IDEATION PHASE

Literature Survey

Deep learning Fundus Image Analysis For Early Detection of Diabetic Retinopathy

---***---

Abstract - *Diabetes Mellitus, often known as Diabetes, is a condition in which a person's body either does not respond to insulin supplied by their pancreas or does not create enough insulin. Diabetics are at a significant risk of acquiring a variety of eye disorders throughout time. Early identification of diabetic eye illness via an automated system has significant advantages over manual detection as a consequence of developments in machine learning techniques. A number of sophisticated research on diabetic eye disease detection have lately been published. This paper provides a thorough review of automated approaches to diabetic retinopathy identification from multiple perspectives, including: accessible datasets, image preprocessing techniques, deep learning models, performance assessment criteria. The study gives a thorough overview of diabetic eye disease detection methodologies, including cutting-edge field approaches, with the goal of providing important knowledge to research communities, healthcare providers, and patients.*

Key Words: Diabetic eye disease, diabetic retinopathy, deep leaning, fundus image, image processing.

1. INTRODUCTION

In many urbanised nations, diabetic retinopathy has been identified as one of the primary causes of blindness. According to the World Diabetes Foundation, by 2030, there will be more than 438 million individuals suffering with diabetes. As a result, early detection and diagnosis can help patients avoid losing their vision as a result of this symptomless diabetes condition [1]. It is a type of metabolic illness that develops as a result of a high amount of glucose in the blood, which causes an insulin secretion malfunction and visual difficulties. Small blood vessel alterations in the retina generate three phases of diabetic retinopathy: pre-proliferative retinopathy, proliferative retinopathy, and non-proliferative retinopathy [2].

The development of exudates on the retina is another hallmark of diabetic retinopathy that has become an important clinical indicator for automated disease identification and diagnosis [3]. Early identification of diabetic retinopathy by an ophthalmologist is more difficult owing to the difficulty in detecting them since they occur in smaller sizes. The early signs of diabetic retinal disease can be recognised utilising machine learning and deep learning algorithms in a real-time automated system. As a result, it can easily reduce the risk of human mistake and the quantity of work that the ophthalmologist has to do [4].

Following that, the severity phases are defined to help the ophthalmologist make suitable treatment and planning decisions following the diagnosis of non-proliferative diabetic retinopathy. This diagnosis is solely based on the retinal fundus image's categorization of dark and bright lesions. Due of the restricted number of specialists available internationally, reading fundus photographs may be more costly. Because of the poor availability of ophthalmologists to patients, the prediction process may be slowed by the lack of a highly trained reader [5].

Furthermore, the automated approach is more reliable and efficient than human identification of diabetic retinopathy severity stages. Effective preprocessing techniques are applied to the recorded diabetic retinopathy pictures in order to eliminate noisy characteristics. The discriminative characteristics are then taken from the pictures by segmenting the region of interest in order to create the appropriate predictions. Despite the fact that several research papers have made significant contributions to the automated diabetic retinopathy screening method. However, due to the complicated structure of lesions, the existence of noise, and the high appearance of interclass similarity, lesions identification remains a difficult task. These characteristics may make it difficult for the automated system to create a flexible and reliable model for lesion segmentation [6].

In the presence of lesions in the target picture, image resolution is also important for disease prediction and categorization of severity stages. A robust solution is required for lesions recognition, segmentation, and disease severity grading, according to the research findings mentioned above. As a result, the goal of this research project is to create a one-of-a-kind platform that can identify and analyse retinal pictures on-board. Rather of exposing the afflicted patient to laser surgery, diabetic retinopathy can be avoided at an earlier stage. Because of a lack of competence and resources, many patients are unable to receive prompt screening and diagnosis [7].

However, as scanning ophthalmoscopes and digital cameras become more common, telemedicine can be used to address on-board diagnosis and prediction of diabetic retinopathy illness. As a result, tele-ophthalmology is recognized as a critical component of the diabetic retinopathy screening procedure. The need for online eye care hospitals to treat diabetes patients is growing as feasible, practical, and clinically validated options become more difficult to come by [8].

Furthermore, the risk of diabetic retinopathy can be reduced by providing good patient information and maintaining a healthy blood glucose level. There are numerous methods used in existing research studies to identify diabetic retinopathy, including dilated eye examination, fluorescein angiography, and optical coherence tomography or fundus photography techniques [9].

