J. Integr. Sci. Technol.2025, 13(1), 1004

Journal of Integrated
SCIENCE& TECHNOLOGY

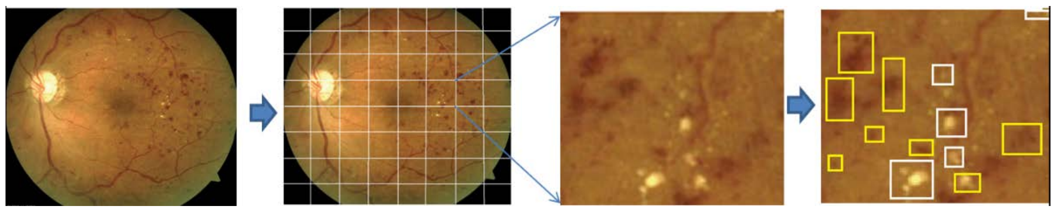
Hybrid convolutional neural network for detection of microaneurysms and exudates in retinal images

Prasad N. Maldhure, S.R. Ganorkar

Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Savitribai Phule Pune University, Pune, India.

Submitted on: 26-Feb-2024, Accepted and Published on: 10-Sep-2024

Article



ABSTRACT

The most common reason for blindness in working-age adults is diabetic retinopathy. The World Health Organization (WHO) estimates that diabetic retinopathy affects about one-third of adults with diabetes. According to the American Diabetes Association (ADA), 4.4 million Americans and 7.7 million Americans, respectively, have diabetic retinopathy that poses a threat to their vision. The development of vision loss brought on by diabetic retinopathy can be stopped or delayed with early detection and treatment. In order to increase performance for the early identification of microaneurysms and exudates, a hybrid neural network (HNN) is a sort of deep learning model that combines convolutional neural networks application on a complete image and an image segmented into 64 parts. In this instance, it is utilized to find exudates and microaneurysms in retinal pictures, both of which are symptoms of diabetic retinopathy. An accurate and effective diagnosis method for diabetic retinopathy (DR) detection is created by training a hybrid convolutional neural network on massive datasets of retinal pictures. Grading diabetic retinopathy is crucial since it aids in figuring out how serious the condition is, informs treatment choices, and tracks the disease's development. The severity of the condition is often determined by the grade of the diabetic retinopathy, which ranges from very mild to severe. HNNs outperformed weighted neural networks (WNN) and convolutional neural networks (CNN) in terms of performance sensitivity, specificity, precision, and accuracy, increasing to 91.91, 87.69, 94.74, and 90.68, respectively.

Keywords: Hybrid Neural Network (HNN), Weighted Neural Networks (WNN), Convolutional Neural Networks (CNN), Medical Image Analysis,

INTRODUCTION

For the prevention and control of diabetic retinopathy, routine eye exams and appropriate diabetes management are crucial.[1] The International Diabetes Federation estimates that diabetic retinopathy, which affects around one-third of patients with diabetes, is the main cause of new occurrences of blindness among working-age adults.[1]

Depending on the population studied and the stage of the illness, diabetic retinopathy prevalence varies. Among persons with diabetes who have been diagnosed, about 28.5% have diabetic retinopathy and about 4.4% have advanced diabetic retinopathy, which could result in significant visual loss. For the prevention and control of diabetic retinopathy, routine eye exams and appropriate diabetes management are crucial.[2] Exudates and microaneurysms are two typical symptoms of diabetic retinopathy, especially in the non-proliferative stage. Microaneurysms are tiny bulges that can form in the retina's blood vessel walls as a result of damage from high blood sugar levels. These lumps may weaken and start to bleed blood and fluid into the retina, which may enlarge and impair vision.[2]

