VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey

This is the Published version of the following publication

Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey

RUBINA SARKI, **KHANDAKAR AHMED**, **(Senior Member, IEEE)**, **HUA WANG**, **(Member, IEEE), AND YANCHUN ZHANG**, **(Member, IEEE)**
Institute for Sustainable Industries and Liveable Cities, Victoria University, Footscray, VIC 3011, Australia

Corresponding author: Rubina Sarki (rubina.sarki@live.vu.edu.au)

ABSTRACT Diabetes Mellitus, or Diabetes, is a disease in which a person's body fails to respond to insulin released by their pancreas, or it does not produce sufficient insulin. People suffering from diabetes are at high risk of developing various eye diseases over time. As a result of advances in machine learning techniques, early detection of diabetic eye disease using an automated system brings substantial benefits over manual detection. A variety of advanced studies relating to the detection of diabetic eye disease have recently been published. This article presents a systematic survey of automated approaches to diabetic eye disease detection from several aspects, namely: i) available datasets, ii) image preprocessing techniques, iii) deep learning models and iv) performance evaluation metrics. The survey provides a comprehensive synopsis of diabetic eye disease detection approaches, including state of the art field approaches, which aim to provide valuable insight into research communities, healthcare professionals and patients with diabetes.

INDEX TERMS Diabetic eye disease, diabetic retinopathy, deep leaning, glaucoma, image processing, macular edema, transfer learning.

I. INTRODUCTION

Diabetic Eye Disease (DED) comprises a group of eye conditions, which include Diabetic Retinopathy, Diabetic Macular Edema, Glaucoma and Cataract [1]. All types of DED have the potential to cause severe vision loss and blindness in patients from 20 to 74 years of age. According to the International Diabetes Federation (IDF) statement, about 425 million citizens worldwide suffered from diabetes in 2017. By 2045, this is forecast to increase to 692 million [2]. Medical, social and economic complications of diabetes impact substantially on public health, with diabetes being the world's fourth-largest cause of death [3]. The effects of diabetes can be observed in different parts of a person's body, including the retina. Fig. 1 shows the normal anatomical structures of the retina. Fig. 2 illustrates a complication of DED in a retina. Serious DED begins with an irregular development of blood vessels, damage of the optic nerve and the formation of hard exudates in the macula region. Four types of DED threaten

eye vision, and they are briefly described in the following subsection.

Diabetic Retinopathy (DR) is caused by damage to blood vessels of the light sensitive tissue (retina) at the back of the eye. The retina is responsible for sensing light and sending a signal to brain. The brain decodes those signals to see the objects around [4]. There are two stages of DR: early DR and advanced DR. In early DR, new blood vessels do not developing (proliferating) and this is generally known as non-proliferative diabetic retinopathy (NPDR). The walls of the blood vessels inside the retina weaken due to NPDR. Narrower bulges (microaneurysms) protrude from the narrower vessel surfaces, often dripping fluid and blood into the eye. Large retinal vessels also start dilating and become irregular in diameter. As more blood vessels become blocked, NPDR progresses from mild to severe. Depending on the severity, the retina's nerve fibres may begin to swell. The central part of the retina (macula) often swells (macular edema); a disease requiring treatment. NPDR is divided into three stages, namely: mild, moderate and severe [5]. Advanced DR is called proliferative diabetic retinopathy (PDR). In this case, the damaged blood vessels leak the transparent

---

The associate editor coordinating the review of this manuscript and approving it for publication was Haiyong Zheng.
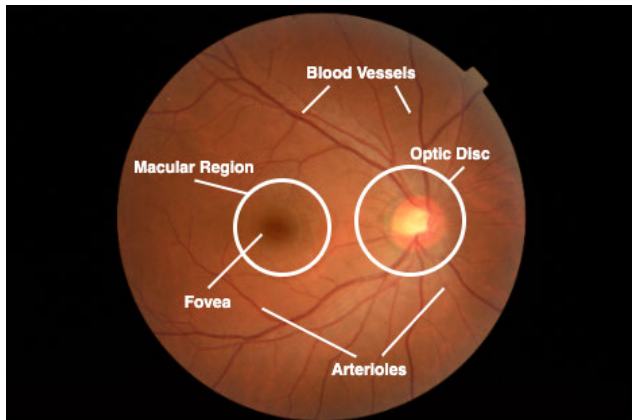
FIGURE 1. Anatomical structures of the retina.

jelly-like fluid that fills the centre of the eye (vitreous) causing the development of abnormal blood vessels in the retina. Pressure can build up in the eyeball because the newly grown blood vessels interrupt the normal flow of the fluid. This can damage the optic nerve that carries images from the eye to the brain, leading to glaucoma.

Glaucoma (Gl) is an ocular disease that damages the optic nerve that links the eye to the brain. When the fluid pressure inside the eye, known as intraocular pressure (IOP), is high, the optic nerve is impaired [6]. An increase in blood sugar doubles the chances of Gl, which leads to blindness and a loss of vision if not detected early. Gl can be classified into three types based on the size of the enlarged optic nerve head or optic disc and Cup-to-Disc Ratio (CDR), or cupping. The stages of Gl are mild, moderate and severe [7].

Diabetic Macular Edema (DME) occurs when fluid builds up in the centre of the retina (macula) due to damage to the blood vessels. The macula is responsible for sharp, straight-ahead vision. Fluid buildup causes swelling and thickening of the macula which distorts vision [4]. The stages of DME can be categorized into mild, moderate and severe based on the following points [8]:

- Retinal thickening of the fovea at or below 500 μ or 1/3 of its disc diameter
- Hard exudates, with subsequent retinal thickening, at or within 500 μ of the fovea
- Retinal thickening at a size that is greater than one disc diameter (1500 μ), and which is within one fovea disc diameter.

Cataract (Ca) is the degeneration of the lens protein due to high sugar level causing blurry lens growth, which in turn leads to blurred vision. Diabetic people are more prone to growing cloudy lenses and developing Ca earlier than non-diabetic people. Usually Ca is graded into four classes: non-cataractous, mild, moderate and severe [9].

Patients suffering from diabetes display a significantly higher predisposition develop DED. As a consequence, early detection of DED has become paramount in preventing vision loss in adults and children. Studies have already shown that 90% of patients with diabetes can avoid DED

development through early detection [10]. Manual detection of DED involves no computer assistance, resulting in longer waiting times between early diagnosis and treatment. Moreover, the initial signs of DED are so minute that even an expert may struggle with its identification.

Advancements in Artificial Intelligence (AI) offer many advantages to automated DED detection over the manual approach. They include a reduction in human error, time-efficiency and detection of minute abnormalities with greater ease. Automated DED detection systems can be assembled through joint image processing techniques using either Machine Learning (ML) or Deep Learning techniques (DL).

In DL approaches, images with DED and without DED are collected. Then, the image preprocessing techniques are applied to reduce noise from the images and prepare for the feature extraction process. The pre-processed images are input to DL architecture for the automatic extraction of features and their associated weights to learn the classification rules. The features weights are optimized recursively to ensure the best classification results. Finally, the optimized weights are tested on an unseen set of images. This type of architecture demands a large number of images for training. Therefore, a limited number of images can severely restrict its performance.

DL techniques require a substantial amount of computational memory and power. Normally, to develop and evaluate the classification model, DL architecture requires a Graphical Processing Unit (GPU). In real world DL applications, this assumption does not always hold. Training images using the DL model can be costly, challenging in terms of *annotated* data collection, and time and power consuming. To account for the above mentioned shortcomings, the approach called Transfer Learning (TL), or Knowledge Transfer (KT), has been introduced by the researchers. In TL, previously derived knowledge (e.g. in terms of features extracted) can be re-adapted to solve new problems. Not only does TL drastically reduce the training time, it also reduces the need for a large amounts of data. The latter point proves particularly convenient in niche applications where high-quality input images annotated by specialists are often limited or expensive to obtain.

A. MOTIVATION
As mentioned above, DL and TL techniques have their advantages and disadvantages however, several researchers have used these methods to build automatic DED detection systems in recent years. Overall, there are very few review studies published in academic databases which simultaneously address all of the types of DED detection. Thus, this literature review is essential to collate the work in the DED detection field.

Ting *et al.* [11] published a review article focusing on eye conditions such as diabetic retinopathy, glaucoma, and age-related macular diseases. They selected papers published between 2016 and 2018 and summarised them in their report. They summarized those papers which used fundus and
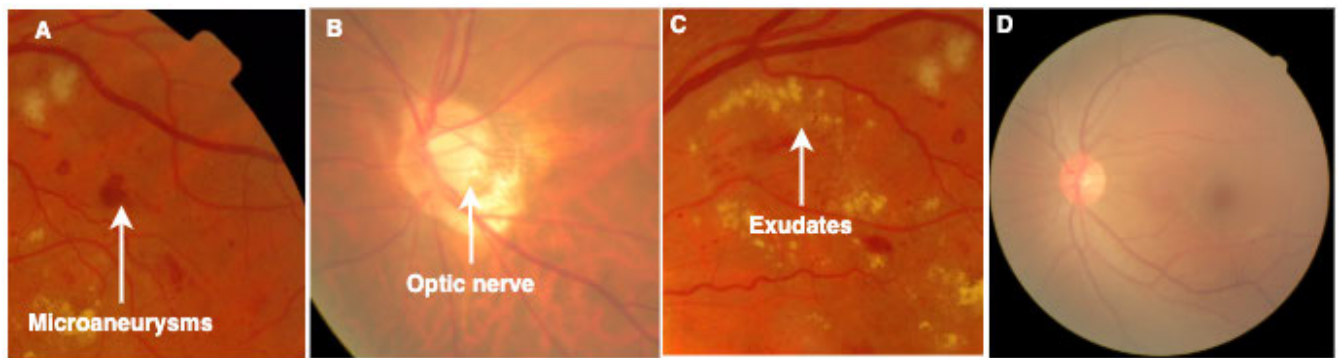
FIGURE 2. Complications of DED in retina; A. Microaneurysms, narrow bulges (Diabetic Retinopathy), B. Optic nerve damage (Glaucoma), C. Exudates with retinal thickening (Diabetic Macular Edema), D. Degeneration of lens (Cataract).

FIGURE 3. Research method flowchart.

optical coherence tomography images, and TL methods. Their research did not include current (2019-2020) publications that incorporated TL methods into their approach, and they omitted the identification of eye cataract disease from their study scope. Similarly, Hogarty *et al.* [12] provided a review of current state articles using AI in Ophthalmology, but their focus lacked comprehensive AI methodologies. Mookiah *et al.* [13], reviewed computer aided DR detection studies, which are largely DR lesion based. Another author, Ishtiaq *et al.* [14], reviewed comprehensive DR detection methods from 2013 to 2018 but their review lacked studies from 2019 to 2020. Hagiwara *et al.* [15], reviewed an article for the computer aided diagnosis of Gl using fundus images. They addressed computer aided systems and systems focused on optical disc segmentation. There are a variety of studies using DL and TL methods for Gl detection that have not discussed in their review paper. It is, therefore, important to review papers that consider existing approaches to DED diagnostics. In fact, most scholars in their review article did not address the period of publication years covered by their studies. Current reviews were too narrow, either in terms of disease (DR, Gl, DME and Ca) or in aspects of methodology (DL and ML). Therefore, to address the limitations of the above-mentioned studies, this article offers a thorough analysis of both DL and TL approaches to automated DED detection published between 2014 and 2020 to cover the current DR detection methods built through DL or TL based approaches.

B. CONTRIBUTION

To provide a structured and comprehensive overview of the state of the art in DED detection systems using DL, the proposed paper surveys the literature from the following perspectives:

1) *Datasets available for DED*
2) *Preprocessing techniques applied to fundus images for DED detection*

3) *DL approaches proposed for DED detection*
4) *Performance measures for DED detection algorithm evaluation.*

The arrangement of this article is as follows. Section II presents the research method followed after surveying the articles. Section III analyses the papers based on the datasets used in their study. Section IV addresses the image-processing techniques used in the prior work. Section V analyses the articles based on the classification methods employed. Section VI discusses the findings and observations. Section VII discusses the gaps and future directions. Finally, Section VIII concludes the paper.

