SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I7P107

Volume 11 Issue 7, 65-76, July 2024

Original Article

HR-NET: Revolutionizing Diabetic Retinopathy Diagnosis with High-Resolution Analysis

A. Jasni[1], I. Sowmy[2]

[1,2]Department of Electronics and Communication Engineering, Noorul Islam Centre for Higher Education, Tamilnadu, India.

[1]Corresponding Author : jasni.A@outlook.com

Received: 06 May 2024 Revised: 06 June 2024 Accepted: 04 July 2024 Published: 27 July 2024

***Abstract** - One of the most prevalent consequences of diabetes and the primary cause of blindness globally is Diabetic Retinopathy, or DR. Early recognition of DR is vital for timely intervention and prevention of vision loss. However, existing methodologies for DR detection often suffer from limitations in accuracy and robustness. This study addresses these challenges by proposing a novel methodology for DR detection using the High-Resolution Network (HR-Net) architecture. The proposed approach begins with the collection of retinal images from a publicly available dataset and applies rigorous preprocessing as well as augmentation methods to boost the dataset's quality and diversity. The preprocessed images are then inputted into the HR-Net model, which leverages its unique architectural features, including high-resolution representation and multi-resolution fusion, to analyze retinal images effectively. Several performance measures are employed to assess the suggested model's effectiveness, and it is contrasted with previous models. Outcomes demonstrate that the proposed framework attains a remarkable accuracy of 95.67%, outperforming state-of-the-art methodologies. The significance of the proposed model lies in its ability to accurately detect DR-related abnormalities in retinal images, thereby facilitating early diagnosis and intervention. The study underscores the potential of HR-Net architecture in enhancing diabetic retinopathy diagnosis and treatment, offering promising prospects for improving clinical outcomes and reducing the burden of vision loss associated with DR.*

***Keywords** - Diabetic retinopathy, HR-Net, Retina, Deep Learning, Hemorrhages.*

1. Introduction

The increased prevalence of diabetes globally can be attributed to higher blood glucose levels. Long-term glucose irregularities cause serious blood vessel damage, which raises the risk of kidney difficulties, DR, gum bleeding, nerve damage, and cardiovascular disease. DR particularly refers to retinal degeneration, which is made worse by elevated blood glucose levels. The World Health Organization (WHO) reports that diabetes is the sixth most prelevant incidence of death, with an estimated 61.3 million cases identified in people between the ages of 20 and 79 [1]. It is anticipated that it will rise to approximately 191 million cases by 2030 [2]. DR causes vision loss that is irreversible once it occurs. Individuals with a history of diabetes have an increased risk of developing the disease; the risk rises with age, irrespective of the type of diabetes (type 1 or type 2) [3].

Blood vessel disruption in the retina caused by high blood sugar deteriorates vision by allowing blood to seep out. But the body is naturally regenerative. When adjacent cells sense blood leakage, they react by activating to eliminate the issue [4]. This results in the atypical formation of fragile new blood vessels that impair vision. For individuals with diabetes, routine eye exams are essential. Fundus photography is a critical diagnostic technique for early disease diagnosis. The magnitude and existence of anomalies such as hemorrhages, microaneurysms, and venous beading usually indicate the intensity of DR. For instance; microaneurysms are little blood clots that are usually circular and have a diameter of 100–120μm. Blood leaks from damaged vessels, cause hemorrhages, whereas neovascularization is the aberrant formation of microscopic blood vessels. Vein dilatation next to blocked arterioles is indicated by venous beading.

If left untreated, DR advances through four clinical phases, starting from mild non-proliferative to moderate and severe non-proliferative stages, as depicted in Figure 1. This progression underscores the significance of early detection and intervention in managing DR effectively. Identifying the stages accurately allows for timely and appropriate treatment, helping to mitigate the risk of vision loss associated with this condition. Thus, understanding the sequential nature of DR stages is critical in guiding clinical decisions and optimizing patient outcomes.

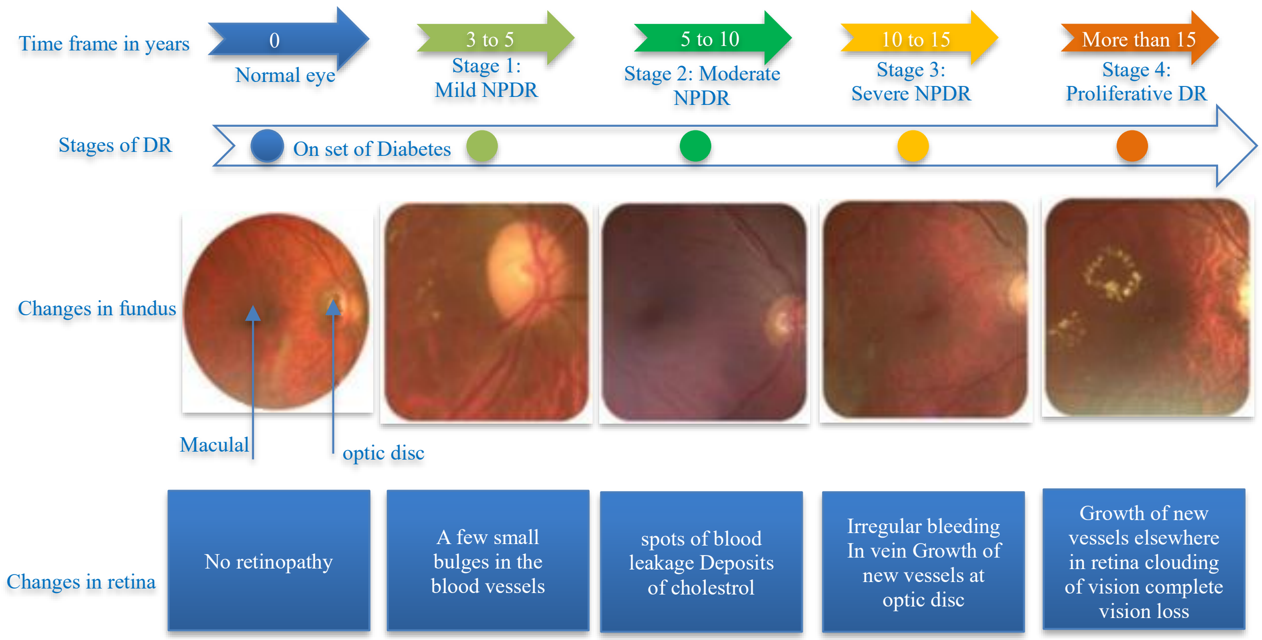
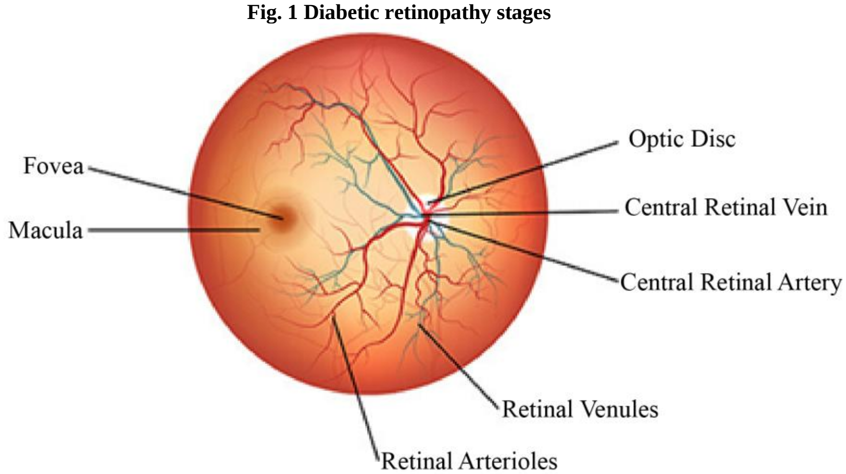
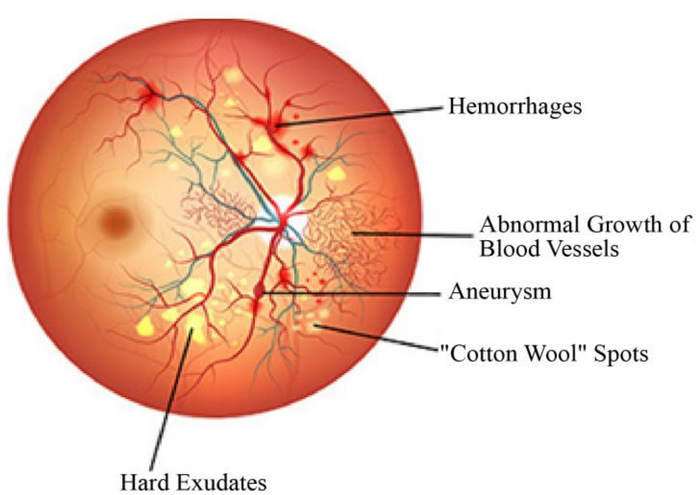