Exudates, often referred to as hard exudates, are white or yellow deposits that can develop in the retina as a result of fluid and lipid leaking from harmed blood vessels.[3] They may indicate macular edoema, a diabetic retinopathy consequence that damages the center region of the retina that is responsible for clear, detailed vision. Exudates and microaneurysms can both be found during a thorough eye exam by an ophthalmologist or optometrist. Depending on the severity of the illness, better diabetes management, laser treatment, or pharmaceutical injections may be used to address these diabetic retinopathy symptoms.[4] Early detection and timely treatment are essential for individuals with diabetic retinopathy to prevent vision

*Corresponding Author: Dr. S.R. Ganokar, Sinhgad College of Engineering, Pune, India
Email: srganorkar.scoe@sinhgad.edu

Cite as: *J. Integr. Sci. Technol., 2025, 13(1), 1004.*
URN:NBN:sciencein.jist.2025.v13.1004
DOI:10.62110/sciencein.jist.2025.v13.1004

loss and preserve eye health. Figure 1 illustrates the presence of exudates and microaneurysms.[5]

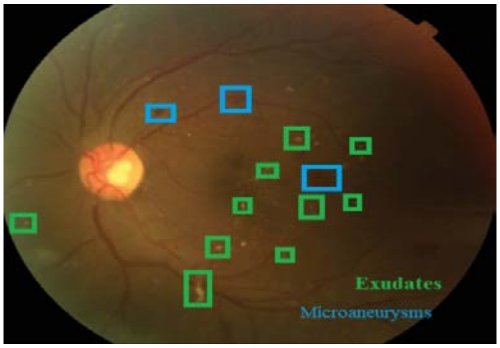


Figure 1. Appearance of Microaneurysm and Exudates

Convolutional neural networks (CNNs) applied to both a full image and its segmented 64 parts represent two distinct neural network approaches that are integrated to enhance the accuracy and efficiency in detecting microaneurysms and exudates. This combination underscores the significance of hybrid neural networks in improving the detection of diabetic retinopathy.[6]

CNNs are used to recognize and analyze image features, and they are efficient at identifying diabetic retinopathy from retinal pictures. However, CNNs are restricted by their incapacity to recognize very small Exudates and Microaneurysms.[7]HNNs, on the other hand, are effective at predicting the temporal changes in diabetic retinopathy across time since they are built to process sequences of picture fragments. It is possible to take advantage of the advantages of both strategies and boost the precision and effectiveness of diabetic retinopathy detection by integrating CNNs into a hybrid neural network.[8] The features of big or medium-sized Microaneurysm and Exudates are extracted from retinal pictures using CNN, whereas the changes in these characteristics for small-sized Microaneurysm and Exudates are modelled using HNN on individual image pieces[9].In general, hybrid neural networks have improved diabetic retinopathy detection and diagnosis, which is crucial for reducing vision loss and blindness in persons with diabetes through early identification.[10]

Related Work

This section offers an extensive review of previous research endeavors aimed at detecting and classifying diabetic retinopathy through diverse machine learning methodologies.

The methodology presented in the paper involves using feature maps generated by ResNet-50 in conjunction with a Random Forest classifier for classification tasks. The approach achieved accuracies of 96% and 75.09% on the respective datasets.[11]

In the paper, the authors present a Convolutional Neural Network (CNN) methodology aimed at accurately evaluating the severity of diabetic retinopathy using digital fundus images. They emphasize that their CNN autonomously identifies intricate features essential for classification, providing diagnoses without requiring user intervention. Trained on the Kaggle dataset using high-performance GPU resources, this network achieves significant performance, particularly in rigorous classification evaluations, with a sensitivity of 95% and an accuracy of 75% on a validation set of 5,000 images selected from an 80,000-image dataset.[12]

Authors in [13]devise three Convolutional Networks (ConvNets) employing a proposed heatmap optimization strategy to detect referable diabetic retinopathy within the Kaggle-train dataset. Without necessitating retraining, these ConvNets undergo assessment for lesion detection using the DiaretDB1 dataset. They surpass previous algorithms that were specifically developed for lesion detection at the image level Particularly noteworthy is the proposed framework, notably "net B," which secures an impressive Az (area under the ROC curve) of 0.9542 in the Kaggle-test, thus evidencing superior performance.