II. RESEARCH METHOD

The overall research method followed is shown in Fig. 3. Initially, a keyword search was conducted using 10 academic databases considering our specific review target. Seven filters were applied to select the primary review target. Afterwards, the selected articles were critically analysed and grouped into three categories based on the following aspects, namely: *(i) papers employing TL*, *(ii) papers proposing a new DL network* and *(iii) papers discussing with DL and ML combined.*

Selection of Articles A systematic review of automated detection methods of various Diabetic Eye Diseases from databases including *IEEE Xplore, MedLine, Scopus, Science Direct, Springer, ACM Digital Library, PubMed, Web of Science and Google Scholar* was performed. The subsequent seven filters applied were: *(i) Target keywords*, *(ii) Publication year*, *(iii) Publication type*, *(iv) Duplicate check*, *(v) Article title, Abstract and Keyword screening for article selection*, *(vi) References of selected articles checked* and *(vii) Final quality assessment of selected article*. Review target keywords were searched using 'AND' Boolean operator and included: *"deep learning", "transfer learning", "image processing", "image classification", "fundus images", "diabetic eye disease", " diabetic retinal disease", "diabetic*
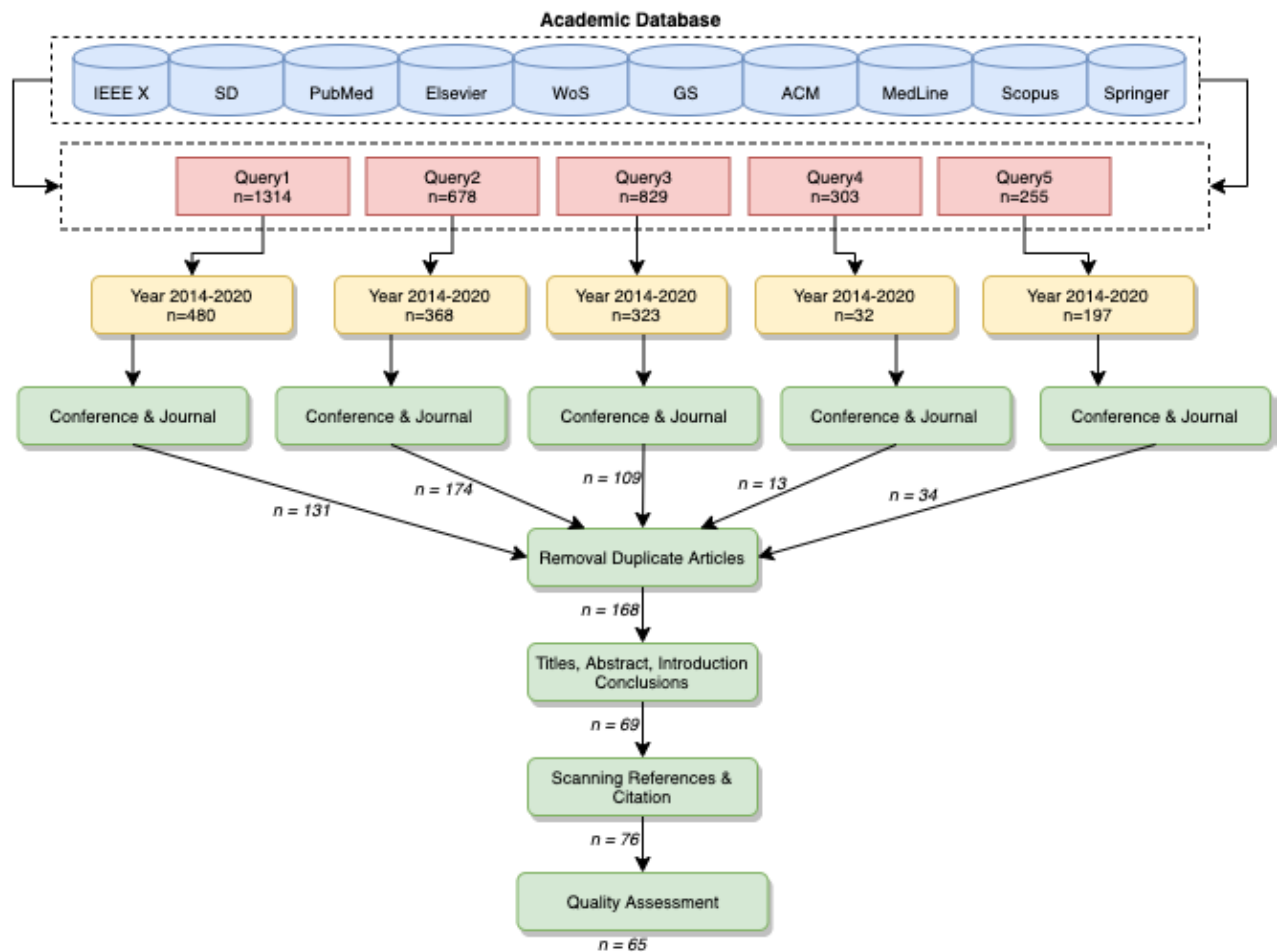
FIGURE 4. Search and filter results: A. Query1 (Q1) = *diabetic eye disease, fundus images, image processing, image classification, deep learning, transfer learning*; B. Query2 (Q2) = *diabetic retinopathy, fundus images, image processing, image classification, deep learning, transfer learning*; C. Query3 (Q3) = *glaucoma, fundus images, image processing, image classification, deep learning, transfer learning*; D. Query4 (Q4) = *diabetic macular edema, fundus images, image processing, image classification, deep learning, transfer learning*; E. Query5 (Q5) = *cataract, fundus images, image processing, image classification, deep learning, transfer learning.*

retinopathy", "glaucoma", "diabetic macular edema" and "cataract".

Papers published between *April* 2014 and *January* 2020 were considered eligible for this study due to rapid advances in the field. We then narrowed our search to *Conference Papers* and *Journal Articles*. After the selection process, we encountered several duplicates as a result of using 10 different databases. After duplicates removal, titles, abstracts and conclusions of the remaining publications were carefully read. 69 articles were obtained focusing on fundus images, DL methods and classification of DED. We studied the bibliography and citation of the selected 69 articles, in which we found 7 more articles for the potential inclusion. Finally, during a quality assessment by reading 76 papers, our selection was narrowed down to 65 studies. The details of the process followed during our systematic review are presented in Fig. 4. We subsequently distributed the final sample of articles into three target groups. The distribution of 65 articles concerning the review target is represented in Table 1. The first group includes papers that use a pretrained network also

referred to as the TL Approach. The second group categorizes articles that use their own built in DL network to detect DEDs. Finally, the third group summarises the articles that use combined DL and ML methods.

III. DIABETIC EYE DISEASE DATASETS

The authors of the selected articles use private and public datasets which are divided into training and testing examples. The most common datasets used for the detection of DR are Kaggle and Messidor [77]. Authors in [19]–[24], [28], [29], [51]–[57], [59], [60], [78] used Kaggle data and [24], [25], [28], [30], [71], [72], [79] used Messidor [77] data. The Kaggle dataset consists of 88,702 images, of which 35,126 are used for training and 53,576 are used for testing. Messidor [77] is the most widely used dataset which consist 1,200 fundus images. The Kaggle and Messidor dataset, is labeled for DR stages. Table 2 describes the datasets included in the chosen articles, listed from the viewpoint of the individual DED analyzed; i.e. DR, Gl, DME, and Ca. The table contains the name of the DED, the name of the

TABLE 1. Selected articles for common objectives (review target).

Served purpose	DED	No. of Articles	References
Employed TL	DR	19	[16]–[34]
	Gl	15	[35]–[48]
	DME	3	[33], [34], [49]
	Ca	1	[50]
Proposed New DL network	DR	11	[51]–[61]
	Gl	5	[62]–[66]
	DME	3	[67]–[69]
	Ca	2	[9], [70]
DL combined with ML	DR	3	[71]–[73]
	Gl	1	[74]
	DME		
	Ca	2	[75], [76]

Legend: DL = Deep Learning, TL = Transfer Learning DED = Diabetic Eye Disease, DR = Diabetic Retinopathy, Gl = Glaucoma, DME = Diabetic Macular Edema, Ca = Cataract.

dataset, the summary of the particular dataset, the sources of the publications that used the dataset and finally, the path where the dataset can be retrieved (if accessible publicly)

IV. IMAGE PREPROCESSING TECHNIQUES IN SELECTED ARTICLES

Images are subjected to numerous image preprocessing steps for visualization enhancement. Once the images are brighter and clearer, a network can extract more salient and unique features. A brief description of the preprocessing techniques used by the researchers addressed in this section. Green channel on the RGB color space provides a better contrast when compared to the other channels. In most of the image preprocessing techniques, green channel extraction is employed. The green channel image produces more information than blue and red channels. For instance, Li *et al.* [24] extracted the green channel of the image for exudates detection, where the exudates reveal better contrast from the background.

Another popular image preprocessing technique is contrast enhancement. The application of contrast enhancement further improves the contrast on a green channel image. To improve the contrast of the image, contrast enhancement is employed to the green channel of the image. For example, again Li *et al.* [24] have enhanced the contrast on the extracted green channel by employing the Contrast Limited Adaptive Histogram Equalization (CLAHE) method. This enhances the visibility of exudates of a green channel image. Normally, after contrast enhancement, illumination correction is implemented to improve the luminance and brightness of the image. A noise removal filter like Gaussian Filtering is then applied to smooth out the image.

The resizing of an image is another popular method of image preprocessing. The image is scaled down to a low resolution image according to the appropriate system. Li *et al.* [24] resized their images with various sizes to the same pixel resolution of 512×512. Similarly, Li *et al.* [25] resized their image to 224×224 pixel resolution, for all the pretrained CNN models that used 224×224 size resolution images. The resolution of an image is resized into the resolution required by the network in use.

Researchers often have to eradicate and mask the blood vessels and optical discs so that they are not classified as wrong DED lesions. Many DED datasets consist of images with a black border, with researchers generally preferring to segment the meaningless black border to focus on the region of interest (ROI). For example, Li *et al.* [24] removed the black border of fundus images using the thresholding method to further focus on the Region Of Interest (ROI).

Image augmentation is applied when there is an image imbalance (as typically observed in real world settings). Images are mirrored, rotated, resized and cropped to produce cases of the selected images for a class where the number of images is lower than the other large proportion of healthy retina images in comparison with DED retina images. Augmentation is a common strategy for enhancing outcomes and preventing overfitting. It is observed that the distribution of the Kaggle dataset is uneven. The Kaggle dataset includes 35,126 fundus images annotated as No DR (25810), Mild DR (2443), Moderate DR(5292), Severe DR(873) and Proliferative DR(708). Thus, Li *et al.* [24], An *et al.* [38], Nguyen *et al.* [31], Xu *et al.* [56], Pires *et al.* [60], Gargeya and Leng [52], Ghosh *et al.* [53], van Grinsven *et al.* [28], Quellec *et al.* [22] used the Kaggle dataset and the adopted augmentation technique to balance the dataset. Sometimes the RGB image is transformed into a greyscale image accompanied by further processing. Grayscale conversion is mostly used in approaches where ML is used.

V. DIABETIC EYE DISEASE CLASSIFICATION TECHNIQUES

In this section, we review the DL based approaches for DED detection. DL is defined as the extension of the ML with a multilayer network for extracting features. In DL architecture the term "deep" refers to the depth of the layers. The classification process is as follows: (i) The annotated dataset is split into testing and training samples for DL architecture, (ii) The dataset is preprocessed using image preprocessing techniques for quality enhancement and (iii) The preprocessed images are fed into DL architecture for features extraction and subsequent classification. Each layer in DL architecture considers the output of the previous layer as its input,

TABLE 2. Datasets available for automatic Diabetes Eye Detection with source (link).