Fig. 1 Diabetic retinopathy stages



(a) Normal



(b) DR

Fig. 2 Schematic illustration of normal and diabetic retinopathy

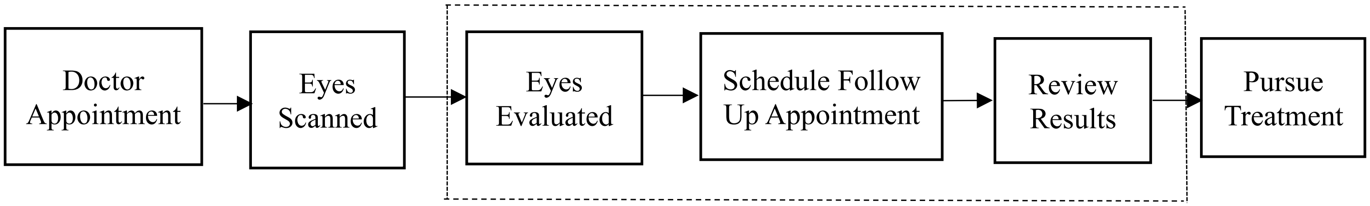


Fig. 3 Process of manual diagnosis

As the condition is asymptomatic or only manifests very modest symptoms, it is challenging to diagnose in the early stages, which keeps a person unaware and eventually impairs vision. Consequently, preventing the complications of this condition depends on early DR detection. Experts and specialists with highly efficient tools and methods are needed for the diagnosis of this sickness in order to contribute to improvements in the prognosis of this disease. Figure 2 depicts the retina in a normal state and the retina having DR, respectively [5].

Therefore, identifying these exudates, scars, and abnormal blood vessels is crucial for diagnosis in DR. As seen in Figure 3, a typical manual diagnosis requires the assistance of qualified professionals and takes seven to fourteen days.

Researchers have extensively explored automating DR diagnosis using computers. Conventional techniques often use a number of feature extraction modules to gather pertinent data from fundus images initially. These manually designed feature-based techniques are time-consuming and frequently produce subpar outcomes. Since classifying the DR requires an accurate automatic detection technique. Machine Learning (ML) approaches were used for feature extraction in most DR studies. Subsequently, the characteristics that were collected are fed into specific classifiers, like the AdaBoost classifier, support vector machines, and random forests. However, when manual feature extraction proved to be challenging, researchers turned to Deep Learning (DL).

The major contributions of the proposed work are as follows:

- To introduce a novel DL framework for early detection and classification of DR from retinal fundus images.
- To enhance detection performance by leveraging the High-Resolution Network (HR-Net) architecture, designed to handle high-resolution visual information.
- To evaluate algorithm performance using standard benchmark metrics, including precision, F1-score, recall, accuracy, as well as AUC showcasing its effectiveness in diabetic retinopathy classification.
- To compare the accuracy of the proposed research with existing models, demonstrating superior proficiency in discerning subtle signs of diabetic retinopathy.

The study is organized into several sections to ensure clarity and coherence in presenting the research findings. In Section 2, a thorough literature review is provided, offering insights into existing studies and methodologies related to the topic under investigation. Section 3 elaborates on the proposed methodology, including experimental settings and procedural details to facilitate reproducibility and understanding.

The results obtained through the implementation of the proposed model are discussed and compared with the present approaches in Section 4, enabling a comprehensive analysis of the model's efficiency. Finally, Section 5 serves as the conclusion of the study.

2. Related Works

Diabetic Retinopathy has become a major global issue that is attracting the attention of academics from all over the world who are trying to figure out the best ways to detect vision abnormalities early on and prevent them from getting worse. Numerous research studies have been undertaken and are still in progress in this area, all with the goal of enhancing patient and medical professional quality of life. This section offers an extensive summary of many research initiatives pertaining to diabetic retinopathy.

To address the global impact of DR, Thippa Reddy Gadekallu et al. (2023) [6] proposed a Deep Neural Network (DNN) approach that incorporates Grey Wolf Optimization (GWO), and Principal Component Analysis (PCA). They employed PCA to reduce dimensionality, the standard scaler method to standardize the dataset, and GWO to optimize the hyperparameters. For binary classification and backpropagation, the DNN used sigmoid activation and RMSprop, respectively. For training and testing, data was divided into an 8:2 ratio and batch training was used. Their model performed better than more conventional algorithms like SVM and Decision Tree, with a sensitivity of 91%.

By developing a DL model enriched with PCA for feature extraction and the Harris Hawks optimization method for classification optimization, Nagaraja Gundluru et al. (2022) [7] tackled the major issue of diabetic retinopathy. Their model, which made use of the Diabetic Retinopathy Debrecen Data Set from the UCI library, demonstrated notable gains in specificity, precision, accuracy, and recall despite prior machine learning and deep learning attempts producing less than ideal outcomes. However, the possibility of overfitting in low-dimensional datasets could limit the model's efficiency.

