

Sheetal J. Nagar¹
Nikhil Gondaliya²

A Comprehensive Review on Diabetic Retinopathy Detection Techniques using Neural Network Architectures



Abstract: - Diabetic retinopathy (DR) is a significant complication arising from diabetes, affecting the eyes and potentially causing vision loss if not identified and addressed promptly. Over the years, there has been a significant advancement in the field of DR detection, primarily driven by advancements in imaging techniques and machine learning algorithms. This review paper presents a comprehensive overview of different techniques and advancements in the detection of diabetic retinopathy using deep learning and several neural network architectures. The comparative study of the existing datasets for the DR detection with the benefits, challenges and possible solutions for each dataset is also provided. The paper discusses the methods, preprocessing, implementation platforms and results of various implementation of CNN architectures like Deep CNN, CNN with Transfer Learning, Capsule Networks and DNN. The objective of this paper is to furnish researchers and clinicians with a thorough understanding of the present status of diabetic retinopathy detection, highlight the strengths and limitations of existing approaches, and identify future research directions in this vital area of healthcare.

Keywords: Image Classification, Diabetic Retinopathy, CNN, Transfer Learning, Capsule Network.

1 INTRODUCTION

DR is a progressive condition characterized by microvascular changes leading to retinal ischemia, increased leakage in the retinal blood vessels, neovascularization, and macular edema [1]. Without intervention, individuals with DR may experience significant visual impairment [4]. In developed nations, DR stands as a leading cause of blindness among individuals of working age, exerting a substantial economic impact on society, particularly on healthcare systems[1,2].

In 2015, diabetic retinopathy contributed to 1.07% of cases of blindness and 1.25% of instances of moderate to severe visual loss. Despite worldwide efforts over the past two decades to reduce visual damage, there has been a significant 25% rise in the prevalence of diabetic retinopathy. This increase is linked to a simultaneous rise in the absolute number of individuals affected by diabetic retinopathy (DR), paralleling the rise in diabetes mellitus (DM) cases. [3].

Proper precautions play a crucial role in slowing down the progression of diabetic retinopathy, thereby preserving vision and maintaining overall eye health. Regular monitoring and early detection can significantly impact the outcome and quality of life for those at risk.

Automated systems for diabetic retinopathy (DR) detection utilize artificial intelligence (AI) and machine learning algorithms. The primary objective of these systems is to aid healthcare professionals in promptly and accurately identifying diabetic retinopathy in patients.

2 MOTIVATION

The progression of diabetic retinopathy unfolds from a mild, non-proliferative stage to moderate and severe non-proliferative diabetic retinopathy (NPDR), eventually advancing to proliferative diabetic retinopathy (PDR). The progression of vision-endangering diabetic retinopathy emerges from diverse mechanisms, encompassing macular edema, macular ischemia, bleeding from newly formed vessels, and the contraction of accompanying fibrous tissue. This process culminates in tractional retinal detachment.

Within the medical domain, image processing assumes a pivotal role in disease diagnosis. The extraction of relevant information from intricate arrays of images presents a substantial challenge. Recognizing that effective management can prevent over 90% of cases of visual impairment, it becomes crucial to systematically categorize,

¹ Gujarat Technological University, Ahmedabad, Gujarat, India

² G H Patel College of Engineering & Technology, Information Technology Engineering, Vallabh Vidyanagar, Gujarat, India
nagar.sheetal@gmail.com and nikhilgondaliya@gcet.ac.in

classify, and stage the severity of DR to establish appropriate therapeutic interventions. [3]. The growing prevalence of diabetes and its consequences Emphasis on Necessity for Precautions. Efficient and automated systems can assist clinicians in the detection and classification of diabetic retinopathy.

Traditionally healthcare professionals visually inspect retinal images and grade them based on the presence and severity of diabetic retinopathy lesions [3]. This method is subjective and demands a considerable amount of time. Neural networks can be trained to exhibit high sensitivity, enabling them to detect subtle and early signs of diabetic retinopathy that may be challenging for human observers to identify.

