IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 16 January 2024, accepted 30 January 2024, date of publication 5 February 2024, date of current version 12 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3361944

SURVEY

Diabetic Retinopathy Classification With Deep Learning via Fundus Images: A Short Survey

SHANSHAN ZHU[1,2,3], CHANGCHUN XIONG[4], QINGSHAN ZHONG[5], AND YUDONG YAO[1,2]

[1]Research Institute of Medical and Biological Engineering, Ningbo University, Ningbo 315211, China
[2]Health Science Center, Ningbo University, Ningbo 315211, China
[3]Key Laboratory of Optoelectronic Science and Technology for Medicine of Ministry of Education, Fujian Provincial Key Laboratory of Photonics Technology, Fujian Normal University, Fuzhou 350117, China
[4]Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China
[5]School of Materials Science and Chemical Engineering, Ningbo University, Ningbo 315211, China

Corresponding author: Shanshan Zhu (zhushanshan@nbu.edu.cn)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LQ23H180004, in part by the General Scientific Research Project of Zhejiang Education Department under Grant Y202146273, in part by the Program for the Introduction of High-End Foreign Experts under Grant GXL20200218001, and in part by the K. C. Wong Magna Fund in Ningbo University.

ABSTRACT Diabetic retinopathy (DR) is a microvascular disease that is associated with diabetes mellitus. DR can cause irreversible vision loss and low vision. DR classification, that is, early DR diagnosis and accurate DR grading, is critical for vision protection and immediate treatment. Deep learning-based automated systems led to significant expectations for DR classification based on fundus images with several advantages. In the past several years, many outstanding studies in this area have been conducted and several review articles have been published. However, the new trends and the future directions are need to furtherly analyzed. Thus, we carefully included and read 94 related articles published from 2018 to 2023 through Web of Science, PubMed, Scopus, and IEEE Xplore. From this review, we found that transfer learning has been used as an outstanding strategy for overcoming the issue of the limited data resources to support DR analysis. CNN models of ResNet and VGGNet with layers of tens or even hundreds are the most popular frameworks used for DR classification. The APTOS 2019 and EyePACS are the most widely used datasets for DR classification. In addition, some lightweight DL architectures like SqueezeNet and MobileNet have been proposed for DR classification tasks, especially for limited data resources and computational capabilities. Although deep learning has achieved or surpassed human-level accuracy in DR classification, there is still a long way to go in real clinical workflows. Further improvements in model interpretability, trustworthiness from ophthalmologists, cost-effective and reliable DR screening systems are needed.

INDEX TERMS Classification, diabetic retinopathy, deep learning, fundus images.

I. INTRODUCTION

Diabetic retinopathy (DR), one of the most feared microvascular complications of diabetes mellitus (DM), is a major cause of irreversible vision impairment and low vision among working adults. Approximately 30% of people with DM have signs of DR, of which 30% have vision-threatening DR [1]. According to the reports by the International Diabetes Federation [2], there are approximately 537 million diabetes in 2021 worldwide. This number will exceed 700 million by 2045 and nearly 30% of them, that is, more than 200 million people will suffer from DR. It is known that DR is a progressive disease with the resulting from long-term diabetes, the risk of incurable vision loss and low vision can be largely reduced by early DR diagnosis and accurate DR grading [3], [4], as shown in Figure 1.

A fundus image is a projection of the fundus captured by a monocular camera on a 2D plane. Unlike optical coherence tomography images and angiographs, fundus images can be acquired in a rapid, non-invasive and cost-effective

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

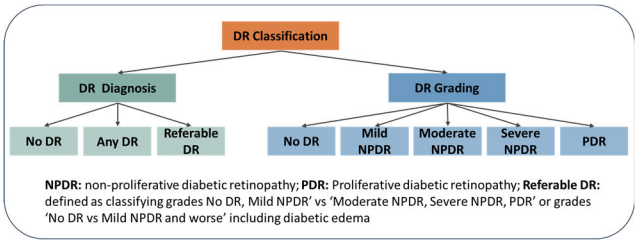

FIGURE 1. DR classification tasks using deep learning.

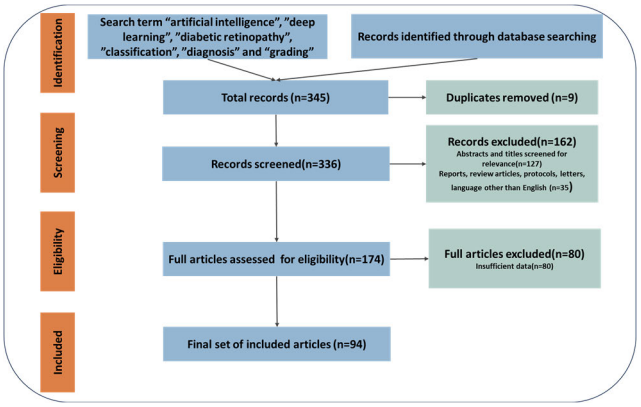

FIGURE 2. The flow diagram of the articles searching and selection.

way, making them more suitable for large-scale screening. Besides, many important biomarkers can be seen in the fundus image, such as optic disc, macula, fovea, blood vessel, and some DR related lesions. Traditionally, DR classification is mainly performed by analyzing lesion features in fundus images obtained from digital fundus photography. However, the interpretation of fundus images requires specialized knowledge and experienced ophthalmologists, and it is time-consuming, labor-intensive, and prone to human errors [5]. Thus, the increased global prevalence of DR and limited availability of professional ophthalmologists have motivated an urgent need to develop fast, cost-effective, and accurate automated systems to assist DR classification.

With the increase in computing power and availability of large amounts of labeled data, the deep learning (DL) technique shows excellent performance in automatic analysis and evaluation of image-related data through the combination of large amounts of data with intelligent algorithms [6]. DL is designed using a multilayer data representation architecture that can automatically extract low-level and high-level features without human interference [7]. Many DL-based algorithms (such as convolution neural networks (CNNs) [8], [9], autoencoders (AE) [10], [11], recurrent neural networks (RNNs) [12], [13], deep belief networks (DBN) [14], [15], and transfer learning [16], [17]) have been designed and applied to fundus images.

It is obvious that the number of papers on fundus images and DL for DR classification are increasing year by year. Recently, several review articles have been published in this topic. In 2019, Asiri et al. [18] published a review of DL techniques applied to DR detection and classification of lesions in fundus images. In 2020, Alyoubi et al. [19] reviewed and analyzed the significant research on DL for DR detection. In 2023, Sebastian et al. [20] presented a review of DL developments in the domain of DR classification including detection and grading based on fundus images. Although these comments cover a lot of work regarding DR lesion detection and classification, a detailed account of the preprocessing methods and the specific DL methods or structures in recent studies has not been included. Image preprocessing is necessary to reduce the heterogeneity resulting from various imaging conditions. Identifying the specific DL structures in recent research can provide new advances on this topic. Therefore, the objective of this paper is to provide a more comprehensive review that analyzes the new trends and highlights the future directions for the DR classification of deep learning in fundus images.

In contrast to previous works, the major contributions of this work be summarized as follows. First, a novel holistic overview is provided by presenting the detailed data preprocessing pipelines and more latest studies within the past three years in the field of DR classification using DL approaches. Second, the databases, DL models and the performance of reported techniques in two classification tasks, i.e., binary classification for DR diagnosis and multi-classification for DR grading are discussed. Third, the limitation and future evolution of the application of DL for DR classification is addressed. Thus, we believe that a more detailed and integrated review is more comprehensive to provide inspiring ideas for researchers in this active area.

The review adopts the PRISMA approach for articles searching and selection [21]. We searched for 345 related papers published from 2018 to 2023 through Web of Science, PubMed, Scopus, and IEEE Xplore using the terms ''artificial intelligence'', ''deep learning'', ''diabetic retinopathy'', ''classification'', ''detection'', and ''grading''. After removing and determining the specific DL tasks for DR, final 94 articles were carefully included. The flow diagram of the articles searching and selection is shown in Figure 2.

This paper is organized as follows. In Section II, a background of DR and related biomarkers on fundus images for DR classification are provided. In Section III, the publicly available DR datasets are described. In Section IV, data preprocessing methods and pipelines are introduced. In Section V, commonly used DL architectures for DR classification are discussed. In Section VI, recent research for DR classification by DL techniques is reviewed and discussed. In Section VII and Section VIII, some of challenges and potential future directions are provided. In Section IX, the conclusion is summarized.

II. DR AND RELATED BIOMARKES

In this section, an overview of DR and the related biomarkers on a fundus image as shown in Figure 3 are provided,

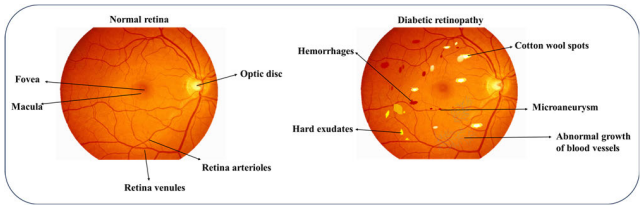

FIGURE 3. Normal retina and diabetic retinopathy on a fundus image.

including blood vessels, optic disc, DR lesions, stages and severity grading of DR.

A. BLOOD VESSELS AND OPTIC DISC

Blood vessel assessments, including assessments of vascular thickness, diameter, length, arteriovenous ratio, bifurcations, and curvature have been proven to be closely related to DR occurrence and its severity [22], [23], [24], [25], especially in the proliferative stage of DR (PDR). The growth of new tiny pathogenic blood vessels in PDR makes accurate segmentation and evaluation of the retinal vasculature much more important [26].

The optic disc (OD), that is, the optic nerve head (ONH), is the brightest region in fundus images, which is oval-shaped with clear boundaries and located exactly 3 mm nasally (medially) to the macula lutea [27]. The optic disc is also called a "low vision spot" of the eye, because it is the only area on the retina without any photoreceptors [27]. During fundus image analysis, the optic disc needs to be removed to avoid the misclassification of OD, because the OD has a similar appearance of pixels to the bright exudates in the fundus image [28].

B. DR LESIONS

In diabetes, large amounts of glucose in the blood can damage retinal blood vessels, resulting in vascular swelling, leakage, and even abnormal growth of new vessels [29]. Various pathological changes in the retina have been observed. Microaneurysms are the earliest symptoms of DR [30]. Possible reasons for microaneurysm formation include vasoproliferative factor release, capillary wall weakness, and increased intraluminal pressure [30]. Microaneurysms present slight widening of the capillary walls and are defined as deep red dots (25 - 100 μm) with sharp margins on the fundus images. Hemorrhages occur when the weak capillaries break. Hemorrhages are identified as red spots similar to microaneurysms. Unlike microaneurysms, hemorrhages are usually larger than 125 μm and have irregular edges [31]. Hard exudates are mainly lipoproteins that leak from the damaged capillaries. They often appear as small white or white-yellow individual dots or continuous flaky spots with sharp margins [32]. Soft exudates or cotton wool spots are lesions of the retinal nerve fiber layer caused by small-artery occlusion [33]. Small-artery occlusion reduces blood flow to the retina, which causes ischemia of the retinal nerve fiber layer and accumulation of axoplasmic debris in retinal

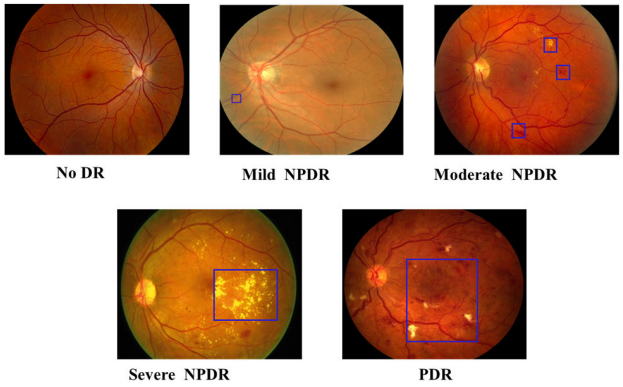

FIGURE 4. The DR grading according to ICDRSS (NPDR: Non-proliferative diabetic retinopathy; PDR: Proliferative diabetic retinopathy).

ganglion cell axons [34]. This accumulated debris appears slightly raised, with small gray-white cloud-like shapes in the superficial layer of the retina [33]. Neovascularization refers to the abnormal formation of new blood vessels on the retinal inner surface, which is a hallmark of PDR [35]. The release of vasoactive factors in ischemic retina can provoke the growth of new vessel, which may lead to vitreous hemorrhages that block vision [36]. Macular edema occurs when vascular permeability increases and abnormal blood vessels leak into the surroundings of the macula, which causes macular swelling and retinal thickening, and even threatens central vision [37].