In Study, employing inclusive and exclusive criteria, the combined sensitivity, specificity, and area under the receiver operating characteristic curve (AUROC) for diabetic retinopathy were assessed to be 0.83 (95% CI: 0.83-0.83), 0.97 (95% CI: 0.95-0.98), and 0.92 (95% CI: 0.92-0.92), respectively. Additionally, the positive and negative likelihood ratios were calculated at 14.11 (95% CI: 9.91-20.07) and 0.10 (95% CI: 0.07-0.16), respectively. Remarkably, the diagnostic likelihood ratio for deep learning models was found to be 136.83 (95% CI: 79.03-236.93).[1]

Authors in [14]innovate a technique geared towards minimizing training time expenses by employing the Extreme Learning Machine (ELM) algorithm as the classifier. In binary classification, this approach achieves accuracy and recall rates of 99.73% and 100%, respectively. Moreover, the proposed model demonstrates accuracy rates of 98.09% and 96.26% for the five stages of diabetic retinopathy classification on the APTOS-2019 and Messidor-2 datasets, respectively.

A notable method leverages ResNet-50 for feature extraction, followed by a Random Forest classifier for final classification. This approach demonstrated impressive accuracies of 96% and 75.09% on two different datasets, effectively combining the deep feature representation capabilities of ResNet-50 with the robust classification power of ensemble methods such as Random Forests.[15]

Another significant advancement involves a Convolutional Neural Network (CNN) model designed for the automatic detection of intricate features necessary for diagnosing diabetic retinopathy (DR) from fundus images. This network, trained on the publicly available Kaggle dataset using high-performance GPUs, achieved a sensitivity of 95% and an accuracy of 75% on a validation subset of 5,000 images drawn from a larger dataset of 80,000 images. This work underscores the potential of CNNs in automating DR grading and highlights their capability to handle large-scale datasets efficiently.[16]

Further advancements include the development of three ConvNets optimized using a heatmap-based method to identify referable DR. These ConvNets were evaluated without retraining on the DiaretDB1 dataset and outperformed previous algorithms specifically trained for lesion detection, achieving an area under the curve (Az) of 0.9542 in the Kaggle-test dataset. The integration of heatmap optimization with ConvNets represents a significant leap in detecting lesions at the image level with high accuracy.[17]

A two-stage deep CNN model has been introduced, with the first stage dedicated to extracting local features and the second stage utilizing these features for global classification into four stages of diabetic retinopathy. This model achieved a Kappa score of 0.767 and an accuracy of 95.90%, showcasing its effectiveness in both detailed feature extraction and classification. The two-stage approach effectively captures both local and global context, enhancing the precision of DR classification. [18]

Attention mechanisms have been incorporated into diabetic retinopathy detection models to enhance feature extraction. Category Attention Blocks (CAB) and Global Attention Blocks (GAB) specifically target critical regions in the images, thereby significantly boosting the model's performance. The CABNet model, for instance, achieved an accuracy of 78.13%, highlighting the importance of attention mechanisms in improving the discrimination capabilities of deep learning models. [19]

Ensemble methods and transfer learning techniques have been investigated to enhance the performance of diabetic retinopathy detection. By combining several pre-trained models fine-tuned on DR datasets and utilizing transfer learning where features from pre-trained CNN models are used for DR grading, these approaches have shown accuracies between 82.84% and 90.01%, varying with the specific architecture and dataset used. This showcases the versatility and robustness of ensemble and transfer learning techniques in medical image analysis. [20]

Innovative models like the Zoom-in Net and hierarchical coarse-to-fine networks (CF-DRNet) have further advanced DR detection. The Zoom-in Net uses an M-Net for initial classification, an A-Net for generating attention maps, and a C-Net for focusing on high-resolution patches, achieving high Kappa scores and accuracies. Similarly, the CF-DRNet employs a two-stage classification process, reflecting the hierarchical nature of DR and providing detailed attention to critical regions in the images [21]

