International Journal of Computer (IJC)
ISSN 2307-4523 (Print & Online)
http://ijcjournal.org/

A Review on Detection of Diabetic Retinopathy using Deep Learning and Transfer Learning based Strategies

Pratik Bijam[a*], Smita Deshmukh[b]

[a,b]Department of Information Technology Terna Engineering College Mumbai, India
[a]Email: pratikbijam@ternaengg.ac.in
[b]Email: smitadeshmukh@ternaengg.ac.in

Abstract

Diabetic Retinopathy (DR) is considered to be one of the most widely observed and a complex variation of diabetes and stands as a leading cause of blindness globally. The occurrence of DR causes impairment in the retinal blood vessels and leads to unusual growth of blood arteries in the eye. Manual examinations and analysis suggests that the prevalence of DR has been enormously growing at an exponential rate and has already registered for more than 160 million cases worldwide. On the other hand, its diagnostic screening is not only challenging, but also computationally expensive at the same time. Due to the highlighting importance of its early diagnosis in terms of treatment, multiple concepts to DR detection have been used in the past few years. However, research in recent times has resulted in the fact that deep learning based CNN structures and Transfer Learning based MedNets have been popularly used in DR detection, due to its superior performance in the medical domain. As a result of such advancements in Deep Learning methodologies, this article proposes a review on automated approaches used to detect diabetic retinopathy using image processing and disease classification techniques. The review is further preceded with a comprehensive analysis on training a model with an already pre-trained network whose primary goal is to generate useful information and provide it to diabetic researchers, medical practitioners and patients.

Keywords: diabetic retinopathy; deep learning; fundus images; glaucoma; transfer learning.

1. Introduction

The primary occurrence of Diabetic Retinopathy is due to the prevalence of diabetes in an individual wherein the blood vessels present in the eye do not react to the insulin being produced by the body and therefore leads to multiple metabolic disorders. In the later stages, the swollen blood vessels begin to damage the capillaries in the body's venules and clump together which results in disrupting the entire process of circulation. This leads to impairment of the eyes and enlargement of retinal capillaries, known as ischemia [1]. Ischemia in the retina stimulates the formation of cytokine protein, which promotes the growth of new blood vessels from the existing ones and leads to deformation of the retinal surface.

----

* Corresponding author.

This ends up creating an internal gravity and finally induces the tissue to expand, resulting in vision loss. Since the primary function of a retina is to sense light and send those particular signals to the brain, its injury directly impacts the eye sight of an individual [2]. According to multiple research scholars, a DR can be detected in either of the two stages; with the first one involving an early detection of DR, followed by an advanced detection of the same. In the early detection of DR, the blood vessels present in the retina tend to become weak and are unable to respond to the vessel surfaces [3]. This process results in dilation of blood vessels and therefore becomes irregular in diameter. Its treatment is however possible, with further classification of the diagnosis as mild, moderate and severe. The second stage of the disease, commonly known as the advanced DR, often tends to create a leak in the blood vessels of the retina which increases the pressure of the eye and damages the optic nerve, leading to glaucoma [4].

According to the worldly statistics, a rough figure of 464 million diabetic retinopathy patients have been recorded, and the numbers are expected to increase to 552 million by the end of 2035 [5]. Thereby making it the fundamental reason behind the occurrence of vision loss globally. Figure below depicts the samples of retinopathy wherein the disease is diagnosed at multiple stages:

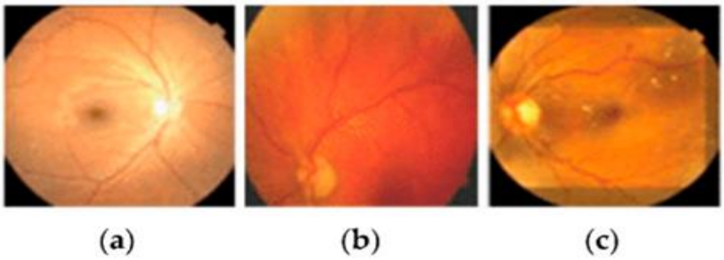

Figure 1: Retinopathies observed in (a) normal (b) mild (c) moderate.