C. DR GRADING

DR grading evaluates vascular changes and identifies DR severity levels. At present, there are various DR grading protocols, including the Early Treatment Diabetic Retinopathy Study (ETDRS) classification [31], International Clinical Diabetic Retinopathy Severity Scale (ICDRSS) [38], International Clinical Diabetic Macular Edema Severity Scale (ICDMESS) [38], and Scottish DR grading protocol [39]. Although the ETDRS grading scheme is the gold standard, its multiple levels and complex implementation make daily clinical and large-scale grading difficult. Owing to its convenience and ease of adoption, the ICDRSS proposed by the Global Diabetic Retinopathy Project Group has attracted much more attention in clinical practice and computer-aided diagnosis (CAD) settings worldwide. According to ICDRSS, DR can be classified into five severity levels [38], that is, No DR, mild NPDR, moderate NPDR, severe NPDR, and PDR as shown in Figure 4; the corresponding lesions at each level are briefly described in Table 1.

III. PUBLIC DATASETS

Datasets are collections of data that can be used to train and test DL models, which is one of the key reasons for the success of DL research. However, in the field of fundus images, high-quality, accurate labeling, and sufficient dataset collection are challenging. One reason is that privacy protection of personal data makes it difficult to acquire and share medical

TABLE 1. The severity level and corresponding lesions of DR according to ICDRSS protocol.

Severity level	Lesions on fundus image
No DR	No abnormalities
Mild NPDR	Microaneurysms only
Moderate NPDR	More than just microaneurysms but less than severe NPDR
Severe NPDR	Any of the following: (1) $>$20 intra-retinal hemorrhages in each of 4 quadrants; (2) definite venous beading in $\geqslant$2 quadrants; (3) prominent intra-retinal microvascular abnormalities in $\geqslant$1 quadrant, or no signs of proliferative retinopathy.
PDR	One or more of the following: neovascularization, vitreous or preretinal hemorrhages.

NPDR: Non-proliferative diabetic retinopathy; PDR: Proliferative diabetic retinopathy

data. In addition, different fundus imaging equipment and settings, imaging characteristics, and operators lead to poor consistency in the image quality and standards. In addition, image labeling is commonly performed by professional ophthalmologists, but labeling standards are significantly inconsistent among different labelers. These drawbacks can be avoided by using public datasets. Herein, we briefly review several publicly available datasets of fundus images for DR classification, as listed in Table 2. Much more detailed information could be obtained in the previous review articles of [40].

Among the above public datasets, MESSIDOR [41], Kaggle EyePACS [42], Kaggle APTOS [43] and MESSIDOR-2 [44] are the most popular for DR classification. The largest one is the Kaggle EyePACS dataset with 88702 fundus images graded according to the ICDRSS [42]. In this dataset, different cameras were used to collect images under various imaging conditions, leading to a good variety of real-world settings. Several fundus image datasets, including DIARETDB1 [45], HEI-MED [46], E-Ophtha [47], DRiDB [48], IDRID [49], and DDR [50], were created for both DR lesion detection and grading tasks. The DRiDB dataset contains 50 fundus images in which the optic disc, vessels, DR lesions, and grading are annotated by several professionals. Although it is small, it provides the largest amount of information and can be used for lesions detection and DR grading. The DDR dataset consists of 12522 images annotated by multiple experts according to ICDRSS. Although it is the second largest dataset for DR classification, there is a significant imbalance in different DR severities; thus, it has not yet been widely used. Because most of these datasets are small, many studies have used different or combinations of them to train and validate DL algorithms [51], [52], [53].

IV. DATA PREPROCESSING

In deep-learning research, preprocessing is a common first step in prepare formatted raw data that the network can accept, which can increase the training performance of DL models. Capturing fundus images using various cameras and settings results in different quality issues (such as size, noise, artifacts, contrast, illumination, and sharpened regions) within the image. These heterogeneities may hide some special details of the DR features during the training process of the DL models. Thus, in this section, we provide a brief review of the preprocessing pipelines for DR classification based on fundus images. The primary operations performed on the fundus images consist of image enhancement, denoising, normalization, and augmentation, as shown in Table 3.

A. IMAGE ENHANCEMENT

Inhomogeneities of contrast or luminosity usually appear in the same image or between different images. Image enhancement techniques, such as contrast enhancement, illumination correction, and color space-based enhancement can address the issues effectively. Contrast enhancement is normally applied to enhance foreground information (such as optic disc, blood vessels, and lesions) in fundus images from the background, so that it is more visible to segment and detect [54], [55], [56]. Histogram-based approaches include histogram equalization, adaptive histogram equalization (AHE) and contrast limited adaptive histogram equalization (CLAHE) [57]. Of these, CLAHE is found to be the most effective, which can not only improve the local contrast to obtain more image details, but also suppress the excessive amplification of noise in near-constant areas of the image. This technique is normally performed by combining color space transformations, such as transforming color images to grayscale, splitting and enhancing individual channels, or enhancing the color space directly. Zhang and Chutatape [58] used CLAHE for local contrast enhancement on sub-images by applying the local mean and standard deviation of intensities to detect hard and soft exudates in fundus images. Yamuna and Maheswari [59] applied CLAHE to gray scale retinal images to detect microaneurysms. Because blood vessels appear more contrast than the background in the green channels, Setiawan et al. [60] split the color fundus image into three individual channels, and performed CLAHE on the green channel, and then merged these three channels to improve the color retinal image quality. Chandni et al. [61] proposed a modified CLAHE using a Laplacian operator and HSV color model to obtain better information on the luminance of the color fundus image and preserve its edges. In addition to contrast enhancement, illumination correction is another way to enhance image contrast by reducing the illumination uniformity error of the retinal images. Gamma correction and logarithmic correction are effective and widely adopted methods for illumination correction [62], [63], [64], [65]. In recent years, DL-based methods have received increasing attention and have been developed for retinal image enhancement [66], [67], [68], [69]. Among these, generative adversarial networks (GAN) and their variations are the most promising. You et al. [66] developed a CycleGAN model for fundus image enhancement that can

TABLE 2. Commonly used public datasets for DR classification.

Year	Dataset (Available link)	Country	Size	Resolution	Annotations	Device used	Tasks
2004	MESSIDOR[41] http://www.adcis.net/en/third-party/messidor/	France	1200	1440×960 2240×1488 2304×1536	Image level	Topcon TRC NW6	DR grading
2007	DIARETDB1[45] https://www.it.lut.fi/project/imageret/	Finland	89	1500×1152	Pixel level	Zeiss FF450+	DR grading Lesion detection
2009	Kaggle EyePACS[42] https://www.kaggle.com/c/diabetic-retinopathy-detection/data	United States	88702	Varying	Image level	Centervue DRS, Optovue iCam, CanonCR1/DGi/CR2, Topcon NW	DR grading
2010	MESSIDOR-2[44] http://www.adcis.net/en/third-party/messidor2/	France	1748	1440×960, 2240×1488, 2304×1536	Image level	Topcon TRC NW6	DR grading
2012	HEI-MED[46] https://www.sciencedirect.com/science/article/pii/S1361841511001010	United States	169	Varying	Pixel level	Zeiss Visucam PRO fundus camera	DR grading Lesion detection
2013	E-Ophtha[47] http://www.adcis.net/en/third-party/e-ophtha/ http://www.adcis.net/en/third-party/e-ophtha/	France	463	2544×1696 1440×960	Pixel level	Canon CR DGI, Topcon TRC NW6	DR diagnosis Lesion detection
2013	DRiDB[48] https://ipg.fer.hr/ipg/resources/image_database#	Croatia	50	768×584	Pixel level	Zeiss VISUCAM 200 DFC	DR grading Lesion detection
2017	IDRID[49] https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid	India	516	4288×2848	Image and pixel level	Kowa VX - 10α fundus camera	DR grading Lesion detection
2019	Kaggle APTOS 2019[43] https://kaggle.com/competitions/aptos2019-blindness-detection	India	5590	Varying	Image level	Variety of common conventional cameras	DR grading
2019	DDR[50] https://github.com/nkicsl/DDR-dataset	China	13673	Varying	Image and Pixel level	Variety of common conventional cameras	DR grading Lesion detection

generate an output image for each input image without paired training samples. They furtherly used a convolutional block attention module (CBAM) [67] to improve the baseline of CycleGAN. Because CBAM can re-weight the high-level features extracted by the convolutional neural network in the channel attention module and spatial attention module, the color and texture of the generated image can be kept consistent with the input images. Quantitative and qualitative analyses prove that the enhanced results by Cycle-CBAM are superior to those by CycleGAN for DR classification tasks. Based on unpaired training settings, Zhao et al. [68] proposed a dynamic retinal image feature constraint to improve the generator in GAN and avoid over-enhancing the exceeding blurry regions, which can effectively reduce the artifacts in the generated images and achieve enhanced results.

B. DENOISING AND NORMALIZATION

Another main challenge in processing is that fundus images suffer from significant noise, such as Gaussian noise and salt and pepper noise. Filtering-based methods [69], [70] are widely used to address this problem. Compared to linear filtering methods (mean filtering, Gaussian filtering, and Wiener filtering), the median filter, as a non-linear filtering technique, is more robust and excellent for removing salt and pepper noise, as well as reducing edge blurring [71], [72]. In addition, to reduce the feature bias during the training process, image intensity normalization is usually conducted.

C. IMAGE AUGMENTATION

It is well known that deep learning models typically depend on high-volume training data to avoid overfitting. However, owing to the lack of rich labeled data and class imbalance, data augmentation is adopted to improve the data size and diversity, which can enhance the robustness and accuracy of the models. Traditional data augmentation approaches for fundus images include geometric and color transformations. Geometric transformation [73], [74], such as rotation, flipping, and cropping, aims to alter the geometry of the image while keeping the CNN invariant to changes in orientation and position. These techniques can be easily and

TABLE 3. Data preprocessing for DR classification.

Preprocessing approaches	Methods	Role
Image enhancement	Contrast enhancement: Histogram equalization (HE), Adaptive Histogram Equalization (AHE), Contrast Limited Adaptive Histogram Equalization (CLAHE)	To enhance the original fundus images' appearance and important information
	Illumination correction: Gamma correction, Logarithmic correction	
	Color space transformation: HSI Color space conversion, HSV color space conversion, Grayscale, Green channel	
	Generative Adversarial Network (GAN)-based methods: CycleGAN, Cycle-CBAM	
Denoising and Normalization	Denoising: Median filtering, Mean filter, Gaussian filtering, Wiener filter Normalization: Intensity normalization	To remove potential noise of images and avoid features biasness
Image augmentation	Geometric transformation: Rotating, shifting, rescaling, cropping, flipping	To increase the size of training data and avoid overfitting
	Color transformation: Brightness transformation, Contrast transformation, Color space transformation	
	GANs-based image generation	

directly performed based on basic image manipulations, but tend to suffer from the padding effect and lose some interesting areas. Another effective strategy, that is, performing augmentation in the color channel space, can generate new fundus images with the aim of keeping the CNN invariant to changes in color and lighting. The major advantage of color transformation is the removal of illumination bias of images; however, this method may lose important color information, and the drawbacks of increased memory, transformation costs, and training time need to be considered. Recently, GANs have become popular for data augmentation, which develop artificial samples with characteristics similar to those of the original dataset. Costa et al. [75] performed fundus image synthesis by combining training with an adversarial autoencoder and GAN. The former was used to generate the vascular tree and the latter was used to generate non-vascular features. Mahapatra and Bozorgtabar [76] proposed an image super resolution (ISR) method based on GANs in which each pixel's importance was defined as a loss in the cost function to generate a high-resolution super-resolved (SR) image. The experimental results showed that the SR images outperformed the methods without weighing pixels. In DR classification, alleviating the severe imbalance of fundus images in different classes is a major challenge. Zheng et al. [77] applied a conditional generative adversarial network (cGAN) to increase the amount of training data and generate synthetic images of label-preserving minority class data. The following results demonstrate that cGAN with a modified U-net (MU-Net) has better performance in exudate detection and generalization properties than cGAN using only MU-Net. Chen et al. [78] proposed dubbed retinal fundus image generative adversarial networks (RF-GANs), which consist of two generation models, that is, RF-GAN1 and RF-GAN2. The images in the source domain were translated into the target domain using RF-GAN1.The translated images were then used to train the semantic segmentation models to extract the structural and lesion masks. Finally, RF-GAN2 was applied to synthesize fundus images using the above masks and DR grading labels.

