Comprehensive Survey of the State-of-the-Art Deep Learning Models for Diabetic Retinopathy Detection and Grading Using Retinal Fundus Photography

Noor Ali Sadek[1], Ziad Tarik Al-dahan[2], Suzan Amana Rattan[3]

Authors Affiliations:

1) Department of Biomedical Engineering, Al- Nahrain University, Bagdad- Iraq.
noor.a.sadek@nahrainuniv.edu.iq
nooraliz810@gmail.com

2) Department of Biomedical Engineering, Al- Nahrain University, Bagdad- Iraq.
ziaddrmt1959@yahoo.co.uk

3) AlKindy College of Medicine, University of Baghdad, Baghdad- Iraq.
suzanamana@kmc.uobaghdad.edu.iq

Abstract

In order to avoid losing sense of sight in a large portion of the working population, Diabetic Retinopathy (DR) identification during broad examination for diabetes is crucial. To prevent blindness in the future, early illness detection and measurement of disease development are essential. DR is diagnosed through medical image analysis. After the success of Deep Learning (DL) in other applications in the real world, it is considered a vital tool for upcoming health sector applications, providing solutions with accurate results for medical image analysis. This review provides a comprehensive survey of the state-of-the-art DL models for DR detection and grading using retinal fundus photography. This review thoroughly examined and summarized 81 relevant publications that were published through IEEE Xplore, Web of Science, PubMed, and Scopus between 2018 and 2023 based on the available database with binary or multiclass CNN classification models as well as the main preprocessing techniques. According to the findings of this review, transfer learning has proven to be an excellent technique for addressing the problems of limited resources for data for DR analysis. CNN models having tens or hundreds of layers are the most frequently utilized frameworks for DR classification. The most extensively utilized datasets for DR categorization are Aptos 2019 and EyePACS. Although DL has attained or surpassed human-level DR classification accuracy, there is still more work to be done in real-world clinical procedures.

Keywords: Diabetic Retinopathy, CNN models, DR Datasets Classification.

دراسة شاملة لنماذج التعلم العميق المتطورة للكشف عن اعتلال الشبكية لمرضى السكري وتصنيفه باستخدام تصويرالشبكية

نور علي صادق١ , زياد طارق الدهان٢, سوزان أمانة رطان٣

الخلاصة

من أجل تجنب فقدان حاسة البصر لدى جزء كبير من السكان العاملين، يعد تحديد اعتلال الشبكية السكري (DR) أثناء الفحص الشامل لمرض السكري أمرًا بالغ الأهمية. ومن أجل الوقاية من العمى في المستقبل، يعد الكشف المبكر عن المرض وقياس تطور المرض أمرًا ضروريًا. يتم تشخيص DR من خلال تحليل الصور الطبية. بعد نجاح التعلم العميق (DL) في التطبيقات الأخرى في العالم الحقيقي، فإنه يعتبر أداة حيوية لتطبيقات القطاع الصحي، حيث يوفر حلولاً ذات نتائج دقيقة لتحليل الصور الطبية. توفر هذه المراجعة مسحًا شاملاً لأحدث نماذج التعلم العميق للكشف عن اعتلال الشبكية وتصنيفها باستخدام تصوير الشبكية. قامت هذه المراجعة بفحص وتلخيص ٨١ منشورًا ذا صلة تم نشرها من خلال المجلات العلمية الرصينة ضمن IEEE Xplore و Web of Science وPubMed وScopus بين عامي ٢٠١٨ و٢٠٢٣ استنادًا إلى قاعدة البيانات المستخدمة مع نماذج CNN ذات التصنيف الثنائي أو التصنيف متعدد الفئات بالإضافة إلى تقنيات المعالجة المسبقة. وفقا لنتائج هذه المراجعة، أثبت نقل التعلم أنه أسلوب ممتاز لمعالجة مشاكل الموارد المحدودة للبيانات للكشف المبكر و تصنيف

اعتلال شبكية السكري. تعد نماذج CNN التي تحتوي على عشرات أو مئات الطبقات الأكثر استخدامًا لتصنيف اعتلال الشبكية. البيانات الأكثر استخدامًا على نطاق واسع في العام لتصنيف مرض اعتلال الشبكية هي Aptos2019 وEyePACS وتستخدم من قبل العديد من الباحثين في هذا المجال. على الرغم من أن DL قد حققت أو تجاوزت لاى حد ما دقة تصنيف وكشف مرض اعتلال الشبكية على المستوى البشري، إلا أنه لا يزال هناك المزيد من العمل الذي يتعين القيام به في الإجراءات السريرية في العالم الحقيقي.