DED	Dataset	Description	Reference	Link
DR	Kaggle	Dataset made available by EyePACS. It consists of 35,126 training images and 53,576 testing images (total of 88,702). These images are labelled with stages.	[19]–[24], [28], [29], [31], [51]–[57], [59], [60], [78]	urlhttps://www.kaggle.com/c/diabetic-retinopathy-detection/data
	Messidor	Dataset consists of 1,200 fundus images. The images were obtained from three ophthalmological branches in France. Of the 1,200 images, 800 images are with pupil dilation, while 400 are without pupil dilation. Images are labelled with DR stages. Deceneiere et al. [77]	[24], [25], [28], [30], [61], [71], [72], [79]	https://www.adcis.net/en/Download-Third-Party/Messidor.html
	Messidor-2	Dataset consists of 1,748 fundus images. Camera used was Topcon TRC NW6 non-mydriatic with 45 degrees field of view. Images are labelled with DR stages. Deceneiere et al. [77].	[16]	urlhttp://www.latim.univ-brest.fr/indexfce0.html
	STARE	Dataset consists of 400 fundus images. The images were taken with Topcon TRV-50 with 35 degrees field of view. Farnell et al. [80]	[17], [30]	http://www.cecas.clemson.edu/~ahoover/stare/
	DR_1	Dataset is presented by the Department of Ophthalmology, Federal University of Sao Paulo, Brazil and consists of 1,014 color fundus images (687 - normal, 327 - abnormal). Abnormal images are further split into 191 with red lesions, 245 with bright lesions and 109 with both red and bright lesions.	[24], [25]	http://www.recod.ic.unicamp.br/site/asdr
	APTOS	The APTOS 2019 repository incorporates 3662 fundus images classified into five levels (normal - 0, mild - 1, moderate - 2, extreme non-proliferative DR - 3, and proliferative DR - 4) as per the severity of DR labelled.	[32]	https://www.kaggle.com/c/aptos2019-blindness-detection/data
	E-optha	This dataset consists 47 images with exudates, 148 with microaneurysms and 268 images with no lesion.	[58]	http://www.adcis.net/en/third-party/e-ophtha/
	DeepDR	Dataset include 2696 images from 748 patients for classification of DR	[81]	https://isbi.deepdr.org/data.html
Gl	RIGA	Dataset contains three different sources: 1) MESSIDOR (dataset consists of 460 original images marked manually by six different ophthalmologists (total of 3220 marked images); 2) Bin Rushed (dataset contains 195 original images marked by six different ophthalmologists (total of 1,365 images); 3) Magrabi Eye Center (dataset contains 95 original images marked by six different ophthalmologists (total of 665 images). Almazroa et al. [82]	[36]	https://deepblue.lib.umich.edu/data/concern/data_sets/3b591905z?locale=en
	ORIGA	A quantified objective benchmarking method was proposed, focusing on optic disc and cup segmentation and Cup-to-Disc Ratio (CDR). ORIGA(-light) contains 650 retinal images annotated by trained professionals from Singapore Eye Research Institute. A wide collection of image signs, critical for Gl diagnosis were annotated Zhang et al. [83].	[62]	Publicly unavailable.
	Drishti-GS	Dataset contains a total of 101 images. The images were divided into 51 testing and 50 training examples. The images were marked by four eye experts and collected from Aravind Eye Hospital. Sivaswamy et al. [84]	[41], [64]	http://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-Dataset2/Dataset_description.php
	BIOMISA	Dataset contains 462 images collected a local hospital. TopCon TRC 50EX camera was used. Hassan et al. [85].	[86]	http://biomisa.org/index.php/glaucoma-database/
	REFUGE	This dataset consist of 1200 color fundus images divided into 1:1:1 ratio for training, validation, and testing. [45]	[65]	http://ai.baidu.com/broad/download?dataset=gon
	ODIR-2019	This dataset consist of eight types of ocular disease, consisting of 207 training classes, 30 off-site testing cases, 58 on-site testing cases.	[87]	https://odir2019.grand-challenge.org/dataset/
	DRIONS	Dataset consists of 110 colour digital retinal images from Ophthalmology Service at Miguel Servet Hospital, Saragossa (Spain). Carmona et al. [88].	[64], [71]	http://www.ia.uned.es/~ejcarmona/DRIONS-DB.html
	RIM-ONE (r1), RIM-ONE (r2), RIM-ONE (r3)	Dataset details: (i) $r1$ 40 Gl and 118 Non-Gl images, (ii) $r2$ 200 Gl and 225 Non-Gl images and (iii) $r3$ 74 Gl and 85 Non-Gl images. Fumero et al. [89]	[40], [64]	http://people.ee.ethz.ch/~cvlsegmentation/driu/downloads.html
	HRF	Contains fundus images by patient condition containing 15 healthy, 15 DR, and 15 glaucoma.	[40]	https://www5.cs.fau.de/research/data/fundus-images/
DME	HEI-MED	Dataset was obtained from Hamilton Eye Institute Macular Edema Data-set (HEI-MED). Dataset contains 169 fundus images to test and train for the detection of exudates and DME. Giancardo et al. [90].	[24]	https://github.com/lgiancaUTH/HEI-MED
	IDRiD	Dataset contains 516 images with both DME and DR cases. The severity of macular edema is based on the existence of hard exudates closer to fovea (macula center) region. This dataset contents 80 hard exudates images. Porwal et al. [91].	[92]	https://idrid.grand-challenge.org/Data/
	DRiDB	The retinal fundus databases including a description of all main anatomical structures including macula, blood vessels and optic disc is annotated. Prentavsic et al. [93].	[68]	https://ipg.fer.hr/ipg/resources/image_database
	CLEOPATRA	CLEOPATRA was a three phase randomised, parallel and single clinical experiment from fifteen opthalmic centres in the United Kingdom. Sivaprasad et al. [94].	[69]	unavailable
	Digifundus Ltd, Finland	Dataset is non-open, anonymous retinal data-set of diabetic patients. Dataset contains 41,122 labelled retinal color fundus images from 14,624 patients. Sahlsten et al. [49].	[49]	Publicly unavailable.
Ca	Beijing Tongren Hospital.	Dataset is composed of 5,620 standard fundus images from Beijing Tongren Eye Center	[9]	Publicly unavailable.
	Not Disclosed by Authors	Study consists of 5,408 preprocessed images as experimental dataset. Dataset contains 1,948 noncataractous images: 1,268 slightly mild, 496 mild, 616 medium, 540 slightly severe and 540 severe images.	[75]	Publicly Unavailable.
	Picture Archiving and Communication System (PACS)	In this dataset each fundus image is manually graded by the ophthalmologist as non, mild, moderate, or severe cataract. There are 767 noncataractous, 246 mild, 128 moderate and 98 severe images (total of 1,239).	[76]	Publicly Unavailable.

Legend: DED = Diabetic Eye Disease, Gl = Glaucoma, DME = Diabetic Macular Edema, Ca = Cataract, M = Messidor, K = Kaggle, NPDR = NonProliferative Diabetic Retinopathy, IRMA = Intraretinal Microvascular Abnormalities.

processes it and passes it onto the next layer. Many authors fine tune the hyperparameters of existing DL algorithms, such as VGG16 or CNN, to improve classification performance. Hyperparameter observed in this study is shown in Table 4

TABLE 3. Image preprocessing techniques employed in selected studies.

GCE	HE	ROI	CLAHE	CE	Re	Au	GSC	BVS	IR	IC	GF	References
✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	Li et al. [24]
✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	X. Li et al. [25], Al-Bander et al. [67]
✓	✗	✗	✗	✗	✓	✓	✗	✗	✗	✓	✗	Zhang et al. [9]
✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	Ran et al. [75]
✓	✓	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	Shaharum et al. [95]
✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	Abbas et al. [71], Yu et al. [58], Diaz et al. [39], De et al. [42], Gomez et al. [44]
✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	Yang et al. [76], Pratap et al. [50], Doshi et al. [51]
✗	✗	✗	✗	✗	✓	✗	✓	✗	✗	✗	✗	Sahlsten et al. [49]
✓	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	Sisodia et al. [78]
✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	Antal et al. [79]
✓	✗	✓	✗	✗	✗	✗	✗	✓	✗	✓	✗	Orlando et al. [72]
✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	Almazroa et al. [82]
✗	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	Chen et al. [62], Ceretinia et al. [40], Perdomo et al [26]
✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	Orlando et al. [41]
✗	✗	✗	✗	✗	✓	✗	✗	✗	✓	✗	✗	Phan et al. [35]
✗	✓	✗	✗	✗	✓	✗	✓	✗	✓	✗	✗	Dong et al. [70]
✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	Asaoka et al. [37], An et al. [38], Nguyen et al. [31], Xu et al. [56]
✗	✗	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	Pal et al. [64]
✗	✓	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	Hemanth et al. [61]
✗	✗	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗	Gondal et al. [19]
✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	Mansour et al. [21]
✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✓	Quellec et al. [22]
✗	✗	✗	✗	✓	✓	✓	✗	✗	✗	✗	✓	Van et al. [28]
✓	✗	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	Umapathy et al. [30]
✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	Diaz et al. [39]
✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	Gargeya et al. [52], Ghosh et al. [53]
✗	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	Jiang et al. [54]
✗	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗	Yang et al. [57]
✗	✓	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	Pires et al. [60]
✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	Singh et al. [47]
✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	Singh et al. [48]

Legend: GCE = Green Channel Extraction, HE = Histogram Equlization, ROI = Region of Interest, CLAHE = Contrast Limited Adaptive Histogram Equlization, CE = Contrast Enhancement, Re = Resize, Au = Augmentation, GSC = Grayscale Conversion, BVS = Blood Vessel Segmentation, IR = Image Rotation, IC = Illumination Correction, GF = Gaussian Filter.

Finally, the last layer of the architecture produces the required result, i.e. classification of DED as for the scope of the study. Out of 65 studies, 38 used TL, 21 used their proposed DL and six used a combination of DL and ML classifiers such as Random Forest(SF), Support Vector Machine(SVM), Backpropagation Neural Network (BPNN).

A. DL APPROACHES EMPLOYING TL

The concept of TL is based on the reuse of the features learned by DL models on the primary task and its adaptation to the secondary task. The idea is to reduce the computational complexity while training Neural Network architecture (resource intensive). Additionally, TL is found to be beneficial in cases where there is insufficient data to train a Neural Network from scratch (high volume of data required). Using TL, the parameters are initialized from the prior learning instead of random generation. Intuitively, the first layers learn to extract basic features such as edges, textures, etc., while the top layers are more specific to the task, e.g. blood vessels and exudates. Therefore, TL is commonly adopted in image recognition applications as the initial features extracted are shared regardless of the tasks. Table. 5 shows the records of works, which applied TL for the detection of DED. The details

regarding a particular type of DED recognition, network architecture and model used were further extracted. Additionally, the classification results were retrieved for the comparison between the studies and state of the art overview. Overall, 38 of the 65 studies adopted the TL approach for the detection of DED through DL (19-DR, 15-Gl, 3-DME and 1-Ca).

Diabetic Retinopathy Abràmoff *et al.* [16] used a CNN model based on *AlexNet* with Random Forests classifier for the detection of DR. Using Messidor-2 data they achieved area under the curve (AUC) of 98.0%, sensitivity of 96.8%, specificity of 87.0% and the predictive negative value was 99.0%. Choi *et al.* [17] used a STARE dataset [80] to perform the binary classification (normal and abnormal) for 10 retinal diseases. They used VGG-19 architecture with SGD optimizer with Random Forests classifier and achieved AUC of 90.3%, sensitivity of 80.3% and specificity of 85.5%. Ting *et al.* [18] used VGGNet architecture to classify DR and other diseases like Age-related Macular Degeneration (AMD) and Gl. They collected dataset from the Singapore National Diabetic Retinopathy Screening Program (SIDRP) from 2010 to 2013 and achieved AUC of 93%, specificity of 91.6% and sensitivity of 90.5%. For the Gl, they achieved AUC of 94.2%, specificity of 87.2% and sensitivity

TABLE 4. Different hyperparameters used in the selected studies.