The diagnostic performance metrics of various models, including sensitivity, specificity, and area under the receiver operating characteristic curve (AUROC), have been assessed. The combined values of 0.83 for sensitivity, 0.97 for specificity, and 0.92 for AUROC highlight the robustness and reliability of these models in clinical settings. Additionally, diagnostic likelihood ratios and positive and negative likelihood ratios provide further evidence of the models' effectiveness in distinguishing between different stages of DR.[22]

METHODOLOGY

A. Dataset

This study utilizes three major datasets—IDRiD, DDR, and EyePACS—to develop and validate a deep learning model for grading diabetic retinopathy (DR). Each dataset is crucial in the training and evaluation stages, significantly enhancing the model's robustness and accuracy.

The Indian Diabetic Retinopathy Image Dataset (IDRiD) contains 508 fundus images representing different stages of diabetic retinopathy and diabetic macular edema. This dataset provides a diverse range of pathological features necessary for training a model to accurately detect and classify diabetic retinopathy. Despite its relatively smaller size, IDRiD is crucial for assessing the model's generalizability and effectiveness, providing a dataset that closely mirrors real-world clinical settings.[22]The Diabetic Retinopathy Detection Challenge (DDR) dataset includes a substantially larger collection of 9,747 images. These images exhibit varying quality and represent different stages of DR, thus providing a comprehensive training ground for the model. The diversity and scale of the DDR dataset are vital for evaluating the model's performance across a broad spectrum of real-world conditions, ensuring its capability to handle various image qualities and DR severities.[1]

The EyePACS dataset is the largest utilized in this study, containing 76,889 images. Recognized extensively in the field of DR detection, EyePACS offers a rich resource for training deep learning models. Its extensive collection of images spanning various DR stages makes it ideal for large-scale model training and validation. The size and diversity of the EyePACS dataset ensure the model's robustness and effectiveness in practical applications, enhancing its ability to generalize across different populations and clinical settings.[22] To improve the quality of the fundus images and ensure accurate feature extraction, the study utilizes Contrast Limited Adaptive Histogram Equalization (CLAHE). This preprocessing technique improves image contrast, making various retinal features more discernible and facilitating accurate DR detection. Furthermore, to address the inherent class imbalance within the datasets, the study utilizes oversampling. This method ensures equal representation of each DR grade category during training, thereby preventing the model from being biased towards majority classes and enhancing its ability to accurately classify images from minority classes.[1]

B. Preprocessing

In diabetic retinopathy, image preprocessing techniques are used to enhance image quality and improve the accuracy of detection and classification algorithms. Contrast enhancement, noise reduction, color normalization, and picture registration are the preprocessing methods. These methods make it simpler to identify and diagnose diabetic retinopathy by removing artefacts and enhancing the clarity of retinal characteristics.

i. Normalization

With the help of color normalization techniques, images are converted to a common color space or have their color and brightness adjusted to match those of a reference image or dataset. By relying on consistent image attributes for accurate analysis, this helps to lessen the effect of color and illumination fluctuations on the accuracy of automated detection and diagnostic algorithms. For a more precise diagnosis of diabetic retinopathy, color normalization is combined with other image processing techniques.

ii. Adaptive histogram equalization

AHE divides the image into small, overlapping parts and carries out histogram equalization on each region independently. This is in contrast to classical histogram equalization, which applies a global change to the entire image. This enables AHE to maintain the overall characteristics of the image while enhancing the contrast of local details. To make subtle retinal features more visible and to help with the detection of retinal illnesses, AHE is utilized in medical image analysis, particularly the study of images of diabetic retinopathy. In figure 2 AHE results of red, green, blue and color images is shown with comparison to conventional histogram equalization.[1]