الكلمات المفتاحية: اعتلال الشبكية السكري، نماذج الشبكات العصبية التلافيفية (CNN)، مجموعات بيانات اعتلال الشبكية، تصنيف اعتلال الشبكية.

1. Introduction

One of the most frequent conditions affecting people in their working years and one of the main causes of visual impairment across the globe is DR (DR) [1, 2]. Retinal microvascular disease is a result of diabetic complications, causing the disease. If diagnosed early on, effective therapy can avoid or delay the development of severe vision impairment. Therefore, it is essential that diabetics undergo routine screenings for DR (DR) with a suitable diabetic eye screening tool. Nevertheless, in its initial stages, the disorder has few signs and those that do present can occasionally be challenging to diagnose. As a result, diagnosis is expensive and time-consuming since skilled evaluators are required. There will be more patients experiencing retinopathy as a result of the rise in diabetes cases [3]. The restrictions that national health systems face around the world, as well as people's ability to work and hence the economy, necessitate the development of inexpensive and efficient patient screening methods. The color fundus images of the retina must be evaluated by a skilled image evaluator in order to use the current DR classification approach. Due to the manual nature of this process, classifications may vary throughout grading centers and may require time. It will also take some time to return the results to patients for treatment [4]. Classification approaches involve detecting lesions linked to abnormalities of the vascular system [5]. It takes a lot of resources to implement this method, even though it works well. Sometimes the required equipment and experience are insufficient, particularly if DR is at its most severe and diabetes is highly prevalent in the community. The more people with diabetes there are, the more ineffective the infrastructure will become to avoid DR-related impairment because the process of classification will take more time from clinicians. Computerized solutions are needed to support, replace, or expedite the diagnosis process because DR examinations have long been thought to be useless [6, 7]. Machine learning has been used extensively for image classification, with effective and accurate outcomes [8, 9,10].

1.1. The Disease bio-marks:

Currently, approximately 422 million people worldwide suffer from diabetes, and this proportion is expected to rise significantly [11,12]. There are two major forms of diabetes: type 1 and type 2. DR has been found to affect more than 25% of patients with type 2 diabetes and over 50% of those with type 1 diabetes [13]. DR is now diagnosed clinically by identifying abnormalities in blood vessels in the retina has traditionally been considered a microvascular disease [14]. It is believed that hyperglycemia plays a role in the development of microvascular damage in the retina. In the early stages of the condition, hyperglycemia dilates retinal blood vessels and alters blood flow. For those who have diabetes, these modifications are assumed to represent metabolic autoregulation, which increases retinal metabolism [15]. Pericyte loss is another component that contributes to early-stage DR. Research conducted both in vivo and in vitro has revealed evidence linking high glucose levels to pericyte apoptosis [15]. Because pericytes give capillaries structural support, their loss results in localized capillary wall out-pouching. The first symptom of DR is a microaneurysm, which is brought on by the dilatation of retinal blood vessels and the outward protrusion of capillary walls [13]. Little red spots in the superficial layers of the retina are warning signs of microaneurysms. The blood-retinal barrier is compromised when the disease progresses to the more severe non-proliferative stages because to pericyte reduction, endothelial cell degeneration, and thickened of the basement membrane [16]. Moreover, retinal ischemia and capillary obstruction are brought on by the notable loss of endothelial cells and pericytes (hypoxia). Cotton wool patches may occur as a result of an obstruction in the blood flow to the RNFL. When microaneurysms rupture in the deeper layers of the retina, such as the outer plexiform and inner nuclear layers, dot and blot hemorrhages may result. Hard exudate and retinal edema are the results of serum lipids, proteins, and proteins leaking out of the arteries due to the breakdown of the blood-retina barrier. When retinal hypoxia becomes more severe, the eye's natural defense mechanisms kick in, trying to compensate by giving tissues more oxygen. Venous anomalies, such as venous dilatation, loops, and beading, are indicative of increasing hypoxia and are almost always observed in the vicinity of regions with retinal ischemia, often referred to as capillary non-perfusion. The best indicator of the onset of PDR is the existence of venous anomalies [15]. New blood vessels begin to form when retinal ischemia continues, causing the extracellular matrix to break down. It is known that the proliferative stage of DR is characterized by neovascularization. Neovascularization or the organizing already-existing capillaries by proliferation within the retinal tissues to