R1	R2	R3	R4	R5	R6	References
VGG19	224×224	-	50	SGD	$1e-6$	[17]
o_O	448×448	36	-	Adam	$1e-4$	[22]
o_O	512×512	-	150	Adam	$1e-2$	[19]
AlexNet	512×512	-	130	-	$1e-2$	[24]
LeNet	48×48	64	30	-	$1e-2$	[96]
OxfordNet	41×41	256	60	-	$1e-5$	[28]
InceptionV4	779×779	-	-	-	-	[29]
VGG19, ResNet152 DenseNet201	$256 \times 256, 512 \times 512$	-	-	-	-	[35]
VGG16	779×779	-	12	Adam, SGD	$1e-3, 1e-4$	[36]
ResNet	224×224	64	-	SGD	$1e-3$	[37]
VGG16, VGG19 ResNet50, Inceptionv3 Xception	$224 \times 224, 299 \times 299$	8	100	SGD	$1e-4$	[39]
OverFeat, VGG-S	$231 \times 231, 224 \times 224$	-	-	-	-	[41]
Standard CNN	231×231	64	50	SGD	$1e-4$	[44]
VGG19, RESNET50 GoogLeNet, DENET	231×231	32	100; 80; 25	SGD	$1e-4$	[44]
InceptionV3	2095×2095	15	-	-	-	[49]
InceptionV3	224×224	-	14	Gradient Descent	$1e-5$	[48]
CNN	224×224	-	600	-	$1e-2$	[54]
CNN	512×512	15	300	Adam	$3 \times 1e-4$	[59]
CNN	128×128	-	250; 150; 70	Adam	$1e-3; 1e-4, 1e-5$	[60]
CNN	64×64	-	20	Adam	$1e-5$	[61]
CNN	256×256	-	100	-	$1e-2; 1e-3; 1e-4$	[63]
CNN	256×256	15	200	Adam	$1e-5$	[65]
CNN	512×512	15	200,250	SGD	$3 \times 1e-4; 3 \times 1e-5$	[67]
CNN	51×51	10	60	Adam	$1e-2$	[69]

Legend: R1 = Model, R2 = Image Size, R3 = Mini Batch Size, R4 = Epoch, R5 = Optimizers, R6 = Initial Learning Rate.

of 96.4%. Last, for the referable DME, they achieved sensitivity of 92%. Quellec *et al.* [22] and Gondal *et al.* [19] used a 26 layered *o_O* solution proposed by Bruggemann and Antony,[1] which ranked second in DR Kaggle competition. Reference [22] achieved AUC of 95.4% on the Kaggle dataset and on the e-ophtha dataset they obtained AUC of 94.9%. Similarly, Gondal [19] used *o_O* solution to detect DR lesions such as red dots, soft exudates, hemorrhages and microaneurysms. They replaced the last dense layer to the global average pooling layer. They achieved AUC of 95.4% on the DIARETDB1 dataset. Gulshan *et al.* [20] detected DR using Inception-v3 on the Kaggle dataset and also datsets collected from three Indian hospitals. They achieved specificity of 98.2% and sensitivity of 90.1% for a moderate and worse stage of DR respectively. Mansour [21] modified AlexNet for the classification of 5 stages of DR. They achieved an accuracy (Acc) of 97.93%, specificity of 93% and sensitivity of 100% on the Kaggle dataset. Roy *et al.* [23] used the Random Forest classifier on the Kaggle dataset and achieved a Kappa Score[2] (KSc) of 86%. Li *et al.* [24] detected exudates using a modified U-Net. U-Net was designed for the segmentation of neuronal membranes. They modified the architecture using unpooling layers rather than deconvolutional

layers of U-Net. The authors trained the model using the e-ophtha dataset and achieved AUC of 96% on DIARETDB1. Li *et al.* [25] used various pretrained CNN models such as AlexNet, GoogLeNet and VGGNet. They achieved an AUC of 98.34%, Accuracy (Acc) of 92.01%, specificity of 97.11% and sensitivity of 86.03%. They achieved an AUC of 97.8% and KSc of 77.59%, following Acc of 95.21%, specificity of 97.80% and sensitivity of 77.79%, respectively. Perdomo *et al.* [26] classified normal DR images and images with exudates using LeNet architecture. Using the e-ophtha dataset the authors achieved an Acc of 99.6%, specificity of 99.6% and sensitivity of 99.8%. Takahashi *et al.* [27] applied a modified GoogLeNet for detecting various stages of DR. They modified GoogLeNet by deleting the five accuracy layers and reduced the batch size to four and achieved an Acc of 81% and Kappa value of 74%. van Grinsven *et al.* [28] used a nine layered CNN, which consisted of five convolution layers with 32 filters inspired by OxfordNet. They achieved AUC of 97.2%, specificity of 91.40% and sensitivity of 91.90% using the Messidor dataset [77]. Sayres *et al.* [29] classified five different stages of DR with an Acc of 88.4%. The accuarcy on the normal images was 96.9% and accuracy on images with mild and worse NPDR was 57.9%. Umapathy *et al.* [30] used images from STARE [80], HRE, MESSIDOR [77] and images acquired from the Retina Institute of Karnataka datasets. The authors proposed two methods for automated detection, Decision Trees classifier

---

[1] https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15617

[2] Diabetic Retinopathy Detection, Evaluation Available; https://www.kaggle.com/c/diabetic-retinopathy-detection/overview/evaluation

TABLE 5. Studies employing TL for automatic DED detection.

DED	Architecture	Model	Ref.	Results
DR	*AlexNet*	CNN	[16]	$AUC = 98.0\%, SE = 96.8\%, SP = 87.0\%$
	VGGNet	CNN	[17]	$AUC = 90.3\%, SE = 80.3\%, SP = 85.5\%$
	VGGNet	CNN	[18]	$AUC = 93.6\%, SE = 90.5\%, SP = 91.6\%$
	o_OSolution	CNN	[19]	$AUC = 95.4\%, SE = 93.6\%, SP = 97.6\%$
	Inception-V3	CNN	[20]	$SE = 90.1\%, SP = 98.2\%$
	AlexNet	CNN	[21]	$ACC = 97.93\%, SE = 100\%, SP = 93\%$
	o_OSolution	CNN	[22]	$AUC = 95.4\%$
	ImageNet	CNN	[23]	$KSc = 86\%$
	U-Net	CNN	[24]	$AUC = 96\%$
	AlexNet, GoogLeNet, VGGNets	CNN	[25]	$AUC = 98.34\%, ACC = 92.01\%, SE = 86.03\%, SP = 97.11\%$
	LeNet	CNN	[26]	$ACC = 99.6\%, SE = 99.8\%, SP = 99.6\%$
	GoogLeNet	CNN	[27]	$ACC = 81\%, PABAK = 74\%$
	OxfordNet	CNN	[28]	$AUC = 97.2\%, SE = 91.90\%, SP = 91.40\%$
	Inception-V4	CNN	[29]	$ACC = 88.4\%$
	Inception-V3	CNN	[30]	$ACC = 88.8\%$
	VGG-16, VGG-19	CNN	[31]	$AUC = 90.4\%, ACC = 82\%, SE = 80\%, SP = 82\%$
	ResNet50	CNN	[33]	$ACC = 96.3\%, AUC = 92.6\%$
	AlexNet	CNN	[34]	$ACC = 90.07\%$
Gl	*VGG19, ResNet152, DenseNet201*	CNN	[35]	$AUC = 90\%$
	VGG-16	CNN	[36]	$ACC = 92.4\%, SE = 91.7\%, SP = 93.3\%$
	ResNet	CNN	[37]	$AUC = 99.7\%$
	VGG-19	CNN	[38]	$AUC = 96.3\%$
	VGG-16,VGG-19, Inception-V3, ResNet50, Xception	CNN	[39]	$AUC = 96.05\%, SE = 93.46\%, SP = 85.80\%$
	GoogLeNet	CNN	[40]	$ACC = 90.0\%, 94.2\%, 86.2\%, 86.4\%, 87.3\%$
	OverFeat, VGG-S	CNN	[41]	$AUC = 76.3\%, 71.8\%$
	VGG-16, VGG-19, ResNet50, InceptionV3, InceptionResNetV2	CNN	[42]	$AUC = 95.7\%, 86.0\%$
	VGG	CNN	[43]	$ACC = 87.6\%, SE = 82.6\%, SP = 93.2\%$
	VGG-19	CNN	[44]	$AUC = 94\%, SE = 87.01\%, SP = 89.01\%$
	ResNet50, 101, 152	CNN	[45]	$AUC = 84.58\%, SE = 72.50\%$
	Xception	CNN	[45]	$AUC = 93.48\%, SE = 85.00\%$
	VGG19	CNN	[45]	$AUC = 88.06\%, SE = 73.18\%$
	ResNet18, CatGAN	CNN	[45]	$AUC = 95.55\%, SE = 89.18\%$
	ResNet	CNN	[45]	$AUC = 95.24\%, SE = 85\%$
	SENet	CNN	[45]	$AUC = 95.87\%, SE = 89.17\%$
	ResNet50	CNN	[45]	$AUC = 98.17\%, SE = 97.60\%$
	ResNet101, 152, DensNet169, 201	CNN	[45]	$AUC = 93.27\%, SE = 92.50\%$
	DeepLabv3	CNN	[45]	$AUC = 95.08\%, SE = 87.50\%$
	ResNet	CNN	[46]	$AUC = 96.5\%$
	InceptionV3	CNN	[47]	-
	InceptionV3	CNN	[48]	$AUC = 92.2\%, AUC = 88.6\%, AUC = 87.9\%$
DME	*Inception-V3*	CNN	[49]	$AUC = 98.7\%, SE = 89.6\%, SP = 97.4\%$
	AlexNet	CNN	[32]	$ACC = 97.9\%$
	ResNet50	CNN	[33]	$ACC = 91.2\%, AUC = 92.4\%$
	AlexNet	CNN	[34]	$ACC = 96.85\%$
Ca	*AlexNet*	CNN	[50]	$ACC = 92.91\%$

Legend: CNN = Convolutional Neural Network, SE = Sensitivity, SP = Specificity, AUC = Area Under Curve, PABAK = Prevalence And Bias-Adjusted Fleiss' Kappa, ACC = Accuracy, KSc = Kappa Score.

and TL. They retrained the last layer of Inception-V3 to classify the normal and DR images. They achieved an Acc of 88.8%. Nguyen *et al.* presented an automated method of DR screening using DL models such as CNN, VGG-16 and VGG-19. The system classifies five categories of DR range 0-4, in which 0 is no DR and 4 is PDR. They obtained an AUC of 90.4%, sensitivity of 80%, specificity of 82% and Acc of 82% respectively.