C. Hybrid Convolutional neural network

In order to extract features from photos or other input data, a hybrid convolutional network is a sort of neural network architecture that incorporates various convolutional layer types.[22] Hybrid convolutional networks combine depth-wise separable convolutional layers, which are intended to be more computationally efficient than standard convolutions and require fewer parameters, with the traditional convolutional layers seen in figure 3.It has been demonstrated that employing hybrid convolutional networks rather than solely conventional convolutional layers improves the accuracy of image classification and object detection tasks. This is due to the fact that combining various convolutional layer types enables a more varied set of characteristics to be recovered from the input data, improving the ability to distinguish between various classes of objects or features. Figure 4 illustrates in detail how hybrid convolutional networks and figure 3 shown the overall training and testing phase.

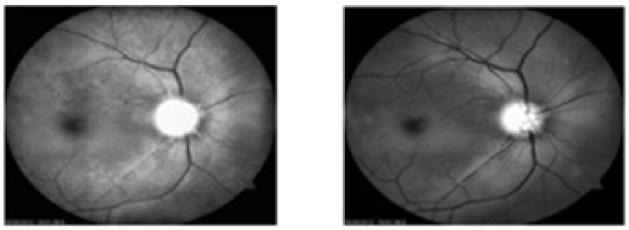


Figure 2.a. Adaptive Histogram Equalization of red and green component.

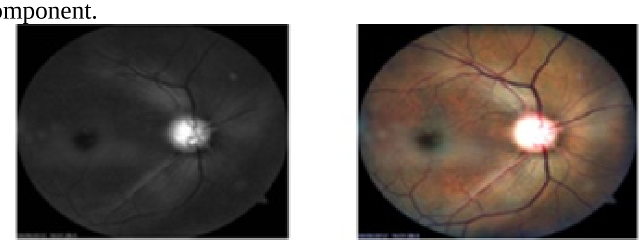


Figure 2.b. Adaptive Histogram Equalization of blue and color image.

Figure 2.c. The difference between adaptive and conventional histogram equalization.

i. Convolutional layer:

Convolutional layers in neural networks are described by the equation:

$$y_{i,j,k} = f\left(\sum l = 1^{C} \sum_{m=1}^{Kh} \sum_{n=1}^{Kw} x_{i} + m - 1, j + n - 1, l.\,wn, n, l, k + b_{k}\right)(1)$$

where x is the input tensor of size $N\times H \times W \times C$,w is the set of learnable convolutional kernels of size $K_h \times K_w \times C \times K$,$b_k$ is the bias term for the k-th output channel, and f is the activation function. The output of the convolutional layer is a feature map f{y} of size $N\times H' \times W' \times C$, Where w is the set of learnable convolutional kernels of size $K_h \times K_w \times C \times K$, b_k is the bias term for the kth output channel, and f is the activation function. x is the input tensor with size $N\times H \times W \times C$. A feature map y of size $N\times H' \times W' \times C$ is what the convolutional layer produces, whereH' and W'are the output's spatial dimensions.[23]

$$H' = H - K_h + 1 \text{ And} W' = W - K_{hw} + 1 \quad (2)$$

During the convolution operation, the kernel is slid over the input tensor, the dot product between the kernel and each local input patch is calculated, and the activation function is then applied to the total. A series of feature maps that capture various facets of the input data serve as the convolutional layer's output and are sent into the network's subsequent layers.

ii. Pooling layer:

The following is a representation of the equation for an image processing pooling layer:

$$yi, j, k = pool(\{x(i + m - 1) \times s, (j + n - 1) \times s, k: 1 \leq m \leq K_h, 1 \leq n \leq K_w\})(3)$$

where pool is the pooling function, which is either maximum or average pooling, and x is the input tensor of size $\times H \times W \times C$. K_h and K_w are the spatial dimensions of the pooling kernel.[23]By performing the pooling operation on specific local patches of the input tensor, the pooling layer lowers the spatial dimension of the input tensor.A feature map y of size$N \times H' \times W' \times C$ is the pooling layer's output, where

$$H' = \lfloor (H - K_h)/s \rfloor + 1, W' = \lfloor (W - K_w)/s \rfloor + 1 \ (4)$$