Detecting and grading DRs in the initial phases is time-consuming and necessitates technical experts. Furthermore, the manual evaluation of DR patients reveals a great deal of discrepancy among practitioners. Approximately 80 percent of DR patients reside in developing or underdeveloped countries, which lack ophthalmologists and a basic DR detection mechanism [6]. Hence, it has been observed and proven that manual screening techniques are still unable to keep pace with the growing need for diagnostic methods of DR worldwide. Due to rapid advancements in the field of computer vision, various automated systems have been evolved by research scholars. However, improvements in computer-aided diagnosis (CAD) systems face significant challenges, including identifying lesions from a retina, subdivision of the optic nerve head, segmentation of arteries, and so on. Although machine learning-based frameworks have demonstrated resilience in DR detection, their effectiveness is heavily reliant on handcrafted features that are yet difficult to generalise. On the other hand, Deep learning (DL) methods have provided automatic feature extraction techniques from fundus images to overcome these drawbacks. In a DL based approach, all the required images are collected from the data repository and image processing methodologies are applied to the obtained dataset. Since the raw data so collected, is imbalanced and contains noise, its dimensionality is reduced and later fed to the hidden neurons for automated extraction of relevant features. Later, all the features are assigned weights and optimized recursively to ensure the best classification results. However, all the techniques that follow the concepts of DL require a certain amount of computational memory and power. This can be challenging at certain times.

Therefore, an approach known as Transfer Learning is adopted that accounts for all the shortcomings and solves the majority of the problem by reducing the large amount of dataset. The concept is primarily based on transferring as much as knowledge possible from one trained network to another model which is designed to perform a similar task. Hence, the major determination of the study is to highlight those methods which can be used for categorizing early stage DR to assist ophthalmologists.

2. Related Works

Multiple research scholars have worked on fundus images and have successfully demonstrated the automation of a system that can detect diabetic retinopathy. They contributed their research on DR lesions and further classified those using DL techniques in an attempt to detect the same. The experiments used in these techniques relied on morphological features that calculated the retina's pixel intensity, Furthermore, when compared to the methods used by clinicians, the proposed method produced better and more accurate results. Hence, it can be observed that an extensive amount of research was performed in the same domain that could detect and classify retinopathy patients into several categories. This section of the review highlights traditional strategies used to classify DR lesions of retina.

A. Binary Classification

The studies conducted under this section depict the categorization of the DR Dataset into two classes. Authors in [7] performed their classification using the CNN network obtained from Kaggle repository. The dataset contained images acquired from DR patients and underwent the process of dimensionality reduction. Once all the images were reduced to a certain dimension of 224*224*3, the concept of augmentation was applied to them. This step was primarily done to increase the dataset through transformations such as rotation and flipping. However, the implementation of the CNN architecture involved 14 input layers, with the activation function being applied to the last layer as Softmax. The conducted experiment generated an accuracy of 94.5 percent.

In a similar study performed by G. Quellec and his colleagues in [8] all the input images obtained from the repository were labelled as positive DR and negative DR. The model was trained on a CNN network and dimensionality reduction was performed in the pre-processing stage, wherein the images were reduced to a size of 448 x 448 pixels. The implementation of the CNN architecture was further combined with the concepts of Transfer Learning based AlexNet model. The trained model achieved an overall accuracy of 95 percent when augmentation was performed using a Gaussian filter. M. T. Esfahan and his colleagues in [9] used the fundamentals of ResNet34 along with the implementation of the CNN model. Since ResNet34 was a heavily trained model, a large number of images were used and resized to 512 x 512 pixels. A similar set of pre-processing techniques were applied and the authors successfully contributed their work into achieving an accuracy of 85 percent.