3 DIABETIC RETINOPATHY DETECTION DATASETS

3.1 Diabetic Retinopathy Datasets Overview

To facilitate the detection of diabetic retinopathy and contribute to the progress of computer-aided diagnosis systems, various datasets related to diabetic retinopathy have been made accessible. These datasets primarily comprise retinal images along with ground truth annotations. This enables researchers to train and evaluate their neural network algorithms for the early identification and classification of diabetic retinopathy.

Diabetic retinopathy dataset available on Zenodo is compact, comprising 757 color fundus images categorized into seven classes as detailed below: 187 images depicting No Diabetic Retinopathy (DR) signs, 4 images representing Mild or early Non-Proliferative Diabetic Retinopathy (NPDR), 80 images depicting Moderate NPDR, 176 images illustrating Severe NPDR, 108 images representing Very Severe NPDR, 88 images showcasing Proliferative Diabetic Retinopathy (PDR), and 114 images displaying Advanced PDR [4].

Diabetic Retinopathy Detection Competition dataset [5] provided by Kaggle, is publicly accessible. This dataset comprises both resized and cropped images and represents a modified version of the Diabetic Retinopathy Kaggle competition dataset. This dataset is currently widely used by researchers. Table 1 displays sample images from five distinct categories within this dataset.

Table 1. Comparison of 5 levels of images from images from eyePACS Dataset

DR 0	DR 1	DR 2	DR 3	DR 4

APTOS Eye Pre-processing in Diabetic Retinopathy dataset is available on Kaggle. This dataset was created for APTOS 2019 Blindness Detection competition [6]. It is one of the largest datasets among all the datasets used by several researchers, with the images belonging to 5 categories.

Eye-diseases-classification dataset is also accessible from Kaggle. This is an almost balanced dataset consists of 1074 Normal images, 1098 Diabetic Retinopathy images, 1038 Cataract and 1007 Glaucoma images of eyes for Diabetic Retinopathy [7].

3.2 Diabetic Retinopathy Datasets Comparisons

The datasets discussed in 3.1 differs in terms of entire size of the dataset, no. of images, no. of classes, range of image size and image dimensions. Thus, advantages, difficulties and probable resolutions linked with each dataset also differs. The benefits, challenges and possible solutions associated with each dataset is represented in Table 2.

Table 2. Benefits, Challenges and Possible Solutions for DR Datasets

Dataset	Benefits	Challenges and Solution
Zenodo [4]	<ul style="list-style-type: none"> - Smallest dataset with few images belonging to each class. - Requires less time for model training thus allows faster experimentation, iteration and model development. - Beneficial when time or computational resources are limited. 	<ul style="list-style-type: none"> - Due to the limited number of samples, the model's performance on a small dataset may have higher variance. - Model performance can fluctuate significantly depending on the specific samples included in the training set; which makes it challenging to determine the true performance of the model. - Techniques such as data augmentation, transfer learning, and regularization can be applied to improve model performance. - Leveraging pre-trained models can help to compensate for the limited sample size and enhance the model's ability to generalize.
eyePACS [5]	<ul style="list-style-type: none"> - The largest datasets in terms of no. of images - Provides sufficient training data for neural networks, resulting in better generalization, improved accuracy, reduced overfitting, enhanced robustness, better representation learning and support for complex models. - Supports models to be adaptable and capable of handling real-world variations 	<ul style="list-style-type: none"> - It is an imbalanced dataset, with maximum no. of normal retinal images and relatively minimum Proliferative DR images. - Imbalanced datasets can lead to biased model performance, especially when evaluating accuracy. - To mitigate the disadvantages of imbalanced datasets, various techniques can be employed, such as resampling methods like under sampling or oversampling, generating synthetic samples, using different cost functions, or applying ensemble methods.
APTOS 2019 Blindness Detection [6]	<ul style="list-style-type: none"> - Provides fundus images with maximum resolution, thus it can have the advantage of improved feature extraction and reduced information loss. 	<ul style="list-style-type: none"> - Larger image sizes may lead to longer training times. Hence, a trade-off between resolution and computational efficiency needs to be considered based on the specific application and available resources.
Eye-diseases-classification [7]	<ul style="list-style-type: none"> - The only balanced dataset, discussed in this paper. - By balancing the dataset, the neural network is less likely to exhibit any biases and can provide fair and unbiased predictions across all the classes. - It enables accurate evaluation of the model's performance using metrics such as accuracy, precision, recall, F1-score or area under the receiver operating 	<ul style="list-style-type: none"> - It is comparatively a small dataset. - Transfer learning can be employed as a solution, leveraging pre-trained models trained on large datasets. By fine-tuning these pre-trained models on the small dataset, the network can benefit from the learned features while adapting them to the specific task. This approach enables effective training even with limited data.

