was extended, the proposed approach attained an AUC of 0.998. The AUCs were above 0.95 for all four classifiers. They found that CNN performed well but is not the most suitable classifier for hard exudate. The mixture of image processing methods and suitable classifiers in hard exudate can perform accurately.

Researchers have been developing new network architectures as an alternative to TL [30-33], creating their DL frameworks to automatically detect diabetic eye diseases. Researchers used their own built DL models with the indicated classifier, the number of layers, a model used, and findings obtained. Two models were proposed in [30], the first of which predicts the DR, whereas the second classifies the five DR phases. Zeng *et al.,* framework's uses a Siamese-like CNN structure trained with TL using the concept of a weight-sharing layer based on Inception V3. The introduced model uses binocular fundus images as inputs and learns their correlation to support prediction and diagnosis. They conducted their experiment on the Kaggle DR competition provided by Eye PACS, which contains 35,126 images (28,104 for the training set and 7024 for the testing set). The framework achieved a kappa score of 0.829. Unfortunately, the model requires paired fundus images, so it may not perform well for those datasets where paired fundus images are not available.

In [31], the authors used a DL approach to forecast the expected DR class and assign scores to distinct pixels to exhibit their relevance in each input sample. They then employed the assigned score to make the final classification decision. They used the Eye PACS dataset hosted on Kaggle, which they split into 75,650 images for training their model and 3000 images for validation. The introduced DL system attained a sensitivity and specificity of more than 90\%. However, the evaluation performance of the learning procedure can be enhanced using suitable measures. According to [32], they developed a new hypertensive retinopathy (DenseHyper) system comprising different multilayer architectures with various trained and dense feature layers integrated within a CNN algorithm to detect hypertensive retinography (HR). To increase the classification accuracy of DenseHyper, they combined it with a learning-based dense feature transform approach. Four datasets with an aggregate of 4270 images and varying techniques were used to achieve the same goal. They achieved significant results: an accuracy of 95\% and an AUC of 0.96. However, the dataset used was too small to generate a highly accurate DL algorithm.

Teresa *et al.,* produced a DL model for a DR-grading computer-aided diagnosis system called DR-GRADUATE [33]. Their model was designed to help decision-making using medically interpretable attention maps and an approximation of the uncertainty of the prediction, leaving the ophthalmologist to estimate how much that result should be trusted. The model achieved a quadratic-weighted Cohen's kappa between 0.71 and 0.84 in five different datasets. Furthermore, low-quality images are commonly associated with higher uncertainties, indicating that images not suitable for analysis lead to less accurate predictions.

Recently, a deep fusion-based model has been effectively discovered in detecting many diseases as in [34]. The model is effective in infectious lung diseases, including the detection of COVID-19. The procedure based on image fusion has provided significant benefits for clinical diagnosis. With respect to [35], the authors clarify that the fusion of optical imaging, radionuclide imaging, computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) has guided comprehensive learning in pathology studies. [36] proposed a haemorrhage detection system based on a 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in DR patients, which achieves better performance in quantitative analysis with great accuracy (0.9771 for the average accuracy). To extract features to form a feature vector, they apply a pre-trained modified CNN model. Convolutional sparse image decomposition fused the feature vector and best features selected by multi-logic regression-controlled entropy variance methods. They're proposed model was assessed on 1509 images from 7 different available datasets (HRF, DRIVE, STARE, MESSIDOR, DIARETDB0, and DIARETDB1). similarly, another model was designed to classify DR in [37] multi-scale feature fusion extraction with adaptive weighting under MobileNetV3 architecture. The proposed model is efficient for fusing different convolutional layers and can significantly fuse the

feature maps of various depths. They achieved results of 77.26% kappa statistic and 97% AUC on the Kaggle dataset, containing 3662 images.

Tables 1, 2 and 3 list each primary study of classification-based models: transfer learning, deep learning with machine learning, and tailored models' information regarding the datasets; the DL or ML methods used; and the performance measure used with their values. Notably, most of the methods used the area under the curve (AUC), accuracy (ACC), sensitivity (Sen), and specificity (Spc) as performance measures to evaluate their models with the highest values of 99.8%, 97.5%, 99.8%, and 99.5% being attained, respectively. A few of the studies used the F1-score, Cohen's kappa (K), and Matthew's correlation coefficient (MCC), with maximum values of 74%, 0.9809, and 0.583 being reported, respectively.

The authors of [21] compared three CNNs models: DenseNet, ResNet50, and VGG16, and applied them to four types of DR lesions, including non-perfusion regions (NP), leakages, microaneurysms, and laser scars. A study [29] reported that the multi-layer perceptron (MLP) and SVM methods achieved classification accuracies of 99.8% and 99.7%, respectively; these performance values are better than the hierarchical ANFIS and the CNNs, which achieved values of 98.8% and 95.1%, respectively. The results suggested that CNN may not be the best choice when the objects of interest do not have well-defined edges. Similarly, in [33], DR-GRADUATE was trained on the Kaggle DR detection training set and evaluated on multiple datasets. In DR grading, a quadratic-weighted Cohen's kappa (K) between 0.71 and 0.84 was achieved for the five different datasets.

Table 1
Comparison of DL diabetes-based eye diseases classification methods for TL models

Study	Dataset	Method	Class	Performance
[21]	Kaggle	DenseNet	Non-Perfusion region:	AUC=87.03% Sen=79.7% Spc=82.7% AUC=94.35%
			Microaneurysms:	Sen=98.0% Spc=77.3% AUC=96.47%
			Leakages:	Sen=84.0% Spc=96.5% AUC=96.53%
			Laser scars:	Sen=80.2% Spc=99.5% AUC=81.40%
		ResNet50	Non-Perfusion region:	Sen=59.0% Spc=87.9% AUC=90.97%
			Microaneurysms:	Sen=97.6% Spc=22.7% AUC=95.85%
			Leakages:	Sen=70.4% Spc=98.8% AUC=91.15%
			Laser scars:	Sen=69.7% Spc=95.4% AUC=71.25%
		VGG16	Non-Perfusion region:	Sen=61.3% Spc=79.6% AUC=55.69%

			Microaneurysms:	Sen=99.8% Spc=14.6% AUC=91.77%
			Leakages:	Sen=69.5% Spc=99.2% AUC=91.15%
			Laser scars:	Sen=57.4% Spc=98.8%
[22]	Kaggle	Inceptionv1, Inceptionv2, Inceptionv3, Xception, VGG16, ResNet-50, DenseNet, AlexNet	Max. on data validation: Max. on data training:	ACC=84.1% F1-score=64% K=0.8836 ACC=84.5% F1-score=74% K=0.9809
[23]	Macula-centred	Transfer learning, ensemble learning	Identification model: Grading model:	AUC=97.7% Sen=97.5% Spc=97.7% AUC=97.8% Sen=98.1% Spc=98.9%
[24]	Retina	AlexNet with SGD and lookahead optimizer AlexNet with Adam and lookahead optimizer	Classification: Validation: Classification: Validation:	ACC=86.88% ACC=97.5% ACC=97.5% ACC=97.5%
[25]	Messidor, Meddidor-2, DRISHTI-GS, Retina	VGG16 with Adam optimizer	Mild multi-classes: Multi-classes:	ACC=85.94% ACC=88.3%

Table 2
Comparison of DL diabetes-based eye diseases classification methods for DL combined with ML

Study	Dataset	Method	Class	Performance
[29]	DIARETDB1	MLP		AUC=99.8%
		SVM		AUC=99.7%
		Hierarchical ANFIS	None	AUC=98.8%
		CNN		AUC=95.1%
[27]	AREDS	CNN	Early/late signs:	ACC=84.2%
			Healthy fundus:	ACC=94.3%
[28]	Kaggle	Random Forest and Decision Tree	None	ACC=93.5%

Table 3
Comparison of DL diabetes-based eye diseases classification methods for tailored DL models

Study	Dataset	Method	Performance
[30]	Kaggle	CNN	AUC=95.1% Sen=82.2% Spc=70.7% K=0.829
[31]	EyePACS	Pixel-wise score	ACC=85.7% Sen=90.6% Spc=85.7% F1-score=71.0% Mcc=0.583
[32]	DR-HAGIS, DRIVE, DiaRetDB0, DiaRetDB1, DR1&DR2, Imam-HR	Dense features transform, deep residual learning	ACC=95% AUC=96% Sen=93% Spc=95%
[33]	Kaggle, Messidor-2,IDRID, DMR, SCREEN-DR	Gaussian-sampling	K=0.84
[36]	HRF, DRIVE, STARE, MESSIDOR, DIARETDB0, DIARETDB1	3D CNN and feature fusion	Sen=97.54% Spc=97.89% AUC=98.22%
[37]	Kaggle APTOS	MobileNetV3 backbone network based on a multi-scale feature fusion	K=77.26% AUC=97%
[38]	Kaggle	Fused using feature fusion VGG-VD-4096 + Inception-V3	AUC= 96.4%
[5]	Messidor, Messidor-2, DRISHTI-GS, Retina datasets	CNN, RMSprop optimizer	Sen=100% Spc=100% ACC=81.33%

3.2 Diabetic Eye Diseases Localization Methods using Deep Learning

Recently, many of the DL models have been based on localization for diabetic eye diseases. This section reviews the state-of-the-art models by categorizing them into two groups based on their techniques. The first subsection describes the DL segmentation-based models, and the second subsection focuses on research employing RCNN for localization. Deep learning algorithms have quickly developed in recent years, and the performance of the DL-based segmentation approach has exceeded that of the conventional segmentation method. CNN can learn without prior awareness and additional pre-processing, which is commonly used in image classification and image recognition. Sreng *et al.,* conducted a comparative study of optic disc segmentation, employing five deep CNNs as the encoder in the DeepLabv3+ architecture [38]. They compared eleven pre-trained CNNs as the glaucoma classifier using transfer learning techniques and compared the eleven pre-trained CNNs as the feature extractors using an SVM classifier. DeepLabv3+ and MobileNet were found to be best for segmentation, whereas the ensemble of approaches performed better than conventional approaches for RIM-ONE, ORIGA, DRISHTI-GS1, and ACRIMA datasets with accuracies of 97.37%, 90.00%, 86.84%, and 99.53%, and AUCs of 100%, 92.06%, 91.67%, and 99.98%, respectively. They attempted to achieve the automatic initial screening of glaucoma based on the quantitative analysis of fundus images to support ophthalmologists. However, the datasets used here included only high-resolution images, but low-quality images and images with effects and noises occur in real-world screening environments.