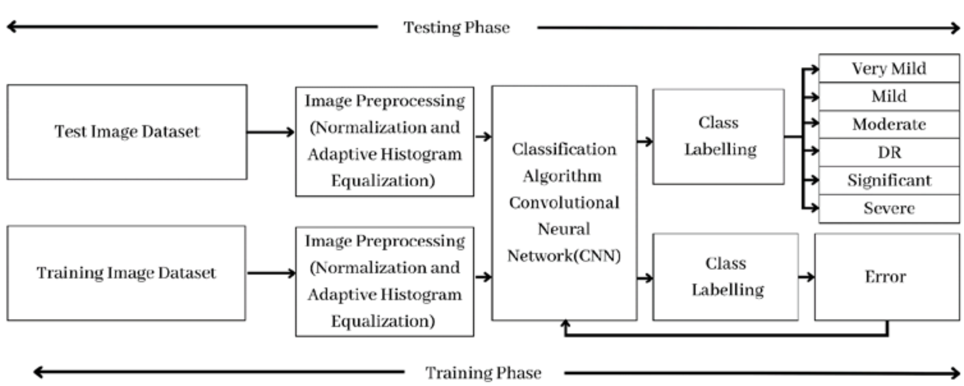


Figure 3. Convolutional Neural Network block diagram

are the output's spatial dimensions. By reducing the spatial dimension of the input tensor, the pooling procedure aids in lowering the computational cost of following layers. Additionally, average pooling serves to make the input more resistant to tiny variations in the input while max pooling helps to extract prominent features from the input.

iii. Normalization layer:

The following equation is used to represent a normalization layer:

$$yi,j,k = \frac{xi,j,k}{\left(\epsilon+\frac{1}{K_h K_w}\sum_{m=1}^{K_h}\sum_{n=1}^{K_w}\left(x_{i+m-1,j+n-1,k}\right)^2\right)^{\alpha}} \quad (5)$$

where K_h and K_w are the spatial dimensions of the normalization kernel, α is a positive constant, and x is the input tensor of size $N \times H \times W \times C$. ϵ is a small positive constant to prevent division by zero.

By normalizing each element by a factor proportionate to the sum of the squares of elements in that element's immediate local neighborhood, the normalization layer performs local response normalization to the input tensor. A tensor with the same shape as the input tensor is the normalization layer's output.Convolutional neural networks employ normalization layers to strengthen the network'scapacity for generalization and lessen the effect of covariate shift. By offering a sort of lateral inhibition where only the most intensely activated features are kept, normalization aids in improving the distinguishability of features.[23]

iv. Fully connected layer:

The equation for a fully connected layer in image processing is represented as:

$$yi = f\left(\sum j = 1^M w_{i,j} \cdot x_j + b_i\right) \ (6)$$

where f is the activation function, which is normally applied element-wise to the output vector, and x is the input feature vector of size M, y is the output feature vector of size N, w is the weight matrix of size $N \times M$, and b is the bias vector of size N.[23]

The fully connected layer performs a linear transformation of the input feature vector by calculating a weighted sum of the input features and adding a bias term.

The network's nonlinearity is then added by passing the resulting vector through an activation function. The output of the completely linked layer is utilized as the network's input for layers below it.Deep neural networks, particularly those for image processing, frequently employ fully connected layers to teach high-level representations of the input data. These layers are frequently positioned at the bottom of the network, combining the information retrieved by preceding layers into a condensed representation that may be applied to classification or regression.

v. Image Splitting:

The Python programming language and image processing library are used to divide an image into 64 parts of similar size. Utilizing the "split" function found in image processing libraries is a typical strategy.[24]