characteristic curve (AUC-ROC).

4 MODEL SPECIFIC SURVEY FOR DR DETECTION

Numerous researchers in the diabetic retinopathy detection domain have designed models employing neural network architectures. A Convolutional Neural Network (CNN) is a fundamental architecture used by several researchers.

The CNN architecture typically consists of multiple convolutional layers followed by pooling layers and fully connected layers [8]. The convolutional layers conduct feature extraction through the convolution of filters across input images to detect relevant patterns and structures. CNNs automatically learn features from the retinal images during the training process. The initial layers of the network capture basic features such as edges, lines, and textures, while deeper layers focus on higher-level features and semantic representations that are pertinent to diabetic retinopathy.

Once the CNN has been trained, it can be used to identify the severity level of unseen retinal images. This is typically done using softmax activation at the output layer, which assigns probabilities to each class. The CNN model, after training, undergoes evaluation on an independent test dataset to measure its performance. Metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are commonly used to evaluate the model's ability to detect and classify DR accurately. If the CNN model performs well during evaluation, it can be deployed as a tool to assist ophthalmologists in diagnosing and monitoring DR. It can analyze retinal images and provide an objective assessment of DR severity, aiding in early intervention and treatment decisions.

Implementation is done in TensorFlow and Keras by several researchers. TensorFlow is an open-source deep learning software library for defining, training and deploying machine learning models [9] Keras is a widely used TensorFlow extension libraries [10], written in Python. A method specific literature survey for DR detection is shown in the Table 5 including preprocessing techniques, datasets, implementation platforms and the results.

Table 3. Method Specific Literature Survey

Model	Authors	Methods, Dataset, Implementation and Results
		Pre-processing:
		- Resized images to dimensions of 256×256
		Method:
		- Extracts (Regions Scoring Map) RSM from an image by a Fully Convolutional Network
		- RSM is designed to identify various lesions, encompassing hard exudates, soft exudates, haemorrhages, and newly formed abnormal blood vessels, while highlighting these specific lesion areas
Deep CNN	Junjun et al. [11]	- The Fully Convolutional Network is structured on the ResNet18 model
		Dataset:
		- EyePACS dataset downloaded from the Kaggle.
		Implementation:
		- Trained TensorFlow models on a single NVIDIA GeForce GTX 1080Ti GPU with 11 GB of VRAM.
		Results:
		- Accuracy : 78.4 %
		- Precision : 84.09 %
Deep	Darshit	Pre-Processing:

CNN	et al. [12]	<ul style="list-style-type: none"> - ImageMagick data pre-processing techniques using OpenCV Python library <p>Method:</p> <ul style="list-style-type: none"> - Convolutional Layer - Max-pooling with a filter size of 2x2 - Dropout Layer - Hidden Layers Leaky Rectifier and Feature Pooling Layers <p>Dataset:</p> <ul style="list-style-type: none"> - EyePACS dataset downloaded from the Kaggle <p>Implementation:</p> <ul style="list-style-type: none"> - Training process NVIDIA GeForce GTX 780 series GPU with 3GB of RAM - Deployed on a machine equipped with an Intel Core i3 processor and 8GB of RAM. - Theano library is used to interface with the GPU. - Python libraries Lasagne and Nolearn for training of the deep network. <p>Results:</p> <ul style="list-style-type: none"> - Quadratic weighted kappa metric score: Achieved a quadratic weighted kappa metric score of 0.386 and the ensemble of three similar models further improved the score to 0.3996
		<hr/> <p>Pre-processing</p> <ul style="list-style-type: none"> - Implemented procedures for Black Box removal and image operations such as resizing, augmentation, green channel extraction, and image enhancement <p>Method:</p> <ul style="list-style-type: none"> - U-Net segmentation technique combined with region merging and Convolutional Neural Network <p>Sairaj Dataset:</p>
CNN	et al. [13]	<ul style="list-style-type: none"> - EyePACS Diabetic Retinopathy database provided by Kaggle - Employed images acquired through a diabetic retinopathy screening program in The Netherlands. - Utilized the DRIVE database containing images of individuals aged between 25 and 90 years. - Incorporated 700 images for each class in the dataset to overcome the problems related to over fitting. - The images are segmented manually. <p>Results:</p> <ul style="list-style-type: none"> - Accuracy : 93.33%.
CNN	Ratul et al. [14]	<hr/> <p>Method:</p> <ul style="list-style-type: none"> - Following layers were used for CNN model architecture: <p>Convolutional Layer, Activation Layer, Max pooling Layer, Dropout Layer, Fully Connected Layer and Classification Layer</p> <p>Dataset:</p> <ul style="list-style-type: none"> - EyePACS dataset downloaded from the Kaggle <p>Implementation:</p> <hr/>

		<ul style="list-style-type: none"> - Using Theano, a Python open-source numerical computation library - Employed ImageMagick's convert tool to remove extraneous spaces from the sides of the images. - Conducted training on a high-performance GPU <p>Results:</p> <ul style="list-style-type: none"> - Accuracy: 95% for Binary classification - Accuracy: 85% accuracy for multiclass classification - kappa score: 0.74
		<p>Method:</p> <ul style="list-style-type: none"> - CNN architecture and data augmentation - Maxpooling - Implemented dropout on dense layers to decrease the effect of overfitting - softmax activation function for classification - Applied L2 regularization on weight and biases - Initialized network with Gaussian initialization to accelerate initial training.
CNN	Harry et al. [15]	<ul style="list-style-type: none"> - Enhanced the model by widely accepted categorical cross-entropy loss function for optimization <p>Dataset:</p> <ul style="list-style-type: none"> - From the Kaggle <p>Implementation:</p> <ul style="list-style-type: none"> - NVIDIA K40c, a High-end GPU for the training of the network - NVIDIA CUDA Deep Neural Network library for GPU learning. <p>Results:</p> <ul style="list-style-type: none"> - Accuracy 75% - Specificity 95% - Sensitivity 30%
		<p>Method:</p> <ul style="list-style-type: none"> - CNN architecture for image classification - MobileNet Architecture using a transfer learning approach coupled with a self-attention mechanism - Utilized Gradient-weighted Class Activation Mapping (Grad-CAM) to accentuate crucial regions in the OCT for prediction
CNN with Transfer Learning	Othmane et al. [16]	<p>Dataset:</p> <ul style="list-style-type: none"> - Messidor dataset <p>Implementation:</p> <ul style="list-style-type: none"> - Keras 2.3.1 with TensorFlow 2.04 <p>Results:</p> <ul style="list-style-type: none"> - Accuracy 98% - Specificity 98% - Sensitivity 98%
DNN	Zhentao et al. [17]	<p>Pre-processing:</p> <ul style="list-style-type: none"> - Size Normalization - Shape Normalization