Anas Bilal et al. (2022) [8] used a two-stage process that included segmenting Blood Vessels (BV) and Optic Discs (OD) using two U-Net models, then extracting features using a hybrid CNN-SVD approach and classifying DR using Inception-V3. Limitations included difficulties in DR classification because of low-resolution images and a lack of training examples in DIARETDB0, even though high accuracies were achieved on the Messidor-2, EyePACS-1and DIARETDB0 datasets. When compared to other models, Inception-V3 performed better.

Wejdan L. Alyoubi et al. (2021) [9] developed an automated diagnostic system to address the irreversible damage that DR causes to retinal blood vessels. Their method identified lesions on the retina and correctly grouped DR images into five stages, outperforming manual techniques. CNN512 obtained 88.6% accuracy on the DDR dataset and 84.1% accuracy on the APTOS Kaggle 2019 dataset by employing deep learning models. On the DDR dataset, YOLOv3 yielded a 0.216 mAP improvement in lesion localization. Combining YOLOv3 with CNN512 yielded an accuracy of 89%, surpassing previous techniques. However, the study's effectiveness was hindered by the unbalanced and low-quality datasets, crucial for efficient screening systems.

Ping-Nan Chen et al. (2021) [10] employed retinal ophthalmoscopy for early detection of DR and utilized deep learning to enhance diagnostic accuracy and streamline workflow. They devised a 2-stage training model to address overfitting issues, where the first stage identified DR presence while the second categorized DR severity. Their model achieved promising accuracies ranging from 84.27% to 92.5% across multiple datasets. However, limitations included the high failure rate of cropping the image at the eyeball edge and potential noise in the dataset due to varying camera qualities. Moreover, large minibatch sizes in training cause data loss from GPU memory.

Zubair Khan et al. (2021) [11] focused on automating Diabetic Retinopathy (DR) diagnosis using a VGG-NiN model, combining VGG16, Spatial Pyramid Pooling layer (SPP), and Network-in-Network (NiN) to categorize DR stages efficiently with minimal learnable parameters. They utilized the EyePACS dataset, resizing images to 1349 × 1024 and randomly cropping them to 1024 × 1024 for preprocessing. The model achieved superior performance compared to existing methods, exhibiting a 52% reduction in parameters while maintaining a micro-AUC of 0.95. The problem of early DR identification was examined by Borys Tymchenko et al. (2021) [12]. They unveiled a deep learning model that uses automatic learning to detect the DR stage from individual fundus images. With an ensemble of three CNN architectures with transfer learning, they were able to obtain excellent sensitivity and specificity. Despite the success, there were still drawbacks, such as the need for labeled datasets and inconsistent expert opinion.

Gangwar et al. (2021) [13] presented a hybrid DL method for the automatic detection of DR. By employing a custom CNN block and a pre-trained Inception-ResNet-v2, they applied Transfer Learning (TL) to achieve test accuracies of 72.33% and 82.18% on the Messidor-1 and APTOS 2019 datasets, correspondingly. One limitation was that the training and testing datasets were not disclosed, making it impossible to compare the results with those of earlier research. Retinal Fundus Image Analysis (RFIA) was suggested by Imran Qureshi et al. (2021) [14] as a DR screening tool to lessen the risk of blindness in diabetic individuals. Their design for Active Deep Learning (ADL) made it easier to recognize DR stages automatically. Convolutional Neural Networks (CNNs) were utilized for feature extraction; nevertheless, there were difficulties because of the need for a lot of labeled data. To solve this problem, they presented ADL-CNN, which makes use of Expected Gradient Length (EGL) active learning. Clinicians were assisted in annotating DR severity levels by the instructive patches and photos that were picked during a two-stage approach. ADL-CNN outperformed other models tested on the EyePACS dataset, demonstrating high classification accuracy, sensitivity, and specificity.

In response to the rising diabetes epidemic in India, Silky Goel et al. (2021) [15] developed a Deep Learning (DL) model utilizing the IDRID dataset for the identification of DR. They stressed the significance of early DR screening because of the disease's progressive effects on vision loss. Their DL model achieved good classification accuracy for DR lesions by using TL on IDRID and the ImageNet dataset. They used different classifiers, using the VGG16 model, to classify lesions with 91% to 95.9% accuracy. The potential need to expand the dataset size or investigate different deep learning models, like ResNet, MobileNet, or EfficientNet, in order to increase feature learning and classification accuracy constitutes the study's drawback.

Gaurav Saxena et al. (2020) [16] trained strong classification models on open datasets to solve the critical issue of DR detection. They employed CNNs and obtained an AUC of 0.92 on the Messidor-2 dataset, with corresponding sensitivities of 88.84% on Messidor-1 and 81.02% on Messidor-2. Notwithstanding the positive outcomes, issues with the quality of the dataset and label noise remained. Deep learning of fundus images was presented by Shu-I Pao et al. (2020) [17] as a useful and affordable method for automatic screening and diagnosis of severe DR. In order to improve referable DR detection, they suggested employing entropy images generated from the green component of fundus images, augmented by Unsharp Masking (UM). They simultaneously processed luminance and green component entropy images using a bi-channel CNN model, which improved detection performance. Thirty-one of the 35,126 fundus images used in the study were chosen for acceptable image quality out of the "Kaggle Diabetic Retinopathy" dataset. At 87.83%, 77.81%, and 93.88%, respectively, the bichannel CNN demonstrated high levels of accuracy, sensitivity, and specificity.

A DNN with Recursive Feature Elimination (RFE) was developed by Ganjar Alfian et al. (2020) [18] to distinguish DR early on the basis of individual risk factors. Their solution outperformed existing algorithms, achieving 82.033% accuracy using a publicly available dataset. They used the model to forecast the development of DR in patients with T1D and T2D and found that it performed better than more conventional methods like SVM and k-Nearest Neighbor (KNN). When compared to DNN alone, the combined DNN plus RFE technique increased prediction accuracy by 9.07%. The dataset's modest size and narrow demographic, however, restricts the study's capacity to be broadly applied. Deep transfer learning utilizing the Inception-v3 model was deployed by Feng Li et al. (2019) [19] to automate the identification of DR in retinal fundus images. They obtained a sensitivity of 96.93%, accuracy of 93.49%, and specificity of 93.45% by using a 10-fold cross-validation technique. Referral recommendations demonstrated potential for the strategy, with a 0.919 j value. The study was retrospective, which limited its real-time usefulness even with its success.