give a new path via non-perfusion zones are examples of intraretinal microvascular abnormalities [15]. Patients with proliferative DR may experience severe visual impairment by tractional separation of the retina or by rupture of aberrant arteries into the vitreous (vitreous hemorrhage). Numerous of these abnormalities could be involved in the vision loss brought on by DR. Firstly, macular edema, which is defined as thickening or enlargement of the macula, can impair central vision. The collapse of the blood-retinal boundary leads to a buildup of fluid in the macula, which is what causes this swelling. Second, proliferative DR's new blood vessels and the fibrous tissue that surrounds them can contract, distorting the retina and causing tractional retinal detachment a condition that results in severe and frequently irreversible visual loss. Third, there's a chance that the newly formed blood vessels can bleed, which could result in vitreous or pre-retinal hemorrhage. The clinical avenues for vision loss depicted in Figure (1) include damage to retinal neurons in addition to these clearly visible vascular alterations.

Figure (1): A patient with DR (right) and a healthy patient (left) experiencing the same eyesight [17].

1.2. DR grading

DR severity levels are determined by DR grading, which evaluates vascular alterations. Many DR grading protocols are currently available, such as the Scottish DR grading protocol [18], the International Clinical Diabetic Retinopathy Severity Scale (ICDRSS) [19], the Early Treatment Diabetic Retinopathy Study (ETDRS) classification [19], and the International Clinical Diabetic Macular Edema Severity Scale (ICDMESS) [20]. Despite the fact that the ETDRS grading system is thought to be the best, daily professional and enormous amounts grading are still challenging because of its intricate implementation and multiple levels. The ICDRSS, created by the Global DR Project Group, has garnered significant attention in therapy and computer-aided diagnostic (CAD) environments across the globe because of its practicality and ease of use. The five DR severity levels that the ICDRSS classifies DR into are no DR, mild NPDR, moderate, NPDR, severe NPDR, and PDR, as shown in Figure 2 [19]. Five phases are used by the World Health Organization (WHO) to classify DR [21]:

DR is absent. As of now, no sickness is visible.

1. Retinopathy from mild diabetes. Early in the course of the disease, microaneurysms are tiny patches of microscopic blood arteries or balloon-like swelling in the retinas. The retina may see fluid leakage from these microaneurysms, refer to this as capillary leakage.

2. Moderate Retinopathy Due to Diabetes. The retinal blood supply may enlarge and warp as the condition worsens. They might also stop being able to move blood. Both disorders result in distinct alterations to the retina's appearance, may exacerbate macula enlargement, and may also increase blood or exudate leakage from capillaries.

3. Severe Retinopathy Due to Diabetes. Several additional blood vessels are obstructed, which results in parts of the retina losing blood flow. The term "capillary non-perfusion" describes this. Growth factors released from these regions instruct the retina to produce newly formed blood vessels.

4. DR with Proliferation. New blood vessels begin to proliferate at this advanced stage due to capillary non-perfusion. These blood vessels grow into the vitreous gel, which fills the eye, and along the inner side of the retina. Because of their fragility, the young blood vessels are more likely to have capillary leakage. Retinal detachment, or the ripping of the retina from underlying tissue, can result from accompanying scar tissue contracting. A retinal detachment may result in irreversible eyesight loss.
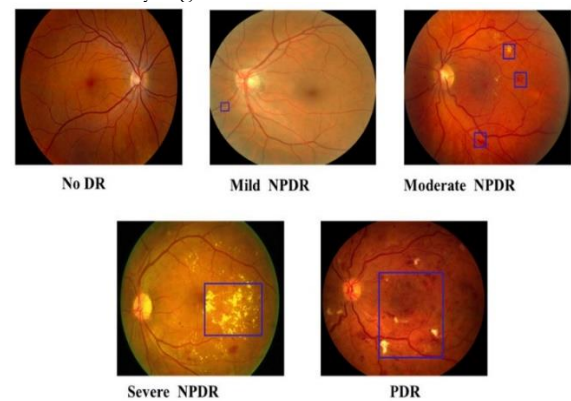
Figure (2): ICDRSS DR grading. (NPDR: Non-proliferative DR; PDR: Proliferative DR) [19].

The main out lines of this review are organized as follows: section 1. introduction that showed general description on DR ocular disease with the disease bio-marks and grading. Section 2 showed the methodology of this paper. Section 3 provides a description of the publicly available datasets with their characteristics. Section 4 included general description of the common preprocessing technique for fundus image enhancements. Section 5 generally explorer the most frequently used DL model with either binary or multi-class classification of DR fundus images with some of the limitations. Finally, section 6 summarized the main conclusions points of this review.