Lu and Chen used the GrabCut method to generate the pseudo ground truths [39], and they trained the network based on a modified U-net model with the generated pseudo ground truth. The

modified U-Net model provided improvements by minimizing the original U-shape structure by adding a two-dimensional convolutional layer. By conducting their experiments on RIM-ONE and DRISHTI-GS, the model segmented the optic disc without the training ground truth marked by medical experts. However, although the technique requires less training, it demonstrates weaker segmentation accuracy than recent approaches because of missed ground truths.

Guo *et al.,* presented a method for automated glaucoma screening that merges clinically measured features and image-based features [40], which improved the UNet neural network and achieved segmentation of the OD and OC based on the region of interest, increasing the field of view feature model and selecting the best feature combination. They trained a gradient boosting decision tree classifier for glaucoma screening. The algorithm achieved excellent glaucoma screening performance with a sensitivity of 0.894, an accuracy of 0.843, and an AUC of 0.901 on the ORIGA dataset. The feature extraction process of the algorithm is slow because it extracts multiple receptive fields' features.

Different DL diabetes-based eye disease methods have been constructed in various studies, which employed RCNN for localization and segmentation [18,41,42]. Glaucoma signs were identified from retinal images [41] using a two-stage structure, the first of which detects and locates the optic disc produced through the RCNN, and the second employs deep CNN to categorize the disc into normal or glaucomatous. The authors formed a rule-based semi-automatic ground truth generation method that provides necessary annotations for training RCNN-based models for automated disc localization. The localization and classification of AUC glaucoma of 0.874 were obtained for ORIGA dataset classification. This classification of glaucoma on the ORIGA dataset showed that reporting only the AUC for class-imbalanced datasets without pre-defined training and testing divides does not accurately reflect the classifier's performance, demonstrating the need for additional performance measures to substantiate the findings. As their model applies a two-stage framework to find and classify glaucoma, the method is computationally complex. The efficiency is influenced by increasing the network's hierarchy, which results in the loss of the discriminative collection of features.

Yuming *et al.,* [42] introduced the object detection-based algorithm to detect glaucoma. End-to-end RCNN was proposed for optic disc and optic cup segmentation, named JointRCNN. They conducted experiments on the ORIGA and the Singapore Chinese Eye Study (SCES) datasets. ORIGA contains 650 retinal fundus images from 482 healthy eyes and 168 glaucoma patients. SCES consists of 1676 images of which only 46 images are from glaucoma patients. Both datasets employed in the experiment had an imbalance between classes. The method achieved excellent optic disc and optic cup segmentation with an AUC of 0.901. Even though the method is robust to glaucoma detection, it is computationally complex, as it manipulates two RCNNs to compute the bounding boxes of the optic cup and optic disc. In [41,42], outstanding performance, as measured by the accuracy, was achieved; however, both of these methods are computationally complex.

Table 4 describes the following for each primary study concerning localization-based models: segmentation and RCNN model information regarding the datasets, the DL or ML methods used, and the performance measure used with their values. Note that most of the studies used the AUC, ACC, Sen, and Spc as performance measures to evaluate their models; few studies used the dice Coefficient (DC) and intersection over union (IoU). The authors of [38] used an ensemble of methods that performed better than the conventional methods for glaucoma classification.

Table 4
Summary of localization-based methods and their corresponding measures in the studies included in this review

Model	Study	Dataset	Method	Class	Performance
Segmentation	[38]	RIM-ONE, ORIGA, DRISHTI-GS1, ACRIMA, REFUGE	DeepLabv3, ensemble method	RIM-ONE:	AUC=100% ACC=97.37%
		ACRIMA, REFUGE		ORIGA:	AUC=92.06% ACC=90.0%
				DRISHTI-GS1:	AUC=91.67% ACC=86.84%
				ACRIMA:	AUC=99.98% ACC=99.53%
				REFUGE:	AUC=95.10% ACC=95.59%
	[39]	RIM-ONE, DRISHTI-GS	U-Net	Fully-supervised1:	AUC=99.56% Sen=89.17% Spc=99.89% IoU=86.37%
				Semi-supervised:	AUC=99.62% Sen=91.49% Spc=99.88% IoU=88.25%
				Fully-supervised2:	AUC=99.72% Sen=92.87% Spc=91.87% IoU=91.87%
	[40]	ORIGA	UNet++		AUC=91.2% ACC=83.7% Sen=90.4% Spc=77.2%
	[43]	DRISHTI-GS	UNet CNN		DC=94.15%
RCNN	[41]	ORIGA, DIARETDB1 OCT&CF1, HRF, Messidor, DRIONS-DB, DRIVE	RCNN + Deep CNN	Random Training: Cross validation:	AUC=86.8% AUC=87.4%
	[42]	ORIGA, SCES	RCNN with attention	ORIGA: SCES:	AUC=85.4% AUC=90.1%
	[50]	ORIGA, DIARETDB, DR-HAGIS, HRF, Messidor	Fast RCNN with fuzzy k-means	MESSIDOR:	AUC=95.8% Sen=96% Spc=95.8%
				ORIGA:	AUC=94.3% Sen=94.1% Spc=94.5%
				HRF	AUC=95.2% Sen=95% Spc=96%
				DR-HAGIS	AUC=89% Sen=94.5% Spc=94.1%

3.3 Diabetic Eye Disease Classification and Localization Methods using the GAN Technique

The generative adversarial network (GAN), proposed by Goodfellow *et al.,* [44], forms an adversarial network of a generation model and a discrimination model. The two networks are optimized until both sides reach a dynamic balance in the training process. By counter-training, the algorithm can quickly learn the distribution of data to create a new sample that conforms to the characteristics of the true sample probability. GAN is generative modelling---a data augmentation technique that creates artificial instances from a dataset to retain similar features to the original set.

Neural networks have been combined with the adversarial idea in medical image processing, and accurate results have been obtained. GANs have proved to be a useful framework for generating anatomically consistent retinal fundus images in synthetic databases. Progress has been rapidly achieved in synthesizing medical images based on the basic concept of adversarial learning. Bellemo *et al.,* [45] explored the possible benefits and drawbacks that need to be resolved before GANs can broadly be applied with retinal imaging. Several GAN architecture extensions have been applied to different neural networks for both the generator and discriminator models. Researchers have employed different GAN techniques to improve the fundus images datasets [46-55].

The researchers in [46-48] applied cycle-consistent GAN (CycleGAN), which is an extension of the GAN for image-to-image translation without paired image data. Yoo *et al.,* found that CycleGAN can improve the retinal image quality and improve the deep learning classifier for disease screening [46]. You *et al.,* presented a retinal image enhancement method, a convolutional block attention module called Cycle-CBAM, which is a simple and efficient attention module to re-weight the high-level features extracted by the CNN [47]. Their method generates visually better images with less noise and particular textural details. It provides superior performance to CycleGAN both in the peak signal-to-noise ratio (PSNR) and structural index similarity (SSIM), which are two tools used for image quality assessment. Cycle-CBAM achieves the migration from low-quality to high-quality fundus pictures. However, CycleGAN's computational cost for training a deep learning network for high resolution is high. CGAN is a type of GAN that uses a generator model to conditionally generate images. Kamran *et al.,* used CGAN to transpose fundus images to fluorescein angiography images, since fluorescein angiography images are necessary in the differential diagnosis of retinal diseases without the requirement for the procedure, which has potential side effects [49]. They constructed a GAN model that consists of a novel residual block capable of producing high-quality fluorescein angiography images, achieving qualitative outcomes as good as real angiograms. Park *et al.,* presented another CGAN called M-GAN, which achieves accurate and correct retinal vessel segmentation via balancing losses within stacked fully deep convolutional networks [50]. To measure their model's accuracy, they applied the IoU, F1-score, and MCC. The results demonstrated that their model obtained accurate performance. They applied DRIVE, STARE, HRF, and CHASE-DB1 datasets to validate their method and compared the result with those of other methods.

Zhou *et al.,* constructed a CGAN model to produce high-resolution fundus images for DR that can be manipulated with arbitrary grading and lesion information [51]. They adopted multi-scale spatial and channel attention, which was devised to enhance the generation capability to synthesize small details. They evaluated the method using an experiment on the EyePACS dataset connected to Kaggle and the FGADR dataset. A CGAN containing two generator modules and four discriminator modules was developed by Tavakkoli *et al.,* [52]. Expert evaluations demonstrated that the model generates high-quality fluorescein angiography images indistinguishable from real angiograms. Many GAN approaches can be found in the literature, each study using their own GAN architectures. One of them is the deep convolution generative adversarial network (DCGAN), which includes deep neural networks inside the GAN, stimulates the GAN training process, and stabilizes the training process.

Yang *et al.,* presented SUD-GAN, which merges DCGAN with short connections and dense blocks to separate blood vessels from fundus images [53]. The authors evaluated their model on the DRIVE and STARE and achieved 0.8340 and 0.9820 sensitivity, and 0.8334 and 0.9897 specificity, respectively.

Guo *et al.,* introduced a technique for producing fundus images based on a combined GAN method called Com-GAN, which creates both regular fundus images and fundus images with hard exudates [54]. They used existing images to train Com-GAN, which consisted of two subnetworks, im-WGAN, and im-CGAN. Before they combined the synthesis images with the original image set to expand the datasets, they performed qualitative and quantitative evaluations on the produced images. They conducted experiments using the DIARETDB1 and e-ophtha EX datasets. The cup-to-disc ratio technique was used for the structural examination of the optic nerve in glaucoma diagnosis. Yang *et al.,* presented a method for cup segmentation employing a combination of the green channel of RGB images and the given optic disc mask as the input to a modified U-Net CNN [43]. They achieved a mean dice coefficient of 94\% in the DRISHTI-GS dataset. Dong-Gun *et al.,* constructed a synthetic image method that reconstructs the vessel image based on prior retinal image data using the multilayer perceptron idea, including ANN [55]. They found that high-resolution vessel images can be derived from images with low resolution using a mathematical analysis employing images with high and low resolution obtained from the same patient to validate their method.