Authors in [10] proposed to automate a model based on CNN architecture and transfer learning based on the VGG-16 model. Further, they conducted their experiments using cross validation techniques and resizing the input image to a dimension of 512 x 512 pixels. The model underwent the process of initialization of respective weights so that the problem of overfitting could be avoided. The experiment was conducted on the dataset obtained from Kaggle repository and the trained model was balanced using augmentation and further classified as DR positive or DR negative. The proposed model reached an accuracy of 98.8 percent.

The study of authors in [11] contributed to three conceptual models of transfer learning, namely; Inception, ResNet-V2 and ResNet152. The primary aim of the authors was to conduct a comparative analysis and derive which model could generate optimized results and classify the dataset as preferred DR or non-referable DR. The authors also combined this model with the CNN architecture and performed dimensionality reduction to resize the image pixels to 520 x 520. Due to the augmentation being performed on the dataset, the model achieved an accuracy of 88.21 percent through ResNet152.

Y. Liu And his colleagues in [12] proposed a method to detect DR images using CNN and data augmentation. The author experimented his study on 60,000 images and resized them through reduction and later fed the input images to the CONV layers with varying kernel sizes. The model was further merged with transfer learning based variants so as to enhance the overall accuracy to 94.23 percent.

However, one of the major disadvantages with binary classification was that it did not take into consideration the stages involved in a conventional DR and therefore overlooked them and generated results. Since, the determination of the exact stage of a patient's DR is necessary for proper treatment of retina, binary classification does not prove to prevent the overall vision loss of a patient so involved.

B. Multi Level Classification

This section of the review briefs on similar studies being performed wherein the obtained dataset is categorized into multiple classes.

An author contributed his work in [13] to detect and diagnose the presence of DR and diabetic macular edema (DME) using CNN as the base architecture. The dataset was obtained from two repositories each containing 1700 and 1500 images respectively. In the initial phase, the images were dimensionally reduced to a desired number of pixels and were fed to the CNN architecture. The output from this layer was further given to Inception V3, wherein all the algorithms were mathematically computed and a linear average value was produced. The results obtained from the model were capable of classifying the disease as moderate DR, mild DR or DME.

M. Abramoff And his colleagues in [14] augmented CNN based images with IDX-DR to diagnose DR based lesions and normal anatomy in the retina. The model was implemented on 1748 images and was further integrated using random forest as a classifier that could detect the five stages of DR. The proposed model accomplished an accuracy of 98 percent. The implementation however had certain limitations and were further eliminated using the hidden layers of CNN.

In a similar work by authors in [15], they proposed to detect and classify all the five stages of a conventional DR based on CNN architecture. In the initial stage, the obtained images from the repository were pre-processed and resized to a dimension of 512 x 512. The overall structure of the model consisted of 21 hidden layers with Softmax being implemented as the activation function in the last year. The model was experimented on 80,000 images and achieved an accuracy of 95 percent. However, the proposed model underwent a major drawback, of not being able to detect multiple lesions in the retinal image.

T. Li And his colleagues in [16] highlighted the implementation of GoogleNet, ResNet, DensNet and VGG-16 and finally conducted a comparative analysis to determine which variant would generate optimized results. As all transfer learning networks are heavily pre-trained, the dataset required for implementation is drastically reduced. In the initial stage all the fundus images were pre-processed and resized to 224 x 224 pixels and fed to the subsequent layers of transfer learning variants. In comparison it was observed that the Inception model generated the best accuracy of 82 percent.

C. Lesion based Classification

This section focuses on research that was done to detect and characterize various types of DR lesions.

Authors in [17], for example, used DL methods combined with domain expertise for feature learning to detect only red lesions of a DR image. The Random Forest method was then used to improve the classification accuracy. The images were obtained from the MESSIDOR repository and a Gaussian filter was processed over the input layers of the model. The images were further resized to 32 x 32 pixels and a total of 8 hidden layers were involved in the CNN architecture. The model achieved an overall accuracy of 78.38 percent. In a similar work by P. Chudzik and his colleagues [18] proposed the implementation of CNN architecture in three datasets containing 100 images from each. The datasets were transformed and resized to the desired pixel number and later underwent the procedure of feature extraction followed by batch normalization. The extracted features were then matched against the input images and morphological function was deployed using three max-pooling layers.