V. DEEP LEARNING METHODS

For DR classification, commonly used DL architectures include convolutional neural networks (CNNs), autoencoders (AEs), recurrent neural networks (RNNs), and deep belief networks (DBNs). In this section, an overview of these architectures is presented in Table 4.

A. CNN AND RELATED MODELS

CNN is the most popular algorithm for DR classification. A commonly used type of CNN includes convolutional, pooling, and fully connected (FC) layers. Several kernels in each convolutional layer are used to generate feature maps; each feature map is down-sampled in the pooling layers to reduce the network parameters to accelerate the training process and avoid overfitting. Feature maps after flattening are used as input of FC layers, and the final CNN output result is obtained

TABLE 4. Commonly used DL architectures in DR classification.

DL techniques	Architectures	Strengths	Weakness	Reference
CNN and related models	AlexNet	ReLU activation function; Overlap pooling; Data augmentation and Dropout to avoid overfitting;	Larger number of parameters (60 million); Be appropriate for handling basic and small data issues on average hardware; Basic and small data issues	[80-82]
	VGGNet	Same small-sized kernels to increase the depth of the network to improve the final performance	Larger number of parameters (138-180 million); High computational cost	[87, 92, 113, 114]
	GoogLeNet (Inception V1)	Inception model to extract features at different scales to increase the 1×1 convolutional kernels for dimensionality reduction reduces the computational complexity; Global average pooling layer instead of FC layers	Larger number of parameters (5 million); Heterogeneous topology between the inception blocks; Representational bottleneck	[80, 81, 92]
	Inception V3	Factorization using 1×n and n×1 convolutional kernels instead of n×n kernels to diminish representational bottleneck; RMSProp optimizer to accelerate the training		[16, 88-90, 115]
	Inception V4	Reduction block for pooling data		[116, 117]
	ResNet	Residual block with a shortcut connection to reduce parameters and accelerate training convergence	High computational cost and requires more powerful hardware support;	[91, 118, 119]
	Inception ResNet	Inception V3 or V4 combines with residual connection to improve computational efficiency	Accuracy is not significantly improved	[120, 121]
AE and related models	Autoencoder	Encoder and Decoder; Strong generalization; Without data annotation	Information loss in the dimensionality reduction; Slow training speed	[100]
	Stacked Autoencoder	Layer-wise unsupervised pre-training and supervised fine-tuning	Mainly be used for features extraction	[101, 122, 123, 124]
DBN and related models	GO-DBN-WKELM	GO algorithm optimizes the kernel parameters of the WKELM and increases the convergence speed	Slow convergence	[104]
	D-MBO-DBN	Number of hidden neurons in DBN is optimized Distance-based Monarch Butterfly Optimization (D-MBO)		[105]
	Two dimensional DBN	Receive multiple two-dimensional inputs to adjust the parameters of the network well and provide more information for learning		[106]
	SNE-O-DBN	A stochastic neighbor embedding feature extraction technique is to remove the noise and reduce the dimensionality		[107]
RNN and related models	RNN	Strong capability to capture contextual data from the sequence	Gradient vanishing; Long training time and high computational cost; Sensitive to hyperparameters	[121,125]
	RNN-LSTM	Introducing memory cells and gating mechanisms to reduce gradient vanishing	Longer training time and high complexity than RNN; Low interpretability	[111, 112]

by passing the outputs of the last FC layer through the activation function. The primary CNN architectures include AlexNet, VGGNet, GoogLeNet, and ResNet.

The initial CNN model of AlexNet only contains eight layers, but it achieves outstanding results in image classification tasks by applying techniques such as ReLU and Dropout [79], [80], [81], [82]. Subsequently, a deeper CNN model called VGGNet [83] is explored, which has two variants, that is, VGG16 and VGG19, according to the depth of the layers. Although this model contains deep weight layers, it adopts plentiful small-sized kernels in the convolutional layer instead of large ones to decrease the number of parameters and computational complication. Since then, small filters have been widely used in CNN architectures. In general, VGG obtains significant results for image classification; however, the extremely high computational cost remains a significant challenge. GoogLeNet (also called Inception V1) is established to address this issue [84]. GoogLeNet utilizes an inception module to enable the network to choose between multiple convolutional filter sizes for each block. In addition, a 1×1 convolutional filter is usually inserted before large kernels to decrease the channel of the feature maps. In GoogLeNet, a global average pooling layer instead of FC layers is used to neglect irrelevant channels, which can greatly decrease the density of connections and the number of parameters. The upgraded Inception V2, Inception V3, and Inception V4 present optimized structures and outperform each other. However, the main shortcoming of GoogLeNet is the heterogeneous topology between the inception blocks and the possibility of valuable information loss in the decreased feature space. With increasing depth, gradient diminishing has become a major issue for CNN. In 2015, He et al. [85] proposed ResNet (Residual Network) by introducing the concept of a bypass pathway to address this issue. ResNet includes a residual block through a shortcut connection inside layers to allow connections of cross-layers, which are parameter-free and data-independent. Furthermore, these shortcut connections can accelerate deep network convergence. Compared with VGG and GoogLeNet, ResNet has better performance in terms of computational speed and model accuracy. Many variants of ResNet with different layers (i.e., ResNet34, ResNet50, ResNet101, and ResNet152) and the variant combining inception and residual blocks, that is, Inception ResNet [86], have been proposed.

Shaban et al. [87] developed a modified version of the VGG19 for DR classification and staging. Five consecutive stages of convolutional layers were designed in which two convolutional layers were added to the first two stages, five convolutional layers were added to the middle two stages, and four convolutional layers were added to the last stage. The last fully connected layer was designed with a three-neuron layer for non-linear classification. Zhang et al. [88] designed an Inception V3 model for automated detection of severe DR based on fundus images. In the preprocessing stage, data augmentation and weighted random sampling were performed to balance the cases of the different classes in the training dataset. Because relevant lesion marks in fundus images usually vary in size, some studies [89], [90] have also applied inception modules to extract features at different resolutions. Cao et al. [91] used ResNet as the backbone network for DR severity classification. The attention module was applied in this architecture to improve the feature extraction and enhance the model performance. CNN models are data-hungry, but it is difficult to obtain large amounts of medical data. The transfer learning technique attempts to solve this issue, in which a CNN that has been well-trained with a large amount of data from a related domain is applied to a new task by fine-tuning it using a relatively small dataset from the target domain. Suedumrong et al. [92] created a new CNN architecture for DR classification using transfer learning from GoogLeNet and VGG16 models, in which the fully connected layer was cut off, and the dense class was added. The experimental results showed that the accuracy of the VGG16 model with fine-tuning outperformed that of GoogLeNet with fine-tuning.

B. AE AND RELATED MODELS

An autoencoder (AE) [93], [94] is a single-hidden layer network that consists of two parts: an encoding part and a decoding part. In the encoding part, a reduced dimensional feature representation is generated from the initial input, and in the decoding part, the initial input is reframed from the information provided by the encoder by minimizing the loss function. Normally, autoencoders include two subtypes: sparse autoencoder [95], [96] and denoising autoencoder [97]. Sparse autoencoders are typically used to extract sparse features from raw data for classification, in which more neurons are in the middle of the network than in the input layer, i.e., ''sparsity penalty'' to penalize neuron activations to fight off overfitting [96]. Denoising autoencoders usually add noise to the input to prevent the output from entering the input without learning features. A stacked Autoencoder (SAE) [98], [99] is a variation of autoencoders that is built with stacks of multiple AEs to form a deep structure. First, an SAE is trained layer-by-layer in an unsupervised manner, then the pre-trained model is fine-tuned using the backpropagation and gradient descent method.

A semi-supervised auto-encoder graph network (SAGN) was proposed for DR grading using limited labeled data [100]. Three major modules, that is, autoencoder feature learning, neighbor correlation mining, and graph representation, were included in this architecture. First, representations were extracted using an encoder-decoder architecture from a small number of labeled fundus images and large number of unlabeled images, and then these representations were reconstructed. Then, neighbor correlations between both the labeled and unlabeled images were explored according to their similarities obtained from the radial basis function. Subsequently, a Graph Convolutional Neural Network (GCN) was used to grade the fundus images according

to the extracted features and corresponding correlations. Atteia et al. [101] combined the power of feature extraction from a pre-trained CNN with feature selection and classification from a stacked autoencoder deep neural network (DFTSA-Net) for DME diagnosis. Four pre-trained networks including ResNet50, SqueezeNet, Inception V3, and GoogLeNet were used to extract features from a small input dataset. Subsequently, a stacked autoencoder neural network was used to select the most informative features through unsupervised learning and trained for classification.

C. DBN AND RELATED MODELS

A deep belief network (DBN) [102] is a generative deep architecture stacked by units of restricted Boltzmann machines (RBMs). An RBM is a probabilistic model that can generate a probability distribution from its input datasets by maximizing the similarity between the input and its projection [103]. Similar to SAEs, the training process of DBNs includes two steps, that is, layer-by-layer pre-trained in an unsupervised manner by the greedy method, and weight fine-tuning using backpropagation algorithms and gradient descent. DBNs have recently demonstrate impressive performance in a broad range of applications, including feature extraction and classification tasks.

A Gannet-optimized DBN-based wavelet kernel extreme learning machine (GO-DBN-WKELM) technique [104] was proposed for DR detection and grading. The DBN was used to extract the reduced-dimensional features from the original datasets, and the extracted images were analyzed using the proposed GO-DBN-WKELM classification model. Because the GO algorithm can optimize the DBM and KELM kernel parameters, the convergence and accuracy of the classifier can be enhanced. A new DR detection model was proposed by Basha and Ramanaiah [105]. Four phases were included in this model: preprocessing, blood vessel segmentation, feature extraction, and classification. For classification purposes, the DBN classifier was used, in which the hidden neurons were optimized using a modified MBO (Monarch butterfly optimization) called Distance-based MBO (D-MBO) algorithm. The introduced D-MBO algorithm divided the subpopulation according to the distance between the current solution and best solution, which can offer improved convergence to increase the accuracy rate. Although abundant types of inputs can provide more useful prior information for learning, standard DBN and its traditional variations, such as multi-resolution DBN, Gaussian DBN, DBN-Softmax and convolutional DBN (CDBN), failed to process such complex types of inputs in parallel. Tehrani et al. [106] presented a two-dimensional DBN based on a Mixed-restricted Boltzmann Machine for DR diagnosis and severity evaluation. This model can receive multiple two-dimensional inputs to adjust the parameters of the network well and provide more information for learning. To remove noise and decrease dimensionality, Rajavel et al. [107] proposed a system with a stochastic neighbor embedding (SNE) feature extraction module followed by an optimized deep belief network (O-DBN) classifier to identify DR severity levels. The SNE-O-DBN approach demonstrated superior performance compared with existing online systems.