Table 5 describes for each primary study considering GAN-based models: CycleGAN, CGAN, DCGAN, and Com-GAN model information regarding the datasets, methods used, and the performance measure used with their values.

Table 5
Summary of GAN-based methods and their corresponding measures

Model	Study	Dataset	Method	Class	Performance
CycleGAN	[46]	Google images	CycleGAN	-	Improved AQE grade values
	[47]	EyePACS	Cycle-CBAM	CycleGAN	PSNR=18.33 SSIM=0.66
				Cycle-CBAM	PSNR=19.27 SSIM=0.68 ACC=53.6% K=0.824
	[48]	DRIVE, DRITSHI-GS	Multiple-channel	-	PSNR=23.011 SSIM=0.8877
CGAN	[49]	Private dataset (Feiz hospital)	Residual block	-	FID=30.3
	[50]	DRIVE, STARE, HRF, CHASE-DB1	Deep residual block M-GAN	Drive:	ACC=97.06% AUC=98.68% F1-score=83.24% MCC=81.63% IoU=71.29
				STARE:	ACC=98.76% AUC=98.73% F1-score=83.7% MCC=83.06% IoU=71.98
				HRF:	ACC=97.61% AUC=98.52% F1-score=79.72% MCC=78.45

				CHASE-DB1:	ACC=97.36% AUC=98.59% F1-score=81.1% MCC=79.79%
	[51]	EyePACS, FGADR	Muti-scale spatial, channel attention	Average:	FID=4.53
				Real synthesized:	ACC=87.98% K=85.81
				Real+Real synthesized:	ACC=89.32% K=87.99
	[52]	Private dataset (Feiz hospital)	Residual block, PatchGAN	-	FID=43 SSIM=0.67
DCGAN	[53]	DRIVE, STARE	U-Net encoder/decoder SUD-GAN	DRIVE:	ACC=95.6% Sen=83.4% Spc=98.2%
				STARE:	ACC=96.63% Sen=83.34% Spc=98.97%
Com-GAN	[54]	DIARETDB1, eophtha EX	im-WGAN, im-CGAN	Normal image:	SSIM=0.77 FID=14.89
				Images with hard exudates:	SSIM=0.74 FID=15.37
Image synthetic approach	[55]	DRIVE, HRF	ANN	-	RMSE=0.335

Most of the researchers used peak signal-to-noise ratio (PSNR), structural index similarity (SSIM), AUC, ACC, Sen, F1-score, and Spc to evaluate the performance of their models; others used Cohen's kappa (K), Frechet inception distance (FID), and MCC. Only one study [46] used the automated equality evaluation (AQE) to evaluate the CycleGAN method. Another study [55] used the root mean square error (RMSE) to evaluate their proposed CNN method.

Despite the success of GAN techniques in the generation of retinal fundus images, their implementation to retinal imaging is recent, and their clinical acceptance seems to be limited or non-existent thus far. So, several limits should be the focus of future research concerning the suggested methods.

i. GAN may work with retinal images that are often lower than the resolution of existing retrieval methods of retinal fundus images. This may contribute to a lack of quality of the synthesized datasets.
ii. The point at which the optic disc and macula appear correctly placed is an important but insufficient condition; plausible diameter and geometry in the clinical context are critical.
iii. Current methods place significance to retinal vascularity, demonstrating that the images produced retain the retinal vessels' morphology. Due to abnormal interruptions, unusual width variation along the same vessel, and lack of differentiation between veins and arteries, the synthetic vessel networks are not clinically acceptable [56].
iv. Although several techniques for evaluation have been explored, such as segmentation methods for image quality, there is no standard evaluation scheme. Furthermore, retinal experts and ophthalmologists should al-ways judge the realism and reliability of synthetic data. A synthetic retinal image can only be considered clinically suitable after clinical examination, and can then be used for more scientific research purposes.
v. Although synthetic retinal images produced with GANs have a consistent overall appearance, retinal lesions cannot be accurately reproduced.

4. Datasets Used in Deep Learning Methods for Detecting Diabetic Eye Diseases

This section is divided into two subsections: the first subsection reviews new DL approaches for data augmentation to enhance the dataset's quality and dataset size for fundus image datasets; the second subsection surveys datasets used in DL-based approaches for diabetic eye detection. Table 6 lists the publicly available datasets involved in the chosen researches. The table includes the name of the dataset, the disease type, a brief description of the particular dataset, and the references of the studies that employed the dataset.

Table 6
Datasets available for automatic Diabetes Eye Detection

Dataset	Diseases Type	Description	Studies
ORIGA	Glaucoma	650 retinal fundus images (from 482 healthy eye and 168 glaucoma patient)	[13,45,47,49,50]
MESSIDOR	DR, DME	Consists of 1200 fundus images, images are labelled with DR	[5,13,30,31,38,41,49,51]
MESSIDOR -2	DR	Consists of 1748 fundus images, images are labelled with DR	[5,30,38]
Kaggle	DR	Competition provided by EyePACS (No DR - Mild – Moderate – Severe - Proliferative DR) total 88,702	[27,31,35,36,38,42,43,52]
Retina	Cataract, Glaucoma	Retina dataset containing four categories: 1) normal 2) cataract 3) glaucoma 4) retina disease.	[5,29,30]
DIARETDB0	DR	160 fundus images for DR with Hypertensive Retinopathy	[37,41]
DIARETDB1	DR	89 fundus images	[13,28,34,37,41,49,53,54]
DRISHTI-GS	Glaucoma	101 fundus images divide into 51 testing and 50 training sample	[5,30,45,46, 48, 51,55]
RIM-ONE	Glaucoma	Total of 131 (39 Glaucoma & 92 Normal) Manual segmentation masks of OD Classification	[45,46]
Macula-centred	DR	13,767 fundus images of 1872 patients	[28]
ACRIMA	Glaucoma	705 fundus images, Classification labels of normal and glaucomatous.	[45]
HRF	DR, Glaucoma	15 healthy, 15 DR, 15 Glaucoma Segmentation masks of FOV	[35,41,49,51,56]
OCT & CFI	Glaucoma	100 fundus images	[49]
REFUGE	Glaucoma	1200 fundus images, Pixel-wise annotations of OD and OC Localization mask of Fovea Classification	[45]
DRIONS-DB	Glaucoma	110 fundus images	[49,51]
DRIVE	Glaucoma, DR	40 fundus images	[37,41,49,51,55,56]
IDRID	DR, DME	516 fundus images with both DR and DMD	[38]
DMR	DR	9939 fundus images	[38]
SCES	Glaucoma	Singapore Chinese Eye Study, is consist of 1676 images	[50]
DR1&DR2	DR	500 fundus images for DR with Hypertensive Retinopathy	[37]
DR-HAGIS	DR, Glaucoma	40 fundus images for 30 DR & 10 glaucoma with Hypertensive Retinopathy (HR)	[37]
CHASE-DB1	Glaucoma	28 fundus images	[51,56]
STARE	Glaucoma	400 fundus images	[41,56]
FGADR	DR	Seg-set (1842 images) and Grade-set (1000 images).	[57]
RIGA	Glaucoma	Contain 665 images from three different sources: MESSIDOR, Bin Rushed, and Magrabi Eye centre.	[31]
HEI-MED	DME	169 fundus images	[31]

Ocular	Glaucoma, Cataract, DR	5,000 classifieds into 8 disease classes.	[58]

Some datasets only contain fundus images for one particular disease, such as RIM-ONE and ORIGA datasets, which are limited to glaucoma. Some datasets have more than one disease, such as MESSIDOR and HRF; some datasets offer different categories for grading diseases, such as Kaggle for DR (No, Mild, Moderate, Severe, and Proliferative DR). There is a large discrepancy in the numbers of images within the datasets. Some datasets contain less than 100 images, such as DRIONS-DB and DRIVE; some contain 1748, such as Messidor-2; and the Kaggle dataset contains 88,702 images. To the best of our knowledge, the datasets on cataracts are few and limited, whereas glaucoma has several datasets. The percentages of the used datasets are illustrated in Figure 1. Overall, 47%, 41%, 9%, and 3% of the datasets are attributed to glaucoma, DR, DME, and cataract diseases, respectively. The following subsection investigates the most commonly used datasets for diabetic eye diseases in DL approaches.

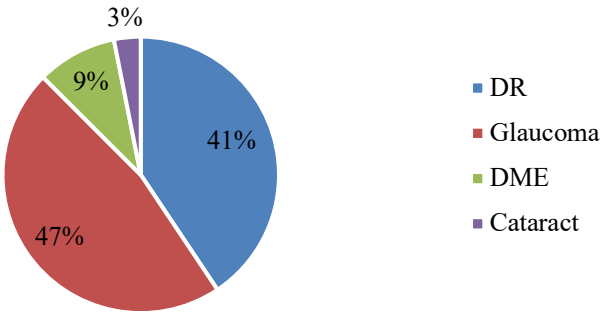

Fig. 1. The percentages of the commonly used Datasets

4.1 DIARETDB1 Dataset [59]

The most common dataset used for DR detection methods is DIARETDB1. The DIARETDB1 dataset only consists of 89 fundus images; the authors of [13,28,34,37,41,49,53,54] used DIARETDB1 data. The DIARETDB1 contains 84 mild non-proliferative signs (microaneurysms) of DR and five recognized as normal, which do not contain any signs of DR. Figure 2 represents an example of data with ground truth where DIARETDB1 data correspond to a good practical condition, where the images are comparable and can be used to evaluate different methods' overall performance.