2. Research Methodology

Related papers were searched for 105 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 because of it was written in other languages or insufficient data …ect.; and determining the specific DL tasks for DR, final 81 articles were carefully included. In contrast to previous publications, the main methods of this work could be

described as follows: First, an overview was presented by showing extensive data preprocessing methodology as well as the most recent works in the field of DR classification employing DL approaches over the last three years. The article discusses databases, DL models, and the performance of approaches such as binary classification for DR diagnosis and multi-classification for DR grading.

3. Publicly available DR Datasets:

One of the main factors contributing to the success of DL research is datasets, which are collections of data that can be utilized to train and test DL models. Nevertheless, significant dataset collection and high-quality, accurate labeling in the field of fundus images are difficult to achieve. One factor is that obtaining and disseminating medical data is challenging due to privacy protection of personal information. Furthermore, inconsistent image standards and quality are caused by a variety of fundus imaging instruments and settings, imaging properties, and operators. In addition, qualified ophthalmologists frequently carry out picture tagging. Here, a brief overview was provided [22] of a number of fundus image datasets for DR classification that are freely accessible, as indicated in Table 1. The earlier review articles contained far more thorough information. The fact that photos were taken using various cameras and under various imaging settings means that this dataset has an adequate variety of real-world scenarios.

Table (1): Popular open-source databases for DR classification.

Year	Dataset	Country	Size	Resolution	Tasks
2004	MESSIDOR [23]	France	1200	2304x1536 2240x1488 1440x960	DR grading
2007	DIARETDBI [27]	Finland	89	1500x1152	DR grading lesion detection
2009	Kaggle EyePACS [24]	United States	88702	Varying	DR grading
2010	MESSIDOR2 [26]	France	1748	1440x960 2240x1488 2304x1536	DR grading
2012	HEI-MED [28]	United States	169	Varying	DR grading lesion detection
2013	E-Ophtha [29]	France	463	2544x1696 1440x960	DR grading lesion detection
2013	DRiDB [30]	Croatia	50	768x584	DR grading lesion detection
2017	IDRiD [31]	India	516	4288x2848	DR grading lesion detection
2019	Kagge APTOS 2019 [25]	India	5590	Varying	DR grading
2019	DDR [32]	China	13673	Varying	DR grading lesion detection

4. Pre-processing of Data:

Preparing structured raw data that the algorithm can accept is a frequent first step in DL research, and it can improve the training efficiency of DL models. A variety of image quality problems (including size, noise, artifacts, contrast, lighting, and sharpened regions) might arise while capturing fundus images with different cameras and settings. During the DL model training process, these heterogeneities could conceal certain unique aspects of the DR features. As shown in Table 2, the primary operations performed on the fundus images were image enhancement, denoising, normalization, and augmentation.

Table (2): A preliminary processing of DR data for classification.

Preprocessing approaches	Methods	Role
Image enhancement	Contrast enhancements: 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: HIS color space conversion, HSV color space conversion, Grayscale, Green channel.	
	Generative Adversarial Network (GAN) based method: Cycle GAN, Cycle-CBAM.	
Denoising and Normalization	Denoising: Median filtering, Mean filter, Gaussian filtering, Wiener filter. Normalization: Intensity normalization.	To remove potential noise of image 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.	
	GAN-based image generation.	

6. DL Related Model for DR Classification:

An incredible history of image processing and interpretation, including clinical imaging, has been

achieved by convolutional neural networks (CNNs), one sort of DL technology. CNNs learn extremely abstract characteristics in training images using labeled data, and these features are subsequently utilized to construct a classification model. When a substantial amount of tagged picture data is available, image classification tasks are best suited for this type of DL. The grade given by the clinicians might serve as the picture label in the setting of DR. This grading will then be used to train the CNN to diagnose DR. After training, the CNNs can forecast images that haven't been seen before. First CNNs were developed with significant help from the backpropagation algorithm. The LeNet-5 CNN architecture was first introduced in a significant milestone, the LeCun et al. publication from 1998 [33]. The layers of this architecture, which are still utilized in neural networks, include completely connected layers and activation functions. However, convolutional and pooling layers are two additional layers that set it apart from a standard neural network at the time. In handwritten character recognition, this enabled CNNs to outperform other automated methods. The amount of computing required and complications like the vanishing gradient problem meant that CNNs could not, regrettably, scale to larger problems. As a result, Support Vector Machines (SVMs) were the preferred approach for computer vision research in the 1990s since they appeared to provide superior outcomes and more robust theoretical underpinnings [34]. Not until 2006, when Geoffrey Hinton released a paper titled "deep belief nets," was the phrase "deep learning" first used [35]. Because interconnected layers of perceptrons allows for the learning of more abstract features, the research of neural networks has become known as "deep learning". Using this technique, people could train neural networks that were deeper and had more layers than perceptrons. DL saw an exponential increase in attention when it eventually produced groundbreaking results on speech recognition tasks in 2012 [36], after its newly discovered feasibility rekindled interest in the discipline. During that year, Geoffrey Hinton and his team of researchers would present their best algorithms of DL for the popular ImageNet Large Scale Visual Recognition Challenge (ILSVRC), that sought for identify photos which belong to the ImageNet dataset [8]. DL has been able to improve image classification problems due to advances in theory, the use of GPUs, and the availability of massive labelled datasets. Websites like Kaggle are making more public data available for analysis [37], and the viability of analyzing this data has increased [38]. These improvements have allowed neural networks particularly and DL in general, better suited for complicated computer vision tasks. CNN architectures were investigated as a result of the groundbreaking CNN AlexNet, which Krizhevsky et al. [8] introduced to the machine learning community in 2012. With the introduction of the backpropagation technique and the growing GPU power, CNNs could now use more than just a few hidden layers. Above all,