Nowadays, visual evaluation, which appears to be a more difficult process, need the use of a retina expert for diagnosis and grading. As a result, most patients would seek the help of a retina expert after losing their eyesight owing to a lack of access to qualified doctors. As a result, a suitable and cost-effective computer-aided diagnosis system is required to treat diabetic retinopathy disease at an early stage.

Figure 1 shows the example images of diabetic retinopathy.

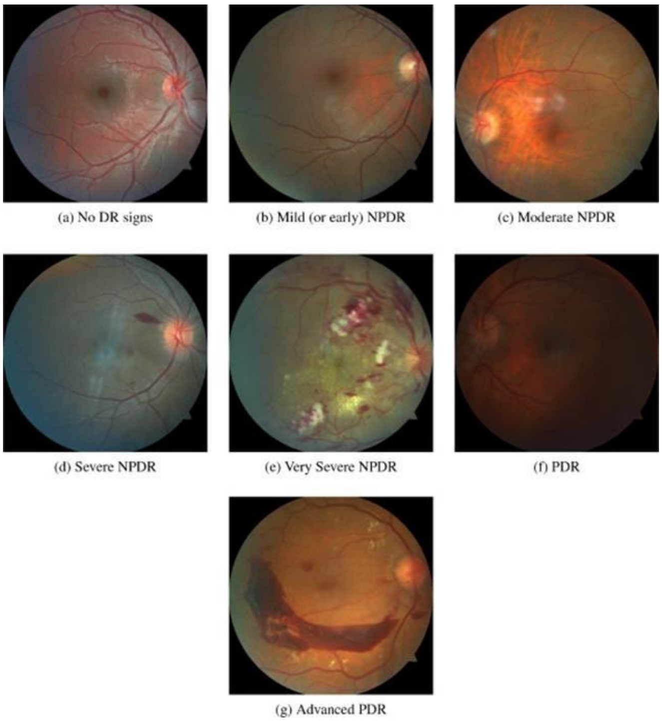

Fig -1: Examples of diabetic retinopathy images

Contribution

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

1) Datasets available for DR

2) Preprocessing techniques applied to fundus images for DR detection

3) DL approaches proposed for DR detection

4) Performance measures for DR detection algorithm evaluation.

The arrangement of this article is as follows. Section II analyses the papers based on the datasets used in their study. Section III addresses the image processing techniques used in the prior work. Section IV analyses the articles based on the classification methods employed. Section V discusses the findings and observations. Finally, Section VI concludes the paper.

2. DIABETIC RETINOPATHY DATASETS

The writers of the chosen publications make use of private and public datasets that are separated into training and testing scenarios. Kaggle and Messidor are the most popular datasets for DR detection. The Kaggle dataset contains 88,702 photos, 35,126 of which are utilised for training and 53,576 for testing. The Messidor dataset, which contains 1,200 fundus pictures, is the most extensively utilised. The DR phases are identified in the Kaggle and Messidor datasets. Table 2 lists the datasets used in the selected studies, organised by the different DEDs studied, i.e. DR, Gl, DME, and Ca. The table includes the DED's name, the dataset's name, a synopsis of the dataset, the sources of publications that utilised the dataset, and lastly, the URL where the dataset may be obtained (if accessible publicly).

Kaggle EyePACS

The Kaggle EyePACS dataset, which contains over 80.000 fundus images and was provided by the EyePACS platform for the Diabetic Retinopathy Detection competition sponsored by the California Healthcare Foundation, is the most widely used and largest public dataset for Diabetic Retinopathy classification. It comprises of a huge number of high-resolution fundus pictures of both eyes taken with various technologies at several primary care facilities throughout California and worldwide under a range of imaging settings. Due to this unpredictability, both the data (e.g. artefacts, blurring, focusing, and exposure issues) and the ground truth labels display noise, which was meant to better replicate a real-world scenario. The photos were rated using the ICDRDSS scale [10] by a certified professional.

Kaggle APTOS 2019

Aravind Eye Hospital in India's rural areas acquired the Kaggle APTOS 2019 Challenge [11] dataset in order to construct robust tools to automatically diagnose Diabetic Retinopathy and increase the hospital's capacity to identify future patients. It has 5590 photos and is the third biggest collection. However, one of its flaws is the enormous class imbalance, particularly in the Severe NPDR class, which only comprises 193 photos. APTOS dataset, like Kaggle

EyePACS dataset, displays variances owing to varying camera settings between centres and noise both in the data (i.e. artefacts, focus difficulties, being under/overexposed) and the labels (due to being gathered in a real world multicentre scenario).