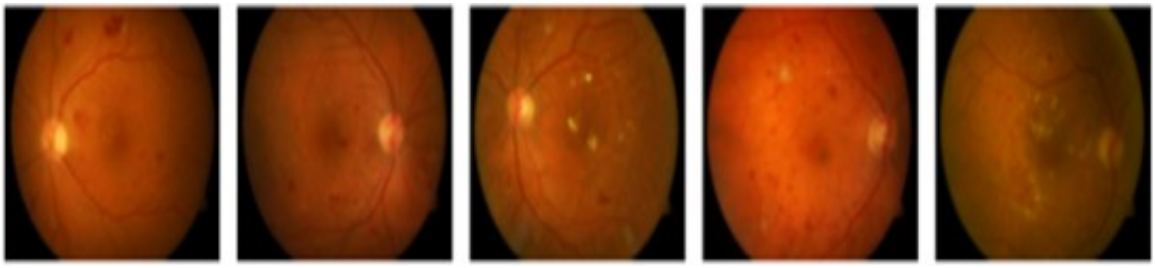

Fig. 2. Architecture of integration IoT with DL in smart cities

4.2 ORIGA Dataset [60]

The Online Retinal fundus Image database for Glaucoma Analysis and research was collected by The Singapore Eye Research Institute and has 482 healthy images and 168 glaucomatous images, as shown in Figure 3. ORIGA provides clinical ground truth of fundus images to benchmark segmentation and classification algorithms. It presents cup-to-disc ratio and labels for each image and is available for online access upon request. This dataset has been used as a primary dataset in some of the recent research on glaucoma detection [13,45,47,49,50].

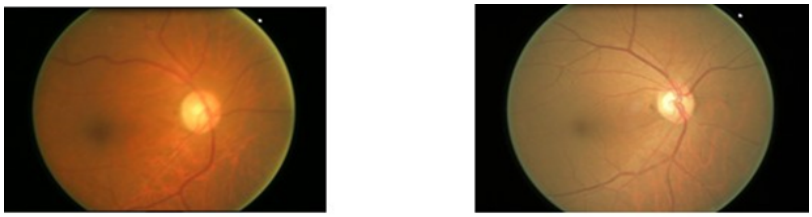

Fig. 3. Samples of ORIGA dataset: (a) glaucomatous image and (b) healthy image

4.3 Kaggle Dataset [61]

The Kaggle dataset was constructed for a DR competition by Eye PACS, which is a large set of high-resolution retina images (total of 88,702 images). The California Healthcare Foundation sponsors the competition. Images are labelled with a caption ID and either left or right and a scale of 0 to 4, indicating No, Mild, Moderate, Severe, and Proliferative DR, respectively, as shown in Figure 4. According to our research of state-of-the-art methods, several studies have applied the Kaggle dataset [27,31,35,36,38,42,43,52].

Fig. 4. Samples of DR from Kaggle: healthy, mild, moderate, severe, and proliferative DR

4.4 MESSIDOR Dataset [62]

MESSIDOR stands for Methods to Evaluate Segmentation and Indexing Techniques in Retinal Ophthalmology, which was a research program funded by the French Ministry of Research and Defence. The dataset contains 1200 images for DR and DME with a resolution of 2304 × 1536, 2240 × 1488, and 1440 × 960 pixels for DR and DME, as shown in Figure 5.

Fig. 5. Sample of MESSIDOR dataset

4.5 MESSIDOR-2 Dataset [62]

This is a public dataset consisting of 1748 colour eye fundus images. Images labelled with DR were acquired with a 45◦ FOV at 1440 × 960, 2240 × 1488, or 2304 × 1536 pixels (px).

4.6 Retina Dataset [63]

This is a recent dataset from the Kaggle datasets, containing four categories for normal, cataract, glaucoma, and retina disease. It includes 601 images divided into 300 retinal images for normal, 100 retinal images for cataracts, 101 retinal images for glaucoma, and 100 retinal images for retina disease, as shown in Figure 6.

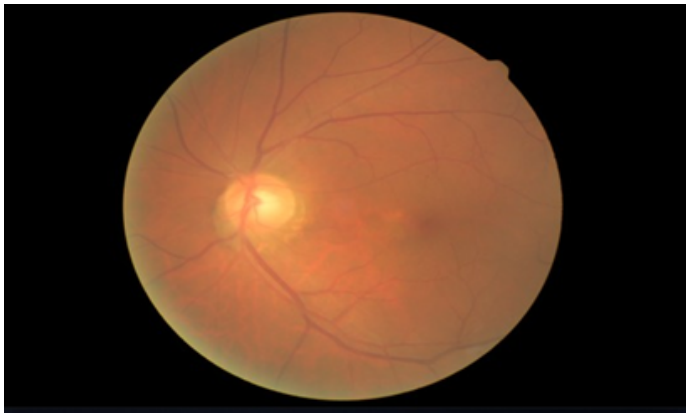

Fig. 6. Sample of cataract image from the Retina Dataset

4.7 IDRiD Dataset [64]

IDRiD stands for Indian Diabetic Retinopathy Image Dataset. This publicly available dataset includes 516 fundus images obtained through a 50◦ FOV expressed in five DR stages. It presents data on the disease severity of DR and DME for every image. Furthermore, it comprises typical DR lesions and normal retinal structures interpreted at the pixel level.

4.8 Ocular Dataset [58]

Ocular Disease Intelligent Recognition -ODIR is an organized ophthalmic database including age, colour fundus pictures from both eyes, and doctors' diagnostic keywords. It is divided into eight

categories, including Diabetes, Glaucoma, Cataract, Age-related Macular Degeneration, Hypertension, Pathological Myopia, Other diseases/abnormalities.

4.9 DMR Dataset [65]

This dataset is formed of 9939 fundus images from 2740 diabetic patients (2720 × 2720 px). Images have a 45◦ field of view (FOV) and are graded by modified Davis grading.

4.10 RIGA Dataset [66]

The University of Michigan released the RIGA dataset to evaluate their segmentation method for glaucoma diagnosis applications. It involves three parts: 460 images from MESSIDOR, 195 images from the Bin Rushed Ophthalmic Centre, and 95 images from the Magrabi Eye Centre.

4.11 HEI-MED Dataset [67]

HEI-MED stands for Hamilton Eye Institute Macular Edema Dataset. It consists of 169 fundus images to train and test image processing algorithms to detect exudates of DR and DME. The dataset is formed of high-quality jpeg images. An expert ophthalmologist manually segmented all images in the dataset.

4.12 DRIVE Dataset [68]

This publicly available dataset is used for blood vessel segmentation. It contains 40 images acquired at a 45◦ FOV. The images have a size of 565 × 584 pixels. Among them, there are seven mild DR images, and the remaining include images of a normal retina.

4.13 SCES Dataset [69]

This dataset was acquired under the Singapore Chinese eye study conducted on 1060 Chinese participants and was graded by a professional grader and retinal specialist. The study was conducted to identify the relative importance of anterior chamber depth in Chinese persons in Singapore.

4.14 E-ophtha Dataset [70]:

This publicly available dataset includes E-ophtha EX and E-ophtha MA. E-ophtha EX includes 47 images with EX and 35 normal images. E-ophtha MA contains 148 images with MA and 233 normal images.

4.15 HRF Dataset [71]

This dataset provides openly available images for blood vessel segmentation. It has 45 images that are 3504× 2336 pixels in size. It contains 15 images of DR, 15 images of healthy eyes, and 15 images of glaucoma.

4.16 STARE Dataset [72]

This publicly available dataset is used for blood vessel segmentation. It contains 20 images acquired at a 35◦ FOV. The images have a size of 700 × 605 pixels. Among them, there are 10 normal images.

4.17 CHASE Dataset [73]

This publicly available dataset is provided for blood vessel segmentation. It contains 28 images with a size of 1280 × 960 pixels and acquired at a 30◦ FOV.

4.18 ROC Dataset [74]

It contains 100 publicly available retina images acquired at a 45◦ FOV. Its size ranges from 768 × 576 to 1389 × 1383 pixels. The images are annotated to detect MA. Only training ground truths are available.

5. Performance Measures of Diabetic Eye Diseases Models Evaluations

The performance of eye disease classification models can be evaluated using several metrics: the accuracy (ACC); loss; recall; positive predictive value (PPV), which is commonly known as the precision; specificity (SPC); negative predictive value (NPV); F1-score; Matthew's correlation coefficient (MCC); and area under the curve (AUC). Correspondingly, a confusion matrix is introduced for each model. Accuracy, in Eq. (1), is the number of examples correctly predicted from the total number of examples. Moreover, the intersection over union (IoU) is calculated as in Eq. (2).

$$\text{Accuracy (ACC)} = \frac{\text{Tp+Tn}}{\text{Tp+Tn+Fp+Fn}} \tag{1}$$

$$\text{Intersection over Union (IoU)} = \frac{\text{Tp}}{\text{Tp+Fp+Fn}} \tag{2}$$

where Tp and Tn are the true positive and negative parameters, respectively. Fp and Fn are the false positive and false negative values. Sensitivity or Recall, given in Eq. (3), is the number of samples actually and predicted as Positive from the total number of samples actually Positive Also known as True positive rate. While the True Negative Rate, as was called Specificity, given in Eq. (4), is the number of samples actually and predicted as Negative from the total number of samples actually Negative.

$$\text{Recall (Sensitivity)} = \frac{\text{Tp}}{\text{Tp+Fn}} \tag{3}$$

$$\text{Specificity (SPC)} = \frac{\text{Tn}}{\text{Tn+Fp}} \tag{4}$$

Eq. (5) shows the Precision, also called Positive Predictive Value [75], which represents the number of samples actually and predicted as Positive from total number of samples predicted as Positive and mAP means the mean average precision (mAP). Whereas the Negative Predictive Value (NPV), [76], is the number of samples actually and predicted as Negative from the total number of

samples predicted as Negative, given in Eq. (6). The Harmonic Mean of Precision and Recall that is known as F1-score is shown in Eq. (7). Finally, Matthew's correlation coefficient range, [77], that allows one to gauge how well the classification model/function is performing.

$$\text{Precision (PPV)} = \frac{\text{Tp}}{\text{Tp+Fp}} \tag{5}$$

$$\text{NPV} = \frac{\text{Tn}}{\text{Tn+Fn}} \tag{6}$$

$$\text{F1} - \text{score} = \frac{2*\text{Tp}}{2*\text{Tp+Fp+Fn}} \tag{7}$$

$$\text{MCC} = \frac{(\text{Tp}\times\text{Tn})-(\text{Fp}\times\text{Fn})}{\sqrt{(\text{Tp+Fp})(\text{Tp+Fn})(\text{Tn+Fp})(\text{Tn+Fn})}} \tag{8}$$

Root Mean Square Error (RMSE) is a standard way to measure the error of a model in predicting quantitative data [78]. Formally it is defined as in Eq. (9). Mean Absolute Error (MAE) measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables, as shown in Eq. (10)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_{predict,i}-y_{data,i})^2}{n}} \tag{9}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{predict,i} - y_{data,i}\right| \tag{10}$$

Where ypredict,i is the predicted energy consumption at time point i, ydata,i is the actual energy consumption at time point i, $\bar{y}$data is the average energy consumption, and n is the total number of data points in the dataset.