The supplied image is divided into an identically sized 8x8 grid of sub-images by this code. The subsequent saving of the sub-images into distinct files with the names "sub image 0.jpg," "sub image 1.jpg," and so on, up to "sub image 63.jpg," is done.

vi. Edge Enhancement:

The Laplacian filter is a method for edge enhancement in image processing and is represented by the equation below:

$$y(i,j) = \sum_{k=-1}^{1}\sum_{l=-1}^{1} h(k,l)\cdot x(i+k,j+l) (7)$$

If $h(k,l)$ is the Laplacian kernel and x is the input image, y is the output image, and y is provided by:

$$\begin{matrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{matrix}$$

The Laplacian filter is a second-order differential operator that emphasizes areas of an image's edge or other rapid intensity shift. The Laplacian filter reduces noise in an image and sharpens edges to enhance them.[24]

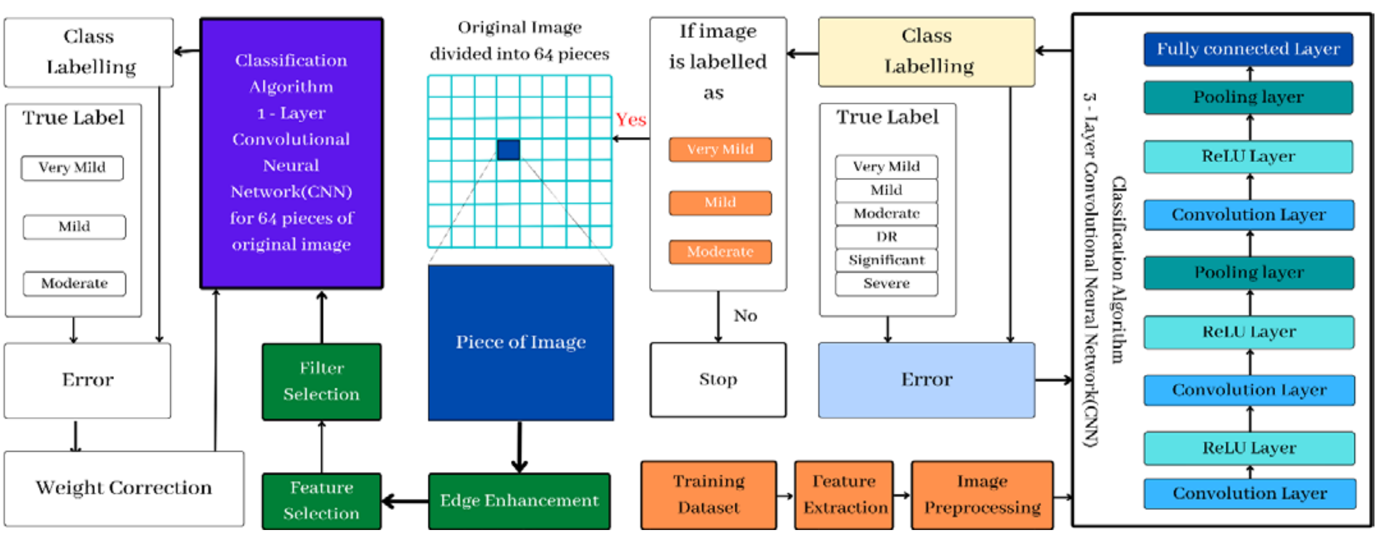


Figure 4. Hybrid Convolutional Neural Network block diagram

RESULT AND DISCUSSION

Common performance indicators used to assess the effectiveness of an HNN (Hybrid Neural Network) model include sensitivity, specificity, precision, and accuracy. These metrics are calculated by comparing a collection of predictions to a set of ground truth labels.

Performance metrics are utilized to assess the accuracy of an HNN model in image classification and to identify areas for improvement. Figure 5 illustrates the percentages for sensitivity, specificity, precision, and accuracy achieved by the hybrid convolutional neural network. Figures 6, 7, 8, and 9 compare the HNN's sensitivity, specificity, precision, and accuracy to those of CNN and WNN. Figure 10 displays a comparison of training and testing accuracy.