Glaucoma A number of research studies have been conducted for the automated detection of Gl using TL. Phan *et al.* [35] applied the Deep Convolutional Neural Network to 3,312 images, which consisted of 369 images of Gl eyes, 256 Gl-suspected images and 2687 images

of non-glaucoma eyes. [3] The AUC achieved was 90%. Ghamdi *et al.* [36] presented a semi-supervised TL CNN model for automatic detection of Gl. They used the RIM-ONE [89] database and achieved an Acc of 92.4%, specificity of 93.3% and sensitivity of 91.7%. Asaoka *et al.* [37] used ResNet architecture and tested two datasets obtained from multiple institutes. They used the method of data augmentation to increase the data volume and measure their accuracy using the area under the receiver operating characteristic curve (AROC). Hence, they obtained

[3]Data Collected; Yamanashi University glaucoma outpatient clinic and Yamanashi Koseiren Hospital

two results, an AROC of 94.8% in an augmented dataset and an AROC of 99.7% in a dataset without augmentation. An *et al.* [38] used TL to detect Gl using color fundus images and 3 dimensional optical coherence tomography (OCT). To evaluate the model AUC the tenfold cross-validation (CV) was used. The Random Forest combined with five separate CNN models improved tenfold CV AUC to 96.3%. Diaz-Pinto *et al.* [39] used five different publicly available datasets resulting in the AUC of 96.05%, specificity of 85.80% and sensitivity of 93.46%. Cerentinia *et al.* [40] used GoogLeNet architecture for the detection of the presence of Gl. They used datasets from various databases and achieved an Acc of 90% from the High Resolution Fundus (HRF) database, 94.2% of accuracy from RIM-ONE(r1) [89], 86.2% of accuracy from RIM-ONE(r2) [89], 86.4% of accuracy from RIM-ONE(r3) [89] and by combining all three databases the accuracy obtained was 87.6%. Orlando *et al.* [41] used two different CNN models from OverFeat and VGG-S to develop an automated Gl detection system. The proposed architecture yielded AUC value for OverFeat and VGG-s of 76.3% and 71.8%, respectively. de Moura Lima *et al.* [42] used VGG-16, VGG-19, ResNet50, InceptionV3 and InceptionResNetV2 to diagnose Gl on RIM-ONE [89] datasets. Promising results were obtained by combining ResNet and Logistic Regression, on RIM-ONE-r2 [89], with AUC of 95.7% and on InceptionResNet with the same classifier yielded AUC of 86% on RIM-ONE-r3 [89]. Li *et al.* [43] used the VGG network to classify glaucoma and non-glaucoma visual fields based on the results of the visual field (VF) study and, for this test, they obtained VF samples from three different ophthalmic centres in mainland China. They obtained an Acc of 87.6%, while the specificity was 82.6% and sensitivity was 93.2%, respectively. In the Gómez-Valverde *et al.* [44] study VGG-19 was used to identify glaucoma and non-glaucoma using two publicly available datasets RIM-ONE [89] and DRISHTI-GS [84] and one private dataset from a screening campaign performed at Hospital de la Esperanza (Parc de Salut Mar) in Barcelona (Spain).

Diabetic Macular Edema Various researchers also investigated the use of a pretrained model to detect DME. Sahlsten *et al.* [49] performed binary classification of Non-Referable DME and Referable DME (NRDME/RDME) and achieved AUC of 98.7%, specificity of 97.4% and sensitivity of 89.6% in binary classification using TL.

Cataract Finally, cataract detection using DL was performed by Pratap and Kokil [50]. Authors have collected data from various sources such as HRF, STARE [80], DIARETDB0, MESSIDOR [77], FIRE, etc. In total, they collected 800 images (200 - normal, 200 - mild, 200 - moderate, 200 - severe). The accuracy achieved was 92.91%.

B. DL APPROACHES EMPLOYING NEW NETWORK

An alternative to TL is the new network development by the researchers. Out of 65 studies, 21 of them have designed their DL architectures for automated detection of DED. Table 6 presents the list of studies, where the researchers have employed their own built DL models with the classifier indicated, number of layers, model used and results obtained.

Diabetic Retinopathy Doshi *et al.* [51] detected the severity of diabetic retinopathy using the 29 layers CNN model and detected five stages of DR, and three CNN achieved an Acc of 39.96% on kappa matrix. Gargeya and Leng [52] identified diabetic retinopathy using the DL approach. They achieved AUC of 94%, specificity of 87% and sensitivity of 93%. Ghosh *et al.* [53] employed a 28 layers CNN for two and five class classification of diabetic retinopathy. Using Softmax they achieved an Acc of 95% for two class and 85% of Acc for five class classification. Jiang *et al.* [54] classified two classes of diabetic retinopathy using fundus images. They used 17 layers deep CNN on the Caffe framework and achieved an Acc of 75.7%. Pratt *et al.* [55] employed a CNN architecture to identify the severity level of DR. They achieved an Acc of 75%, specificity of 30% and sensitivity of 95% using Softmax classifier. Xu *et al.* [56] employed a 16 layer model for early detection of DR. Using Softmax classifier they achieved an Acc of 94.50%. Yang *et al.* [57] employed local and global CNN architectures. Local CNN (10 layers) was used for lesion detection and the global CNN (26 layers) for grading DR. The authors achieved an AUC of 0.9687, specificity of 89.80% and sensitivity of 95.90%. Yu *et al.* [58] detected exudates using 16 layers CNN. With the Softmax classifier, they achieved an Acc of 91.92%, specificity of 96% and sensitivity of 88.85%. de La Torre *et al.* [59] used 17 layered CNN architecture obtaining specificity of 90.8% and sensitivity of 91.1%. Pires *et al.* [60] proposed 16 layer CNN architecture. They used Messidor-2 and DR2 dataset to test the model. With the neural networks classifier, they achieved AUC of 96.3% in the DR2 dataset and AUC of 98.2% in Messidor-2 and with the Random Forests classifier, they achieved AUC of 96.1% in DR2 dataset and AUC of 97.9% in Messidor-2.

Hemanth *et al.* [61] proposed a hybrid method based on using both image processing and DL for improved results. using 400 retinal fundus images within the MESSIDOR [77] database and average values for different performance evaluation parameters were obtained an Acc 97%, sensitivity (recall) 94%, specificity 98%, precision 94%, FScore 94% and geometric mean (GMean) 95%.

Glaucoma Chen *et al.* [62] developed six layer CNN model. With the Softmax classifier they achieved an AUC of 83.1% and 88.7% in ORIGA [83] and SCES datasets. Raghavendra *et al.* [63] build an eighteen layer CNN framework to diagnose Gl using 1426 fundus images in where 589 were normal and 937 were with glaucoma. They achieved an Acc of 98.13%, sensitivity of 98% and specificity of 98.3%. Pal *et al.* [64] introduced a novel multi-model DL network named G-EyeNet for glaucoma detection using DRI-ONS [88] and Drishti-GS [84] datasets. Their experimental findings revealed an AUC of 92.3%.

Diabetic Macular Edema Al-Bander *et al.* [67] proposed a CNN system to grade the severity of DME using fundus images using the MESSIDOR [77] dataset of 1200 images.

TABLE 6. Studies employing new network for automatic DED detection.

DED	Classifier	Model	Layers	Ref.	Results
DR	*Softmax*	CNN	29	[51]	$KSc = 39.96\%$
	Decision Trees	CNN	6	[52]	$AUC = 94\%, SE = 93\%, SP = 87\%$
	Softmax	CNN	28	[53]	$ACC = 85\%, KSc = 75.4\%, Prec = 88.20\%, SE = 95\%$
	Softmax	CNN	17	[54]	$ACC = 75.70\%$
	Softmax	CNN	13	[55]	$SE = 95\%, SP = 30\%$
	Softmax	CNN	16	[56]	$ACC = 94.5\%$
	Softmax	CNN	10	[57]	$AUC = 96.87\%, SE = 95.90\%, SP = 89.90\%$
	Softmax	CNN	16	[58]	$ACC = 91.92\%, SE = 88.85\%, SP = 96\%$
	Softmax	CNN	17	[59]	$SE = 91.1\%, SP = 90.8\%$
	Softmax	CNN	16	[60]	$AUC = 96.1\%$
	Softmax	CNN	8	[61]	$ACC = 97\%, SE = 94\%, SP = 98\%, Prec = 94\%, FSc = 94\%, GMean = 95\%$
Gl	*Softmax*	CNN	6	[62]	$AUC = 83.1\%, 88.7\%$
	Softmax	CNN	18	[63]	$ACC = 98.13\%, SE = 98\%, SP = 98.3\%$
	Softmax	CNN	6	[64]	$AUC = 92.3\%$
	Softmax	CNN	6	[65]	$ACC = 90\%, SE = 96\%, SP = 84\%$
	Softmax	CNN	12	[66]	$AUC = 8.31\%, 88.7\%$
DME	*Softmax*	CNN	13	[67]	$ACC = 88.8\%, SE = 74.7\%, SP = 96.5\%$
	Softmax	CNN	10	[68]	$SE = 78\%, PPV = 78\%, FSc = 78\%$
	Softmax	CNN	10	[69]	$SE = 87.58\%, SP = 98.73\%$
Ca	*Softmax*	CNN	8	[9]	$AUC = 93.52\%$
	Softmax	CNN	5	[70]	$ACC = 94.07\%, 81.91\%$

Legend: CNN = Convolutional Neural Network, SSAE = Stacked Sparse Auto Encoder, ACC = Accuracy, SE = Sensitivity, SP = Specificity, AUC = Area Under Curve, FSc = F-Score, KSc = Kappa Score, Prec = Precision, PPV = Positive Predictive Value, GMean = Geometric mean.

They obtained an Acc of 88.8%, sensitivity of 74.7% and specificity of 96.5% respectively. Prentašić and Lončarić [68] introduced a novel supervised CNN based exudate detection method using the DRiDB dataset [93]. The proposed network consists of 10 alternating convolutional and max-pooling layers. They achieved sensitivity of 78%, Positive Predictive Value (PPV) of 78% and FSc of 78% respectively. Tan *et al.* [69] used the CLEOPATRA [94] image dataset. They obtained sensitivity of 87.58% and specificity of 98.73% respectively.

Cataract Zhang *et al.* [9] proposed eight layers of DCNN architecture. With the Softmax classifier, they achieved an Acc of 93.52% and 86.69%. Dong *et al.* [70] used a Softmax classifier with five layer CNN architecture and achieved an Acc of 94.07% and 81.91%, respectively.

C. APPROACHES EMPLOYING COMBINED DL AND ML

Out of 65 studies, six proposed a combination of DL and ML classifiers. Table 7 shows the studies in which the authors applied a combination of DL and ML classifiers namely: Random Forest (RF), Support Vector Machine(SVM) and Backpropagation Neural Network (BPNN) based architectures for DED detection. Abbas *et al.* [71] developed a DL Neural Network (DLNN) to discover the severity degree of DR in fundus images using studying Deep Visual Features (DVFs). For feature extraction, they used Gradient Location Orientation Histogram (GLOH) and Dense Color Scale Invariant Feature Transform (DColor-SIFT). They converted the features through the use of Principle Component Analysis (PCA). Afterwards, a three layer deep neural network was used to learn these features and subsequently, an SVM classifier was applied for the classification of DR fundus

images into five severity stages, including no-DR, moderate, mild, severe NPDR (Nonproliferative Diabetic Retinopathy) and PDR (Proliferative Diabetic Retinopathy). They obtained sensitivity of 92.18%, specificity of 94.50% and AUC of 92.4% on three publicly available datasets (Foveal Avascular Zone Messidor [77], DIARETDB1) and one extraordinary dataset (from the, Hospital Universitario Puerta del Mar, HUPM, Cádiz, Spain). Orlando *et al.* [72] combined ML and DL for the detection of lesions (red). They used three public datasets, namely Messidor [77], DIARETDB1 and e-ophtha. They extracted intensity and shape as features using knowledge transferred LeNet architecture, which consists of 10 layers. They achieved AUC of 93.47% and sensitivity of 97.21%, respectively. Arunkumar and Karthigaikumar [73] employed a Deep Belief Network (DBN) for diabetic retinal image classification. At first, with three hidden layers, the deep features were extracted with Deep Belief Network (DBN), then those features were decreased by applying the Generalised Regression Neural Network (GRNN) technique and finally, the retinal images were classified using SVM. On they publicly available ARIA dataset, the authors achieved an Acc of 96.73%, specificity of 97.89% and sensitivity of 79.32%, respectively. Al-Bander *et al.* [74] used CNN for feature extraction and SVM for Gl and non Gl classification. They achieved an Acc of 88.2%, specificity of 90.8% and sensitivity of 85%, respectively. Ran *et al.* [75] used a 17 layer DCNN feature extractor, which adopts a residual network to learn more detailed features of fundus images. The DCNN contains three modules, namely shallow, residual and pooling. Here, the shallow and residual modules extract features on a deep, medium and shallow level and the final feature vectors for Random Forests are output from the

TABLE 7. Studies employing combined DL and ML for automatic DED detection.