Authors in [19] detected red DR lesions by integrating the model with transfer learning based LeNet architecture. The RGB image obtained from the dataset underwent the process of noise removal through the implementation of a Gaussian filter. Morphological methods were employed and random forest was used as the classifier. The method however had certain drawbacks and were further eliminated using a data augmentation layer that added more images to the dataset. H. Wang and his colleagues in [20] proposed to detect multiple instances of DR lesions in retinal images. The author conducted his experiments on the HEI-MED and E- Optha dataset and integrated the concepts of CNN and random forest. Crop rotation, changing the camera aperture, and using morphological constructions were used to process the dataset. The implementation of the model however included 8 hidden layers and generated an overall accuracy of 92 percent.

D. Retina Dataset

There are numerous publicly available sets of data for detecting DR and vessels in the retina. These datasets are frequently used to train, validate, and test systems, as well as test the efficiency of one system to that of others.

Retinal imaging includes fundus colour images and optical coherence tomography (OCT). A wide range of publicly available fundus image datasets are prevalent. The following are the Fundus image datasets:

- Kaggle: It contains 88,702 slightly elevated images gathered from various cameras with resolutions varying from 433 x 289 pixels to 5184 x 3456 pixels. However, many of the images on Kaggle are of poor quality and have inaccurate classifications
- Messidor: This dataset contains 1200 fundus colour images obtained at a 45-degree FOV and analysed to four different DR phases
- Indian Diabetic Retinopathy Image dataset (IDRiD): This dataset includes 516 fundus images with a 50-degree FOV that have been documented to five DR stages
- APTOS: Blood vessel segmentation is performed using this publicly available dataset. It includes 20 images taken with a 35-degree field of view. The images are 700 x 605 pixels in size. There are ten normal images among them

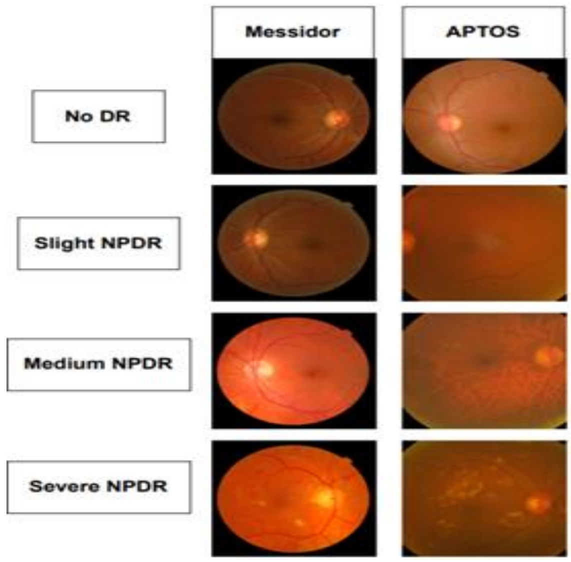

Figure 2: Dataset image samples of Messidor (left) and APTOS (right).

3. Overview Of Methodology

A. Image Processing

The dataset collected in the initial phase includes images from multiple devices that are responsible to capture the DR Dataset with different cameras and their respective adjustments. Apart from this, the image also contains noise in the form of varying pixel intensities and hence needs to be removed to prevent inconsistency. Therefore, to counteract this issue, a step of pre-processing the data is followed and a standard set of images are adopted. First, an interpolation of the fundus images are done over a 4 x 4 pixel value and a sample of the retinal images are formed. This interpolation is highly preferred as it sustains quality and further locks the aspect ratio. An overview of the Pre-processing methods that the researchers covered is presented in this section. When compared to the other channels, the RGB colour space's green channel offers more distinction.