Messidor & Messidor 2

The Messidor dataset [12] contains 1200 retina fundus pictures taken between 2005 and 2006 by three ophthalmology institutions in France. The first 800 photos were taken with pupil dilation, whereas the latter 400 were taken without it. The Messidor 2 dataset includes 1058 photos from the original Messidor dataset, as well as 690 additional photographs gathered in the Brest University Hospital's Ophthalmology department between 2009 and 2010. Unlike the Kaggle EyePACS dataset, both dataset's pictures are of excellent quality, with no discernible noise. The datasets provide an image-level medical diagnostic of the severity of Diabetic Retinopathy for each image, but no pixel-level lesion segmentation information. Their own grading system, on the other hand, did not follow the widely recognized ICDRS procedure, limiting its validity and usefulness.

IDRiD

The IDRiD [13] dataset contains 516 high-quality photos taken with a Kowa VX - 10 fundus camera at an ophthalmology clinic in Nanded, India. Prior to the picture capturing operation, all individuals' eyes were dilated in both directions. For all 516 photos, it includes image-level grading on the severity of Diabetic Retinopathy according to the ICDRS scale, as well as grading on the risk of Diabetic Macula Edema (DME). It also includes pixel-by-pixel annotations on the pertinent lesions (such as Hard and Soft Exudates, Microaneurysms, and Haemorrhages) as well as the optical-disc structure for 81 photos in the dataset.

DDR

With 12522 photos, the DDR [14] dataset is the second biggest when it comes to the classification challenge, although it is a relatively new dataset that hasn't been frequently utilised yet. The data was collected between 2016 and 2018 in 147 hospitals throughout China's 23 provinces and annotated by numerous specialists using a majority voting schema according to the ICDRDSS scale. Furthermore, a sixth grade was added to separate low-quality photos into a distinct group. However, there is a significant disparity between the healthy/moderate DR classifications and the others, such as mild, severe, and proliferative DR, which might result in overfitting. Regarding the relevant DR lesions, 757 pictures from the dataset were labelled at the pixel level for lesion segmentation and detection purposes, as well as bounding boxes surrounding them.

Table -1: Datasets of retina fundus images

Name	Size	Resolution	Annotation	Task
EyePACS [10]	88702	Varying	Image Level	DR Grading
APTOS 2019[11]	5590	Varying	Image Level	DR Grading
Messidor [12]	1200	1440 × 960, 2240 × 1488 2304 × 1536	Image Level	DR Grading
Messidor 2 [12]	1748	Varying	Image Level	DR Grading
IDRiD [13]	576	4288 × 2848	Image & Pixel Level	DR Grading
DDR [14]	12522	Varying	Image & Pixel Level	DR Grading
E-Ophtha [15]	463	Varying	Pixel Level	Healthy vs. Diseased Exudates and Micro aneurysms Detection
DiaRetDB 1 [16]	89	1500 × 1152	Pixel Level	Lesion Segmentation
DRiDB [17]	50	768 × 584	Pixel Level	MAS, HMs, HEs, SEs, OD and Macula Detection and Vessel Extraction

E-Ophtha

The E-Ophtha [15] dataset contains 463 photos, including 268 images pertaining to healthy people, 148 images pertaining to patients with micro aneurysms or other tiny red lesions, and 47 images pertaining to exudates. It has been used to forecast DR in a binary job automatically (healthy vs diseased). However, because the dataset contains fewer pictures than larger datasets (Kaggle and Messidor), it is generally utilised in the research to create segmentation techniques rather than classification systems.

DiaRetDB1

DiaRetDB1 [16] is made up of 89 fundus photos that were obtained in a controlled setting at Kuopio University and rated by four specialists. However, because not only is the data sample small and from a single clinical location, but all of the photos were obtained in a controlled setting with no substantial changes in the collecting technique, their distribution does not reflect a typical population.

DRiDB

DRiDB [17] is a collection of 50 fundus photographs with comments on the anatomy of the retina's optic disc and vessels, as well as any existing diseases, neovascularizations, and disease grades, all of which were determined by numerous specialists. Despite the fact that it is a tiny dataset, it is the most instructive.