DED	Model	Layers	Features	Ref.	Classifier	Results
DR	CNN	3	DColor-SIFT, GLOH	[71]	Softmax	$AUC = 92.4\%, SE = 92.18\%, SP = 94.50\%$
	CNN	10	Shape, Intensity	[72]	RF	$AUC = 93.47\%, SE = 97.21\%$
	DBN	3	Shape, Intensity	[73]	SVM	$ACC = 96.73\%, SE = 79.32\%, SP = 97.89\%$
Gl	CNN	23	-	[74]	RF	$ACC = 88.2\%, SE = 85\%, SP = 90.8\%$
Ca	DCNN	17	Shallow, residual, pooling	[75]	RF	$ACC = 90.69\%$
	CNN	2	Wavelet, Sketch, Texture	[76]	SVM, BPNN	$ACC = 93.2\%, 84.5\%$

Legend: CNN = Convolutional Neural Network, DBN = Deep Belief Network, RF = Random Forests, SVM = Support Vector Machine, BPNN = Back-Propagation Neural Network, SE = Sensitivity, SP = Specificity, AUC = Area Under Curve, Acc = Accuracy, DColor-SIFT = Dense Color Scale-Invariant Feature Transform, GLOH = Gradient Location Orientation Histogram.

pooling module. The authors detected six classes of cataract with an Acc of 90.69%. Last, [57] proposed 2 layers stacking architecture with Support Vector Machine and backpropagation neural network classifier. The ensemble classifier achieved an Acc of 93.2% and 84.5%.

D. ANALYSIS AND REVIEW OF PERFORMANCE EVALUATION METRICS

Detailed description of performance measures, namely: specificity, sensitivity, accuracy, area under curve (AUC), precision, f-score, and positive predictive value can be found in [97]. Likewise, Kappa Score, PABAK Index discussions can be found in [98], respectively. In the majority of listed academic papers, the authors used specificity, sensitivity, accuracy and AUC as their assessment metrics to evaluate the efficiency of the classifier. The combined effect of performance metrics found to be used frequently was Sensitivity, Specificity and Accuracy. This variation was used 12 times out of a total 60 trials, accompanied by 12 uses of and Sensitivity, Specificity, AUC and two use of Sensitivity, Specificity, Accuracy and AUC. Instead of Sensitivity, some researchers used Recall. We accommodated Recall under Sensitivity, rather than using it as another success indicator. The performance measurements frequently used include Sensitivity (32 times), Specificity (25 times), Accuracy (26 times), and AUC (25 times). Other performance metrics not commonly used by research groups were: F-Score (twice), Precision (twice), PABAK (once), Kappa Score (3 times), Positive Predictive Value (once) and GMean (once).

VI. DISCUSSION AND OBSERVATIONS

AI is one of the most intriguing technologies used in the material science toolset in recent decades. This compendium of statistical techniques has already shown that it is capable of significantly accelerating both fundamental and applied research. ML, already has a rich history in biology [99], [100] and chemistry [101], and it has recently gained prominence in the field of solid state materials science. Presently, DL models in ML are effectively used in imaging for classification, detection [102], segmentation [103] and prepossessing. The most famous and commonly employed DL architecture in the selected 65 studies is CNN, which is used in 64 cases, while DBN is implemented once. We can infer that CNN is

currently the most preferred deep neural network, particularly for the detection of DED as well as the diagnosis of other pathological indications from the medical images.

We have noticed that DL performed well on binary classification tasks (eg. DR and Non DR), whereas its performance significantly dropped when the number of classes increased. As an example, Ghosh *et al.* [53] obtained an Acc of 95% on DR and Non DR classification task and Acc of 85% on a multi class problem (five stages of severity), with 10% loss in Acc. Else, Choi *et al.* [17] classified 10 distinct retinal diseases and achieved an Acc of 30.5%. Also, Dong *et al.* [70] performed cataract classification based on two features, namely: i) features extracted using DL and ii) features extracted using wavelet. Classification for the binary problem (Cataract and Non Cataract) achieved an Acc of 94.07% and 81.91%. Then, the authors performed the classification of four classes of cataract and obtained an Acc of 90.82% and 84.17%. This shows that features extracted using wavelet increased an Acc of the Softmax classifier by 4% in the four class problem. Still, the overall highest accuracy was observed for a binary classification task.

This study reveals the research gap for more rigorous approaches to the development of multiclass DED classification problems. Furthermore, we have observed that binary classification is mostly conducted between the normal and the affected DED cases. For instance, Ghosh *et al.* [53] and Choi *et al.* [17] classified DR and Non-DR. Also, Al-bander *et al.* [74] and Phan *et al.* [35], identified glaucomatous and nonglaucomatous retinal images, while Dong *et al.* [70] detected cataract and noncataract conditions. The methods used in these articles are effectively identifying the vast proportion of severe cases where pathological signs are more prominent. Thus, there is a need for classifiers that perform equally well for mild stages of DED developments, where the lesions are tiny and difficult to detect.

Early detection of DED or mild DED is especially necessary to take effective preventive steps and to avoid possible blindness due to deterioration condition over time. As we can see, DL has shown an extensive capacity in the field of health care and especially in the field of DED detection. However, there are some limitations in its large-scale implementation. In terms of the validation of the proposed methods, the authors predominantly used Accuracy, Specificity

TABLE 8. Performance metrices employed in selected studies.

ACC	AUC	SE	SP	KSc	FSc	Prec	PPV	Pabak	GMean	References
✓	✗	✓	✓	✗	✗	✗	✗	✗	✗	[21], [26], [36], [43], [58], [63], [65], [67], [70], [73], [74]
✗	✓	✓	✓	✗	✗	✗	✗	✗	✗	[16]–[19], [28], [39], [44], [52], [57], [71], [72]
✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	[20], [55], [59], [69]
✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	[23], [51]
✓	✗	✓	✗	✓	✗	✓	✗	✗	✗	[53]
✗	✗	✓	✗	✗	✓	✗	✓	✗	✗	[68]
✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	[61]
✓	✗	✗	✗	✗	✗	✗	✗	✓	✗	[27]
✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	[25], [31]
✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	[29], [30], [32], [34], [34], [50], [54], [56], [70], [75], [76]
✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	[9], [22], [24], [35], [37], [38], [41], [42], [48], [60], [62], [64]

Legend: ACC = Accuracy, AUC = Area Under Curve, SE = Sensitivity, SP = Specificity, KSc = Kappa Score, FSc = F-Score , Prec = Precision, PPV = Positive Predictive Value, Pabak = Prevalence And Bias Adjusted Fleiss' Kappa, GMean = Geometric Mean.

and Sensitivity to report their classification performance. For instance, Perdomo *et al.* [26] used LeNet CNN to detect exudates and reported accuracy (99.6%), specificity (99.6%) and sensitivity (99.8%) for the approach proposed. Another widely used metric was AUC, accuracy and sensitivity. This combination is appropriate in DL methods where image classes are imbalanced. However, data imbalance has been solved using geometric transformation (augmentation techniques) or re-sampling images from each class. For example, Chen *et al.* [62] used augmentation to overcome the overfitting on image data and obtained AUC (83.10%), and AUC (88.7%)on the ORIGA and SCES datasets. Other metrics have been used to measure performance such as Ksc used by Roy *et al.* [23], AROC used by Asaoka *et al.* [37], and PABAK used by Takahashi *et al.* [27].

VII. GAPS AND FUTURE RESEARCH DIRECTIONS

This segment poses numerous research issues that researchers have not been able to solve in previous DED detection studies. Significant research is, therefore, still needed to improve the effectiveness of various techniques for DED detection. The research challenges that need to be addressed are set out below.

Developing stronger DL models While DL has already shown extremely promising success in the field of medical imaging and retinal disease diagnosis, it will be difficult to further refine and create more effective deep neural networks. Yet another solution may be to increase computational power by increasing the capacity of the network [104], [105] while addressing the risk of over-fitting. Another approach could be to create a different object based model rather than an image based model. For example, if researchers are interested in detecting a particular eye malformation (e.g. exudates only), they would design such a deep convolutional neural network that only learns with exudates and other malformation that are not of interest would not be learned by the model. It is argued in [106] that object based identification is more effective than image based identification.

Training on minimal data DL software typically involves a large number of retinal fundus images for learning. If the training set is small, it may not produce satisfactory results in terms of accuracy. There are two solutions available. First, using a range of enhancement methods including rotating, shifting, cropping and colour setting. Second, employ feeble learning algorithms to retrieve training data. Further, investigations shows Generative Adversal Network (GAN) for the generation of training, so that the DL architecture can be trained with robustness and more distinctive features [107].

Similar DL architecture for medical imaging Mostly in DL, several TL frameworks (such as GoogLeNet, AlexNet, VGGNet, and LeNet) for object recognition are available for retraining on a new collection of images such as medical images. Nevertheless, as far as classification efficiency is concerned, these architectures are less suitable for medical images. For example, Choi *et al.* [17] used VGGNet for DR diagnosis utilising eye fundus images and obtained nearly 85.5% specificity. This is because such TL frameworks are designed for objects such as animals, flowers, etc. As a result, such frameworks may be unsuitable for real time medical images. Potential study could implement a TL architectural design which has been learned on appropriate medical images rather than objects, functioning as a generic architecture, and eventually retrained to improve the accuracy of medical image classification.

Automated Choosing the optimum values for DL Architectures Neural networks have provided promising results primarily in the area of computer vision and particularly in DED detection, but the complexity of modulating is not well known and is considered to be a black box. For instance, several researchers fine tune the constraints of current DL techniques, such as CNN or AlexNet, in order to enhance classification efficiency. In certain instances, however, the history of DL architectures is not well known and is perceived to be a black box. It is, therefore, still hard to determine the effective model and optimum values for the number of hidden layers and modules in various layers.

Domain specific knowledge is also necessary for the selection of attributes for the number of epochs, learning rate and regularization. Thus, in the future, automated optimization algorithms could be proposed to find optimal rates for various DL architectures on various DED datasets and other similar resources for medical images.

Integrating DL, cloud computing and telehealth In particular, rural regions suffer from a lack of human capital, notably in healthcare. In such cases, therefore, telehealth can play an important role in overcoming this drawback. Neural networks, cloud computing and telehealth may be combined in the future to diagnose DED from eye fundus images. For example, in rural communities, the patient could use his or her mobile phone with a retinal camera to capture an image of the eye fundus. This image could also be transferred to a cloud computing platform where the DED detection model (constructed through a ML or a DL approach) can be applied. This configured system will therefore identify DED from the image file and return the results of the detection and prescription to the patient.

VIII. CONCLUSION

This review paper provides a comprehensive overview of the state of the art on Diabetic Eye Disease (DED) detection methods. To achieve this goal, a rigorous systematic review of relevant publications was conducted. After the final selection of relevant records, following the inclusion criteria and quality assessment, the studies have been analyzed from the perspectives of 1) Datasets used, 2) Image preprocessing techniques adopted and 3) Classification method employed. The works were categorized into the specific DED types, i.e. DR, Gl, DME and Ca for clarity and comparison. In terms of classification techniques, our review included studies that 1) Adopted TL, 2) Build DL network architecture and 3) Used combined DL and ML approach. Details of the findings obtained are included in Section VI.

We have also identified several limitations associated with our study. First, we narrowed down the review conducted from *April 2014 - January 2020* due to rapid advances in the field. Second, we limited our review to DL based approaches due to their state of the art performance, in particular on the image classification task. Finally, our review focused on a collection of predefined keywords that provides a thorough coverage of the DED area of detection but may not be exhaustive. Furthermore, we hope that our research can be further expanded in the future to include an all encompassing and up-to-date overview of the rapidly developing and challenging field of DED detection.

REFERENCES

[1] National Eye Institute. (Sep. 2015). *Facts About Diabetic Eye Disease.* [Online]. Available: https://nei.nih.gov/health/diabetic/retinopathy

[2] N. G. Forouhi and N. J. Wareham, "Epidemiology of diabetes," *Medicine*, vol. 38, no. 11, pp. 602–606, 2010.