CNNs were no longer as vulnerable to the vanishing gradient problem after being trained by the gradient-based learning method [38]. The creation of AlexNet and other benchmarking CNN designs, such as VGG [38], laid the groundwork for CNN research to handle progressively difficult computer vision applications. These days, CNNs are frequently employed in numerous cutting-edge image classification applications [39, 40, 41, 42]. Given that CNNs have continued to develop and become increasingly effective in recent years, research in this area is still relatively young. The ideal architectures and hence, the layering for computer vision applications is still up for debate among researchers. Increasing depth has been the main emphasis of CNN model development in an effort to increase classification accuracy. There seems to be a direct correlation between improving accuracy and layer depth (Table 3).

Table (3): The key features of some CNN models.

Model input	Size trainable (pixels)	Parameters (millions)	Layers count
AlexNet	256x256	60	8
VGG	224x244	2.5	19
GoogLeNet	244x244	12	22
InceptionV3	299x299	23.8	159

Overfitting, though, could result from the depth and parameter increases [43]. In addition to being generalizable to different datasets, the architecture of the VGG has demonstrated resilience against overfitting [44]. Longer training periods and increased processing power are also necessary because to the rise in layer depth and parameters. Because of the convolution layers' nature, the architectures' parameters grow exponentially with the size of the image. Table 4 presents a summary of popular architectures. Summaries of the most frequently employed DL algorithms for binary and classification of multiple classes may be found in Tables 5 and 6, respectively.

Even with impressive progress, DL in DR classifications still faces numerous obstacles. Initially, there is a deficiency of more efficient labeled data. The data utilized for DR classification at the moment are primarily taken from publicly available databases, and the majority of them only cover one ethnicity or the target condition, making it impossible for them to accurately represent people's varied ethnic backgrounds as well as real clinical status. One typical issue with several public databases is insufficient data, or there may be imbalances across classifications. In clinical practice, labeling and grading consistency can be challenging to get. Achieving appropriate labeling and grading in clinical practice can be difficult at times. Secondly, it is challenging to analyze DL models for clinical use. DL models, which use multi-layer non-linear structures to extract visual characteristics for recognition and classification, are often criticized for their lack of transparency.

Table (4): Frequently utilized DL algorithms in DR classification.

DL techniques	Architectures	Strength	Weakness	Ref.

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.	[45-47]
	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.	[48, 53, 54, 55]
	GoogLeNet (Inception V1)	Inception model to extract features at different scales to increase the 1x1 convolutional kernels for dimensionality reduction reduces the computational complexity	Lager number of parameters (5 million); heterogeneous topology between the inception blocks.	[45, 46, 51]
	Inception V3	Factorization using 1xn and nx1 convolutional kernels instead of nxn kernels to diminish representational bottleneck; RMSProp optimizer to accelerate the training.		[56, 49-51, 57]
	Inception V4	Reduction blook for pooling data		[58, 59]
	ResNet	Residual block with a shortcut connection to reduce parameters and accelerate training convergence	High computational cost and requires more powerful hardware support.	[52, 60, 61]
	Inception ResNet	Inception v3 or v4 combines with residual connection to improve computational efficiency	Accuracy is not significantly improved	[62, 63]

Table (5): A list of the main DL models based on binary DR classification.