		<ul style="list-style-type: none"> - Color Normalization - Augmentation <p>Method:</p> <ul style="list-style-type: none"> - Inception-V3 network as a core model. - Softmax output layer for classification in four categories <p>Dataset:</p> <ul style="list-style-type: none"> - The development was undertaken by the authors from three clinical departments—the Ophthalmology Department, Health Management Center, and Endocrinology & Metabolism Department—in Sichuan Provincial People's Hospital, a prominent institution in Ophthalmology in Sichuan Province in China - Introduced initial diagnostic services for diabetic retinopathy to multiple hospitals <p>Results:</p> <ul style="list-style-type: none"> - Accuracy: 88.72% for a four-degree classification - During the clinical evaluation, the system demonstrated a concordance rate of 91.8% when compared with analysed by ophthalmologists.
		<p>Method:</p> <ul style="list-style-type: none"> - Convolutional layer for extracting features from the input image - Primary capsule layer responsible for processing the extracted features - dynamic routing is utilized to manage the features for classification using class capsule layer - Softmax layer to transform the results in probabilities associated with each class of diabetic retinopathy <p>Dataset:</p> <ul style="list-style-type: none"> - Messidor dataset <p>Implementation</p> <ul style="list-style-type: none"> - Implementation and training using NVIDIA CUDA architecture. - The system includes uses nVIDIA GPU with DDR6 Memory of 8GB, AMD Ryzen processor 2700x processor with eight cores and 32 GB RAM. - Keras API in TensorFlow <p>Results:</p> <ul style="list-style-type: none"> - Accuracy for healthy retina 97.98%, stage 1 fundus images 97.65%, stage 2 fundus images 97.65% and stage 3 fundus images 98.64%
Capsule Networks	G. Kalyani et al. [18]	
Capsule Networks	Gaurav et al. [19]	<p>Preprocessing:</p> <ul style="list-style-type: none"> - Most of the dark space is removed by identifying the center and radius of the eye retina image. - The images found entirely black or very close to being entirely black, were manually excluded from the dataset. <p>Method:</p> <ul style="list-style-type: none"> - DRDNet with a layer of primary capsules network. - The hidden layer comprises two primary components: the convolution layer and the capsule layer. - ReLU convolution layer - Capsule layer - Digicaps layer

- Sigmoid activation function

Dataset:

- From the Kaggle competition and APTOS 2019 Blindness Detection competitions is used.

Implementation:

- The model is trained using Python on Nvidia DGX II.
 - The implementation of the proposed framework utilized specific Python modules, including Keras version 2.2.4, TensorFlow-gpu version 1.14, Seaborn, and Scikit-learn

Results:

- Accuracy: 80.59%

5 CONCLUSION

Neural network based identification of diabetic retinopathy holds immense potential in improving the efficiency and accessibility of DR screening programs. Overall, the combination of advanced imaging techniques, machine learning algorithms and CNN architectures have shown promising results in the identification of severity level of Diabetic Retinopathy, contributing to improved patient care and potentially reducing the risk of vision loss associated with this condition. Further research can be carried out to overcome the existing challenges and translate these technological advancements into routine clinical practice, ultimately benefiting patients with diabetes and reducing the burden of vision loss.