[3] S. A. Tabish, "Is diabetes becoming the biggest epidemic of the twenty-first century?" *Int. J. Health Sci.*, vol. 1, no. 2, pp. V–VIII, 2007.

[4] H. Kolb. (2011). *Simple Anatomy of There Tina by Helga Kolb.* Accessed: Jan. 31, 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK11533/

[5] 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. 98, no. 5, pp. 786–806, 1991.

[6] R. Varma, P. P. Lee, I. Goldberg, and S. Kotak, "An assessment of the health and economic burdens of glaucoma," *Amer. J. Ophthalmol.*, vol. 152, no. 4, pp. 515–522, Oct. 2011.

[7] A. Almazroa, R. Burman, K. Raahemifar, and V. Lakshminarayanan, "Optic disc and optic cup segmentation methodologies for glaucoma image detection: A survey," *J. Ophthalmol.*, vol. 2015, pp. 1–28, Sep. 2015.

[8] F. C. Gundogan, U. Yolcu, F. Akay, A. Ilhan, G. Ozge, and S. Uzun, "Diabetic macular edema," *Pakistan J. Med. Sci.*, vol. 32, no. 2, p. 505, 2016.

[9] L. Zhang, J. Li, i. Zhang, H. Han, B. Liu, J. Yang, and Q. Wang, "Automatic cataract detection and grading using deep convolutional neural network," in *Proc. IEEE 14th Int. Conf. Netw., Sens. Control (ICNSC)*, May 2017, pp. 60–65.

[10] P. Vashist, N. Gupta, S. Singh, and R. Saxena, "Role of early screening for diabetic retinopathy in patients with diabetes mellitus: An overview," *Indian J. Community Med.*, vol. 36, no. 4, p. 247, 2011.

[11] D. S. W. Ting, L. R. Pasquale, L. Peng, J. P. Campbell, A. Y. Lee, R. Raman, G. S. W. Tan, L. Schmetterer, P. A. Keane, and T. Y. Wong, "Artificial intelligence and deep learning in ophthalmology," *Brit. J. Ophthalmology*, vol. 103, no. 2, pp. 167–175, 2019.

[12] D. T. Hogarty, D. A. Mackey, and A. W. Hewitt, "Current state and future prospects of artificial intelligence in ophthalmology: A review," *Clin. Experim. Ophthalmol.*, vol. 47, no. 1, pp. 128–139, Jan. 2019.

[13] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude, "Computer-aided diagnosis of diabetic retinopathy: A review," *Comput. Biol. Med.*, vol. 43, no. 12, pp. 2136–2155, 2013.

[14] U. Ishtiaq, S. A. Kareem, E. R. M. F. Abdullah, G. Mujtaba, R. Jahangir, and H. Y. Ghafoor, "Diabetic retinopathy detection through artificial intelligent techniques: A review and open issues," *Multimedia Tools Appl.*, pp. 1–44, 2019.

[15] Y. Hagiwara, J. E. W. Koh, J. H. Tan, S. V. Bhandary, A. Laude, E. J. Ciaccio, L. Tong, and U. R. Acharya, "Computer-aided diagnosis of glaucoma using fundus images: A review," *Comput. Methods Programs Biomed.*, vol. 165, pp. 1–12, Oct. 2018.

[16] M. D. Abràmoff, Y. Lou, A. Erginay, W. Clarida, R. Amelon, J. C. Folk, and M. Niemeijer, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Invest. Ophthalmol. Vis. Sci.*, vol. 57, no. 13, pp. 5200–5206, 2016.

[17] J. Y. Choi, T. K. Yoo, J. G. Seo, J. Kwak, T. T. Um, and T. H. Rim, "Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database," *PLoS ONE*, vol. 12, no. 11, Nov. 2017, Art. no. e0187336.

[18] D. S. W. Ting, C. Y.-L. Cheung, G. Lim, G. S. W. Tan, N. D. Quang, A. Gan, H. Hamzah, "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *J. Amer. Med. Assoc.*, vol. 318, no. 22, pp. 2211–2223, 2017.

[19] W. M. Gondal, J. M. Köhler, R. Grzeszick, G. A. Fink, and M. Hirsch, "Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 2069–2073.

[20] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.

[21] R. F. Mansour, "Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy," *Biomed. Eng. Lett.*, vol. 8, no. 1, pp. 41–57, Feb. 2018.

[22] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Med. Image Anal.*, vol. 39, pp. 178–193, Jul. 2017.

[23] P. Roy, R. Tennakoon, K. Cao, S. Sedai, D. Mahapatra, S. Maetschke, and R. Garnavi, "A novel hybrid approach for severity assessment of diabetic retinopathy in colour fundus images," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1078–1082.

[24] G. Li, S. Zheng, and X. Li, "Exudate detection in fundus images via convolutional neural network," in *International Forum on Digital TV and Wireless Multimedia Communications*. Singapore: Springer, 2017, pp. 193–202.

[25] X. Li, T. Pang, B. Xiong, W. Liu, P. Liang, and T. Wang, "Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification," in *Proc. 10th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2017, pp. 1–11.

[26] O. Perdomo, J. Arevalo, and F. A. González, "Convolutional network to detect exudates in eye fundus images of diabetic subjects," in *Proc. 12th Int. Symp. Med. Inf. Process. Anal.*, Jan. 2017, Art. no. 101600.

[27] H. Takahashi, H. Tampo, Y. Arai, Y. Inoue, and H. Kawashima, "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0179790.

[28] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.

[29] R. Sayres, A. Taly, E. Rahimy, K. Blumer, D. Coz, N. Hammel, J. Krause, A. Narayanaswamy, Z. Rastegar, D. Wu, S. Xu, S. Barb, A. Joseph, M. Shumski, J. Smith, A. B. Sood, G. S. Corrado, L. Peng, and D. R. Webster, "Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy," *Ophthalmology*, vol. 126, no. 4, pp. 552–564, Apr. 2019.

[30] A. Umapathy, A. Sreenivasan, D. S. Nairy, S. Natarajan, and B. N. Rao, "Image processing, textural feature extraction and transfer learning based detection of diabetic retinopathy," in *Proc. 9th Int. Conf. Biosci., Biochem. Bioinf.*, 2019, pp. 17–21.

[31] Q. H. Nguyen, R. Muthuraman, L. Singh, G. Sen, A. C. Tran, B. P. Nguyen, and M. Chua, "Diabetic retinopathy detection using deep learning," in *Proc. 4th Int. Conf. Mach. Learn. Soft Comput.*, Jan. 2020, pp. 103–107.

[32] N. E. M. Khalifa, M. Loey, M. H. N. Taha, and H. N. E. T. Mohamed, "Deep transfer learning models for medical diabetic retinopathy detection," *Acta Inf. Medica*, vol. 27, no. 5, p. 327, 2019.

[33] 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.

[34] 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.

[35] S. Phan, S. Satoh, Y. Yoda, K. Kashiwagi, and T. Oshika, "Evaluation of deep convolutional neural networks for glaucoma detection," *Jpn. J. Ophthalmol.*, vol. 63, pp. 276–283, Feb. 2019.

[36] M. Al Ghamdi, M. Li, M. Abdel-Mottaleb, and M. A. Shousha, "Semi-supervised transfer learning for convolutional neural networks for glaucoma detection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 3812–3816.

[37] R. Asaoka, M. Tanito, N. Shibata, K. Mitsuhashi, K. Nakahara, Y. Fujino, M. Matsuura, H. Murata, K. Tokumo, and Y. Kiuchi, "Validation of a deep learning model to screen for glaucoma using images from different fundus cameras and data augmentation," *Ophthalmol. Glaucoma*, vol. 2, no. 4, pp. 224–231, Jul. 2019.

[38] G. An, K. Omodaka, K. Hashimoto, S. Tsuda, Y. Shiga, N. Takada, T. Kikawa, H. Yokota, M. Akiba, and T. Nakazawa, "Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images," *J. Healthcare Eng.*, vol. 2019, pp. 1–9, Feb. 2019.

[39] A. Diaz-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: An extensive validation," *Biomed. Eng. OnLine*, vol. 18, no. 1, p. 29, Dec. 2019.

[40] A. Cerentinia, D. Welfera, M. C. d'Ornellasa, C. J. P. Haygertb, and G. N. Dottob, "Automatic identification of glaucoma sing deep learning methods," in *Proc. 16th World Congr. Med. Health Informat. Precision Healthcare Through Informat. (MEDINFO)*, vol. 245, 2018, p. 318.

[41] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "Convolutional neural network transfer for automated glaucoma identification," in *Proc. 12th Int. Symp. Med. Inf. Process. Anal.*, vol. 10160, 2017, p. 101600U.

[42] A. C. de Moura Lima, L. B. Maia, R. M. P. Pereira, G. B. Junior, J. D. S. de Almeida, and A. C. de Paiva, "Glaucoma diagnosis over eye fundus image through deep features," in *Proc. 25th Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Jun. 2018, pp. 1–4.

[43] F. Li, Z. Wang, G. Qu, D. Song, Y. Yuan, Y. Xu, K. Gao, G. Luo, Z. Xiao, D. S. C. Lam, H. Zhong, Y. Qiao, and X. Zhang, "Automatic differentiation of glaucoma visual field from non-glaucoma visual filed using deep convolutional neural network," *BMC Med. Imag.*, vol. 18, no. 1, p. 35, Dec. 2018.

[44] J. J. Gómez-Valverde, A. Antón, G. Fatti, B. Liefers, A. Herranz, A. Santos, C. I. Sánchez, and M. J. Ledesma-Carbayo, "Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning," *Biomed. Opt. Express*, vol. 10, no. 2, pp. 892–913, 2019.

[45] J. I. Orlando *et al.*, "REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Med. Image Anal.*, vol. 59, Jan. 2020, Art. no. 101570.

[46] N. Shibata, M. Tanito, K. Mitsuhashi, Y. Fujino, M. Matsuura, H. Murata, and R. Asaoka, "Development of a deep residual learning algorithm to screen for glaucoma from fundus photography," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, Dec. 2018.

[47] A. Singh, S. Sengupta, and V. Lakshminarayanan, "Glaucoma diagnosis using transfer learning methods," *Appl. Mach. Learn.*, vol. 11139, Sep. 2019, Art. no. 111390U.

[48] J. M. Ahn, S. Kim, K.-S. Ahn, S.-H. Cho, K. B. Lee, and U. S. Kim, "A deep learning model for the detection of both advanced and early glaucoma using fundus photography," *PLoS ONE*, vol. 13, no. 11, Nov. 2018, Art. no. e0207982.

[49] J. Sahlsten, J. Jaskari, J. Kivinen, L. Turunen, E. Jaanio, K. Hietala, and K. Kaski, "Deep learning fundus image analysis for diabetic retinopathy and macular edema grading," 2019, *arXiv:1904.08764*. [Online]. Available: http://arxiv.org/abs/1904.08764

[50] T. Pratap and P. Kokil, "Computer-aided diagnosis of cataract using deep transfer learning," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101533.

[51] D. Doshi, A. Shenoy, D. Sidhpura, and P. Gharpure, "Diabetic retinopathy detection using deep convolutional neural networks," in *Proc. Int. Conf. Comput., Anal. Secur. Trends (CAST)*, 2016, pp. 261–266.

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

[53] R. Ghosh, K. Ghosh, and S. Maitra, "Automatic detection and classification of diabetic retinopathy stages using CNN," in *Proc. 4th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2017, pp. 550–554.

[54] Y. Jiang, H. Wu, and J. Dong, "Automatic screening of diabetic retinopathy images with convolution neural network based on caffe framework," in *Proc. 1st Int. Conf. Med. Health Informat.*, 2017, pp. 90–94.

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

[56] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, no. 12, p. 2054, Nov. 2017.

[57] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2017, pp. 533–540.

[58] S. Yu, D. Xiao, and Y. Kanagasingam, "Exudate detection for diabetic retinopathy with convolutional neural networks," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 1744–1747.