Researchers have mostly evaluated the GAN fundus images using PSNR and SSIM to measure the image's quality, since the evaluation of fundus images is a challenging task and human opinion is influenced by environmental and psychological parameters. PSNR determines the ratio of the maximum possible power of a signal to the power of reducing noise that affects its representation accuracy. The SSIM was designed based on three factors: contrast distortion, luminance distortion, and loss of correlation.

Automated quality evaluation (AQE) assesses the retinal image quality of the most important human factors for evaluating image quality, which are image sharpness and illumination [78]. The dice coefficient is used in evaluating the accuracy of the area segmentation of the correct classifications of points inside the segmented region. The dice coefficient is similar to Jaccard, as shown in Eq. (11).

$$Dice\ Coefficient\ (DC) = \frac{2\text{Tp}}{2\text{Tp+Fp+Fn`}} \tag{11}$$

The Frechet inception distance (FID) is applied to assess the performance of image generation models. FID is used to estimate the similarity of the real images to the synthesized images in GAN. FID is calculated by measuring the Frechet distance within a couple of Gaussians fitted to the feature representations of the inception network. Another measure is Cohen's kappa, which is a statistic that is used to measure the inter-rater reliability for categorical items, as shown in Eq. (12). he quadratic

weighted kappa broadly adopts the multi-class classification model to measure the agreement between two ratings that each classify N items into C mutually exclusive categories [79].

$$Cohen's\ Kappa\ (K)\ = 1 - \frac{1-p_0}{1-p_e} \tag{12}$$

Some researchers applied Cohen's kappa score, whereas others used weighted kappa, which allows disagreements to be weighed separately and is particularly beneficial when codes are ordered. Human experts have reviewed the results, meaning that some models were evaluated by professional ophthalmologists to determine the effect of the detection model and the synthesized images in GAN. Table 7 describes the different performance measures used to evaluate diabetic eye diseases models.

Table 7
Different performance measures for diabetic eye disease models in the studies considered in this review

Model Technique	Categories	Performance Measurements																	
		ACC	Sen	AUC	SPC	F1-score	PPV	SSIM	PSNR	AQE	RMSE	IoU	MCC	MAP	DC	FID	K	HE	Study
Classification based model	TL	-	√	√	√	-	-	-	-	-	-	-	-	-	-	-	-	-	[26]
		√	-	-	-	√	-	-	-	-	-	-	-	-	-	-	√	-	[27]
		√	√	√	√	√	√	-	-	-	-	-	-	-	-	-	√	√	[28]
		√	-	-	-	-	-	-	-	-	-	-	-	-	-	-	√	-	[29]
		√	√	-	√	-	√	-	-	-	-	-	-	-	-	-	-	-	[30]
	DL with ML	√	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[32]
		√	√	√	-	√	√	-	-	-	√	-	-	-	-	-	-	√	[33]
		√	-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[34]
	Tailored DL	-	√	√	√	-	-	-	-	-	-	-	-	-	-	-	√	-	[35]
		-	√	-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	[36]
		√	√	√	√	-	-	-	-	-	-	-	-	-	-	-	-	-	[37]
		-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	√	√	[38]
		√	√	√	√	√	√	-	-	-	-	-	-	-	-	-	-	-	[41]
		√	-	√	-	-	-	-	-	-	-	-	-	-	-	-	√	-	[42]
		√	√	-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	[5]
Localization based model	Segmentation	√	-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	√	[45]
		√	√	√	√	-	-	-	-	-	-	√	-	-	√	-	-	√	[46]
		√	√	√	√	-	-	-	-	-	√	-	-	-	-	-	-	-	[47]
	RCNN	√	√	√	√	-	-	-	-	-	-	-	-	√	-	-	-	-	[13]
		-	-	√	-	-	-	-	-	-	-	-	-	-	-	-	-	-	[50]
		-	√	√	√	√	√	-	-	-	-	√	-	-	-	-	-	-	[49]
GAN based models	CycleGAN	-	-	-	-	-	-	-	-	√	-	-	-	-	-	-	-	-	[75]
		-	-	-	-	-	-	√	√	-	-	-	-	-	-	-	-	-	[55]
		√	-	-	-	-	-	√	√	-	-	-	-	-	-	-	√	-	[52]
	CGAN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	√	-	-	[76]
		√	√	√	√	-	-	√	-	-	-	√	√	-	-	-	-	-	[56]
		-	-	-	-	-	-	√	√	-	-	-	-	-	-	√	-	-	[77]
		√	-	-	-	-	-	-	-	-	-	-	-	-	-	√	√	-	[57]
	Other GANs	√	√	√	√	-	√	-	-	-	-	-	-	-	-	-	-	-	[79]
		√	√	-	√	-	-	√	-	-	-	-	-	-	-	√	-	-	[53]
		-	-	-	-	-	-	-	-	-	-	-	-	-	√	-	-	-	[48]
		-	-	-	-	-	-	-	-	-	√	-	-	-	-	-	-	-	[80]

Positive predictive value (PPV), peak signal-to-noise ratio (PSNR), structural index similarity (SSIM), automated quality evaluation (AQE), root mean square error (RMSE), mean average precision (MAP), Frechet inception distance (FID), human expert (HR)

The most commonly used evaluation measures of DED models, per Table 6, are the accuracy, sensitivity, area under the curve (AUC), and specificity, as shown in Figure 7. Overall, 16%, 14%, 14%, and 12% of the reviewed studies used accuracy, sensitivity, AUC, and specificity, respectively, to evaluate the DED models. More recent work is using deep learning and machine learning in many areas as in the studies [81-93].

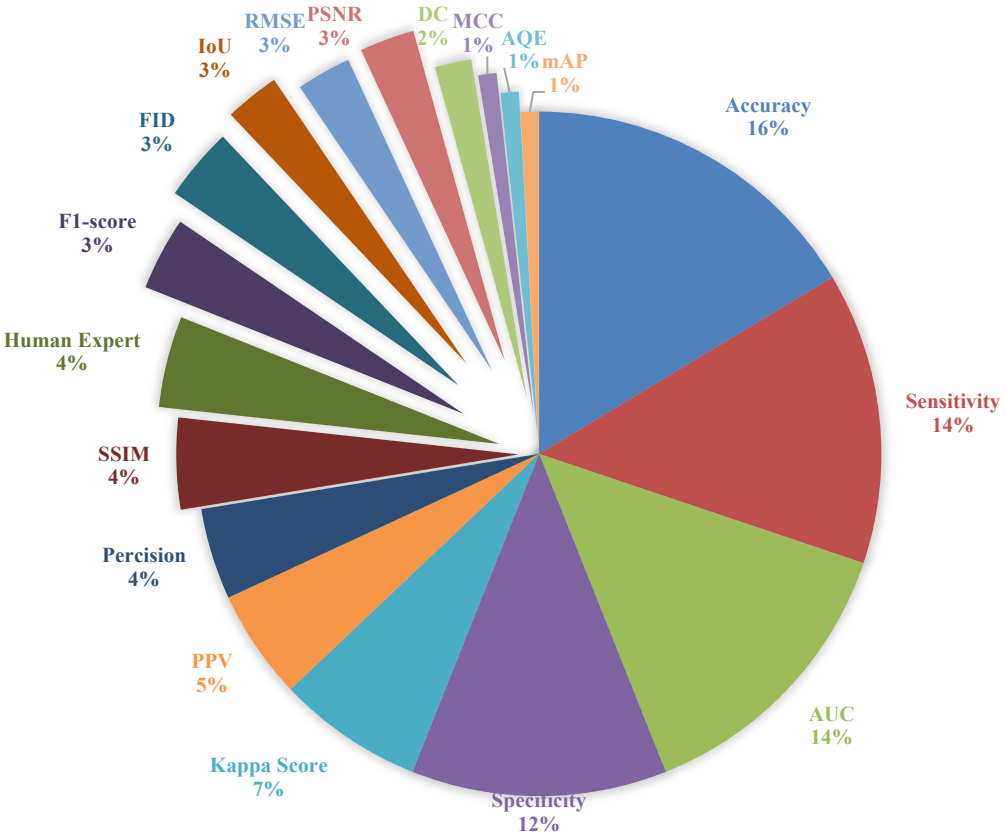

Fig. 7. The percentages of the Diabetic eye diseases performance measures

7. Conclusion

This review paper delivered a comprehensive overview of the state-of-the-art studies on diabetic eye disease (DED) methods that used machine learning and deep learning for detecting, classifying, and grading diabetic eye diseases. We focused on the common four eye diseases: diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma, and cataracts. We classified the model techniques reported in recent studies into three main categories: classification-based, localization-based, and generative adversarial network (GAN) models. Furthermore, we investigated the current and commonly used datasets that are available containing fundus images of diabetic eye diseases. We reviewed studies that employed different GAN techniques to improve the fundus images dataset or increase the size of the images in the datasets. Finally, we described the performance measures used in the previous studies for evaluating various models of diabetic eye diseases.