One of the most widely implemented Pre-processing techniques is contrast enhancement. The difference on a green channel image is further enhanced by the use of contrast enhancement. This method is applied to the green channel of the image to increase contrast. For instance, Li and his colleagues in [21] have improved the difference on the derived green channel by utilising the Contrast Limited Adaptive Histogram Equalization (CLAHE) method. This elevates the exposure of exudates of a green Chromecast. Typically, lighting correction is used to increase the picture's luminance and sharpness after contrast enhancement. The image is then smoothed using a noise removal filter, such as Gaussian Filtering. Another well-liked technique for image Pre-processing is image resizing. According to the efficient framework, the image is reduced in size to a low resolution image. Li and his colleagues images were rescaled from different sizes to the same 512 512 pixel resolution. For all the fully convolutional CNN models that used 224 x 224 shape image resolutions, Li and his colleagues [21] also resized their image to desired pixels. An image's resolution is reduced to match the resolution needed by the active network. In order to avoid misclassifying blood vessels and optical discs as DED lesions, researchers frequently have to remove and mask them. Researchers typically prefer to stratify the irrelevant black border to emphasize on the region of interest because many DED datasets incorporate images with a black border (ROI). For instance, authors in [21] used the adaptive threshold process to reduce the black border from fundus images to sharpen the perspective on the Region of Interest (ROI). The retinal images typically had a dark background and were yellowish in colour. Because the fundus details did not blend into the background, they can be removed to lower noise. Darkness expanded into the image's details when the fundus images' black backgrounds were equalised. With regard to this, many Pre-processing techniques were used to remove the black background by adjusting the pixel endpoint to zero and non-zero for all bright areas. When there is a disparity in the image, image augmentation is used. To generate cases of the chosen images for a class where the number of pictures is relatively low than the other significant proportion of healthy retinal fundus images in contrast with DED retina images, images were rotated, resized, and cropped. A typical way for improving results and avoiding overfitting is augmentation.

B. Disease Classification Techniques

This section of the paper highlights the classification techniques used to diagnose DR.

The definition of DL is primarily observed to be as the additional framework of ML with a multilayer network for feature extraction. The word "deep" in DL architecture refers to the thickness of the layers. Following is the classification procedure of the concept: The manually labelled dataset is divided into testing and training samples for the DL structure, the dataset is normalised for improving quality using image pre - processing techniques, and the Pre-processed images are then fed into the DL design for feature extraction and eventual classification. In a DL architecture, each layer treats the output of the layer below as its input, operates it, and then passes it on to the layer above. To enhance classification performance, many authors tweak the model parameters of existing DL algorithms, like VGG16 or CNN.

- Deep Learning: In recent times, the application of DL has been widely accepted and is fundamentally used for medical image analysis that includes classification, detection, collection and segmentation of images. Hence, its implementation in DR detection and classification has proved to generate optimized and better results. The basic advantage of a DL is its ability to learn the features of input data and integrate with respective algorithms in combination with heterogeneous sources [22]. This process of feature extraction is performed on multiple levels of abstraction. Therefore, the shift of focus from ML to DL dramatically enhanced state-of-art pattern recognition, image processing, computer vision, object detection and many other relative domains. By using the backpropagation algorithm to determine how a model should upgrade its internal parameters to quantify each layer's interpretation from the recognition of the previous layer, DL reveals complex pattern structure in large datasets. Its performance is also likely to increase with an increased number of training data. The step involved in a conventional DL is as follows:

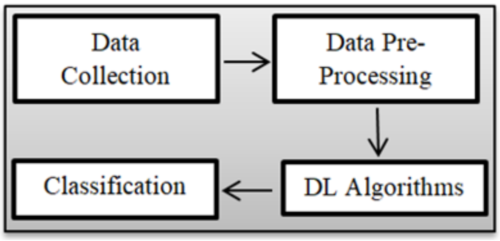

Figure 3: Classification Steps in DL.

- CNN: The working concept of a typical CNN has always been attached to the fundamentals of Supervised Learning; wherein it receives input images, extracts features and finally classifies them on a certain criteria.

These networks are a subcategory of neural networks and hence possess all the features that a neural network exhibits. Its execution takes place in two blocks; wherein the first block is responsible for feature extraction and the second block for classification purposes using ML algorithms. In order to execute these blocks, CNN uses two operations, pooling and convolution over a series of layers [23].