[59] J. de La Torre, A. Valls, and D. Puig, "A deep learning interpretable classifier for diabetic retinopathy disease grading," *Neurocomputing*, vol. 396, pp. 465–476, Jul. 2019.

[60] R. Pires, S. Avila, J. Wainer, E. Valle, M. D. Abramoff, and A. Rocha, "A data-driven approach to referable diabetic retinopathy detection," *Artif. Intell. Med.*, vol. 96, pp. 93–106, May 2019.

[61] D. J. Hemanth, O. Deperlioglu, and U. Kose, "An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 707–721, Feb. 2020.

[62] X. Chen, Y. Xu, D. W. K. Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 715–718.

[63] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudigar, J. H. Tan, and U. R. Acharya, "Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images," *Inf. Sci.*, vol. 441, pp. 41–49, May 2018.

[64] A. Pal, M. R. Moorthy, and A. Shahina, "G-eyenet: A convolutional autoencoding classifier framework for the detection of glaucoma from retinal fundus images," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 2775–2779.

[65] A. Sharma, M. Agrawal, S. D. Roy, and V. Gupta, "Automatic glaucoma diagnosis in digital fundus images using deep CNNs," in *Advances in Computational Intelligence Techniques*. Singapore: Springer, 2020, pp. 37–52.

[66] L. K. Singh *et al.*, "Automated glaucoma type identification using machine learning or deep learning techniques," in *Advancement of Machine Intelligence in Interactive Medical Image Analysis*. Singapore: Springer, 2020, pp. 241–263.

[67] B. Al-Bander, W. Al-Nuaimi, M. A. Al-Taee, B. M. Williams, and Y. Zheng, "Diabetic macular edema grading based on deep neural networks," in *Proc. 3rd Int. Workshop Ophthalmic Med. Image Anal. (OMIA)*. Iowa City, IA, USA: Univ. of Iowa, 2016, pp. 121–128.

[68] P. Prentašić and S. Lončarić, "Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion," *Comput. Methods Programs Biomed.*, vol. 137, pp. 281–292, Dec. 2016.

[69] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, "Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network," *Inf. Sci.*, vol. 420, pp. 66–76, Dec. 2017.

[70] Y. Dong, Q. Zhang, Z. Qiao, and J.-J. Yang, "Classification of cataract fundus image based on deep learning," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2017, pp. 1–5.

[71] Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, "Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features," *Med. Biol. Eng. Comput.*, vol. 55, no. 11, pp. 1959–1974, Nov. 2017.

[72] 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.

[73] 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.

[74] B. Al-Bander, W. Al-Nuaimy, M. A. Al-Taee, and Y. Zheng, "Automated glaucoma diagnosis using deep learning approach," in *Proc. 14th Int. Multi-Conf. Syst., Signals Devices (SSD)*, Mar. 2017, pp. 207–210.

[75] J. Ran, K. Niu, Z. He, H. Zhang, and H. Song, "Cataract detection and grading based on combination of deep convolutional neural network and random forests," in *Proc. Int. Conf. Netw. Infrastruct. Digit. Content (ICNIDC)*, Aug. 2018, pp. 155–159.

[76] J.-J. Yang, J. Li, R. Shen, Y. Zeng, J. He, J. Bi, Y. Li, Q. Zhang, L. Peng, and Q. Wang, "Exploiting ensemble learning for automatic cataract detection and grading," *Comput. Methods Programs Biomed.*, vol. 124, pp. 45–57, Feb. 2016.

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

[78] D. S. Sisodia, S. Nair, and P. Khobragade, "Diabetic retinal fundus images: Preprocessing and feature extraction for early detection of diabetic retinopathy," *Biomed. Pharmacol. J.*, vol. 10, no. 2, pp. 615–626, Jun. 2017.

[79] B. Antal and A. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowl.-Based Syst.*, vol. 60, pp. 20–27, Apr. 2014.

[80] D. J. J. Farnell, F. N. Hatfield, P. Knox, M. Reakes, S. Spencer, D. Parry, and S. P. Harding, "Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators," *J. Franklin Inst.*, vol. 345, no. 7, pp. 748–765, Oct. 2008.

[81] 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, 2019.

[82] A. A. Almazroa, S. Alodhayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlaim, M. Alkatee, K. Raahemifar, and V. Lakshminarayanan, "Retinal fundus images for glaucoma analysis: The RIGA dataset," in *Proc. Med. Imag. : Imag. Informat. Healthcare, Res., Appl.*, Mar. 2018, Art. no. 105790.

[83] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 3065–3068.

[84] J. Sivaswamy, S. R. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-GS: Retinal image dataset for optic nerve head(ONH) segmentation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 53–56.

[85] T. Hassan, M. U. Akram, M. F. Masood, and U. Yasin, "Biomisa retinal image database for macular and ocular syndromes," in *Proc. Int. Conf. Image Anal. Recognit.* Cham, Switzerland: Springer, 2018, pp. 695–705.

[86] T. Khalil, M. U. Akram, S. Khalid, and A. Jameel, "Improved automated detection of glaucoma from fundus image using hybrid structural and textural features," *IET Image Process.*, vol. 11, no. 9, pp. 693–700, Sep. 2017.

[87] M. T. Islam, S. A. Imran, A. Arefeen, M. Hasan, and C. Shahnaz, "Source and camera independent ophthalmic disease recognition from fundus image using neural network," in *Proc. IEEE Int. Conf. Signal Process., Inf., Commun. Syst. (SPICSCON)*, Nov. 2019, pp. 59–63.

[88] E. J. Carmona, M. Rincón, J. García-Feijoó, and J. M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," *Artif. Intell. Med.*, vol. 43, no. 3, pp. 243–259, Jul. 2008.

[89] F. Fumero, S. Alayón, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," in *Proc. 24th Int. Symp. Comput.-Based Med. Syst. (CBMS)*, Jun. 2011, pp. 1–6.

[90] 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.

[91] 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, 2018.

[92] P. Prasanna, P. Samiksha, K. Ravi, K. Manesh, D. Girish, S. Vivek, and M. Fabrice. (2018). *Indian Diabetic Retinopathy Image Dataset (IDRiD)*, doi: 10.21227/H25W98.

[93] P. Prentašić, S. Lončarić, Z. Vatavuk, G. Benčić, M. Subašić, T. Petković, L. Dujmović, M. Malenica-Ravlić, N. Budimlija, and R. Tadić, "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.

[94] S. Sivaprasad, J. C. Vasconcelos, A. T. Prevost, H. Holmes, P. Hykin, S. George, C. Murphy, J. Kelly, G. B. Arden, "Clinical efficacy and safety of a light mask for prevention of dark adaptation in treating and preventing progression of early diabetic macular oedema at 24 months (CLEOPATRA): A multicentre, phase 3, randomised controlled trial," *Lancet Diabetes Endocrinol.*, vol. 6, no. 5, pp. 382–391, May 2018.

[95] S. M. Shaharum *et al.*, "Automatic detection of diabetic retinopathy retinal images using artificial neural network," in *Proc. 10th Nat. Tech. Seminar Underwater Syst. Technol.* Singapore: Springer, 2019, pp. 495–503.

[96] O. Perdomo, S. Otalora, F. Rodríguez, J. Arevalo, and F. A. González, "A novel machine learning model based on exudate localization to detect diabetic macular edema," in *Proc. 3rd Int. Workshop Ophthalmic Med. Image Anal. (OMIA)*. Iowa City, IA, USA: Univ. of Iowa, 2016, pp. 137–144.

[97] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manage.*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

[98] G. Chen, P. Faris, B. Hemmelgarn, R. L. Walker, and H. Quan, "Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa," *BMC Med. Res. Methodol.*, vol. 9, no. 1, Dec. 2009.

[99] H. Zheng, R. Wang, Z. Yu, N. Wang, Z. Gu, and B. Zheng, "Automatic plankton image classification combining multiple view features via multiple kernel learning," *BMC Bioinf.*, vol. 18, no. S16, p. 570, Dec. 2017.

[100] N. Tang, F. Zhou, Z. Gu, H. Zheng, Z. Yu, and B. Zheng, "Unsupervised pixel-wise classification for chaetoceros image segmentation," *Neuro-computing*, vol. 318, pp. 261–270, Nov. 2018.

[101] J. H. Noordik, *Cheminformatics Developments: History, Reviews and Current Research*. IOS Press, 2004.

[102] G. Bargshady, X. Zhou, R. C. Deo, J. Soar, F. Whittaker, and H. Wang, "Enhanced deep learning algorithm development to detect pain intensity from facial expression images," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113305.

[103] D. Pandey, X. Yin, H. Wang, M.-Y. Su, J.-H. Chen, J. Wu, and Y. Zhang, "Automatic and fast segmentation of breast region-of-interest (ROI) and density in MRIs," *Heliyon*, vol. 4, no. 12, Dec. 2018, Art. no. e01042.

[104] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: http://arxiv.org/abs/1409.1556

[105] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[106] W. Ouyang, P. Luo, X. Zeng, S. Qiu, Y. Tian, H. Li, S. Yang, Z. Wang, Y. Xiong, C. Qian, Z. Zhu, R. Wang, C.-C. Loy, X. Wang, and X. Tang, "DeepID-net: Multi-stage and deformable deep convolutional neural networks for object detection," 2014, *arXiv:1409.3505*. [Online]. Available: http://arxiv.org/abs/1409.3505

[107] H. Chen and P. Cao, "Deep learning based data augmentation and classification for limited medical data learning," in *Proc. IEEE Int. Conf. Power, Intell. Comput. Syst. (ICPICS)*, Jul. 2019, pp. 300–303.

KHANDAKAR AHMED (Senior Member, IEEE) received the M.Sc. degree in Network and e-Business Centred Computing (NeBCC) under the joint consortia of the University of Reading, U.K.; the Aristotle University of Thessaloniki, Greece; and the Charles III University of Madrid (UC3M), Spain, and the Ph.D. degree from RMIT University, Melbourne, VIC, Australia. He is currently a Lecturer with the Discipline of IT, College of Engineering and Science, Victoria University. He has extensive industry engagement as a CI in multiple research projects related to the Internet-of-Things, smart cities, machine learning, cybersecurity, and health informatics. He has published more than 50 refereed scholarly articles.

HUA WANG (Member, IEEE) received the Ph.D. degree in computer science from the University of Southern Queensland (USQ), in 2004. He was a Professor with USQ, from 2011 to 2013. He is currently a full-time Professor with the Centre for Applied Informatics, Victoria University. He has authored or coauthored over 250 peer-reviewed research articles mainly in artificial intelligence, data analysis, data mining, access control, privacy, and web services, as well as their applications in the fields of e-health and e-environment.

YANCHUN ZHANG (Member, IEEE) has been a Professor and the Director of the Centre for Applied Informatics, Victoria University, since 2004. He has published over 300 research papers in international journals and conference proceedings, including the *ACM Transactions on Computer-Human Interaction* (TOCHI), the IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING (TKDE), *VLDB* Journal, SIGMOD, and ICDE conferences as well as health/medicine journals. His research interests include databases, data mining, web services, and e-health. He is a Founding Editor and the Editor-In-Chief of *World Wide Web Journal* (Springer) and *Health Information Science and Systems* journal (Springer). He served as an expert panel member at various international funding agencies, such as the Australia Research Council; the National Natural Science Fund of China (NSFC); the Royal Society of New Zealand's Marsden Fund, from 2015 to 2017; the Performance-Based Research Fund (PBRF) Panel of New Zealand; the Medical Research Council, U.K.; and the NHMRC of Australia on Built Environment and Prevention Research.

RUBINA SARKI received the M.E. degree in human–computer interaction and robotics from the University of Science and Technology, South Korea. She is currently pursuing the Ph.D. degree in computer science with the School of Engineering and Science, Victoria University, Melbourne, VIC, Australia. She worked as a Researcher with the Korea Institute of Science and Technology (KIST). Her research interests include the development of automated disease classification and detection systems, deep learning, data analysis, image analysis, machine learning, and automated disease detection systems. She was a recipient of the International Postgraduate Research Scholarship.

• • •