References

- [1] Lihteh Wu, Priscilla Fernandez-Loaiza, Johanna Sauma, Erick Hernandez-Bogantes, Marissé Masis, Classification of diabetic retinopathy and diabetic macular Edema, World Journal of Diabetes (2013)
- [2] Williams R, Airey M, Baxter H, Forrester J, Kennedy-Martin T, Girach A., Epidemiology of diabetic retinopathy and macular oedema: a systematic review, pp. 963-983. Eye (2004)
- [3] World Health Organization, Strengthening diagnosis and treatment of diabetic retinopathy in the South-East Asia Region. New Delhi. World Health Organization, Regional Office for South-East Asia; (2020), <https://www.who.int/publications/i/item/9789290227946>, last accessed 2022/08/23.
- [4] Dataset from fundus images for the study of diabetic retinopathy, <https://zenodo.org/record/4647952#.YGNjXVUzbIU>, last accessed 2022/08/28.
- [5] Diabetic Retinopathy Detection Competition Dataset Resized/Cropped, <https://www.kaggle.com/tanlikesmath/diabetic-retinopathy-resized>, last accessed 2023/06/15.
- [6] APTOS: Eye Preprocessing in Diabetic Retinopathy, <https://www.kaggle.com/code/ratthachat/aptos-eye-preprocessing-in-diabetic-retinopathy/data>, last accessed 2023/06/15.
- [7] eye-diseases-classification, <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification>, last accessed 2023/06/15.
- [8] NEILI ZAKARIA, FEZARI MOHAMED, REDJATI ABDELGHANI, KENNETH SUNDARAJ, VGG16, ResNet-50, and GoogLeNet Deep Learning Architecture for Breathing Sound Classification: A Comparative Study, IEEE (2021).
- [9] Peter Goldsborough, A Tour of TensorFlow, ResearchGate (2016).
- [10] Tom Hope, Yehezkel S. Resheff, and Itay Lieder, Learning TensorFlow A Guide to Building Deep Learning Systems, O'Reilly (2017).
- [11] Junjun Pan, Zhifan Yong, Dong Sui, Hong Qin, Diabetic Retinopathy Detection Based on Deep Convolutional Neural Networks for Localization of Discriminative Regions, 2018 International Conference on Virtual Reality and Visualization, pp. 46–52. IEEE (2018)
- [12] Darshit Doshi, Aniket Shenoy, Deep Sidhpura, Dr. Prachi Gharpure, Diabetic Retinopathy Detection using Deep Convolutional Neural Networks, International Conference on Computing, Analytics and Security Trends, pp. 261–266. IEEE (2016).

- [13] Sairaj Burewar, Anil Balaji Gonde, Santosh Kumar Vipparthi, Diabetic Retinopathy Detection by Retinal segmentation with Region merging using CNN, 13th International Conference on Industrial and Information Systems, pp. 136–142. IEEE (2018).
- [14] Ratul Ghosh, Kuntal Ghosh, Sanjit Maitra, Automatic Detection and Classification of Diabetic Retinopathy stages using CNN, International Conference on Signal Processing and Integrated Networks, pp. 550–554. IEEE (2017).
- [15] Harry Pratt, Frans Coenen, Deborah M Broadbent, Simon P Harding, Yalin Zheng, Convolutional Neural Networks for Diabetic Retinopathy, International Conference On Medical Imaging Understanding and Analysis, pp. 200–205. ELSEVIER, Loughborough, UK (2016).
- [16] Othmane Daanouni, Bouchaib Cherradi and Amal Tmiri, Self-Attention Mechanism for Diabetic Retinopathy Detection, Emerging Trends in ICT for Sustainable Development, Advances in Science, Technology & Innovation, pp. 95–104. Springer Nature Switzerland AG (2021).
- [17] Zhentao Gao, Jie Li, Jixiang Guo, Yuanyuan Chen, Zhang Yi, And Jie Zhong, Diagnosis of Diabetic Retinopathy Using Deep Neural Networks, VOLUME 7, pp. 3360–3370, IEEE Access (2019).
- [18] G. Kalyani, B. Janakiramaiah, A. Karuna, L. V. Narasimha Prasad, Diabetic retinopathy detection and classification using capsule networks, Complex & Intelligent Systems, Springer (2021)
- [19] Gaurav Kumar, Shraban Kumar Chatterjee, Chiranjoy Chattopadhyay, DRDNET: DIAGNOSIS OF DIABETIC RETINOPATHY USING CAPSULE NETWORK, Sixth International Conference on Multimedia Big Data, pp. 379–385, IEEE (2020).