With the goal of automating the diagnosis of DR, Zhentao Gao et al. (2018) [20] developed a dataset of DR fundus images. This dataset was used to train deep convolutional neural networks, which produced a classification accuracy of 88.72% for four-degree severity. When compared to ophthalmologists, the models' consistency rate was 91.8% when they were implemented on a cloud platform, offering diagnostic services to numerous hospitals. The dataset included 4476 images that were preprocessed and had data augmentation performed from three clinical departments. The use of a moderate-sized dataset and different sample sizes for each category in the training, as well as test sets, were constraints, notwithstanding the successful pilot deployment.

3. Materials and Methods

Since DR affects a greater number of people, it is critical to identify them effectively. The proposed methodology for DR detection begins with the collection of retinal images sourced from Kaggle. Following this, image preprocessing and augmentation are employed to boost the dataset. The pre-processed and augmented images are then divided into a training and a testing set in an 80:20 ratio. These images are subsequently input into the proposed HR-Net model, which initiates with a stem block and progresses through three stages, each comprising multiple residual blocks. These stages progressively downsample the spatial resolution of the feature maps. Subsequently, global average pooling is employed, followed by a dense layer with sigmoid activation for binary classification. To prove its higher accuracy, the proposed model's performance is measured using a variety of performance indicators and contrasted with the models that already exist. Figure 4 depicts the overall framework, and the individual steps are described below.

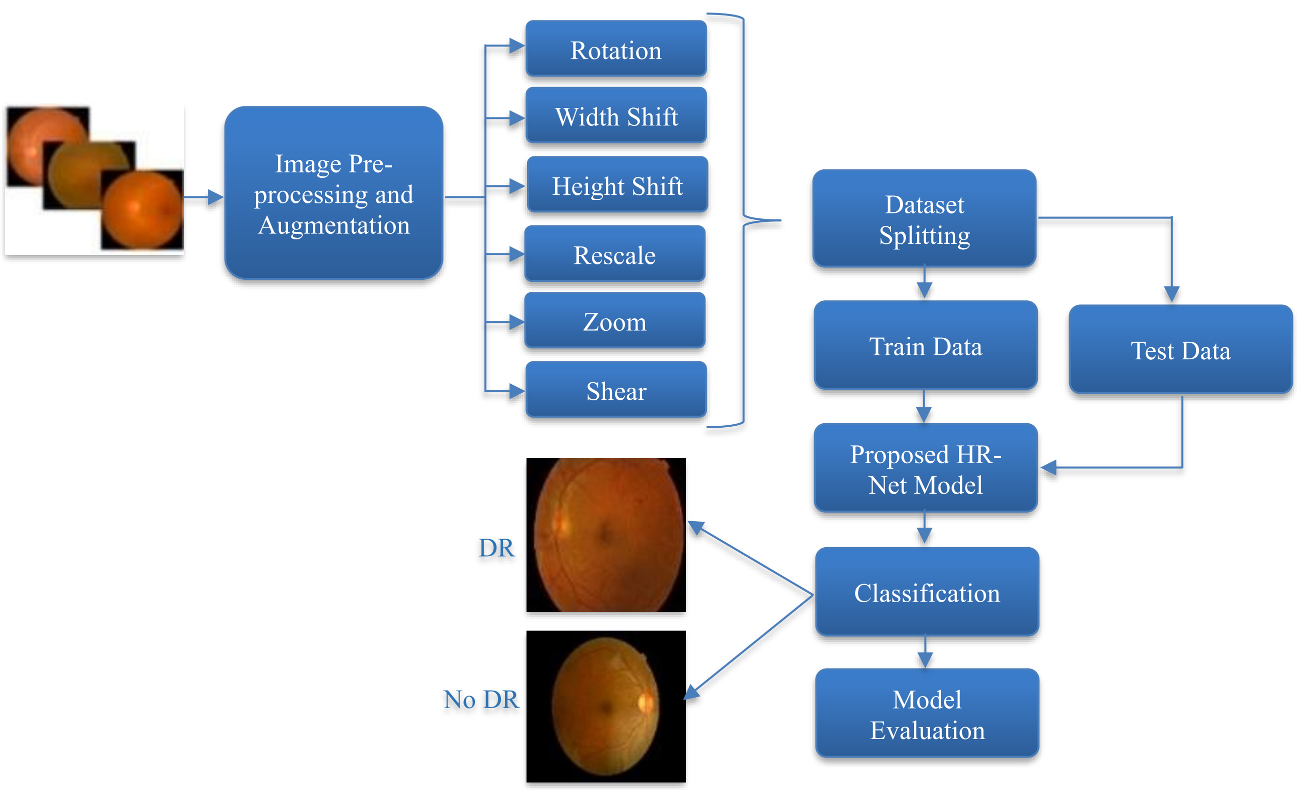


Fig. 4 Block schematic for the proposed approach

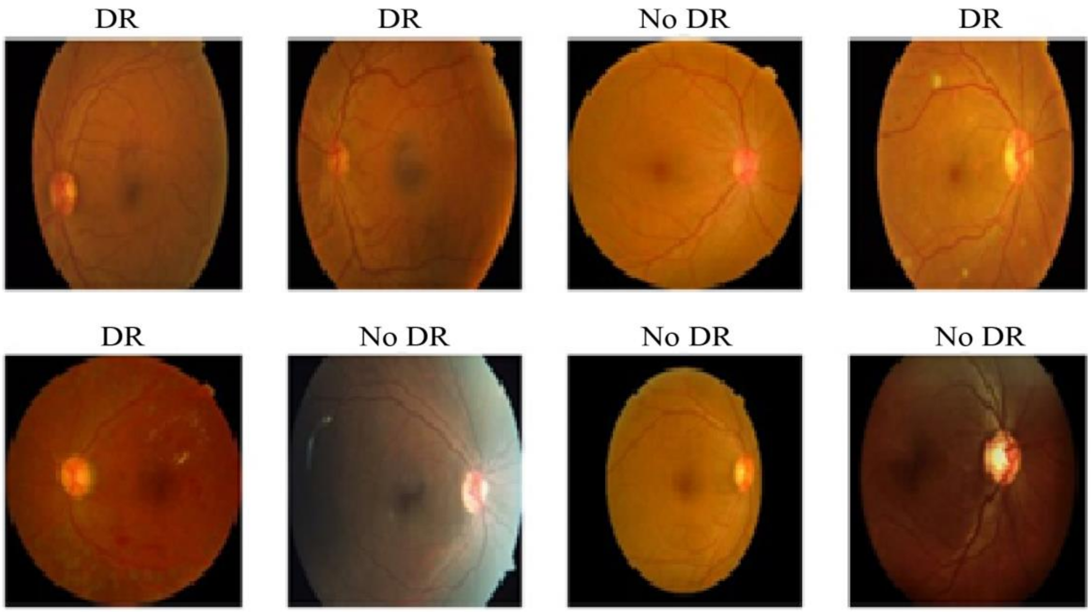


Fig. 5 Sample images from the dataset

3.1. Dataset

The dataset was obtained from Kaggle, a publicly accessible repository that may be accessed at https://www.kaggle.com/datasets/pkdarabi/diagnosis-of-diabetic-retinopathy/code. It consists of an enormous collection of sharp retinal images taken in a variety of imaging scenarios. A medical expert reviewed each image and issued a rating ranging from 0 to 1, indicating the presence or absence of DR. In particular, a score of 0 denotes the existence of DR, whereas a score of 1 denotes its absence. Figure 5 displays some of the dataset's sample images.