Based on the review conducted in this study, we found that the generation of retinal fundus images in GAN techniques still faced several limitations to clinical acceptance. Some issues include the point at which the optic disc and macula appear correctly placed, and retinal lesions cannot reproduce accurately. So, the results should be judged by retina experts or ophthalmologists on the reality and reliability of the synthetic data. Due to the limited availability of cataract datasets, the cataract disease did not have sufficient studies investigating cataract classification and prediction.

Furthermore, most studies studying cataracts were independent studies to classify the cataract disease separately, not combine the cataract disease with other diseases to test the classification's accuracy with other diseases.

To conclude the most advanced method in Diabetic Eye Diseases, we found that 25% of the studies used commonly CNN architectures such as AlexNet, VGG16, DensNet (Based on included studies in this review). Furthermore, we found that RCNN, Fast RCNN, and RCNN with attention were the most recent method segmentation studies which Region-Based Convolutional Neural Networks. The convolutional residual block is a stack of layers mainly employed with GAN techniques for generating synthesis fundus images.

Acknowledgment
Researchers would like to thank the Deanship of Scientific Research, Qassim University for funding publication of this project.

References

[1] Vashist, Praveen, Sameeksha Singh, Noopur Gupta, and Rohit Saxena. "Role of early screening for diabetic retinopathy in patients with diabetes mellitus: an overview." *Indian Journal of Community Medicine* 36, no. 4 (2011): 247-252. https://doi.org/10.4103/0970-0218.91324

[2] Ting, Daniel Shu Wei, Louis R. Pasquale, Lily Peng, John Peter Campbell, Aaron Y. Lee, Rajiv Raman, Gavin Siew Wei Tan, Leopold Schmetterer, Pearse A. Keane, and Tien Yin Wong. "Artificial intelligence and deep learning in ophthalmology." *British Journal of Ophthalmology* 103, no. 2 (2019): 167-175. https://doi.org/10.1136/bjophthalmol-2018-313173

[3] Hogarty, Daniel T., David A. Mackey, and Alex W. Hewitt. "Current state and future prospects of artificial intelligence in ophthalmology: a review." *Clinical & experimental ophthalmology* 47, no. 1 (2019): 128-139. https://doi.org/10.1111/ceo.13381

[4] Hagiwara, Yuki, Joel En Wei Koh, Jen Hong Tan, Sulatha V. Bhandary, Augustinus Laude, Edward J. Ciaccio, Louis Tong, and U. Rajendra Acharya. "Computer-aided diagnosis of glaucoma using fundus images: A review." *Computer methods and programs in biomedicine* 165 (2018): 1-12. https://doi.org/10.1016/j.cmpb.2018.07.012

[5] Mookiah, Muthu Rama Krishnan, U. Rajendra Acharya, Chua Kuang Chua, Choo Min Lim, E. Y. K. Ng, and Augustinus Laude. "Computer-aided diagnosis of diabetic retinopathy: A review." *Computers in biology and medicine* 43, no. 12 (2013): 2136-2155. https://doi.org/10.1016/j.compbiomed.2013.10.007

[6] Ishtiaq, Uzair, Sameem Abdul Kareem, Erma Rahayu Mohd Faizal Abdullah, Ghulam Mujtaba, Rashid Jahangir, and Hafiz Yasir Ghafoor. "Diabetic retinopathy detection through artificial intelligent techniques: a review and open issues." *Multimedia Tools and Applications* 79 (2020): 15209-15252. https://doi.org/10.1007/s11042-018-7044-8

[7] Sarki, Rubina, Khandakar Ahmed, Hua Wang, and Yanchun Zhang. "Automatic detection of diabetic eye disease through deep learning using fundus images: a survey." *IEEE access* 8 (2020): 151133-151149. https://doi.org/10.1109/ACCESS.2020.3015258

[8] Alyoubi, Wejdan L., Wafaa M. Shalash, and Maysoon F. Abulkhair. "Diabetic retinopathy detection through deep learning techniques: A review." *Informatics in Medicine Unlocked* 20 (2020): 100377. https://doi.org/10.1016/j.imu.2020.100377

[9] Islam, Md Mohaimenul, Hsuan-Chia Yang, Tahmina Nasrin Poly, Wen-Shan Jian, and Yu-Chuan Jack Li. "Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis." *Computer Methods and Programs in Biomedicine* 191 (2020): 105320. https://doi.org/10.1016/j.cmpb.2020.105320

[10] Sebastian, Anila, Omar Elharrouss, Somaya Al-Maadeed, and Noor Almaadeed. "A survey on deep-learning-based diabetic retinopathy classification." *Diagnostics* 13, no. 3 (2023): 345. https://doi.org/10.3390/diagnostics13030345

[11] Bilal, A., G. Sun, and S. Mazhar. "Survey on recent developments in automatic detection of diabetic retinopathy." *Journal Français d'Ophtalmologie* 44, no. 3 (2021): 420-440. https://doi.org/10.1016/j.jfo.2020.08.009

[12] Kandel, Ibrahem, and Mauro Castelli. "Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review." *Applied Sciences* 10, no. 6 (2020): 2021. https://doi.org/10.3390/app10062021

[13] Shabbir, Amsa, Aqsa Rasheed, Huma Shehraz, Aliya Saleem, Bushra Zafar, Muhammad Sajid, Nouman Ali, Saadat Hanif Dar, and Tehmina Shehryar. "Detection of glaucoma using retinal fundus images: A comprehensive review." *Mathematical Biosciences and Engineering* 18, no. 3 (2021): 2033-2076. https://doi.org/10.3934/mbe.2021106

[14] Mirzania, Delaram, Atalie C. Thompson, and Kelly W. Muir. "Applications of deep learning in detection of glaucoma: a systematic review." *European Journal of Ophthalmology* 31, no. 4 (2021): 1618-1642. https://doi.org/10.1177/1120672120977346

[15] Ran, An Ran, Clement C. Tham, Poemen P. Chan, Ching-Yu Cheng, Yih-Chung Tham, Tyler Hyungtaek Rim, and Carol Y. Cheung. "Deep learning in glaucoma with optical coherence tomography: a review." *Eye* 35, no. 1 (2021): 188-201. https://doi.org/10.1038/s41433-020-01191-5

[16] Li, Tao, Wang Bo, Chunyu Hu, Hong Kang, Hanruo Liu, Kai Wang, and Huazhu Fu. "Applications of deep learning in fundus images: A review." *Medical Image Analysis* 69 (2021): 101971. https://doi.org/10.1016/j.media.2021.101971

[17] Cheloni, Riccardo, Stefano A. Gandolfi, Carlo Signorelli, and Anna Odone. "Global prevalence of diabetic retinopathy: protocol for a systematic review and meta-analysis." *BMJ open* 9, no. 3 (2019): e022188. https://doi.org/10.1136/bmjopen-2018-022188

[18] Lu, Lei, Ying Jiang, Ravindran Jaganathan, and Yanli Hao. "Current advances in pharmacotherapy and technology for diabetic retinopathy: a systematic review." *Journal of Ophthalmology* 2018 (2018). https://doi.org/10.1155/2018/1694187

[19] Soomro, Taha, Neil Shah, Magdalena Niestrata-Ortiz, Timothy Yap, Eduardo M. Normando, and M. Francesca Cordeiro. "Recent advances in imaging technologies for assessment of retinal diseases." *Expert Review of Medical Devices* 17, no. 10 (2020): 1095-1108. https://doi.org/10.1080/17434440.2020.1816167

[20] World Health Organization. "World report on vision." (2019).

[21] Threatt, Jennifer, Jennifer F. Williamson, Kyle Huynh, Richard M. Davis, and Kathie Hermayer. "Ocular disease, knowledge and technology applications in patients with diabetes." *The American journal of the medical sciences* 345, no. 4 (2013): 266-270. https://doi.org/10.1097/MAJ.0b013e31828aa6fb

[22] Nazir, Tahira, Aun Irtaza, Ali Javed, Hafiz Malik, Dildar Hussain, and Rizwan Ali Naqvi. "Retinal image analysis for diabetes-based eye disease detection using deep learning." *Applied Sciences* 10, no. 18 (2020): 6185. https://doi.org/10.3390/app10186185

[23] Musat, Ovidiu, Corina Cernat, Mahdi Labib, Andreea Gheorghe, Oana Toma, Madalina Zamfir, and Ana Maria Boureanu. "Diabetic macular edema." *Romanian journal of ophthalmology* 59, no. 3 (2015): 133.