3. PREPROCESSING TECHNIQUES APPLIED TO FUNDUS IMAGES FOR DR DETECTION

Contrast enhancement

To begin, contrast enhancement is a popular preprocessing method used in any image processing or analysis pipeline to distinguish the foreground from the background. The histogram equalization [18] is a basic approach for contrast enhancement in fundus pictures that raises the overall contrast of the image while ignoring the local changes throughout the image. Adaptive Histogram Equalization is a more advanced contrast adjustment method that takes into account local differences surrounding a specified region of each pixel. Contrast Limited Adaptive Histogram Equalization (CLAHE) is more widely employed by the scientific community when it comes to fundus imaging.

Denoising & normalization

[19] uses Non-Local Means Denoising (NLMD) to reduce any picture noise. However, while the better the denoising algorithm is, the more noise it will remove, it will also impair the image's fine features (i. e. the image becomes blurry). Additionally, picture intensity normalization is used to reduce bias and long training durations in the network, as well as to standardize the data to a certain scale (e.g., each image has a mean value of 0 and a standard deviation of 1 in terms of its pixels' intensity).

Color space transformation

Apart from contrast improvement, normalisation, and noise reduction, the model's performance has been improved by converting the colour picture into another colour model or just using one of the RGB channels. Lin et al. [20] used entropy pictures to convert the data, resulting in the DLS outperforming models trained on standard datasets. Furthermore, due to its rich information and strong contrast in compared to the other two colour channels, the green channel of the fundus colour picture is frequently extracted.

Cropping and resizing

Additionally, the files may include photos with varying resolutions and aspect ratios. Uninformative dark space portions might also be seen in the photos. The photographs may be cropped, rescaled, and enlarged to a certain resolution to standardise the image size and remove such dark space areas. Bravo et al. [21] used two distinct cropping approaches in their trials. In one experiment, they cropped the pictures so that the retina encircled the cropped image, but in the other experiment, they clipped the retina's greatest square image.

Augmentation

Although DL has been shown to operate effectively when raw data is put into a single-model pipeline, it has also been claimed that using specific preprocessing approaches, such as those described in this section, improves performance [22], particularly for fundus pictures. In addition, data augmentation techniques are utilised to improve the model's resilience and accuracy owing to a shortage of rich and balanced datasets. Rotating, shifting (translation), rescaling, shearing, and flipping pictures, as well as the usage of Generative Adversarial Networks for image synthesis, are examples of such approaches in the case of imaging datasets. The majority of the publications analysed employ some form of augmentation to enhance the number of available photos and hence speed up the model's training.

4. DL APPROACHES PROPOSED FOR DR DETECTION

Diabetic retinopathy is the major cause of blindness in people in their working years. Carson et al [22] used colour fundus pictures to illustrate the application of convolutional neural networks (CNNs) for diabetic retinopathy stage detection. They also looked at multinomial classification models, demonstrating that due to the CNN's failure to recognize subtle illness signs, the majority of mistakes arise when mild disease is misclassified as normal. They observed that using contrast limiting adaptive histogram equalization and assuring dataset consistency through expert class label verification improved detection of subtle characteristics.

Sheikh et al. [23] constructed a unique deep convolutional neural network that conducts early-stage detection by recognizing all microaneurysms (MAs), which are the earliest indicators of DR, as well as properly assigning labels to retinal fundus pictures that are graded into five categories. They used the largest publicly accessible Kaggle diabetic retinopathy dataset to test the network. In terms of both computing time and space, the suggested design is both simple and efficient.

Wan et al. [24] figured out a mechanism to categorise a series of fundus photos automatically. They use convolutional neural networks (CNNs) to solve three fundamental problems in DR detection: classification, segmentation, and detection. We use AlexNet, VGGNet, GoogleNet, and ResNet, together with transfer learning and hyper-parameter tweaking, to see how well these models do with DR image categorization. They used the Kaggle platform, which is open to the public, to train these models.

A diabetic retinopathy deep learning interpretable classifier was reported by De et al [25]. On the one hand, it accurately separates retina pictures into distinct severity degrees. On the other hand, by giving a score to each point in the hidden and input spaces, this classifier is capable of explaining the classification findings. The contribution of each pixel to the final categorization is indicated by these scores. They suggested a novel pixel-wise score propagation model to get these scores, which separates the observed output score into two components for each neuron. The resulting visual maps can be easily comprehended by an ophthalmologist using this procedure.