3.2. Image Preprocessing and Augmentation

Image preprocessing and augmentation are pivotal stages in data preparation for deep learning models. Due to the presence of considerable noise in the dataset's images, preprocessing was essential. The image preprocessing phase in this study relied on the OpenCV library [21], leveraging its versatile capabilities for efficient image manipulation. Specifically, the library's functions, including cv2.warpAffine and cv2.warpPerspective, were utilized for image transformation tasks, along with noise reduction techniques. OpenCV emerged as a crucial tool due to its extensive array of built-in features tailored for rapid and effective image processing.

By employing these functions, the preprocessing procedure successfully eliminated prominent black borders often present in fundus images, enhancing the overall quality and suitability of the images for further evaluation. Preprocessing encompasses tasks like resizing, normalization, and noise reduction aimed at enhancing image consistency and quality. The images were standardized to dimensions of 256x256 pixels in width and height to facilitate uniformity.

The size of the training dataset is one of the key components in the effective processing of DL models. Therefore, a large dataset is needed for deep learning network training in order to prevent overfitting and problems with generalization. A data expander becomes indispensable, serving to augment existing datasets without the need for additional data collection. This process, facilitated by the built-in functionalities of the Keras library, effectively expands the dataset's volume. To eliminate noise from fundus images and enlarge the retinal dataset at various sizes, data augmentation strategies are utilized. By transforming existing images in different ways, augmentation techniques generate new training samples. [22]. Moreover, to enhance efficiency and streamline the learning process, fundus images are cropped to smaller sizes, eliminating extraneous elements and focusing solely on the relevant regions of interest. This approach optimizes data utilization while maintaining the integrity and relevance of the information processed by the DL model.

Augmentation techniques include rotation, width and height shifting, rescaling, shearing, and zooming. Rotation allows the image to be rotated within a specified range, while width and height shifting shifts the image horizontally and vertically. Rescaling normalizes the pixel values to a range between 0 and 1. Shearing introduces shear deformation while zooming changes the scale of the image.

The fill mode 'nearest' is used to fill in any empty pixels created by the transformations with the nearest available pixel value. Overall, these augmentation methods help create a more varied and comprehensive dataset, which eventually enhances the model's efficiency and capacity for generalization. Figure 6 shows the primary data augmentation processes performed.

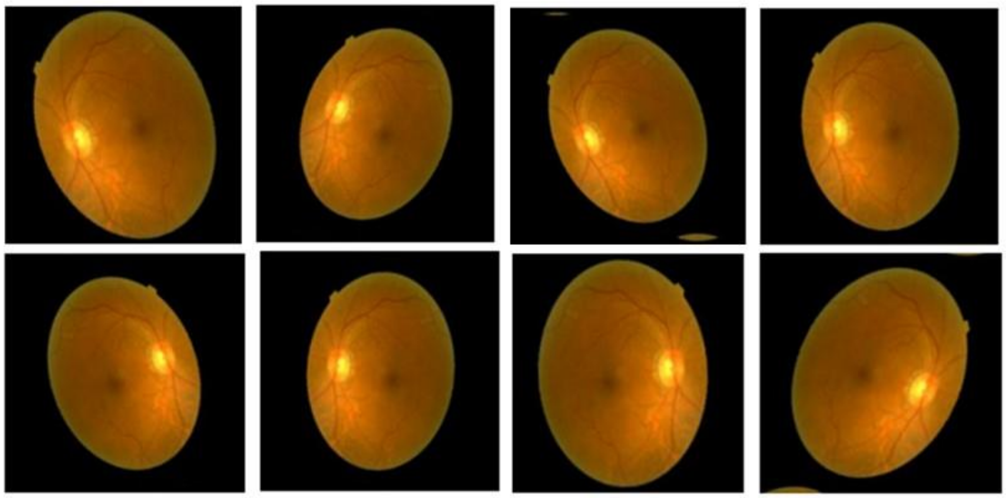


Fig. 6 Augmented retinal images

Following data preprocessing and augmentation, the retinal images within each dataset were partitioned into training and testing sets, adhering to an 80:20 ratio. Subsequently, the prepared data was then inputted into the HR-Net architecture for model training.

3.3. Proposed HR-Net Model

The High-Resolution Network (HR-Net) is a DNN framework designed specifically to handle and exploit high-resolution visual information effectively [23]. It is particularly useful in computer vision tasks where maintaining fine-grained details in images is essential for accurate analysis and decision-making. The architecture of HR-Net typically consists of a hierarchical structure that progressively refines feature representations at different scales, allowing it to gather high-level contextual information and fine-grained features simultaneously.

At the core of the HR-Net architecture is the concept of maintaining high-resolution representations throughout the network. Unlike traditional Convolutional Neural Networks (CNNs) that downsample the spatial resolution of feature maps as they progress through the network, HR-Net retains high-resolution feature maps at each stage. As a result, the network can preserve detailed information throughout the processing pipeline, making it particularly well-suited for tasks that require precise localization and fine-grained analysis.

The architecture typically begins with a stem block, which serves as the initial feature extractor. Multiple stages follow this, each containing a series of residual blocks or parallel pathways responsible for processing feature maps at different resolutions. These blocks operate in parallel, allowing HR-Net to maintain high-resolution representations throughout the network. Within each block, convolutional units perform feature extraction, capturing both local as well as global information from the input data. To accommodate multi-scale processing, HR-Net includes down-sampling and up-sampling operations, which enable the model to aggregate information across different levels of detail. Down-sampling reduces the spatial resolution of feature maps, facilitating the extraction of high-level features, while up-sampling restores the resolution, preserving fine-grained details. This hierarchical architecture, coupled with parallel processing and multi-scale feature extraction, enables HR-Net to analyze complex images effectively.

Figure 7 illustrates the HR-Net architecture, typically featuring four main stages: high-resolution representation, multi-resolution fusion, high-to-low resolution transformation, and prediction. The network processes input images at their original resolution in the high-resolution representation stage, maintaining fine details. In the multi-resolution fusion stage, features from various resolutions are combined to capture both local and global information efficiently. As the network transitions to the high-to-low resolution transformation stage, it gradually reduces spatial resolution while retaining crucial information. This ensures a balance between computational efficiency and information preservation. Finally, in the prediction stage, the refined feature representations guide tasks such as classification, segmentation, or object detection.