The initial two layers of the network architecture are used to execute the first block, or feature extraction, while the fully connected layer produces the final output by mapping the features that have been extracted from the earlier layers. This ultimate result typically constitutes the second block of execution, or classification. Since all of the mathematical operations in the network take place in the first layer, which happens to be the convolutional layer, this layer is crucial to the overall implementation of the work. Additionally, a grid pattern is used to carry out the entire CNN process. The pixels of images are stored in this grid pattern's grid parameters, which are two-dimensional arrays referred to as kernels. These kernels serve as the model's actual feature extractors, which is what gives CNNs their high level of image processing efficiency. The output from one layer is fed as the yield input to the next layer, with the result that all the layers in this network have a tendency to gradually increase their level of complexity. Training is the term for the parameter optimization process carried out in kernels to reduce the discrepancy between output values and input labels. Back-propagation optimization algorithms are used to carry out this process.

The implementation framework of CNN is depicted in Figure below:

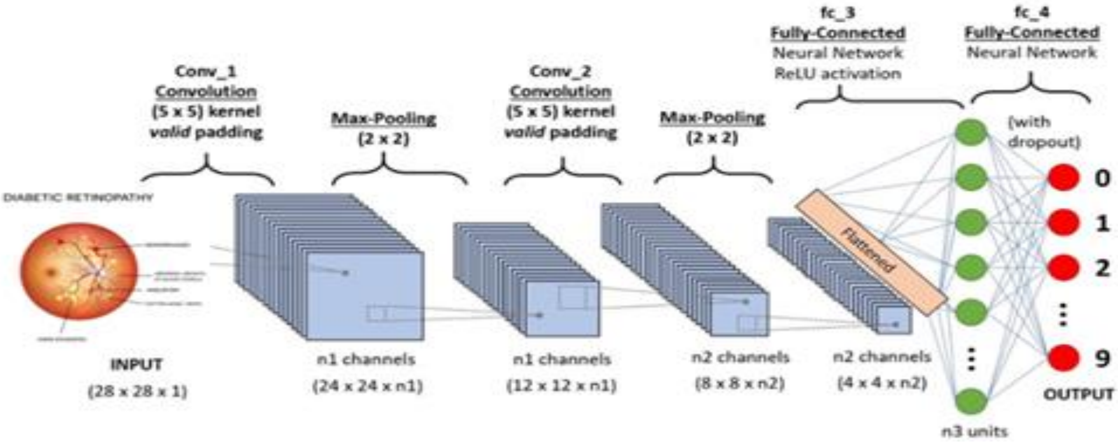

Figure 4: Architecture of CNN.

- Transfer Learning: It is a concept where the information extracted and implemented on a given dataset can be applied to a new dataset with a much smaller population to train provided; both the datasets work on a similar task of CNN architecture.

In a conventional CNN, this process is executed by training initial parameters on large datasets. After successful training a CNN is capable of extracting significant features. Based on the potential of a CNN to extract features, a specific model is selected for transfer learning. This approach is called feature extraction. The primary objective of this strategy is to retain both: architectural framework of a CNN model and the neuron weights. This concept is generally used to compensate for the computational cost of developing a neural network from scratch. The second strategy involves, adopting one of the many transfer learning based variant models such as Alexnet, Densenet, Mobilenet, Inception and VGG-16, wherein certain parametric adjustments are made to the model, to achieve optimal results [24].

The idea behind TL is based on adapting and reusing the features that DL models have learned on the primary task. While trying to train neural network architecture, the goal is to eliminate computational complexity. Furthermore, it has been discovered that using TL is advantageous when there is not enough data to completely train a neural network. Instead of using random generation, TL initialises the parameters based on prior learning. Intuitively, the top layers are more focused on the task at hand, such as blood vessels and exudates, while the initial layers learn to extract basic features like edges, textures, etc.