[24] Shah, Shaheen P., Claire E. Gilbert, Hessom Razavi, Elizabeth L. Turner, and Robert J. Lindfielda. "Acuité visuelle pré-opératoire des patients subissant une opeacute; ration de la cataracte et eacute; tat de deacute; veloppement des pays une eacute; tude mondiale." *Bulletin of the World Health Organization* 89, no. 10 (2011): 749-756. https://doi.org/10.2471/BLT.10.080366

[25] Pan, Xiangji, Kai Jin, Jing Cao, Zhifang Liu, Jian Wu, Kun You, Yifei Lu *et al.,* "Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning." *Graefe's Archive for Clinical and Experimental Ophthalmology* 258 (2020): 779-785. https://doi.org/10.1007/s00417-019-04575-w

[26] Samanta, Abhishek, Aheli Saha, Suresh Chandra Satapathy, Steven Lawrence Fernandes, and Yu-Dong Zhang. "Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset." *Pattern Recognition Letters* 135 (2020): 293-298. https://doi.org/10.1016/j.patrec.2020.04.026

[27] Zhang, Wei, Jie Zhong, Shijun Yang, Zhentao Gao, Junjie Hu, Yuanyuan Chen, and Zhang Yi. "Automated identification and grading system of diabetic retinopathy using deep neural networks." *Knowledge-Based Systems* 175 (2019): 12-25. https://doi.org/10.1016/j.knosys.2019.03.016

[28] Syarifah, Mas Andam, Alhadi Bustamam, and Patuan P. Tampubolon. "Cataract classification based on fundus image using an optimized convolution neural network with lookahead optimizer." In *AIP Conference Proceedings*, vol. 2296, no. 1. AIP Publishing, 2020. https://doi.org/10.1063/5.0030744

[29] Sarki, Rubina, Khandakar Ahmed, Hua Wang, and Yanchun Zhang. "Automated detection of mild and multi-class diabetic eye diseases using deep learning." *Health Information Science and Systems* 8, no. 1 (2020): 32. https://doi.org/10.1007/s13755-020-00125-5

[30] Oh, Kangrok, Hae Min Kang, Dawoon Leem, Hyungyu Lee, Kyoung Yul Seo, and Sangchul Yoon. "Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images." *Scientific reports* 11, no. 1 (2021): 1897. https://doi.org/10.1038/s41598-021-81539-3

[31] Grassmann, Felix, Judith Mengelkamp, Caroline Brandl, Sebastian Harsch, Martina E. Zimmermann, Birgit Linkohr, Annette Peters, Iris M. Heid, Christoph Palm, and Bernhard HF Weber. "A deep learning algorithm for prediction of

age-related eye disease study severity scale for age-related macular degeneration from color fundus photography." *Ophthalmology* 125, no. 9 (2018): 1410-1420. https://doi.org/10.1016/j.ophtha.2018.02.037

[32] Malik, Sadaf, Nadia Kanwal, Mamoona Naveed Asghar, Mohammad Ali A. Sadiq, Irfan Karamat, and Martin Fleury. "Data driven approach for eye disease classification with machine learning." *Applied Sciences* 9, no. 14 (2019): 2789. https://doi.org/10.3390/app9142789

[33] Theera-Umpon, Nipon, Ittided Poonkasem, Sansanee Auephanwiriyakul, and Direk Patikulsila. "Hard exudate detection in retinal fundus images using supervised learning." *Neural Computing and Applications* 32, no. 17 (2020): 13079-13096. https://doi.org/10.1007/s00521-019-04402-7

[34] Zeng, Xianglong, Haiquan Chen, Yuan Luo, and Wenbin Ye. "Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network." *IEEE access* 7 (2019): 30744-30753. https://doi.org/10.1109/ACCESS.2019.2903171

[35] de La Torre, Jordi, Aida Valls, and Domenec Puig. "A deep learning interpretable classifier for diabetic retinopathy disease grading." *Neurocomputing* 396 (2020): 465-476. https://doi.org/10.1016/j.neucom.2018.07.102

[36] Abbas, Qaisar, and Mostafa EA Ibrahim. "DenseHyper: an automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning." *Multimedia Tools and Applications* 79, no. 41 (2020): 31595-31623. https://doi.org/10.1007/s11042-020-09630-x

[37] Araújo, Teresa, Guilherme Aresta, Luís Mendonça, Susana Penas, Carolina Maia, Ângela Carneiro, Ana Maria Mendonça, and Aurélio Campilho. "DR| GRADUATE: Uncertainty-aware deep learning-based diabetic retinopathy grading in eye fundus images." *Medical Image Analysis* 63 (2020): 101715. https://doi.org/10.1016/j.media.2020.101715

[38] Sreng, Syna, Noppadol Maneerat, Kazuhiko Hamamoto, and Khin Yadanar Win. "Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images." *Applied Sciences* 10, no. 14 (2020): 4916. https://doi.org/10.3390/app10144916

[39] Lu, Zheng, and Dali Chen. "Weakly supervised and semi-supervised semantic segmentation for optic disc of fundus image." *Symmetry* 12, no. 1 (2020): 145. https://doi.org/10.3390/sym12010145

[40] Guo, Fan, Weiqing Li, Jin Tang, Beiji Zou, and Zhun Fan. "Automated glaucoma screening method based on image segmentation and feature extraction." *Medical & Biological Engineering & Computing* 58 (2020): 2567-2586. https://doi.org/10.1007/s11517-020-02237-2

[41] Bajwa, Muhammad Naseer, Muhammad Imran Malik, Shoaib Ahmed Siddiqui, Andreas Dengel, Faisal Shafait, Wolfgang Neumeier, and Sheraz Ahmed. "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning." *BMC medical informatics and decision making* 19 (2019): 1-16. https://doi.org/10.1186/s12911-019-0842-8

[42] Jiang, Yuming, Lixin Duan, Jun Cheng, Zaiwang Gu, Hu Xia, Huazhu Fu, Changsheng Li, and Jiang Liu. "JointRCNN: a region-based convolutional neural network for optic disc and cup segmentation." *IEEE Transactions on Biomedical Engineering* 67, no. 2 (2019): 335-343. https://doi.org/10.1109/TBME.2019.2913211

[43] Makhzani, Alireza, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow, and Brendan Frey. "Adversarial autoencoders." *arXiv preprint arXiv:1511.05644* (2015).

[44] Bellemo, Valentina, Philippe Burlina, Liu Yong, Tien Yin Wong, and Daniel Shu Wei Ting. "Generative adversarial networks (GANs) for retinal fundus image synthesis." In *Computer Vision–ACCV 2018 Workshops: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers 14*, pp. 289-302. Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-21074-8_24

[45] Yoo, Tae Keun, Joon Yul Choi, and Hong Kyu Kim. "CycleGAN-based deep learning technique for artifact reduction in fundus photography." *Graefe's Archive for Clinical and Experimental Ophthalmology* 258 (2020): 1631-1637. https://doi.org/10.1007/s00417-020-04709-5

[46] You, Qijing, Cheng Wan, Jing Sun, Jianxin Shen, Hui Ye, and Qiuli Yu. "Fundus image enhancement method based on CycleGAN." In *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 4500-4503. IEEE, 2019. https://doi.org/10.1109/EMBC.2019.8856950

[47] Yu, Zekuan, Qing Xiang, Jiahao Meng, Caixia Kou, Qiushi Ren, and Yanye Lu. "Retinal image synthesis from multiple-landmarks input with generative adversarial networks." *Biomedical engineering online* 18 (2019): 1-15. https://doi.org/10.1186/s12938-019-0682-x

[48] Kamran, Sharif Amit, Khondker Fariha Hossain, Alireza Tavakkoli, Stewart Zuckerbrod, Salah A. Baker, and Kenton M. Sanders. "Fundus2Angio: a conditional GAN architecture for generating fluorescein angiography images from retinal fundus photography." In *Advances in Visual Computing: 15th International Symposium, ISVC 2020, San Diego, CA, USA, October 5–7, 2020, Proceedings, Part II 15*, pp. 125-138. Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-64559-5_10

[49] Park, Kyeong-Beom, Sung Ho Choi, and Jae Yeol Lee. "M-gan: Retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks." *IEEE Access* 8 (2020): 146308-146322. https://doi.org/10.1109/ACCESS.2020.3015108
[50] Zhou, Yi, Boyang Wang, Xiaodong He, Shanshan Cui, and Ling Shao. "DR-GAN: conditional generative adversarial network for fine-grained lesion synthesis on diabetic retinopathy images." *IEEE journal of biomedical and health informatics* 26, no. 1 (2020): 56-66. https://doi.org/10.1109/JBHI.2020.3045475
[51] Tavakkoli, Alireza, Sharif Amit Kamran, Khondker Fariha Hossain, and Stewart Lee Zuckerbrod. "A novel deep learning conditional generative adversarial network for producing angiography images from retinal fundus photographs." *Scientific Reports* 10, no. 1 (2020): 21580. https://doi.org/10.1038/s41598-020-78696-2
[52] Yang, Tiejun, Tingting Wu, Lei Li, and Chunhua Zhu. "SUD-GAN: deep convolution generative adversarial network combined with short connection and dense block for retinal vessel segmentation." *Journal of digital imaging* 33 (2020): 946-957. https://doi.org/10.1007/s10278-020-00339-9
[53] Guo, Jifeng, Zhiqi Pang, Fan Yang, Jiayou Shen, and Jian Zhang. "Study on the method of fundus image generation based on improved GAN." *Mathematical Problems in Engineering* 2020 (2020): 1-13. https://doi.org/10.1155/2020/6309596
[54] Lima, Arthur Azevedo, Alana C. de Carvalho Araújo, Alan C. de Moura Lima, Jefferson Alves de Sousa, João Dallyson S. de Almeida, Anselmo Cardoso de Paiva, and Geraldo Braz Júnior. "Mask overlaying: a deep learning approach for individual optic cup segmentation from fundus image." In *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 99-104. IEEE, 2020. https://doi.org/10.1109/IWSSIP48289.2020.9145459
[55] Lee, Dong-Gun, Yonghun Jang, and Yeong-Seok Seo. "Intelligent Image Synthesis for Accurate Retinal Diagnosis." *Electronics* 9, no. 5 (2020): 767. https://doi.org/10.3390/electronics9050767
[56] Costa, Pedro, Adrian Galdran, Maria Ines Meyer, Meindert Niemeijer, Michael Abràmoff, Ana Maria Mendonça, and Aurélio Campilho. "End-to-end adversarial retinal image synthesis." *IEEE transactions on medical imaging* 37, no. 3 (2017): 781-791. https://doi.org/10.1109/TMI.2017.2759102
[57] Ramani, R. Geetha, and J. Jeslin Shanthamalar. "Improved image processing techniques for optic disc segmentation in retinal fundus images." *Biomedical Signal Processing and Control* 58 (2020): 101832. https://doi.org/10.1016/j.bspc.2019.101832
[58] Saman, Gule, Neelam Gohar, Salma Noor, Ambreen Shahnaz, Shakira Idress, Neelam Jehan, Reena Rashid, and Sheema Shuja Khattak. "Automatic detection and severity classification of diabetic retinopathy." *Multimedia Tools and Applications* 79 (2020): 31803-31817. https://doi.org/10.1007/s11042-020-09118-8
[59] Kälviäinen, R. V. J. P. H., and H. Uusitalo. "DIARETDB1 diabetic retinopathy database and evaluation protocol." In *Medical image understanding and analysis*, vol. 2007, p. 61. Citeseer, 2007.
[60] Zhang, Zhuo, Feng Shou Yin, Jiang Liu, Wing Kee Wong, Ngan Meng Tan, Beng Hai Lee, Jun Cheng, and Tien Yin Wong. "Origa-light: An online retinal fundus image database for glaucoma analysis and research." In *2010 Annual international conference of the IEEE engineering in medicine and biology*, pp. 3065-3068. IEEE, 2010.
[61] Kumar, Athota Manoj, Atchukola Sai Gopi Kanna, and Ramya G. Franklin. "Diabetic retinopathy detection." In *International Conference on Emerging Trends and Advances in Electrical Engineering and Renewable Energy*, pp. 463-470. Singapore: Springer Nature Singapore, 2020. https://doi.org/10.1007/978-981-15-8685-9_48
[62] Decencière, Etienne, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Béatrice Cochener, Caroline Trone, Philippe Gain *et al.,* "Feedback on a publicly distributed image database: the Messidor database." *Image Analysis and Stereology* 33, no. 3 (2014): 231-234. https://doi.org/10.5566/ias.1155
[63] Dataset, Retina. "Cataract Dataset." (2020). https://www.kaggle.com/jr2ngb/cataractdataset
[64] Porwal, Prasanna, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau. "Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research." *Data* 3, no. 3 (2018): 25. https://doi.org/10.3390/data3030025
[65] Takahashi, Hidenori, Hironobu Tampo, Yusuke Arai, Yuji Inoue, and Hidetoshi Kawashima. "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy." *PloS one* 12, no. 6 (2017): e0179790. https://doi.org/10.1371/journal.pone.0179790
[66] Almazroa, Ahmed, Sami Alodhayb, Essameldin Osman, Eslam Ramadan, Mohammed Hummadi, Mohammed Dlaim, Muhannad Alkatee, Kaamran Raahemifar, and Vasudevan Lakshminarayanan. "Retinal fundus images for glaucoma analysis: the RIGA dataset." In *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10579, pp. 55-62. SPIE, 2018. https://doi.org/10.1117/12.2293584
[67] Giancardo, Luca, Fabrice Meriaudeau, Thomas P. Karnowski, Yaqin Li, Seema Garg, Kenneth W. Tobin Jr, and Edward Chaum. "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets." *Medical image analysis* 16, no. 1 (2012): 216-226. https://doi.org/10.1016/j.media.2011.07.004