Target class	Dataset	DL model	Model performance			Year	Ref.
			Accuracy	Sensitivity	Specificity		
Any DR/ No DR	EyePACS	VGG16	-	93%	85%	2018	[64]
PDR/ Non PDR	EyePACS	WP-CNN	94.23%	90.94%	95.74%	2019	[65]
Any DR/No DR	MESSIDOR-1	CNN	-	90.4%	91.0%	2020	[66]
Any DR/No DR	MESSIDOR-1, and APTOS 2019	Pre-trained Inception-ResNet-V2	82.18%	-	-	2020	[67]
Sever DR/ Non-sever DR	APTOS 2019	Inception V3	-	92.5%	90.7%	2022	[50]
Any DR/ No DR	DRR	pre-trained ResNet 18	-	86.53%	86.72%	2022	[68]
Any DR/ No DR	APTOS 2019	CNN	94.6%	86%	96%	2022	[69]
Any DR/ No DR	EyePACS and APTOS 2019	LB CNN based on ResNet 18	97.4%	94.6%	96.6%	2022	[70]
Any DR/ No DR	APTOS 2019	DenseNet169 and convolutional block attention module	97%	97%	98.3%	2022	[71]

Table (6): A description of the main DL models using multi-class DR classification.

Target class	Dataset	DL model	Model performance			Year	Ref.
			Accuracy	Sensitivity	Specificity		
4 classes: Normal, and Non PDRs.	EyePACS-1, MESSIDOR-2, and DIARETDB0	Pre-trained VGG16	96.6%	-	-	2022	[72]
5 classes: Normal, Non PDRs.	APTOS 2019, MESSIDOR -2	Supervised learning with Xception encoder	84.36%	-	-	2022	[73]
5 classes: Normal, Non PDRs.	APTOS 2019	Ensemble method based on pre-trained VGG16 VGG19	87%	-	-	2022	[74]
4 classes: Normal, and Non PDRs.	MESSIDOR, STARE and DRIVE	Deep LSTM-RFO algorithm	97.89%	98.47%	97.43%	2022	[75]
5 classes: Normal, Non PDRs and PDR.	Kaggle DR 2015 and APTOS 2019	PCNN	97.27%	-	-	2023	[76]
5 classes: Normal, Non PDRs and PDR	Kaggle	Revised ResNet50	74.32%	-	-	2023	[77]
5 classes: Normal, Non PDRs and PDR	APTOS 2019	Tripple- DRNet	92.08%	-	-	2023	[78]
5 classes: Normal, Non PDRs and PDR	APTOS 2019 and DDR	DenseNet-121- rendered model	98.36%	-	-	2023	[79]
5 classes: Normal, Non PDRs and PDR	IDRiD	SqueezeNet and DCNN	91.1%	89.8%	91.3%	2023	[80]

5 classes: Normal, Non PDRs and PDR	APTOS 2019 and EyePACS	MobileNet V3-Small	98.4%	-	-	2023	[81]

The exact process of architecture and its clinical implications were not fully understood. Third, the DL used in clinical practice involved ethical constraints. Errors in diagnostic results utilizing DL approaches remain unclear. Even with impressive progress, deep learning in DR classifications still faces numerous obstacles. Initially, there is a deficiency of more efficient labeled data. The data utilized for DR classification at the moment are primarily taken from publicly available databases, and the majority of them only cover one ethnicity or the target condition, making it impossible for them to accurately represent people's varied ethnic backgrounds as well as real clinical status. One typical issue with several public databases is insufficient data, or there may be imbalances across classifications. In clinical practice, labeling and grading consistency can be challenging to get. Achieving appropriate labeling and grading in clinical practice can be difficult at times. Secondly, it is challenging to analyze DL models for clinical use. DL models, which use multi-layer non-linear structures to extract visual characteristics for recognition and classification, are often criticized for their lack of transparency. The exact process of architecture and its clinical implications were not fully understood. Third, the DL used in clinical practice involved ethical constraints. Errors in diagnostic results utilizing deep learning approaches remain unclear.

6. Conclusion:

Our paper summarizes current developments in fundus image-based DL classifications research. Important outcomes can be attained. First, it is commonly acknowledged that transfer learning is an excellent way to address the issue of not having enough data samples available for model training. There are numerous pretrained networks accessible to support DR analysis thanks to transfer learning techniques. Improved training time and robustness of the model are achievable. Using features from the model that has already been trained can improve the model's resilience and training time. The most often used datasets for DR classification are EyePACS and APTOS 2019. Third, two lightweight DL architectures, SqueezeNet and MobileNet, have been presented for data-restricted and computationally-intensive DR classification problems. With these architectures, complex image processing parameters can be reduced dramatically without sacrificing model accuracy. Though human accuracy in the diagnosis and assessment of DR cases has been achieved by DL, real-world clinical processes are still a long way off. Further developments are required in the areas of the model's interpretability, the affordability and dependability of DR screening techniques, and the credibility of ophthalmologists.