D. RNN AND RELATED MODELS

Recurrent neural networks (RNNs) [108] are a type of feed-forward neural network trained to generate an output by combining the current input with the information of the previous iterations. An RNN model usually contains numerous successive recurrent layers that are sequentially modeled to map a sequence with others. Owing to the addition of loops, RNNs have a strong capability to capture contextual data from the sequence, and these data are effectively used in classification process. However, RNN suffers from the problem of a vanishing gradient, which may hinder the learning curve of the model, especially in long sequences. Long short-term memory (LSTM) model [109], [110] is a popular RNNs variation for tackling this problem by introducing memory cells and gating mechanisms. These gates enable cells to record information over any time interval by controlling the flow of information into and out of the cell. After jointly connecting the previous state, available memory, and current input, the LSTM network can selectively activate and update the cells, and each recurrent unit can adaptively capture information at different time scales. LSTM has proven to be successful in capturing order dependencies in non-linear sequence classification problems.

Özçelik and Altan [111] proposed a model that combined chaotic swarm intelligence optimization with recurrent long short-term memory, which can avoid non-linear dynamics, especially chaoticity in fundus images when determining the severity of DR. There were four stages in the proposed model. First, two-dimensional stationary wavelet transform (2D-SWT) was used to reveal the characteristic features of the images. Then, the statistical- and entropy-based feature functions were applied to different matrices of the 2D-SWT to extract 96 features. Subsequently, a chaotic-based wrapper approach was used to select the features that maintained high classification performance. At last, the recurrent neural network-long short-term memory (RNN-LSTM) architecture was created by selecting optimum feature vectors to achieve the high classification performance. In addition, Spoorthi and Rekha [112] designed a deep learning approach by combining a DCNN and RNN-LSTM to classify DR into four stages, in which a DCNN was used as the feature extractor and LSTM was used as the classifier. The output from the max-pooling layer in the DCNN was passed to the next LSTM layer.

VI. DR CLASSIFICATION

The main aims of DR classification can be divided into binary classification for DR diagnosis and multi-class classification for DR grading. In this section, we provide a short review of recent literature on published DL models used in DR classification tasks. And three standard metrics, i.e., accuracy,

TABLE 5. A summary of the major DL models for binary DR classification.

Target class	Dataset (Total Size)	DL model	Best Performance			Year	Reference
			Accuracy	Sensitivity	Specificity		
Any DR/No DR	EyePACS (35126)	VGG16	-	93%	85%	2018	Seth *et al*. [126]
RDR / Non-RDR	EyePACS (60000)	WP (weighted paths)-CNN	94.23%	90.94%	95.74%	2019	Liu *et al*. [127]
Any DR/ No DR; RDR/Non- RDR; Sight threatening DR (STDR)/ Non-STDR	MESSIDOR-1 (1533)	CNN	-	90.4%; 94.7%; 91.7%	91.0%; 97.4%; 93.0%	2020	Shah *et al*. [128]
Clinically Significant Macular Edema (CSME)/ Non-CSME	MESSIDOR, UoA-DR, and IDRiD (1903)	Pre-trained ResNet50, AlexNet, GoogLeNet, Inception V3, Inception-ResNet-v2, ResNet18, ResNet101, VGG16, VGG19, DenseNet, and SqueezeNet	91.6%	80.9%	95.2%	2020	Chalakkal *et al*. [129]
Any DR/ No DR	MESSIDOR-1, and APTOS 2019 (6679)	Pre-trained Inception-ResNet-v2	82.18%	-	-	2020	Gangwar *et al.* [130]
Severe DR /Non-severe DR	APTOS 2019 (4192)	Inception V3	-	92.5%	90.7%	2022	Zhang *et al*. [88]
Any DR/ No DR	DDR (12519)	Pre-trained ResNet18	-	86.53%	86.72%	2022	Feng *et al*. [131]
Any DR/ No DR	APTOS 2019 (3600)	CNN	94.6%	86%	96%	2022	Padmanayana *et al*. [132]
Any DR/ No DR	EyePACS and APTOS 2019 (52924)	Local binary convolutional neural network (LBCNN) based on ResNet18	97.4%	94.6%	96.6%	2022	Macsik *et al*. [133]
Any DR/ No DR	APTOS 2019 (3662)	DenseNet169 and Convolutional Block Attention Module	97%	97%	98.3%	2022	Farag *et al*. [134]

sensitivity, and specificity are used to evaluate the performance of DL models in these studies.

A. BINARY CLASSIFICATION FOR DIAGNOSIS

Binary classification is primarily used to differentiate between healthy and diseased individuals. Any DR is defined as the presence of NPDR, PDR, DME, or a combination thereof. Most models have been trained for binary classification tasks, such as the presence or absence of any DR, referable DR (RDR) or non-referable cases, and severe DR or non-severe DR. Table 5 summarizes the major DL models for binary DR classification.

Many significant studies present from 2018 to 2020 [126], [127], [128], [129], [130]. In [126], VGG16 was proposed for DR identification using the loss function of binary cross entropy. They compared two experiments, that is, the VGG16 combined with a linear SVM and VGG16 combined with a softmax function as an output fully-connected layer, and a higher sensitivity of 93% and specificity of 85% were achieved by the former. Using the same EyePACS dataset, Liu et al. [127] designed a CNN model by applying multiple weighted paths into a convolutional neural network, named WP-CNN (weighted paths-CNN), for referable and non-referable DR identification. This binary classification experiment finally obtained an accuracy of 91.05%, sensitivity of 89.3%, and specificity of 90.89%. Chalakkal et al. [129] developed a simplified approach for screening clinically significant macular edema (CSME) using a combination of pre-trained DCNNs and a meta-heuristic feature selection approach. The results indicate that Inception-ResNet-v2 yielded the best performance. Gangwar and Ravi [130] used transfer learning on pre-trained Inception-ResNet-v2, and a

custom block of CNN layers was added on top to design a hybrid model to detect DR. The performance of this model was evaluated on the MESSIDOR-1 and ATOS 2019 datasets, and a high accuracy of 82.18% was achieved for the latter one.

In 2022, a large number of related research in DR classification using DL method has emerged [88], [131], [132], [133], [134]. In [88], Inception V3 was adopted to identify severe DR and non-severe cases based on the recognition of DR lesions. Owing to the large imbalance between cases of severe DR and non-severe DR, a weighted random sampling strategy was used to balance the positive and negative cases in the training set. The Kaggle public dataset for DR grading was used, and a sensitivity of 92.5% and specificity of 90.7% were achieved. Padmanayana and Anoop [132] designed a CNN architecture to classify images of DR or non-DR, and different optimizers, such as Adagrad, Adam, and RMSPROP with momentum, were used to compare the performance of the model. Testing on the APTOS 2019 dataset, the highest accuracy of 94.6%, the sensitivity of 86%, and the specificity of 96% were respectively obtained. An additional private hospital dataset was also used for testing, an accuracy of 94.6%, sensitivity of 88% and specificity of 97% were obtained. Macsik et al. [133] proposed a new local binary convolutional neural network (LBCNN) deterministic filter generation approach, in which fewer learnable parameters and less memory were useful for enhancing the performance of the standard CNN. These experiments were evaluated on the EyePACS and APTOS datasets, and the performance of the LBCNN with 24 fixed filters outperformed that of all the other DL models in this study. Farag et al. [134] proposed a model by combing DenseNet169's encoder to construct a visual embedding and convolutional block attention module (CBAM) to enhance its discriminative power. They applied their algorithm to APTOS dataset for binary and multi-class classification tasks. On the binary classification for DR or No DR, accuracy of 97%, sensitivity of 97% and specificity of 98.3% were achieved. Moreover, this network also showed high accuracy of 82% for severity grading.

B. MULTI-CLASS CLASSIFICATION FOR GRADING

Multi-class classification is commonly defined as assigning a fundus image to different disease stages according to the most severe grade in both eyes of each patient. As described in Section II, DR is graded according to a five-level protocol: no DR, mild NPDR, moderate NPDR, severe NPDR, and PDR. In this section, we present research regarding published DL models for multi-class classification of DR. Table 6 summarizes the major DL models for multi-class DR classification.

In 2018, Wan et al. [135] adopted transfer learning and hyper parameter tuning on the pre-trained models of AlexNet, VGGNet, GoogleNet, and ResNet for DR classification. The models were fine-tuned using the EyePACS dataset. The best results, with an accuracy of 95.68%, sensitivity of 86.47%, and specificity of 97.43% were obtained from the pre-trained VGGNet.

In 2019, Hagos and Kant [89] used a small subset of the EyePACS dataset to train Inception V3 for 5-class classification. The inception modules in this model enabled different-sized feature extraction from the input images in one of the convolution layers. A high accuracy of 90.9% was achieved. Qummar et al. [117] trained an ensemble architecture of five deep CNN models (ResNet50, Inception V3, Xception, Dense121, and Dense169) to classify the DR stages. This ensemble architecture can encode rich features to improve classification performance. The experimental results show that the proposed model can effectively identify five stages of DR, and its performance outperforms that of other common models trained on the same Kaggle dataset. Zhang et al. [120] built an automatic grading system to evaluate the DR severity using two strategies. The first strategy consisted of two phases. In the first phase, a binary classification was performed to identify abnormal and normal images via Xception, and in the second phase, a ternary classification was used to evaluate DR severity based on the above abnormal images using ResNet50. The other strategy was a four-class model based on all fundus images using ResNet50 and DenseNet. A high sensitivity of 98.1% and specificity of 98.9% were achieved by the DenseNet. Harangi et al. [136] graded DR stages by combining hand-crafted features with AlexNet. The Kaggle dataset was used for training and the IDRiD dataset was used for testing. Classification accuracies of 90.07% and 96.85% were achieved for the 5-class DR and the 3-class DME tasks, respectively.

In 2020, Tymchenko et al. [137] applied a multistage approach to transfer learning to detect the stages of DR. They built a deep ensemble CNN architecture by combining 3 CNN architectures (EfficientNet-B4, EfficientNet-B5, and SE-ResNeXt50) and transfer learning. To train the encoder based on a small amount of training data, ImageNet-pre-trained CNNs were used for the initialization. The training process was performed on the APTOS 2019, IDRiD, and MESSIDOR datasets and the model was tested on the EyePACS dataset with a sensitivity of 99.3% and specificity of 99.3% too. Mishra et al. [138] used a pre-trained DenseNet with quadratic weighted kappa (QWK) on the APTOS 2019 dataset to automatically detect the DR stage and finally got an accuracy of 96.1%. Comparing the accuracy of 73.26% from VGG16 architectures trained without QWK and ImageNet, the QWK can significantly enhance the accuracy of DenseNet architecture. Tu et al. [139] proposed a feature separation and union network (SUNet) for simultaneous DR and DME grading. SUNet contained a feature-blending block with two parts: feature separation and feature union. In the feature-separation part, task-specific features for lesion detection and DR/DME grading can be learned, whereas in the feature-union part, these features can be aggregated. Thus, irrelevant features can be extracted and used for related tasks to improve the performance of each task. Experiments on the IDRiD dataset demonstrate that SUNet significantly outperformed some existing models, such as VGG19 and ResNet34, for both DR and DME grading.

TABLE 6. A summary of the major DL models for multi-class DR classification.