Rakhlin [26] uses deep Convolutional Neural Networks (CNNs) to diagnose eye fundus pictures, which have proved innovative in a variety of disciplines of computer vision, including medical imaging. They utilised a publicly available Kaggle data set to train the models. They employed a subset of Kaggle data that had been omitted from training and the Messidor-2 reference standard for testing. When done by skilled optometrists, the findings are comparable to contemporary state-of-the-art models trained on far larger data sets and outperform the average outcomes of diabetic retinopathy screening.

Manoj et al [27] developed the prediction network using the publicly available Kaggle dataset using a convolutional neural network-based technique for autonomous detection of diabetic retinopathy. The network was trained using around 35,000 photos. The network was trained to identify the laterality of the eye using 8,810 photos.

Xiaomeng et al. [28] proposed a unique cross-disease attention network (CANet) to simultaneously assess DR and DME by examining the intrinsic link between the illnesses using just image-level supervision. The disease-specific attention module, which selectively learns relevant characteristics for individual diseases, and the disease-dependent attention module, which captures the intrinsic link between the two diseases, are two of our major achievements. We use a deep network to combine these two attention modules to provide disease-specific and disease-dependent characteristics, as well as to maximize overall performance for grading DR and DME. Our network is tested on two publicly available benchmark datasets: the ISBI 2018 IDRiD challenge dataset and the Messidor dataset.

A technique based on the Radon transform (RT) and the visibility graph (VG) was suggested by Mohammadpoory et al [29] to automatically identify grades 0 (normal), 1, 2, and 3 of the DR from fundus pictures. Feature extraction and classification are the two steps of the proposed technique. For the first time in the realm of image processing, the VG approach was used to extract features in this work. Then, for classification purposes, these attributes were handed to the error-correcting output codes (ECOC) approach. The proposed approach was simple to use, with a 97.92 percent accuracy, 95.83 percent sensitivity, and 98.61 percent specificity. The VG-based approach can be a simple, inexpensive, and effective test for automatic DR stage grading, and it can also be used in other image processing applications.

Abdelsalam [30] developed a comprehensive and reliable automated system for detecting DR individuals early. Two stages are necessary for the process to work: 1) Employing created custom algorithms to reconstruct, improve, and re-continue blood vessels, and 2) using an Artificial Neural Network (ANN) as an automated classifier between diabetics without diabetic retinopathy (DR) and Mild to Moderate Non-Proliferative Diabetic Retinopathy (NPDR) individuals. The overall accuracy of the method was 97 percent. The classification system's performance characteristics for normal vs diabetes were 97.5 percent sensitivity, 96.67 percent specificity, and 95.2 percent accuracy. The sensitivity, specificity, and precision scores for a diabetic without DR against non-proliferative DR (mild to moderate) were 96.67 percent, 96.67 percent, and 96.67 percent, respectively. 3.33 percent of the time, there was a misclassification mistake.

Xianglong et al [31] used a transfer learning strategy to build a unique convolutional neural network model with a Siamese-like design. Unlike earlier studies, the proposed approach takes binocular fundus pictures as inputs and learns their association to aid prediction. The suggested binocular model achieves an area under the receiver operating curve of 0.951 with a training set of just 28 104 pictures and a test set of 7024 images, which is 0.011 greater than the existing monocular model. A binocular model for five-class DR detection is also trained and assessed on a 10% validation set to further validate the efficiency of the binocular architecture. The result reveals that it gets a kappa score of 0.829, which is greater than the non-ensemble model currently in use.

Multiple weighted pathways were implemented to a convolutional neural network termed the WP-CNN by Liu et al [32], which was driven by ensemble learning. Back propagation is used to improve various path weight coefficients in WP-CNN, and the output features are averaged for redundancy reduction and quick convergence. The experiment findings reveal that WP-CNN obtains an accuracy of 94.23 percent with sensitivity of 90.94 percent, specificity of 95.74 percent, area under the receiver operating curve of 0.9823, and F1-score of 0.9087 while using an efficient training convergence rate.

Ensemble learning approaches

By integrating the benefits of different AI frameworks, ensemble learning has also played a key role in establishing robust and strong AI frameworks for DR categorization. Because of the knowledge gain produced by their complementarity, ensemble learning has been shown to outperform solo models. That means that various base models might acquire different degrees of semantic representations implicitly, either due of changes in their design or because of the training technique.

Zhang et al. [33] created two ensemble models, one for illness identification (binary classification) and the other for disease grading (quinary classification). The various models were built using a combination of pre-trained networks for feature extraction and a bespoke standard dense neural network for classification. In all tasks, the ensemble models beat individual models, with a sensitivity of 98.10 percent and a specificity of 98.56 percent. The authors also point out that the 'stronger' the base learner (pre-trained network), the better the overall performance. Furthermore, in several circumstances, a dual ensemble (ensemble of ensembles) outperformed a single ensemble.