For the proposed model, HR-Net ensures the preservation of fine details crucial for DR diagnosis through its ability to maintain high-resolution representations throughout the network. With multi-resolution fusion capabilities, it captures both local and global features, enabling a comprehensive analysis of retinal abnormalities associated with DR. The input images have dimensions of 224x224 pixels with three channels. The model commences with the initial stem block, comprising a Conv2D layer with 64 filters, utilizing a 3x3 kernel, a stride of 2, and employing the same padding for maintaining spatial dimensions.

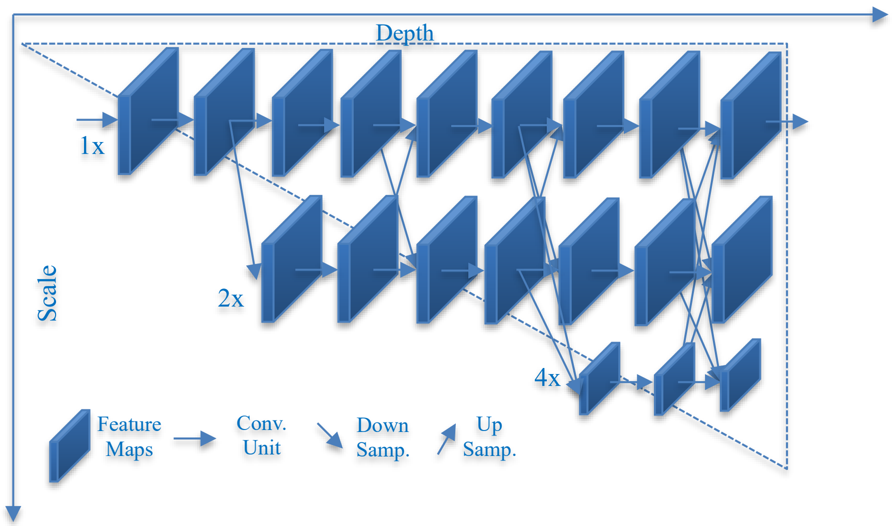


Fig.7 General layout of HR-Net

Batch normalization and ReLU activation are applied subsequently to enhance feature extraction. Following this, a MaxPooling2D layer with a 3x3 pool size, a stride of 2, and the same padding is employed to further down-sample the feature maps. Subsequently, the model progresses through three stages (Stage 1, Stage 2, and Stage 3), each comprising two residual blocks tailored to capture distinct features with specified numbers of filters and strides. A GlobalAveragePooling2D layer is then utilized to aggregate the features across spatial dimensions. Ultimately, a dense layer with a single neuron and sigmoid activation function is employed for binary classification, facilitating the detection of DR-related abnormalities in retinal images. This detailed architecture ensures effective feature extraction and classification, contributing to the accurate diagnosis of DR.

4. Results and Discussion

4.1. Hardware and Software Setup

The research employed a high-performance computational setup comprising an Intel Core i7 processor, 32GB of RAM, and the powerful NVIDIA GeForce GTX 1080Ti GPU. Python was the chosen language for implementation, leveraging a diverse array of libraries for image processing and familiarizing with the system to develop a convolutional neural network like HR-Net. For image manipulation tasks such as rotation and resizing, the OpenCV library was employed, while NumPy [24] was utilized for mathematical operations required during implementation. TensorFlow [25] and Scikit-learn [26] were instrumental in effectively managing deep learning models and defining the model architecture. Keras, recognized for its intuitive interface and advanced functionalities, played a pivotal role in designing complex Neural Network architectures. This framework efficiently utilizes computing resources, supporting various environments, including CPU, GPU, and TPU. For enhanced computational prowess and streamlined model training, Google Colab was utilized. This cloud-based Python notebook environment not only offers access to robust computational resources but also facilitates collaborative development, making it an ideal platform for training models.

In configuring the proposed model, hyperparameters play a pivotal role in determining its behavior and performance during the training phase. Unlike model parameters learned from the data, hyperparameters are predetermined by the user. The neural network model is optimized using the Adam optimizer and encompasses 2776449 trainable parameters. Training is guided by the binary cross-entropy loss function, with data processed in batches of 64 samples per iteration. Throughout 30 epochs, representing the number of times the entire training dataset is iterated, the framework is trained. These hyperparameter selections collectively define the training configuration tailored to optimize the model's performance for the specified task. Table 1 shows an overview of the model configuration in the proposed approach.

4.2. Performance Evaluation

The proposed methodology for DR detection utilizes various essential performance evaluation metrics to gauge the efficiency of the HR-Net model, as detailed in Table 2. These indicators provide an extensive evaluation of the model's functionality, providing insights into its predictive capabilities across diverse aspects.

Table 1. Model configurations

Model Parameters	Values
Trainable Parameters	2776449
Optimizer	Adam
Batch size	64
Epochs	30
Activation Function	Sigmoid, ReLU
Loss Function	Binary Cross-Entropy

Table 2. Performance metrics

Performance Parameters	Equations
Accuracy	(TP + TN) / (TP + TN + FP + FN)
Precision	TP / (TP + FP)
Recall	TP / (TP + FN)
F1 Score	2 * (Precision * Recall) / (Precision + Recall)
where TP-True Positives (correctly identified as having DR), TN-True Negatives (correctly identified as not having DR), FP-False Positives (incorrectly predicted as DR when they are not), and FN-False Negatives (incorrectly predicted as Not having DR when they are DR)	

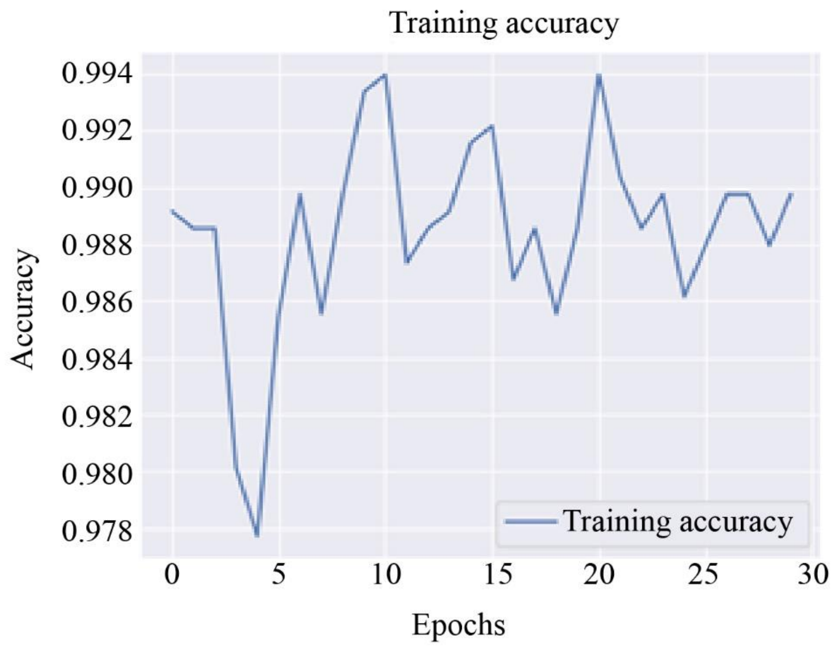


Fig. 8 Accuracy plot of the proposed model

Fig. 9 Loss plot of the proposed model

4.3. Experimental Results

The accuracy and loss plots serve as crucial visual aids in comprehending the performance and learning dynamics of the DR detection framework. The accuracy plot visually tracks the model's accuracy during its training iterations on both the training and validation datasets. Accuracy measures how accurately the predictions of the model align with the actual labels of the data.