Target class	Dataset (Total Size)	DL model	Best Performance			Year	Reference
			Accuracy	Sensitivity	Specificity		
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS (36126)	Pre-trained VGGNet AlexNet, VGGNet, GoogleNet, and ResNet	95.68%	86.47%	97.43%	2018	Wan *et al.* [135]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS (2500)	Pre-trained Inception V3	90.9%	-	-	2019	Hagos *et al.* [89]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS (35126)	An ensemble architecture of ResNet50, Inception V3, Xception, Dense121, and Dense169	-	99%	-	2019	Kummar *et al.* [117]
4-class: Normal NPDR, NPDR to PDR, PDR	Custom-developed by Sichuan Academy of Medical Sciences and Sichuan Provincial Peoples Hospital (3833)	ResNet50, Xception, and DenseNet	-	98.1%	98.9%	2019	Zhang *et al.* [120]
5 classes of DR and 3 classes of risk of DME	APTOS2019 and IDRiD (23216)	AlexNet	90.07% 96.85%	-	-	2019	Harangi *et al.* [136]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS, APTOS 2019, IDRiD and MESSIDOR (36739)	EfficientNet-B4, EfficientNet-B5, SE-ResNeXt50, transfer learning	-	99.3%	99.3%	2020	Tymchenko *et al.* [137]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (5590)	Pre-trained DenseNet	96.1%	-	-	2020	Mishra *et al.* [138]
5 classes of DR and 3 classes of risk of DME	IDRiD (516)	SUNet	65.05% 81.55%	-	-	2020	Tu *et al.* [139]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (3662)	Pre-trained VGG19	95.4%	-	-	2021	Islam *et al.* [140]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (3662)	EfficientNet-B3, transfer learning	98%	-	-	2021	Sugeno *et al.* [141]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACs and APTOS 2019 (1440)	Pre-trained AlexNet, GoogleNet, Inception V4, Inception-ResNet -v2, and ResNeXt50	97.53%	-	-	2021	Tariq *et al.* [142]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	DDR and APTOS 2019 (4419)	CNN512	88.6%	-	-	2021	Alyoubi *et al.* [143]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS (35126)	ResNet, Attention Mechanism	91.3%	-	-	2022	Cao *et al.* [91]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	EyePACS (88702)	Pre-trained GoogLeNet, pre-trained VGG16	71.65%	-	-	2022	Suedumrong *et al.* [92]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (3662)	DenseNet169 and Convolutional Block Attention Module	82%	-	-	2022	Farag *et al.* [134]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	IDRiD (516)	Pre-trained ResNet50, VGG16, and VGG19	82.5%	-	-	2022	Albahli *et al.* [144]

TABLE 6. *(Continued.)* **A summary of the major DL models for multi-class DR classification.**

4-class: Normal, mild NPDR, moderate NPDR, severe NPDR	EyePACS-1, MESSIDOR-2, and DIARETDB0 (11841)	Pre-trained VGG16	96.60%	-	-	2022	Bilal *et al.* [145]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019, MESSIDOR-2 (5406)	Supervised contrastive learning with Xception Encoder	84.36%	-	-	2022	Islam *et al.* [146]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (3662)	Ensemble method based on pre-trained VGG16 VGG19	87%	-	-	2022	Kale *et al.* [147]
4-class: Normal, mild NPDR, moderate NPDR, severe NPDR	MESSIDOR, STARE and DRIVE (1570)	Deep LSTM-RFO algorithm	97.89%	98.47%	97.43%	2022	Priya *et al.* [148]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	Kaggle DR 2015 and APTOS 2019 (38646)	Parallel convolutional neural network (PCNN)	97.27 %	-	-	2023	Nahiduzzaman *et al.* [149]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	Kaggle (35126)	Revised ResNet50	74.32%	-	-	2023	Lin *et al.* [150]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 (3545)	Triple-DRNet	92.08%	-	-	2023	Jian *et al.* [151]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 and DDR (16184)	DenseNet-121-rendered model	98.36%	-	-	2023	Alwakid *et al.* [152]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	IDRiD (516)	SqueezeNet and DCNN	91.1%	89.8%	91.3%	2023	Beevi *et al.* [153]
5-class: Normal, mild NPDR, moderate NPDR, severe NPDR, PDR	APTOS 2019 and EyePACS (40690)	MobileNet V3-Small	98.4%	-	-	2023	Wahab Sait *et al.* [154]

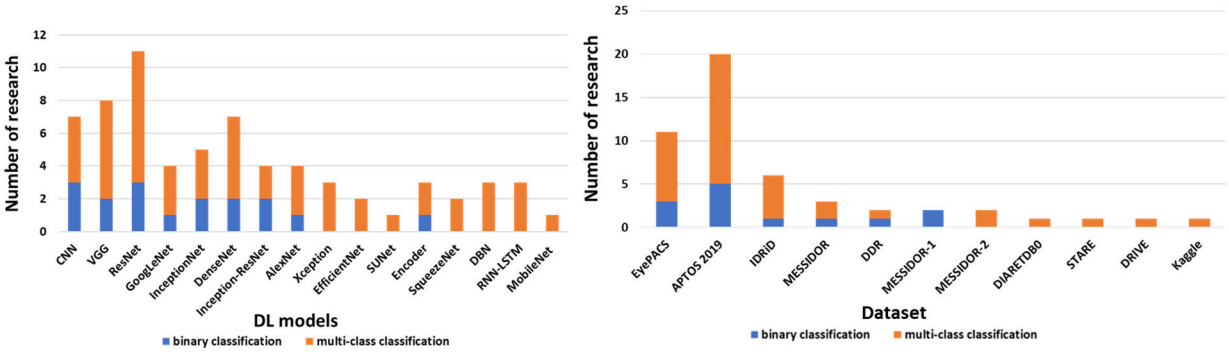

FIGURE 5. General DL models and datasets used in DR classification between 2018-2023.

In 2021, Islam et al. [140] developed a customized CNN model based on a pre-trained VGG19 combined with a channel-wise attention-like module to detect DR severity. The features extracted from VGG19 were passed into this module to obtain detailed semantic features of different DR severities. The fundus images were preprocessed and down-sampled to overcome the problem of dataset imbalance before feature extraction and classification. The results show that proposed model achieved the best evaluation accuracy of 95.4%. Sugeno et al. [141] developed a DR severity grading system using a pre-trained EfficientNet-B3 on the APTOS 2019 dataset. In this model, the blurred and duplicated images were removed from the input dataset by a numerical threshold, which made a classification accuracy of 98% for severity grading. Tariq et al. [142] applied transfer learning and AlexNet, GoogleNet, Inception V4, Inception-ResNet-v2, and ResNeXt50 for a five-level classification of DR. Because ResNeXt50 adopted a squeeze and excitation (SE) block for

each non-identity branch of a residual block, the best classification accuracy of 97.53% was achieved by the pre-trained Se-ResNeXt50. Alyoubi et al. [143] built a deep CNN512 model for DR grading, in which the entire image was used as an input to identify its stages. The accuracies of 88.6% and 84.1% were achieved for the DDR and APTOS 2019 datasets, respectively.

In 2022, Cao et al. [91] proposed a ResNet-based network-workable scheme for DR classification. In this model, the feature information of the hidden layer can be enhanced by modifying the structure of the residual blocks and by adding an attention mechanism. A higher accuracy of 91.3% was achieved, which outperformed those of the original models. Suedumrong et al. [92] used transfer learning combined with GoogLeNet and VGG16 to identify five stages of DR. The experimental results show that VGG16 with fine-tuning can achieve a higher accuracy of 71.65% than that of GoogLeNet. Albahli and Yar [144] adopted three pre-trained models, that is, ResNet50, VGG16, and VGG19, to identify both DR severity and the risk of ME. The best accuracy of 82.5% was achieved by ResNet50 performed on original images. Bilal et al. [145] designed a novel two-stage framework for DR grading based on lesions with no signs (normal), microaneurysms (mild), hemorrhages (moderate), and exudates (severe). First, the optic disc and blood vessels were segmented and extracted using two distinct U-Net models. The preprocessed fundus images were then used as inputs for the pre-trained VGGNet. The proposed model was evaluated on the EyePACS-1, MESSIDOR-2, and DIARETDB0 datasets, and the corresponding accuracies were 96.60%, 93.95%, and 92.25%, respectively. Islam et al. [146] proposed a supervised contrastive learning (SCL) method, that is, a two-stage training method with a supervised contrastive loss function to identify DR and its severity. The APTOS 2019 and MESSIDOR-2 datasets were used to validate the performance of the model. Finally, a better accuracy of 84.36% for grading five DR stages on the APTOS 2019 dataset was achieved. Kale and Sharma [147] designed an ensemble model by stacking two convolutional models that were trained on a DR dataset. The first model was trained using transfer learning on VGG19 and the second model was trained using transfer learning on VGG16. This ensemble model demonstrated better results, with accuracy of 87% in identifying DR stages. Priya et al. [148] proposed a deep long short-term memory (LSTM) in a neural network with Red Fox optimization (deep LSTM-RFO) algorithm for classifying normal, mild, moderate, and severe NPDR stages. The proposed method was tested on three datasets such as MESSIDOR, STARE, and DRIVE. The MESSIDOR dataset showed 97.59% accuracy, 97.61% sensitivity, and 97.03% specificity. Likewise, the STARE provided 97.89% accuracy, 98.47% sensitivity, and 97.43% specificity. Finally, 94.03% accuracy, 94.26% sensitivity and 94.74% specificity were obtained by using the DRIVE dataset. In addition, this approach showed less execution time and lower computational cost than some existing methods.

In 2023, Nahiduzzaman et al. [149] proposed a parallel CNN (PCNN) for feature extraction, in which the time required to extract distinctive features can be reduced using fewer parameters and layers. These features were used for DR severity classification using the extreme learning machine (ELM) technique. The model was validated on the Kaggle DR 2015 and APTOS 2019 datasets with accuracies of 91.78% and 97.27%, respectively. Lin and Wu [150] compared the performance of the revised ResNet50, Xception, AlexNet, VGGNet, VGG16, and ResNet50 for DR grading. In the revised ResNet50, the standard operation procedure (SOP) was used for fundus image processing, and an adaptive learning rating was adopted to adjust the weight of the layers. These techniques effectively enhanced the accuracy of the revised ResNet50. Jian et al. [151] proposed the triple-cascade network model (Triple-DRNet) for DR grading. This model consisted of three individual subnetworks: DR-Net, PDR-Net, and NPDR-Net. The first network was designed to distinguish between DR and No DR, the second network was designed to distinguish between PDR and NPDR, and the third network was designed to distinguish mild, moderate, and severe NPDR. Finally, Triple-DRNet was evaluated on the APTOS 2019 dataset, and its accuracy reached 92.08%. Alwakid et al. [152] used CLAHE and the enhanced super resolution generative adversarial networks (ESRGAN) to generate high-quality images for the APTOS and DDR datasets. And a pre-trained DensNet-121 was used for DR stages classification. Finally, an impressive 98.7% accuracy for APTOS dataset and 79.6% accuracy for DDR dataset were achieved when using these preprocessing approaches. Beevi [153] used SqueezeNet and DCNN to provide a two-level classification strategies of DR. First, Squeezenet was used to classified the fundus image into the normal or abnormal class of DR, in which Fractional War Strategy Optimization (FrWSO), i.e., War Strategy Optimization (WSO) combined with Fractional Calculus (FC) tuned the SqueezeNet. Second, DCNN was used to determine the severity level for the abnormal images, in which the Fractional War Royale Optimization (FrWRO) algorithm by combing Battle Royale Optimization (BRO) with FrWSO can adjust the weight of the DCNN. The results show that the accuracy, sensitivity, and specificity of the proposed approach were 91.1%, 89.8%, and 91.3% for categorizing the severity level of DR. Due to dataset imbalances or limited computational resources in some DR detection system, Wahab Sait [154] developed a lightweight deep-learning (DL)-based DR-severity grading system named MobileNet V3-Small model. You Only Look Once (Yolo) V7 technique was used for feature extraction process and a tailored quantum marine predator algorithm (QMPA) was used for selecting appropriate features. The proposed model was evaluated on the APTOS and EyePacs datasets. The outcomes reveal that the proposed model achieved an accuracy

of 98% and 98.4% in the APTOS and EyePacs datasets, respectively.

VII. CHALLENGES

Despite striking advances, many challenges remain for deep learning in DR classifications. First, lack of high-quality labeled data. Currently, the data currently used for DR classification are mostly public datasets and most of the data only include the target disease or only one single race, which means that it cannot reflect the real clinical status and obvious ethnic differences among people. Insufficient data or imbalances between different classes are common problems in some public datasets. Consistency of labeling and grading in clinical practice is difficult to achieve. Second, interpretation of DL models and clinical understanding is difficult. As DL models perform recognition and classification based on the features of images extracted by multi-layer non-linear structures, DL systems are often criticized for being non-transparent. The detailed mechanisms within the architecture and their clinical significance are not well understood. Third, DL applications in clinical practice have ethical limitations. It is unclear who is responsible for the erroneous diagnostic results when using DL methods.