4. Analysis Of Evaluation Metrics

To evaluate the effectiveness of DL methods for classification, a variety of performance metrics are used. Accuracy, sensitivity, specificity, and area under the ROC curve are frequently used metrics in DL (AUC). The percentage of anomalous images labelled as abnormal is known as sensitivity, and the percentage of natural images categorised as normal is known as specificity [25]. Sensitivity and specificity are plotted to produce an AUC graph. Figure below depicts the proportion of performance metrics used in multiple researches.

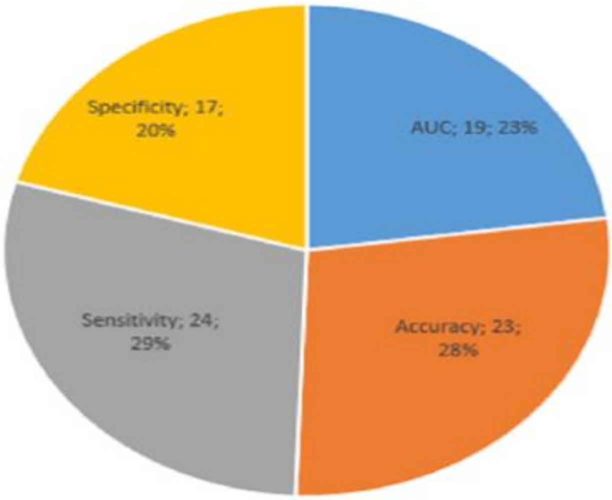

Figure 5: Percentage of performance metrics.

The proportion of correctly classified images is known as accuracy. The formulae for each measurement are mentioned below.

- Specificity: TN / (TN FP)
- Sensitivity: TP/ (TP FN)
- Accuracy = Specificity / Sensitivity

5. Conclusion And Future Direction

It is clear and obvious that ML techniques are significantly less scalable with regard to high-dimensional data and require more time for model analysis and training than DL techniques. When there are more features and data, ML models come to suboptimal conclusions while DL models work to produce the best results possible. This paper identifies and reviews a significant number of DL models and their frameworks, for understanding of the working principle, their evolution and integration on using hybrid techniques, and how such models can be transitioned on scarcity of data and resources, to produce effective models and outcomes. This is based on the enormous applications of DL in recently proposed models.

Automated diagnostic methods drastically cut down on the time needed to make diagnoses, saving ophthalmologists' time and money while also enabling patients to receive treatment sooner. Hence, automated DR detection systems are crucial for spotting DR at an early stage. The phases of DR are determined by the kind of lesions that develop on the retina. A thorough, systematic review of relevant publications was carried out to reach this objective. We have characterised the publicly accessible common fundus DR datasets and provided a brief introduction to deep learning methods. Due to its effectiveness, CNN has been used by the majority of scholars for the classification and detection of DR images. Multiple papers were reviewed for the current study. Deep learning techniques were used in all of the aforementioned studies in the current work to automate the diabetic retinopathy screening system. Recent growth in the number of diabetic patients has made the need for trustworthy diabetic retinopathy screening systems critical. The issue of choosing trustworthy features for ML is solved by using DL in DR detection and classification.

Future research works that need to be highlighted are mentioned below:

- Training with little data for learning purposes, deep learning software frequently uses a lot of retinal fundus images. If the training dataset is limited, it may not produce satisfactory performance in terms of accuracy. There are two potential answers. Utilizing various enhancement techniques, such as rotating, shifting and cropping and using weak learning algorithms to get training data, second. Further research reveals the use of Generative Adversarial Network (GAN) for training generation, enabling the DL architecture to be trained with greater robustness and distinguishing features.
- Combining telemedicine services, cloud computing, and DL, especially in the area of healthcare, remote regions struggle with a lack of human capital. In such cases, therefore, telemedicine can play a significant role in countering this limitation. In the future, RD may be assessed from eye fundus images using neural networks, cloud computing, and remote monitoring.