The accuracy plot provides information about the model's capacity to differentiate between retinal images with and without signs of DR throughout the training phase. Ideally, during early epochs, the training accuracy rises simultaneously, indicating the model's ability to generalize its learnings beyond the training data, as shown in Figure 8. This suggests that the model is grasping underlying patterns rather than merely memorizing the training samples.

The difference between the predicted outcomes and the actual labels of the data is represented numerically as the model's loss, which is depicted in the loss plot. As training progresses, the loss ideally decreases, signifying iterative improvements in the model's predictions as it learns to minimize errors, as depicted in Figure 9.

A useful technique for assessing how well the model performs in identifying retinal images is the confusion matrix. It offers a tabular display of the model's predictions compared to the actual ground truth labels across different classes. The matrix typically consists of rows denoting the actual classes (true labels) and columns denoting the predicted classes. Each cell in the matrix contains the count of instances where predictions match or differ from the ground truth labels.

The confusion matrix typically consists of four quadrants, as shown in Figure 10. The diagonal components of the matrix indicate scenarios in which the model's predictions coincide with the ground truth labels, while the off-diagonal elements signify misclassifications.

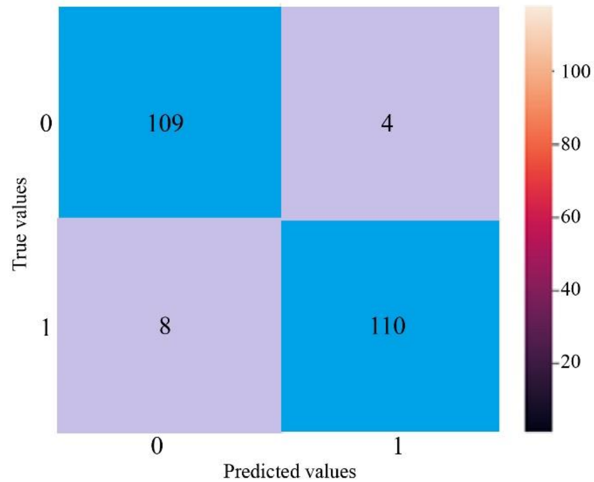


Fig. 10 Confusion matrix

From the Confusion matrix, various performance measures are calculated to assess the effectiveness of the model, as depicted in Figure 11, providing a comprehensive understanding of its strengths and weaknesses in DR detection.

A graphical depiction of the model's performance at various discrimination thresholds is the Receiver Operating Characteristic (ROC) curve. On the ROC curve, the True Positive Rate (sensitivity) is graphed against the False Positive Rate (specificity - 1) for various threshold values. The percentage of correctly recognized positive cases (DR-positive images) among all actual positive instances is known as sensitivity, and the percentage of correctly identified negative cases (DR-negative images) among all actual negative cases is known as specificity.

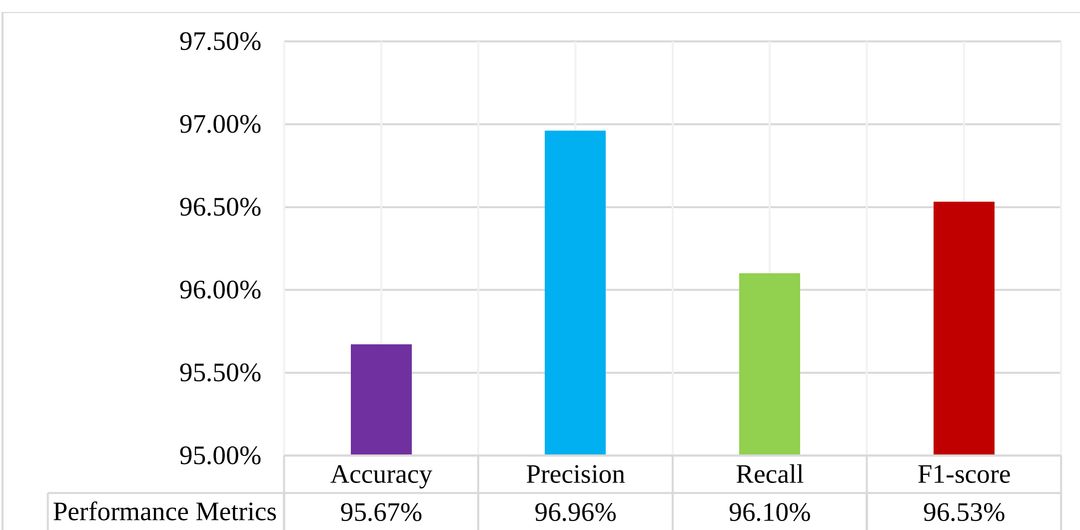


Fig. 11 Performance metrics

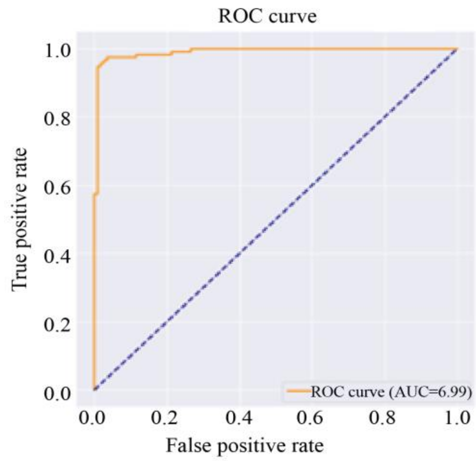


Fig. 12 ROC curve

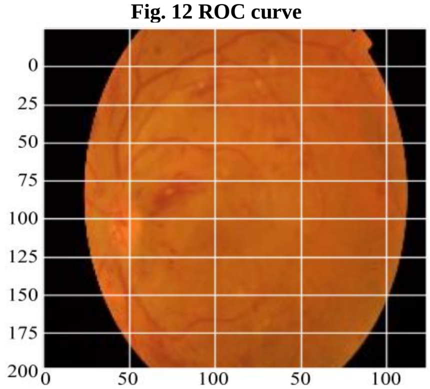


Fig. 13 Prediction output

An Area Under the Curve (AUC) value of 0.99, attained in this study, as shown in Figure 12, indicates an exceptional performance of the DR detection model. The AUC quantifies the overall discriminative ability of the model across all possible thresholds, with a value closer to 1 indicating higher discriminatory power. In this case, the high AUC value suggests that the model achieves near-perfect classification performance, effectively distinguishing between retinal images with and without signs of DR across a range of threshold values. This indicates the robustness and effectiveness of the proposed DR detection methodology.