VIII. FUTURES DIRECTIONS

In this section, we provide some directions for future research. First, building a more sophisticated dataset is one of the main directions. Enhancing the variety of fundus images regarding image capture devices, ethnic groups, and multiple categorical retinal diseases as well as the balance between different classes during data collection can improve the generalization ability of DL models. Ophthalmologists need to work together to develop a disease specific consensus and subsequently provide a comprehensive standard for labeling fundus image in diagnosis and grading. Second, the successful adoption of deep learning in DR classification requires medical professionals to understand its underlying principles and techniques. Thus, closer collaboration between AI experts and ophthalmologists could make DL techniques more transparent and easier to understand. Third, there is also a trend to develop cost-effective DR detection systems by designing more lightweight architectures and using micro portable devices, such as Raspberry Pi or smartphones, to assist large-scale DR screening at primary health centers with a lower cost.

IX. CONCLUSION

In this study, we provide a comprehensive review of recent advances in deep learning-based research on DR classification based on fundus images. Some key findings can be obtained. First, there is an obvious trend that transfer learning is an outstanding strategy for overcoming the issue of the limited data samples available during model training. With the help of transfer learning techniques, a number of pre-trained networks are accessible to support DR analysis. Both the training time and robustness of the model can be improved by training with parameters from the pre-trained model. Second, CNN models of ResNet and VGGNet are the most popular frameworks used for DR classification. The depth of the ResNet- and VGGNet-based networks can reach tens or even hundreds of layers, which can provide outstanding classification results. The APTOS 2019 and EyePACS are the most widely used datasets for DR classification. Third, some lightweight DL architectures like SqueezeNet and MobileNet have been proposed for DR classification tasks, especially for limited data resources and computational capabilities. These architectures can greatly reduce parameters while ensure model accuracy in the complicated image analysis. Although deep learning has achieved or surpassed human-level accuracy in DR diagnosis and grading, there is still a long way to go in real clinical workflows. Further improvements in model interpretability, trustworthiness from ophthalmologists, and cost-effective and reliable DR screening systems are needed.

REFERENCES

[1] D. S. W. Ting, G. C. M. Cheung, and T. Y. Wong, "Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review," *Clin. Experim. Ophthalmol.*, vol. 44, no. 4, pp. 260–277, May 2016.

[2] *International Diabetes Federation*, International Diabetes Federation Diabetes Atlas, Brussels, Belgium. Accessed: Dec. 2021. [Online]. Available: https://www.diabetesatlas.org/en/

[3] H. Safi, S. Safi, A. Hafezi-Moghadam, and H. Ahmadieh, "Early detection of diabetic retinopathy," *Surv. Ophthalmol.*, vol. 63, no. 5, pp. 601–608, Sep. 2018.

[4] L. Hill and L. E. Makaroff, "Early detection and timely treatment can prevent or delay diabetic retinopathy," *Diabetes Res. Clin. Pract.*, vol. 120, pp. 241–243, Oct. 2016.

[5] S. Vujosevic, S. J. Aldington, P. Silva, C. Hernández, P. Scanlon, T. Peto, and R. Simó, "Screening for diabetic retinopathy: New perspectives and challenges," *Lancet Diabetes Endocrinol.*, vol. 8, no. 4, pp. 337–347, Apr. 2020.

[6] Y. Xu et al., "Artificial intelligence: A powerful paradigm for scientific research," *Innovation*, vol. 2, no. 4, Nov. 2021, Art. no. 100179.

[7] L. Alzubaidi, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, Mar. 2021.

[8] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Proc. Comput. Sci.*, vol. 90, pp. 200–205, Jan. 2016.

[9] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017.

[10] D. Maji, A. Santara, S. Ghosh, D. Sheet, and P. Mitra, "Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 3029–3032.

[11] A. G. Roy and D. Sheet, "DASA: Domain adaptation in stacked autoencoders using systematic dropout," in *Proc. 3rd IAPR Asian Conf. Pattern Recognit. (ACPR)*, Nov. 2015, pp. 735–739.

[12] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "DeepVessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2016, pp. 132–139.

[13] L. Ravala and G. K. Rajini, "Automatic diagnosis of diabetic retinopathy from retinal abnormalities: Improved Jaya-based feature selection and recurrent neural network," *Comput. J.*, vol. 65, no. 7, pp. 1904–1922, Jun. 10, 2021.

[14] S. S. Athalye and G. Vijay, "Taylor series-based deep belief network for automatic classification of diabetic retinopathy using retinal fundus images," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 3, pp. 882–901, May 2022.

[15] R. Arunkumar and P. Karthigaikumar, "Multi-retinal disease classification by reduced deep learning features," *Neural Comput. Appl.*, vol. 28, no. 2, pp. 329–334, Feb. 2017.

[16] F. Li, Z. Liu, H. Chen, M. Jiang, X. Zhang, and Z. Wu, "Automatic detection of diabetic retinopathy in retinal fundus photographs based on deep learning algorithm," *Transl. Vis. Sci. Technol.*, vol. 8, no. 6, p. 4, Nov. 2019.

[17] A. S. Jadhav, P. B. Patil, and S. Biradar, "Optimal feature selection-based diabetic retinopathy detection using improved rider optimization algorithm enabled with deep learning," *Evol. Intell.*, vol. 14, no. 4, pp. 1431–1448, Dec. 2021.

[18] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101701.

[19] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informat. Med. Unlocked*, vol. 20, Jun. 2020, Art. no. 100377.

[20] A. Sebastian, O. Elharrouss, S. Al-Maadeed, and N. Almaadeed, "A survey on deep-learning-based diabetic retinopathy classification," *Diagnostics*, vol. 13, no. 3, p. 345, Jan. 2023.

[21] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Brit. Med. J.*, vol. 372, p. n71, Mar. 2020.

[22] M. B. Sasongko, T. Y. Wong, T. T. Nguyen, C. Y. Cheung, J. E. Shaw, and J. J. Wang, "Retinal vascular tortuosity in persons with diabetes and diabetic retinopathy," *Diabetologia*, vol. 54, no. 9, pp. 2409–2416, May 2011.

[23] A. E. Kiely, "Computer-assisted measurement of retinal vascular width and tortuosity in retinopathy of prematurity," *Arch. Ophthalmol.*, vol. 128, no. 7, p. 847, Jul. 2010.

[24] C. Y.-L. Cheung, E. Lamoureux, M. K. Ikram, M. B. Sasongko, J. Ding, Y. Zheng, P. Mitchell, J. J. Wang, and T. Y. Wong, "Retinal vascular geometry in Asian persons with diabetes and retinopathy," *J. Diabetes Sci. Technol.*, vol. 6, no. 3, pp. 595–605, May 2012.

[25] M. B. Sasongko, J. J. Wang, K. C. Donaghue, N. Cheung, P. Benitez-Aguirre, A. Jenkins, W. Hsu, M.-L. Lee, and T. Y. Wong, "Alterations in retinal microvascular geometry in young type 1 diabetes," *Diabetes Care*, vol. 33, no. 6, pp. 1331–1336, Mar. 2010.

[26] S. Akbar, M. Sharif, M. U. Akram, T. Saba, T. Mahmood, and M. Kolivand, "Automated techniques for blood vessels segmentation through fundus retinal images: A review," *Microsc. Res. Techn.*, vol. 82, no. 2, pp. 153–170, Feb. 2019.

[27] L. Wang, H. Liu, Y. Lu, H. Chen, J. Zhang, and J. Pu, "A coarse-to-fine deep learning framework for optic disc segmentation in fundus images," *Biomed. Signal Process. Control*, vol. 51, pp. 82–89, May 2019.

[28] T. Walter, J. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imag.*, vol. 21, no. 10, pp. 1236–1243, Oct. 2002.

[29] T. Y. Wong, C. M. G. Cheung, M. Larsen, S. Sharma, and R. Simó, "Diabetic retinopathy," *Nat. Rev. Dis. Primers.*, vol. 2, Mar. 2016, Art. no. 16012.

[30] M. K. Durzhinskaya, "Microaneurysms as a biomarker of diabetic retinopathy," *Vestnik Oftal'Mologii*, vol. 137, no. 5, p. 300, 2021.

[31] Early Treatment Diabetic Retinopathy Study Research Group, "Grading diabetic retinopathy from stereoscopic color fundus photographs—An extension of the modified Airlie House classification: ETDRS report number 10," *Ophthalmology*, vol. 127, no. 4S, pp. S99–S119, Apr. 2020.

[32] H.-P. Hammes, "Diabetic retinopathy: Hyperglycaemia, oxidative stress and beyond," *Diabetologia*, vol. 61, no. 1, pp. 29–38, Sep. 2017.

[33] T. Bek and H. Lund-Andersen, "Cotton-wool spots and retinal light sensitivity in diabetic retinopathy," *Brit. J. Ophthalmol.*, vol. 75, no. 1, pp. 13–17, Jan. 1991.

[34] D. McLeod, "Why cotton wool spots should not be regarded as retinal nerve fibre layer infarcts," *Brit. J. Ophthalmol.*, vol. 89, no. 2, pp. 229–237, Feb. 2005.

[35] M. U. Akram, A. Tariq, and S. A. Khan, "Detection of neovascularization for screening of proliferative diabetic retinopathy," in *Proc. Int. Conf. Image Anal. Recognit.*, Jun. 2012, pp. 372–379.

[36] M. E. Hartnett, "Studies on the pathogenesis of avascular retina and neovascularization into the vitreous in peripheral severe retinopathy of prematurity (an American ophthalmological society thesis)," *Trans. Amer. Ophthalmolog. Soc.*, vol. 108, pp. 96–119, Dec. 2010.

[37] R. Klein, S. E. Moss, B. E. Klein, M. D. Dams, and D. L. DeMets, "The Wisconsin epidemiologic study of diabetic retinopathy. XI: The incidence of macular edema," *Ophthalmology*, vol. 96, no. 10, pp. 1501–1510, Oct. 1989.

[38] C. P. Wilkinson, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, Sep. 2003.

[39] S. Zachariah, W. Wykes, and D. Yorston, "Grading diabetic retinopathy (DR) using the Scottish grading protocol," *Community Eye Health*, vol. 28, no. 92, pp. 72–73, Jan. 2015.

[40] K. C. Pathak, R. B. Shah, R. R. Tharakan, B. N. Patel, and D. C. Jariwala, "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," *Comput Biol Med.*, vol. 135, Jul. 2021, Art. no. 104599.

[41] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, and A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed image database: The messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, 2014.

[42] Kaggle. (2015). *Diabetic Retinopathy Detection*. [Online]. Available: https://kaggle.com/competitions/diabetic-retinopathy-detection.

[43] Kaggle. (2019). *Aptos 2019 Blindness Detection*. [Online]. Available: https://kaggle.com/competitions/aptos2019-

[44] M. D. Abrámoff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell, P. Massin, B. Cochener, P. Gain, L. Tang, M. Lamard, D. C. Moga, G. Quellec, and M. Niemeijer, "Automated analysis of retinal images for detection of referable diabetic retinopathy," *JAMA Ophthalmol.*, vol. 131, no. 3, pp. 351–357, Mar. 2013.

[45] (2007). *Diaretdb1 Dataset*. [Online]. Available: http://www.it.lut.fi/project/imageret/diaretdb1/

[46] L. Giancardo, F. Meriaudeau, T. P. Karnowski, Y. Li, S. Garg, K. W. Tobin, and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Med. Image Anal.*, vol. 16, no. 1, pp. 216–226, Jan. 2012.

[47] *E-Ophtha*. Accessed: Apr. 2013. [Online]. Available: http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html

[48] P. Prentašic, S. Loncaric, Z. Vatavuk, G. Bencic, M. Subašic, T. Petkovic, L. Dujmovic, M. Malenica-Ravlic, N. Budimlija, and R. Tadic, "Diabetic retinopathy image database(DRiDB): A new database for diabetic retinopathy screening programs research," in *Proc. 8th Int. Symp. Image Signal Process. Anal. (ISPA)*, Sep. 2013, pp. 711–716.

[49] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, Jul. 2018.

[50] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," *Inf. Sci.*, vol. 501, pp. 511–522, Oct. 2019.