[68] Staal, Joes, Michael D. Abràmoff, Meindert Niemeijer, Max A. Viergever, and Bram Van Ginneken. "Ridge-based vessel segmentation in color images of the retina." *IEEE transactions on medical imaging* 23, no. 4 (2004): 501-509. https://doi.org/10.1109/TMI.2004.825627
[69] Sng, Chelvin C., Li-Lian Foo, Ching-Yu Cheng, John C. Allen Jr, Mingguang He, Gita Krishnaswamy, Monisha E. Nongpiur, David S. Friedman, Tien Y. Wong, and Tin Aung. "Determinants of anterior chamber depth: the Singapore Chinese Eye Study." *Ophthalmology* 119, no. 6 (2012): 1143-1150. https://doi.org/10.1016/j.ophtha.2012.01.011
[70] Decenciere, Etienne, Guy Cazuguel, Xiwei Zhang, Guillaume Thibault, J-C. Klein, Fernand Meyer, Beatriz Marcotegui *et al.,* "TeleOphta: Machine learning and image processing methods for teleophthalmology." *Irbm* 34, no. 2 (2013): 196-203. https://doi.org/10.1016/j.irbm.2013.01.010
[71] Budai, Attila, Rüdiger Bock, Andreas Maier, Joachim Hornegger, and Georg Michelson. "Robust vessel segmentation in fundus images." *International journal of biomedical imaging* 2013 (2013). https://doi.org/10.1155/2013/154860
[72] Hoover, A. D., Valentina Kouznetsova, and Michael Goldbaum. "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response." *IEEE Transactions on Medical imaging* 19, no. 3 (2000): 203-210. https://doi.org/10.1109/42.845178
[73] Owen, Christopher G., Alicja R. Rudnicka, Robert Mullen, Sarah A. Barman, Dorothy Monekosso, Peter H. Whincup, Jeffrey Ng, and Carl Paterson. "Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program." *Investigative ophthalmology & visual science* 50, no. 5 (2009): 2004-2010. https://doi.org/10.1167/iovs.08-3018
[74] Dataset, R. O. C. "ROC Dataset." (2020). http://roc.healthcare.uiowa.edu
[75] Manski, Charles F. "Bounding the accuracy of diagnostic tests, with application to COVID-19 antibody tests." *Epidemiology* 32, no. 2 (2021): 162-167. https://doi.org/10.1097/EDE.0000000000001309
[76] Özkaya, Umut, Şaban Öztürk, and Mucahid Barstugan. "Coronavirus (COVID-19) classification using deep features fusion and ranking technique." *Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach* (2020): 281-295. https://doi.org/10.1007/978-3-030-55258-9_17
[77] Amasyali, Kadir, and Nora M. El-Gohary. "A review of data-driven building energy consumption prediction studies." *Renewable and Sustainable Energy Reviews* 81 (2018): 1192-1205. https://doi.org/10.1016/j.rser.2017.04.095
[78] Setiadi, De Rosal Igantius Moses. "PSNR vs SSIM: imperceptibility quality assessment for image steganography." *Multimedia Tools and Applications* 80, no. 6 (2021): 8423-8444. https://doi.org/10.1007/s11042-020-10035-z
[79] Bartling, Herman, Peter Wanger, and Lene Martin. "Automated quality evaluation of digital fundus photographs." *Acta ophthalmologica* 87, no. 6 (2009): 643-647. https://doi.org/10.1111/j.1755-3768.2008.01321.x
[80] Cohen, Jacob. "Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit." *Psychological bulletin* 70, no. 4 (1968): 213. https://doi.org/10.1037/h0026256
[81] Islam, Md Shahidul, and Shamsuddin Ahmed. "Work Standardization in Lean Manufacturing for Improvement of Production Line Performance in SME." *Malaysian Journal on Composites Science and Manufacturing* 13, no. 1 (2024): 68-81. https://doi.org/10.37934/mjcsm.13.1.6881
[82] Sobran, Nur Maisarah Mohd, and Zool Hilmi Ismail. "A Systematic Literature Review of Unsupervised Fault Detection Approach for Complex Engineering System." *Journal of Advanced Research in Applied Mechanics* 103, no. 1 (2023): 43-60. https://doi.org/10.37934/aram.103.1.4360
[83] Aziz, Nur Farhana Abdul, Norsuzila Yaacob, Azita Laily Yusof, and Murizah Kassim. "A Review of Wildfire Studies Using Machine Learning Applications." *Journal of Advanced Research in Applied Mechanics* 114, no. 1 (2024): 13-32.
[84] Thaim, Ahmad Imran Mohd, Norazlianie Sazali, Kumaran Kadirgama, Ahmad Shahir Jamaludin, Faiz Mohd Turan, and Norhaida Ab Razak. "Smart Glove for Sign Language Translation." *Journal of Advanced Research in Applied Mechanics* 112, no. 1 (2024): 80-87. https://doi.org/10.37934/aram.112.1.8087
[85] Ghazalli, Shuwaibatul Aslamiah, Hazlina Selamat, Nurulaqilla Khamis, and Mohammad Fadzli Haniff. "Short Review on Palm Oil Fresh Fruit Bunches Ripeness and Classification Technique." *Journal of Advanced Research in Applied Mechanics* 106, no. 1 (2023): 37-47. https://doi.org/10.37934/aram.106.1.3747
[86] Arip, Afifuddin Arif Shihabuddin, Norazlianie Sazali, Kumaran Kadirgama, Ahmad Shahir Jamaludin, Faiz Mohd Turan, and Norhaida Ab Razak. "Object Detection for Safety Attire Using YOLO (You Only Look Once)." *Journal of Advanced Research in Applied Mechanics* 113, no. 1 (2024): 37-51. https://doi.org/10.37934/aram.113.1.3751
[87] Albelaihi, Arwa, and Dina M. Ibrahim. "DeepDiabetic: An Identification System of Diabetic Eye Diseases Using Deep Neural Networks." *IEEE Access* (2024). https://doi.org/10.1109/ACCESS.2024.3354854

[88] Albahli, Saleh, Nudrat Nida, Aun Irtaza, Muhammad Haroon Yousaf, and Muhammad Tariq Mahmood. "Melanoma lesion detection and segmentation using YOLOv4-DarkNet and active contour." *IEEE access* 8 (2020): 198403-198414. https://doi.org/10.1109/ACCESS.2020.3035345

[89] Hoe, Ho Kok, and Tamon Yotsuyanagi. "Development of Smart Mini Manufacturing System Model for Industry 4.0 Application." *Journal of Advanced Research in Applied Mechanics* 114, no. 1 (2024): 94-108.

[90] Aldayri, Amnah, and Waleed Albattah. "A deep learning approach for anomaly detection in large-scale Hajj crowds." *The Visual Computer* (2023): 1-15. https://doi.org/10.1007/s00371-023-03124-1

[91] Alboqomi, Athar Ibrahim, and Rehan Ullah Khan. "Sky Pixel Detection in Outdoor Urban Scenes: U-Net with Transfer Learning." *International Journal of Advanced Computer Science & Applications* 15, no. 2 (2024). https://doi.org/10.14569/IJACSA.2024.0150225

[92] Jamil, Amirah Hanani, Fitri Yakub, Azizul Azizan, Shairatul Akma Roslan, Sheikh Ahmad Zaki, and Syafiq Asyraff Ahmad. "A review on Deep Learning application for detection of archaeological structures." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 26, no. 1 (2022): 7-14. https://doi.org/10.37934/araset.26.1.714

[93] Zamani, Hadi, and Muhamad Kamal Mohammed Amin. "Classification of phishing websites using machine learning techniques." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 5, no. 2 (2016): 12-19.