Following the classification and prediction process, a random image from the test dataset was chosen, and the proposed HR-Net model accurately identified it as a DR image, as shown in Figure 13. This successful classification underscores the effectiveness of the model in discerning DR-related features and patterns within retinal images. Such precision in classification highlights the potential of the HR-Net architecture in aiding clinicians by swiftly and accurately identifying DR cases, contributing to early intervention and management strategies.

This outcome exemplifies the model's robustness and reliability in real-world applications, offering promising prospects for enhancing diabetic retinopathy diagnosis and treatment.

Table 3 presents a comparative analysis of the proposed framework's accuracy alongside various existing studies. Remarkably, the proposed model achieves the highest accuracy of 95.67%, demonstrating its superior performance compared to the other methodologies considered.

Table 3. Performance comparison of the proposed model with existing studies

Author	Methodology	Accuracy (%)
Gangwar et al. (2021) [13]	Inception-ResNet-v2	82.18
Zubair khan et al. (2021) [11]	VGG-NiN	85
Ping-Nan Chen et al. (2021) [10]	NASNet-Large deep CNN	85-92
Nagaraja Gundluru et al. (2022) [7]	XG Boost -PCA-HHO	88
Wejdan L. Alyoubi et al. (2021) [9]	CNN512 and YOLOv3	89
Anas Bilal et al. (2022) [8]	Hybrid CNN-SVD	93.52
Proposed Model		**95.67**

5. Conclusion

The proposed model presents a novel methodology for DR detection utilizing the High-Resolution Network (HR-Net) architecture. Through rigorous preprocessing and augmentation techniques, the proposed approach enhances the quality and diversity of the retinal image dataset, laying a solid foundation for accurate DR diagnosis. Leveraging the unique features of HR-Net, including high-resolution representation and multi-resolution fusion, the proposed framework achieves exceptional performance in detecting DR-related abnormalities in retinal images. The comparative analysis shows that the suggested model is better than current approaches, with an astounding accuracy of 95.67%. These outcomes highlight the significance of HR-Net architecture in advancing diabetic retinopathy diagnosis and treatment, offering promising avenues for early intervention and improved clinical outcomes. Moving forward, the proposed methodology holds great potential for real-world application, contributing to the early detection and management of DR, thus mitigating the risk of vision loss and enhancing the standard of life for individuals affected by diabetes.

Acknowledgements

The author expresses profound appreciation to the supervisor for providing guidance and unwavering support throughout this study.

References

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] David R. Whiting et al., "IDF Diabetes Atlas: Global Estimates of the Prevalence of Diabetes for 2011 and 2030," *Diabetes Research and Clinical Practice*, vol. 94, no. 3, pp. 311-321, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[3] Diabetic Retinopathy, Mayoclinic. [Online]. Available: https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611

[4] Shirin Hajeb Mohammad Alipour, Hossein Rabbani, and Mohammad Reza Akhlaghi, "Diabetic Retinopathy Grading by Digital Curvelet Transform," *Computational and Mathematical Methods in Medicine*, vol. 2012, no. 1, pp. 1-11, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[5] Howard E. LeWine, Retinopathy, Harvard Medical School, 2023. [Online]. Available: https://www.health.harvard.edu/a_to_z/retinopathy-a-to-z

[6] Thippa Reddy Gadekallu et al., "Deep Neural Networks to Predict Diabetic Retinopathy," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 5407-5420, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Nagaraja Gundluru et al., "Enhancement of Detection of Diabetic Retinopathy Using Harris Hawks Optimization with Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, pp. 1-13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Anas Bilal et al., "AI-Based Automatic Detection and Classification of Diabetic Retinopathy Using U-Net and Deep Learning," *Symmetry*, vol. 14, no. 7, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Wejdan L. Alyoubi, Maysoon F. Abulkhair, and Wafaa M. Shalash, "Diabetic Retinopathy Fundus Image Classification and Lesions Localization System Using Deep Learning," *Sensors*, vol. 21, no. 11, pp. 1-22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Ping-Nan Chen et al., "General Deep Learning Model for Detecting Diabetic Retinopathy," *BMC Bioinformatics*, vol. 22, pp. 1-14, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Zubair Khan et al., "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408-61416, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Borys Tymchenko, Philip Marchenko, and Dmitry Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection," *arXiv*, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Akhilesh Kumar Gangwar, and Vadlamani Ravi, "Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning," *Evolution in Computational Intelligence, Advances in Intelligent Systems and Computing*, vol. 1176, pp. 679-689, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Imran Qureshi, Jun Ma, and Qaisar Abbas, "Diabetic Retinopathy Detection and Stage Classification in Eye Fundus Images Using Active Deep Learning," *Multimedia Tools and Applications*, vol. 80, pp. 11691-11721, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] Silky Goel et al., "Deep Learning Approach for Stages of Severity Classification in Diabetic Retinopathy Using Color Fundus Retinal Images," *Mathematical Problems in Engineering*, vol. 2021, no. 1, pp. 1-8, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Gaurav Saxena et al., "Improved and Robust Deep Learning Agent for Preliminary Detection of Diabetic Retinopathy Using Public Datasets," *Intelligence-Based Medicine*, vol. 3-4, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Shu-I Pao et al., "Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network," *Journal of Ophthalmology*, vol. 2020, no. 1, pp. 1-7, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Ganjar Alfian et al., "Deep Neural Network for Predicting Diabetic Retinopathy from Risk Factors," *Mathematics*, vol. 8, no. 9, pp. 1-19, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Feng Li et al., "Automatic Detection of Diabetic Retinopathy in Retinal Fundus Photographs Based on Deep Learning Algorithm," *Translational Vision Science & Technology*, vol. 8, no. 6, pp. 1-13, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] Zhentao Gao et al., "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks," *IEEE Access*, vol. 7, pp. 3360-3370, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] OpenCV Library. [Online]. Available: https://opencv.org/

[22] Connor Shorten, and Taghi M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, pp. 1-48, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Jingdong Wang et al., "Deep High-Resolution Representation Learning for Visual Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3349-3364, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] NumPy. [Online]. Available: https://numpy.org/

[25] An End-to-End Platform for Machine Learning, Tensorflow. [Online]. Available: https://www.tensorflow.org/

[26] Scikit-Learn Machine Learning in Python, Scikit-Learn. [Online]. Available: https://scikit-learn.org/stable/