[51] A. Serener and S. Serte, "Geographic variation and ethnicity in diabetic retinopathy detection via deeplearning," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 28, no. 2, pp. 664–678, Mar. 2020.

[52] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.

[53] X. Li, X. Hu, L. Yu, L. Zhu, C.-W. Fu, and P.-A. Heng, "CANet: Cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1483–1493, May 2020.

[54] C. González-Gonzalo, V. Sánchez-Gutiérrez, P. Hernández-Martínez, I. Contreras, Y. T. Lechanteur, A. Domanian, B. van Ginneken, and C. I. Sánchez, "Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration," *Acta Ophthalmolog.*, vol. 98, no. 4, pp. 368–377, Nov. 2019.

[55] G. García, J. Gallardo, A. Mauricio, J. López, and C. D. Carpio, "Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images," in *Proc. Int. Conf. Artif. Neural Netw. (ICANN)*, vol. 10614, Oct. 2017, pp. 635–642.

[56] P. Khojasteh, B. Aliahmad, and D. K. Kumar, "Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms," *BMC Ophthalmol.*, vol. 18, no. 1, p. 288, Nov. 2018.

[57] Y. Qi, "A comprehensive overview of image enhancement techniques," *Arch. Comput. Methods Eng.*, vol. 29, pp. 583–607, Apr. 2021.

[58] Z. Xiaohui and O. Chutatape, "Detection and classification of bright lesions in color fundus images," in *Proc. Int. Conf. Image Process. (ICIP)*, Nov. 2004, pp. 139–142.

[59] T. Yamuna and S. Maheswari, "Detection of abnormalities in retinal images," in *Proc. IEEE Int. Conf. Emerg. Trends Comput., Commun. Nanotechnol. (ICECCN)*, Mar. 2013, pp. 236–240.

[60] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color retinal image enhancement using CLAHE," in *Proc. Int. Conf. ICT Smart Soc.*, Jun. 2013, pp. 1–3.

[61] P. R. R. Chandni, J. Justin, and R. Vanithamani, "Fundus image enhancement using EAL-CLAHE technique," *Adv. Data Inf. Sci.*, vol. 318, pp. 613–624, Feb. 2022.

[62] P. S. Reddy, H. Singh, A. Kumar, L. K. Balyan, and H.-N. Lee, "Retinal fundus image enhancement using piecewise gamma corrected dominant orientation based histogram equalization," in *Proc. Int. Conf. Commun. Signal Process. (ICCSP)*, Apr. 2018, pp. 124–128.

[63] B. Bataineh and K. H. Almotairi, "Enhancement method for color retinal fundus images based on structural details and illumination improvements," *Arabian J. Sci. Eng.*, vol. 46, no. 9, pp. 8121–8135, Feb. 2021.

[64] S. Gross, M. Klein, and D. Schneider, "Segmentation of blood vessel structures in retinal fundus images with logarithmic Gabor filters," *Current Med. Imag. Rev.*, vol. 9, no. 2, pp. 138–144, May 2013.

[65] B. Yang, H. Zhao, L. Cao, H. Liu, N. Wang, and H. Li, "Retinal image enhancement with artifact reduction and structure retention," *Pattern Recognit.*, vol. 133, Jan. 2023, Art. no. 108968.

[66] Q. You, C. Wan, J. Sun, J. Shen, H. Ye, and Q. Yu, "Fundus image enhancement method based on CycleGAN," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 4500–4503.

[67] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 3–19.

[68] H. Zhao, B. Yang, L. Cao, and H. Li, "Data-driven enhancement of blurry retinal images via generative adversarial networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent (MICCAI)*, vol. 11764, Oct. 2019, pp. 75–83.

[69] K. G. Lee, S. J. Song, S. Lee, H. G. Yu, D. I. Kim, and K. M. Lee, "A deep learning-based framework for retinal fundus image enhancement," *PLoS ONE*, vol. 18, no. 3, Mar. 2023, Art. no. e0282416.

[70] M. Karunakaran and R. Neduncheliian, "Comparision of various noises and filters for fundus images using pre-processing techniques," *Int. J. Pharma Bio Sci.*, vol. 5, pp. B499–B508, Jan. 2014.

[71] A. F. M. Hani, T. A. Soomro, I. Faye, N. Kamel, and N. Yahya, "Denoising methods for retinal fundus images," in *Proc. 5th Int. Conf. Intell. Adv. Syst. (ICIAS)*, Jun. 2014, pp. 1–6.

[72] A. A. G. Elseid, M. E. M. Elmanna, and A. O. M. Hamza, "Evaluation of spatial filtering techniques in retinal fundus images," *Amer. J. Artif. Intel.*, vol. 2, no. 2, pp. 16–21, Oct. 2018.

[73] S. A. Khowaja, P. Khuwaja, and I. A. Ismaili, "A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification," *Signal, Image Video Process.*, vol. 13, no. 2, pp. 379–387, Sep. 2018.

[74] M. S. Patil, S. Chickerur, C. Abhimalya, A. Naik, N. Kumari, and S. Maurya, "Effective deep learning data augmentation techniques for diabetic retinopathy classification," *Proc. Comput. Sci.*, vol. 218, pp. 1156–1165, Jan. 2023.

[75] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abràmoff, A. M. Mendonça, and A. Campilho, "End-to-end adversarial retinal image synthesis," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 781–791, Mar. 2017.

[76] D. Mahapatra and B. Bozorgtabar, "Retinal vasculature segmentation using local saliency maps and generative adversarial networks for image super resolution," 2017, *arXiv:1710.04783*.

[77] R. Zheng, L. Liu, S. Zhang, C. Zheng, F. Bunyak, R. Xu, B. Li, and M. Sun, "Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network," *Biomed. Opt. Exp.*, vol. 9, no. 10, p. 4863, Sep. 2018.

[78] Y. Chen, J. Long, and J. Guo, "RF-GANs: A method to synthesize retinal fundus images based on generative adversarial network," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–17, Nov. 2021.

[79] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[80] Z. Zhao, K. Zhang, X. Hao, J. Tian, M. C. Heng Chua, L. Chen, and X. Xu, "BiRA-Net: Bilinear attention net for diabetic retinopathy grading," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1385–1389.

[81] C. K. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," *AMIA Summits Transl. Sci. Proc.*, vol. 2018, pp. 147–155, May 2018.

[82] C.-H. Hua, T. Huynh-The, K. Kim, S.-Y. Yu, T. Le-Tien, G. H. Park, J. Bang, W. A. Khan, S.-H. Bae, and S. Lee, "Bimodal learning via trilogy of skip-connection deep networks for diabetic retinopathy risk progression identification," *Int. J. Med. Informat.*, vol. 132, Dec. 2019, Art. no. 103926.

[83] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[84] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*.

[85] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 770–778.

[86] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," 2016, *arXiv:1602.07261*.

[87] M. Shaban, Z. Ogur, A. Mahmoud, A. Switala, A. Shalaby, H. Abu Khalifeh, M. Ghazal, L. Fraiwan, G. Giridharan, H. Sandhu, and A. S. El-Baz, "A convolutional neural network for the screening and staging of diabetic retinopathy," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0233514.

[88] X. Zhang, F. Li, D. Li, Q. Wei, X. Han, B. Zhang, H. Chen, Y. Zhang, B. Mo, B. Hu, D. Ding, X. Li, W. Yu, and Y. Chen, "Automated detection of severe diabetic retinopathy using deep learning method," *Graefe's Arch. Clin. Experim. Ophthalmol.*, vol. 260, no. 3, pp. 849–856, Mar. 2022.

[89] M. T. Hagos and S. Kant, "Transfer learning-based detection of diabetic retinopathy from small dataset," *arXiv:1905.07203*.

[90] S. Rêgo, M. Dutra-Medeiros, F. Soares, and M. Monteiro-Soares, "Screening for diabetic retinopathy using an automated diagnostic system based on deep learning: Diagnostic accuracy assessment," *Ophthalmologica*, vol. 244, no. 3, pp. 250–257, Oct. 2020.

[91] J. Cao, J. Chen, X. Zhang, Q. Yan, and Y. Zhao, "Attentional mechanisms and improved residual networks for diabetic retinopathy severity classification," *J. Healthcare Eng.*, vol. 2022, pp. 1–10, Mar. 2022.

[92] C. Suedumrong, K. Leksakul, P. Wattana, and P. Chaopaisarn, "Application of deep convolutional neural networks VGG-16 and GoogLeNet for level diabetic retinopathy detection," in *Proc. Future Technol. Conf.*, vol. 2, Nov. 2021, pp. 56–65.

[93] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[94] C.-Y. Liou, W.-C. Cheng, J.-W. Liou, and D.-R. Liou, "Autoencoder for words," *Neurocomputing*, vol. 139, pp. 84–96, Sep. 2014.

[95] M. T. García-Ordás, C. Benavides, J. A. Benítez-Andrades, H. Alaiz-Moretón, and I. García-Rodríguez, "Diabetes detection using deep learning techniques with oversampling and feature augmentation," *Comput. Methods Programs Biomed.*, vol. 202, Apr. 2021, Art. no. 105968.

[96] Y. Zhu, X. Wu, J. Qiang, Y. Yuan, and Y. Li, "Representation learning via an integrated autoencoder for unsupervised domain adaptation," *Frontiers Comput. Sci.*, vol. 17, no. 5, Jan. 2023, Art. no. 175334.

[97] H. Wang, X. Shi, and D. Y. Yeung, "Relational stacked denoising autoencoder for tag recommendation," in *Proc. AAAI Conf. Artif. Intell.*, Feb. 2015, pp. 3052–3058.

[98] X. Wang, Y. Qin, Y. Wang, S. Xiang, and H. Chen, "ReLTanh: An activation function with vanishing gradient resistance for SAE-based DNNs and its application to rotating machinery fault diagnosis," *Neurocomputing*, vol. 363, pp. 88–98, Oct. 2019.

[99] S. Kumar, A. Mallik, and S. S. Sengar, "Community detection in complex networks using stacked autoencoders and crow search algorithm," *J. Supercomput.*, vol. 79, no. 3, pp. 3329–3356, Sep. 2022.

[100] Y. Li, Z. Song, S. Kang, S. Jung, and W. Kang, "Semi-supervised autoencoder graph network for diabetic retinopathy grading," *IEEE Access*, vol. 9, pp. 140759–140767, 2021.

[101] G. Atteia, N. Abdel Samee, and H. Zohair Hassan, "DFTSA-net: Deep feature transfer-based stacked autoencoder network for DME diagnosis," *Entropy*, vol. 23, no. 10, p. 1251, Sep. 2021.

[102] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[103] N. Le Roux and Y. Bengio, "Representational power of restricted Boltzmann machines and deep belief networks," *Neural Comput.*, vol. 20, no. 6, pp. 1631–1649, Jun. 2008.

[104] S. Krishnamoorthy, Y. Weifeng, J. Luo, and S. Kadry, "GO-DBN: Gannet optimized deep belief network based wavelet kernel ELM for detection of diabetic retinopathy," *Expert Syst. Appl.*, vol. 229, Nov. 2023, Art. no. 120408.

[105] S. S. Basha and K. V. Ramanaiah, "Study on distance-based monarch butterfly oriented deep belief network for diabetic retinopathy and its parameters," in *Advances in Computational Intelligence Techniques*. Singapore: Springer, Feb. 2020, pp. 129–155.

[106] A. A. Tehrani, A. M. Nickfarjam, H. Ebrahimpour-komleh, and D. Aghadoost, "Multi-input 2-dimensional deep belief network: Diabetic retinopathy grading as case study," *Multimedia Tools Appl.*, vol. 80, no. 4, pp. 6171–6186, Oct. 2020.

[107] R. Rajavel, B. Sundaramoorthy, K. Gr, S. K. Ravichandran, and K. Leelasankar, "Cloud-enabled diabetic retinopathy prediction system using optimized deep belief network classifier," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 10, pp. 14101–14109, Jul. 2022.