7. References:

[1] I Kocur and S Resnikoff. Visual impairment and blindness in Europe and their prevention. British Journal of Ophthalmology, 86(7):716–722, 2002.

[2] Jennifer Evans, Cleone Rooney, Fred Ashwood, Nairupa Dattani, and Richard Wormald. Blindness and partial sight in England and wales: April 1990-march 1991. Health Trends, 28(1):5–12, 1996.

[3] Strategic Policy Sector et al. State of the nation 2012. Science Technology and Innnovation Council, 2013.

[4] NHS. Diabetic eye screening: Your screening results. https://www.nhs.uk/conditions/diabetic-eye-screening/#your-screening-results, 2018. [Online; accessed 23/01/2024].

[5] Bo Wu, Weifang Zhu, Fei Shi, Shuxia Zhu, and Xinjian Chen. Automatic detection of microaneurysms in retinal fundus images. Computerized Medical Imaging and Graphics, 55:106 – 112, 2017. Special Issue on Ophthalmic Medical Image Analysis.

[6] G S Scotland, P McNamee, A D Fleming, K A Goatman, S Philip, G J Prescott, P F Sharp, G J Williams, W Wykes, G P Leese, and J A Olson. Costs and consequences of automated algorithms versus manual grading for the detection of referable diabetic retinopathy. British Journal of Ophthalmology, 94(6):712–719, 2010.

[7] A Tufail, V Kapetanakis, S Salas-Vega, C Egan, C Rudisill, C Owen, and Rudnicka R. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technology Assessment, 20(92):172, 2016.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, Advances in Neural Information Processing Systems 25, pages 1097–1105. Curran Associates, Inc., 2012.

[9] C. Szegedy, Wei Liu, Yangqing Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 1–9, June 2.

[10] Daniel Shu Wei Ting, Gemmy Chui Ming Cheung, and Tien Yin Wong. Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review. Clinical & Experimental Ophthalmology, 44(4):260–277, 2015.

[11] World Health Organization. World health organization - global report on diabetes, 2016.

[12] Deborah Broadbent, JA Scott, Jiten Vora, and S P Harding. Prevalence of diabetic eye disease in an inner city population: The Liverpool Diabetic Eye Study. Eye (London, England), 13 (Pt 2):160–5, 05 1999.

[13] Rohini Mathur, Ian Douglas, Krishnan Bhaskaran, and Liam Smeeth. Diabetic eye disease: A UK incidence and prevalence study. British Medical Journal, 2017. [Online; accessed 25/12/2023].

[14] Wei Wang and Amy C. Y. Lo. Diabetic retinopathy: Pathophysiology and treatments. International Journal of Molecular Sciences, 19(6), 2018.

[15] Abdhish R Bhavsar and Neal H Atebara. Diabetic retinopathy. Medscape, 2019.

[16] Donald S. Fong, Lloyd P. Aiello, Frederick L. Ferris, and Ronald Klein. Diabetic retinopathy. Diabetes Care, 27(10):2540–2553, 2004.

[17] Health Jade. Vision in diabetic retinopathy patient, 2018. Online; accessed 24/07/2023.

[18] 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.

[19] 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.

[20] 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.

[21] National Eye Insitute. Structure of the human eye. https://nei.nih.gov/ health/diabetic/retinopathy, 2018. [Online; accessed 24/04/2023].

[22] 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.

[23] 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.

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

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

[26] 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.

[27] (2007). Diaretdb1 Dataset. [Online]. Available: http://www.it.lut.fi/ project/imageret/diaretdb1/

[28] 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.

[29] E-Ophtha. Accessed: Apr. 2023. [Online]. Available: http://www. adcis.net/en/Download-Third-Party/E-Ophtha.html

[30] 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.

[31] 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.

[32] 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.

[33] Yann LeCun, Bernhard E. Boser, John S. Denker, Donnie Henderson, R. E. Howard, Wayne E. Hubbard, and Lawrence D. Jackel. Handwritten digit recognition with a backpropagation network. In D. S. Touretzky, editor, Advances in Neural Information Processing Systems 2, pages 396–404. Morgan-Kaufmann, 1990.

[34] Aurlien Gron. Hands-On Machine Learning with Scikit-Learn and Tensor Flow Concepts, Tools, and Techniques to Build Intelligent Systems. O'REILLY, 2017.