References

[1] Mellitus D., "Diagnosis and classification of diabetes mellitus. Diabetes Care", 2005, 28, S5–S10

[2] Fong, D.S.; Aiello, L.; Gardner, T.W.; King, G.L.; Blankenship, G.; Cavallerano, J.D.; Ferris, F.L.; Klein, R. "Retinopathy in diabetes. Diabetes Care", 2004, 27, s84–s87

[3] Wilkinson, C.; Ferris, F.L.; Klein, R.E.; Lee, P.P.; Agardh, C.D.; Davis, M.; Dills, D.; Kampik, A.; Pararajasegaram, R.; Verdaguer, J.T. "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales", Ophthalmology 2003, 110, 1677–1682

[4] Kempen, J.H.; O'Colmain, B.J.; Leske, M.C.; Haffner, S.M.; Klein, R.; Moss, S.E.; Taylor, H.R.; Hamman, R.F. "The prevalence of diabetic retinopathy among adults in the United States.", Arch. Ophthalmol. 2004, 122, 552–563

[5] Grzybowski, A.; Brona, P.; Lim, G.; Ruamviboonsuk, P.; Tan, G.S.W.; Abramoff, M.; Ting, D.S.W. "Artificial intelligence for diabetic retinopathy screening: A review.", Eye 2020, 34, 451–460

[6] Ting, D.S.W.; Cheung, C.M.G.; Wong, T.Y. "Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review.", Clin. Exp. Ophthalmol. 2016, 44, 260–277

[7] K. Xu, D. Feng, H. Mi "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image Molecules", 22 (12) (2017), p. 2054

[8] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, M. Lamard, "Deep image mining for diabetic retinopathy screening Med Image Anal", 39 (2017), pp. 178-193

[9] M.T. Esfahani, M. Ghaderi, R. Kafiyeh, "Classification of diabetic and normal fundus images using new deep learning method Leonardo Electron J Pract Technol", 17 (32) (2018), pp. 233-248

[10] R. Pires, S. Avila, J. Wainer, E. Valle, M.D. Abramoff, A. Rocha, "A data-driven approach to referable diabetic retinopathy detection Artif Intell Med", 96 (2019), pp. 93-106

[11] H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma, W. Qian "An interpretable ensemble deep learning model for diabetic retinopathy disease classification", 2019, 41st Annual International conference of

the IEEE engineering in medicine and biology society (EMBC) (2019), pp. 2045-2048

[12] Y.P. Liu, Z. Li, C. Xu, J. Li, R. Liang, "Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network", Artif Intell Med, 99 (2019), p. 101694

[13] V. Gulshan, et al. "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs", Am Med Assoc, 316 (2016), pp. 2402-2410

[14] M.D. Abràmoff, et al. "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning Investig", Ophthalmol Vis Sci, 57 (13) (2016), pp. 5200-5206

[15] H. Pratt, F. Coenen, D.M. Broadbent, S.P. Harding, Y. Zheng, "Convolutional neural networks for diabetic retinopathy", Procedia Comput Sci, 90 (2016), pp. 200-205

[16] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, H. Kang "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening", Inf Sci, 501 (2019), pp. 511-522

[17] J.I. Orlando, E. Prokofyeva, M. del Fresno, M.B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images Comput Methods", Progr Biomed, 153 (2018), pp. 115-127

[18] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, A., "Hunter Microaneurysm detection using fully convolutional neural networks Comput Methods", Progr Biomed, 158 (2018), pp. 185-192

[19] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, A., "Hunter Microaneurysm detection using fully convolutional neural networks Comput Methods", Progr Biomed, 158 (2018), pp. 185-192

[20] H. Wang, et al. "Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening Comput Methods", Progr Biomed, 191 (2020), p. 105398

[21] 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

[22] A. V Vasilakos, Y. Tang, Y. Yao, "Neural networks for computer-aided diagnosis in medicine : a review Neurocomputing", 216 (2016), pp. 700-708

[23] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, H. Kang "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening", Inf Sci, 501 (2019), pp. 511-522

[24] Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. "A comprehensive survey on transfer learning", Proc IEEE. 2020;109:43–76

[25] 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