[108] S. Das, A. Tariq, T. Santos, S. S. Kantareddy, and I. Banerjee, "Recurrent neural networks (RNNs): Architectures, training tricks, and introduction to influential research," in *Machine Learning for Brain Disorders*. New York, NY, USA: Humana, Jul. 2023, Ch. 4.

[109] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.

[110] S. H. Kim, LeeHanBeom, S. W. Chun, K. Daeyeon, and S. J. Lee, "Prediction of blood glucose in diabetic inpatients using LSTM neural network," *Medicine*, vol. 47, no. 12, pp. 1120–1125, Dec. 2020.

[111] Y. B. Özçelik and A. Altan, "Overcoming nonlinear dynamics in diabetic retinopathy classification: A robust AI-based model with chaotic swarm intelligence optimization and recurrent long short-term memory," *Fractal Fractional*, vol. 7, no. 8, p. 598, Aug. 2023.

[112] K. V. Spoorthi and B. S. Rekha, "Diabetic retinopathy prediction using deep learning," in *Proc. IEEE Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solutions (CSITSS)*, Dec. 2021, pp. 1–6.

[113] G. Kumar, S. Chatterjee, and C. Chattopadhyay, "DRISTI: A hybrid deep neural network for diabetic retinopathy diagnosis," *Signal, Image Video Process.*, vol. 15, no. 8, pp. 1679–1686, Apr. 2021.

[114] P. Saranya, S. Prabakaran, R. Kumar, and E. Das, "Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning," *Vis. Comput.*, vol. 38, no. 3, pp. 977–992, Jan. 2021.

[115] A. Bora, S. Balasubramanian, B. Babenko, S. Virmani, S. Venugopalan, A. Mitani, G. de Oliveira Marinho, J. A. Cuadros, P. Ruamviboonsuk, G. S. Corrado, L. H. Peng, D. R. Webster, A. V. Varadarajan, N. Hammel, Y. Liu, and P. Bavishi, "Predicting risk of developing diabetic retinopathy using deep learning," *Lancet Digit. health*, vol. 3, no. 1, pp. e10–e19, Nov. 2020.

[116] F. Li, Y. Wang, T. Xu, L. Dong, L. Yan, M. Jiang, X. Zhang, H. Jiang, Z. Wu, and H. Zou, "Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs," *Eye*, vol. 36, no. 7, pp. 1433–1441, Jul. 2021.

[117] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019.

[118] M. Ashikur, M. Arifur, and J. Ahmed, "Automated detection of diabetic retinopathy using deep residual learning," *Int. J. Comput. Appl.*, vol. 177, no. 42, pp. 25–32, Mar. 2020.

[119] M. S. Sallam, A. L. Asnawi, and R. F. Olanrewaju, "Diabetic retinopathy grading using ResNet convolutional neural network," in *Proc. IEEE Conf. Big Data Anal. (ICBDA)*, Nov. 2020, pp. 73–78.

[120] W. Zhang, J. Zhong, S. Yang, Z. Gao, J. Hu, Y. Chen, and Z. Yi, "Automated identification and grading system of diabetic retinopathy using deep neural networks," *Knowl.-Based Syst.*, vol. 175, pp. 12–25, Jul. 2019.

[121] C. Bhardwaj, S. Jain, and M. Sood, "Deep learning–based diabetic retinopathy severity grading system employing quadrant ensemble model," *J. Digit. Imag.*, vol. 34, no. 2, pp. 440–457, Mar. 2021.

[122] J. D. Bodapati, "Stacked convolutional auto-encoder representations with spatial attention for efficient diabetic retinopathy diagnosis," *Multimedia Tools Appl.*, vol. 81, no. 22, pp. 32033–32056, Apr. 2022.

[123] A. M. Dayana and W. R. S. Emmanuel, "An enhanced swarm optimization-based deep neural network for diabetic retinopathy classification in fundus images," *Multimedia Tools Appl.*, vol. 81, no. 15, pp. 20611–20642, Mar. 2022.

[124] Y. E. A. Mustaf and B. H. Ismail, "Classification of diabetic retinopathy using stacked autoencoder-based deep neural network," *J. Comput. Sci. Intell. Technol.*, vol. 1, no. 1, pp. 9–14, 2020.

[125] K. Gunasekaran, R. Pitchai, G. K. Chaitanya, D. Selvaraj, S. A. Sheryl, H. S. Almoallim, S. A. Alharbi, S. S. Raghavan, and B. G. Tesemma, "A deep learning framework for earlier prediction of diabetic retinopathy from fundus photographs," *BioMed Res. Int.*, vol. 2022, pp. 1–15, Jun. 2022.

[126] S. Seth and B. Agarwal, "A hybrid deep learning model for detecting diabetic retinopathy," *J. Statist. Manage. Syst.*, vol. 21, no. 4, pp. 569–574, Jun. 2018.

[127] Y.-P. Liu, Z. Li, C. Xu, J. Li, and R. Liang, "Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101694.

[128] P. Shah, D. K. Mishra, M. P. Shanmugam, B. Doshi, H. Jayaraj, and R. Ramanjulu, "Validation of deep convolutional neural network-based algorithm for detection of diabetic retinopathy—Artificial intelligence versus clinician for screening," *Indian J. Ophthalmol.*, vol. 68, pp. 398–405, Feb. 2020.

[129] R. Chalakkal, F. Hafiz, W. Abdulla, and A. Swain, "An efficient framework for automated screening of clinically significant macular edema," *Comput. Biol. Med.*, vol. 130, Mar. 2021, Art. no. 104128.

[130] A. K. Gangwar and V. Ravi, "Diabetic retinopathy detection using transfer learning and deep learning," in *Evolution in Computational Intelligence*, vol. 1176. Singapore: Springer, Sep. 2020, pp. 679–689.

[131] X. Feng, S. Zhang, L. Xu, X. Huang, and Y. Chen, "Robust classification model for diabetic retinopathy based on the contrastive learning method with a convolutional neural network," *Appl. Sci.*, vol. 12, no. 23, p. 12071, Nov. 2022.

[132] B. K. Anoop, "Binary classification of DR-diabetic retinopathy using CNN with fundus colour images," *Mater. Today, Proc.*, vol. 58, pp. 212–216, Jan. 2022.

[133] P. Macsik, J. Pavlovicova, J. Goga, and S. Kajan, "Local binary CNN for diabetic retinopathy classification on fundus images," *Acta Polytechnica Hungarica*, vol. 19, no. 7, pp. 27–45, 2022.

[134] M. M. Farag, M. Fouad, and A. T. Abdel-Hamid, "Automatic severity classification of diabetic retinopathy based on DenseNet and convolutional block attention module," *IEEE Access*, vol. 10, pp. 38299–38308, 2022.

[135] S. Wan, Y. Liang, and Y. Zhang, "Deep convolutional neural networks for diabetic retinopathy detection by image classification," *Comput. Electr. Eng.*, vol. 72, pp. 274–282, Nov. 2018.

[136] B. Harangi, J. Toth, A. Baran, and A. Hajdu, "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 2699–2702.

[137] B. Tymchenko, P. Marchenko, and D. Spodarets, "Deep learning approach to diabetic retinopathy detection," in *Proc. 9th Int. Conf. Pattern Recognit. Appl. Methods*, Mar. 2020, pp. 501–509.

[138] K. N. S. Sree, D. V. Sree, G. H. Lakshmi, and S. Ramesh, "Diabetic retinopathy detection using deep learning," in *Proc. 2nd Int. Conf. Electron. Sustain. Commun. Syst. (ICESC)*, Aug. 2021, pp. 1670–1674.

[139] Z. Tu, S. Gao, K. Zhou, X. Chen, H. Fu, Z. Gu, J. Cheng, Z. Yu, and J. Liu, "SUNet: A lesion regularized model for simultaneous diabetic retinopathy and diabetic macular edema grading," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1378–1382.

[140] M. Islam, M. Hasan, and M. Nahiduzzaman, "Severity grading of diabetic retinopathy using deep convolutional neural network," *Int. J. Innova. Res. Sci. Eng. Technol.*, vol. 6, p. 7, Feb. 2021.

[141] A. Sugeno, Y. Ishikawa, T. Ohshima, and R. Muramatsu, "Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104795.

[142] H. Tariq, M. Rashid, A. Javed, E. Zafar, S. S. Alotaibi, and M. Y. I. Zia, "Performance analysis of deep-neural-network-based automatic diagnosis of diabetic retinopathy," *Sensors*, vol. 22, no. 1, p. 205, Dec. 2021.

[143] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, p. 3704, May 2021.
[144] S. Albahli and G. N. Ahmad Hassan Yar, "Automated detection of diabetic retinopathy using custom convolutional neural network," *J. X-Ray Sci. Technol.*, vol. 30, no. 2, pp. 275–291, Mar. 2022.
[145] A. Bilal, G. Sun, S. Mazhar, A. Imran, and J. Latif, "A transfer learning and U-Net-based automatic detection of diabetic retinopathy from fundus images," *Comput. Methods Biomechanics Biomed. Eng., Imag. Visualizat.*, vol. 10, no. 6, pp. 663–674, Nov. 2022.
[146] M. R. Islam, L. F. Abdulrazak, M. Nahiduzzaman, M. O. F. Goni, M. S. Anower, M. Ahsan, J. Haider, and M. Kowalski, "Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105602.
[147] Y. Kale and S. Sharma, "Detection of five severity levels of diabetic retinopathy using ensemble deep learning model," *Multimedia Tools Appl.*, vol. 82, no. 12, pp. 19005–19020, Dec. 2022.
[148] R. Pugal Priya, T. Saradadevi Sivarani, and A. Gnana Saravanan, "Deep long and short term memory based red fox optimization algorithm for diabetic retinopathy detection and classification," *Int. J. Numer. Methods Biomed. Eng.*, vol. 38, no. 3, Mar. 2022, Art. no. e356.
[149] M. Nahiduzzaman, M. R. Islam, M. O. F. Goni, M. S. Anower, M. Ahsan, J. Haider, and M. Kowalski, "Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier," *Expert Syst. Appl.*, vol. 217, May 2023, Art. no. 119557.
[150] C.-L. Lin and K.-C. Wu, "Development of revised ResNet-50 for diabetic retinopathy detection," *BMC Bioinf.*, vol. 24, no. 1, p. 157, Apr. 2023.
[151] M. Jian, H. Chen, C. Tao, X. Li, and G. Wang, "Triple-DRNet: A triple-cascade convolution neural network for diabetic retinopathy grading using fundus images," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106631.
[152] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Deep learning-enhanced diabetic retinopathy image classification," *Digit. Health*, vol. 9, Aug. 2023, Art. no. 20552076231194942.
[153] S. Zulaikha Beevi, "Multi-level severity classification for diabetic retinopathy based on hybrid optimization enabled deep learning," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104736.
[154] A. R. Wahab Sait, "A lightweight diabetic retinopathy detection model using a deep-learning technique," *Diagnostics*, vol. 13, no. 19, p. 3120, Oct. 2023.

SHANSHAN ZHU received the Ph.D. degree from the College of Medicine and Biological Information Engineering, Northeastern University, China, in 2020. She is currently an Assistant Research Fellow with the Research Institute of Medical and Biological Engineering, Ningbo University, China. Her research interests include optical spectroscopy, machine learning, and diabetes and complications analysis based on biomedical engineering technology.

CHANGCHUN XIONG is currently pursuing the master's degree with the Faculty of Electrical Engineering and Computer Science and the Research Institute of Medical and Biological Engineering, Ningbo University, under the supervision of Prof. Yudong Yao and Shanshan Zhu. His current research interests include machine learning and big data analysis.

QINGSHAN ZHONG is currently pursuing the master's degree with the School of Materials Science and Chemical Engineering and the Research Institute of Medical and Biological Engineering, Ningbo University, under the supervision of Dr. Shanshan Zhu. His current research interests include machine learning and optical spectroscopy.

YUDONG YAO received the Ph.D. degree in electronic information engineering from Southeast University, in 1988. He is currently a Professor with the Research Institute of Medical and Biological Engineering, Ningbo University, China. His research interests include deep learning, medical imaging processing, machine learning, and health care big data analysis. He is a fellow of the American Institute for Medical and Biological Engineering.