Because, according to the authors, each distinct lesion type is best recognised at various training iterations, Quellec et al. [34] built a CNN model that was exported at numerous iterations during the training method. They then used ensemble learning (Random Forest Classifier) to merge the stored models to predict DR's severity score.

Jiang et al. [35] used the Adaboost classifier on three models based on the InceptionV3, ResNet152, and Inception-Resnet-V2 architectures to create an ensemble model. They used a proprietary dataset created in partnership with Beijing Tongren Eye Centre to train the algorithm. Sensitivity = 85.57 percent, Specificity = 90.85 percent, Accuracy = 88.21 percent, and AUC = 0.946 were achieved by the ensemble model, which outperformed the individual models. InceptionV3 did better in terms of Specificity, with a score of 91.46 percent.

5. PERFORMANCE MEASURES FOR DR DETECTION ALGORITHM EVALUATION

Some specific measures, such as Sensitivity, Specificity, Accuracy, Precision, and F1 score, can be used to characterize the classification model's performance. Furthermore, the Receiver Operating Characteristic (ROC) curve depicts the classifier's performance by graphing its Sensitivity against its Specificity at various thresholds for the classification outcome (i.e. at which probability a given sample is considered as a positive or negative outcome). Finally, AUC calculates the area beneath the whole ROC curve, providing an overall performance metric across all categorization criteria. The performance of the reviewed papers is shown in Table 2.

Table -2: Performance of the reviewed classification models.

References	Classification	Best architecture	Accuracy
[22]	2-class (EyePACS) 5-class (EyePACS)	GoogLeNet	-
[23]	2 Diseased (EyePACS) 2 Risk (EyePACS) 5-class (EyePACS)	Custom CNN	-
[24]	5-class (EyePACS)	VGGNet	95.68 %
[25]	2-class (Messidor-2) 5-class (Messidor-2)	Custom CNN	91%
[26]	2-class (EyePACS) 2-class (Messidor)	Modified VGGNet	-
[27]	5-class (EyePACS)	Custom CNN	-
[28]	2-class (Messidor)	Custom CNN	92.6 %
[29]	2-class	Radon transform	97.92 %
[30]	2- class	ANN	97%
[31]	2-class	WP-CNN	94.23 %

6. CONCLUSIONS

Diabetic retinopathy is a serious complication of diabetes that causes progressive retinal deterioration and, in severe cases, blindness. To avoid degeneration and retinal damage, it is vital to recognize and treat it as soon as possible. In recent years, there has been a surge in interest in employing deep learning to detect diabetic retinopathy, and as various DL systems develop and are integrated into clinical practice, clinicians will be able to treat patients more successfully and efficiently. The current state of research on the application of deep learning in the diagnosis of diabetic retinopathy is examined in this article. Ophthalmologists still need to improve their performance, interpretability, and trustworthiness, despite the fact that deep learning has paved the way for more accurate diagnosis and treatment.

REFERENCES

J. Kaur and D. Mittal, Estimation of severity level of non-proliferative diabetic retinopathy for clinical aid, Biocybernetics and Biomedical Engineering 38 (2018) 708-732.

R. O. Fite et al., Diabetic retinopathy in Ethiopia: A systematic review and meta-analysis, Diabetology & Metabolic Syndrome 13 (2019) 1885-1891.

S. Joshi and P. T. Karule, A review on exudates detection methods for diabetic retinopathy, Biomedicine & Pharmacotherapy 97 (2018) 1454-1460.

U. Ishtiaq et al., Diabetic retinopathy detection through artificial intelligent techniques: A review and open issues, Multimedia Tools and Applications 79 (2020) 21-22.

J. Lin et al., Retinal image quality assessment for diabetic retinopathy screening: A survey, Multimedia Tools and Applications 79 (2020) 16173-16199.

P. Porwal, et al, IDRiD: Diabetic retinopathy – Segmentation and grading challenge ID, Medical Image Analysis 59 (2019).

S. Stolte and R. Fang, A survey on medical image analysis in diabetic retinopathy, Medical Image Analysis 64 (2020) 101742.

D. Kalogeropoulos et al., The role of tele-ophthalmology in diabetic retinopathy screening, Journal of Optometry 13 (2020) 262-268.