[35] Geoffrey E. Hinton and Simon Osindero. A fast learning algorithm for deep belief nets. Neural Computation, 18, 2006.

[36] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine, 29(6):82–97, Nov 2012.

[37] Kaggle: Platform for predictive modelling and analytics competitions. Available at https://www.kaggle.com/. 245

[38] Nada Elgendy and Ahmed Elragal. Big data analytics: A literature review paper. In Petra Perner, editor, Advances in Data Mining. Applications and Theoretical Aspects, pages 214–227. Springer International Publishing, 2014.

[39] Md. Zahangir Alom, Tarek M. Taha, Christopher Yakopcic, Stefan Westberg, Mahmudul Hasan, Brian C. Van Esesn, Abdul A. S. Awwal, and Vijayan K. Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. Computing Research Repository, abs/1803.01164, 2018.

[40] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.

[41] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385, 2015.

[42] Gao Huang, Zhuang Liu, and Kilian Q. Weinberger. Densely connected convolutional networks. Computing Research Repository, abs/1608.06993, 2016.

[43] Md. Zahangir Alom, Tarek M. Taha, Christopher Yakopcic, Stefan Westberg, Mahmudul Hasan, Brian C Van Esesn, Abdul Awwal, and Vijayan Asari. The history began from alexnet: A comprehensive survey on deep learning approaches. Arxiv, 03 2018.

[44] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large scale image recognition. Computing Research Repository, abs/1409.1556, 2014.

[45] 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.

[46] 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.

[47] 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.

[48] 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.

[49] 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.

[50] M. T. Hagos and S. Kant, "Transfer learning-based detection of DR from small dataset," arXiv:1905.07203.

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

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

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

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

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

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

[57] 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 DR using deep learning," Lancet Digit. health, vol. 3, no. 1, pp. e10–e19, Nov. 2020.

[58] 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 DR and diabetic macular oedema in retinal fundus photographs," Eye, vol. 36, no. 7, pp. 1433–1441, Jul. 2021.

[59] 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 DR detection," IEEE Access, vol. 7, pp. 150530–150539, 2019.

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

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

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

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

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

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

[66] 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 DR—Artificial intelligence versus clinician for screening," Indian J. Ophthalmol., vol. 68, pp. 398–405, Feb. 2020.

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

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

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

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

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

[72] A. Bilal, G. Sun, S. Mazhar, A. Imran, and J. Latif, "A transfer learning and U-Net-based automatic detection of DR from fundus images," Comput. Methods Biomechanics Biomed. Eng., Imag. Visualizat., vol. 10, no. 6, pp. 663–674, Nov. 2022.

[73] 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 DR and its severity levels from fundus images," Comput. Biol. Med., vol. 146, Jul. 2022, Art. no. 105602.

[74] Y. Kale and S. Sharma, "Detection of five severity levels of DR using ensemble deep learning model," Multimedia Tools Appl., vol. 82, no. 12, pp. 19005–19020, Dec. 2022.

[75] R. Pugal Priya, T. Saradadevi Sivarani, and A. Gnana Saravanan, "Deep long and short term memory based red fox optimization algorithm for DR detection and classification," Int. J. Numer. Methods Biomed. Eng., vol. 38, no. 3, Mar. 2022, Art. no. e356.

[76] M. Nahiduzzaman, M. R. Islam, M. O. F. Goni, M. S. Anower, M. Ahsan, J. Haider, and M. Kowalski, "DR identification using parallel convolutional neural network based feature extractor and ELM classifier," Expert Syst. Appl., vol. 217, May 2023, Art. no. 119557.

[77] C.-L. Lin and K.-C. Wu, "Development of revised ResNet-50 for DR detection," BMC Bioinf., vol. 24, no. 1, p. 157, Apr. 2023.

[78] M. Jian, H. Chen, C. Tao, X. Li, and G. Wang, "Triple-DRNet: A triplecascade convolution neural network for DR grading using fundus images," Comput. Biol. Med., vol. 155, Mar. 2023, Art. no. 106631.

[79] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Deep learning enhanced DR image classification," Digit. Health, vol. 9, Aug. 2023, Art. no. 20552076231194942.

[80] S. Zulaikha Beevi, "Multi-level severity classification for DR based on hybrid optimization enabled deep learning," Biomed. Signal Process. Control, vol. 84, Jul. 2023, Art. no. 104736.

[81] A. R. Wahab Sait, "A lightweight DR detection model using a deep-learning technique," Diagnostics, vol. 13, no. 19, p. 3120, Oct. 2023.