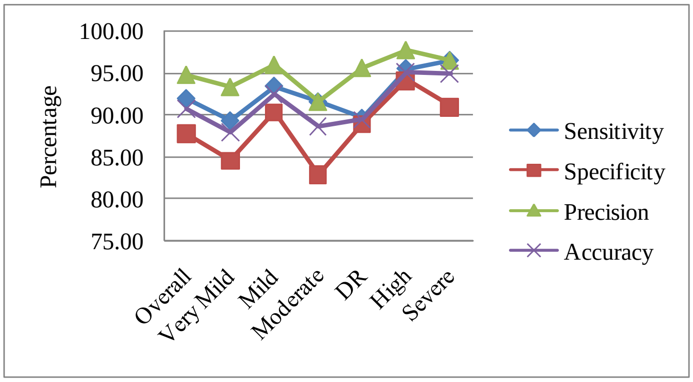


Figure 5. Percentages of sensitivity, specificity, accuracy, and precision for a hybrid convolutional neural network

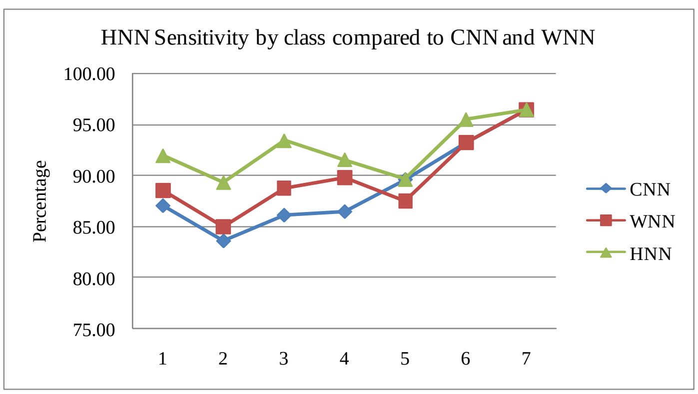


Figure 6. HNN Sensitivity by class compared to CNN and WNN.

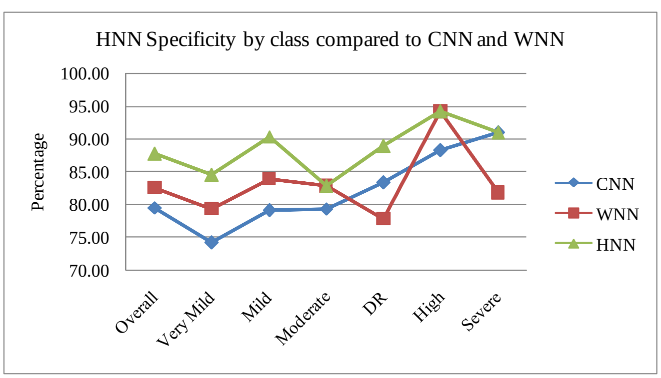


Figure 7. HNN Specificity by class compared to CNN and WNN.

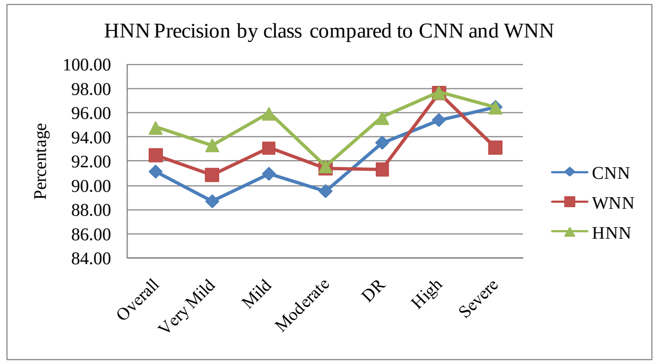


Figure 8. HNN Precision by class compared to CNN and WNN.

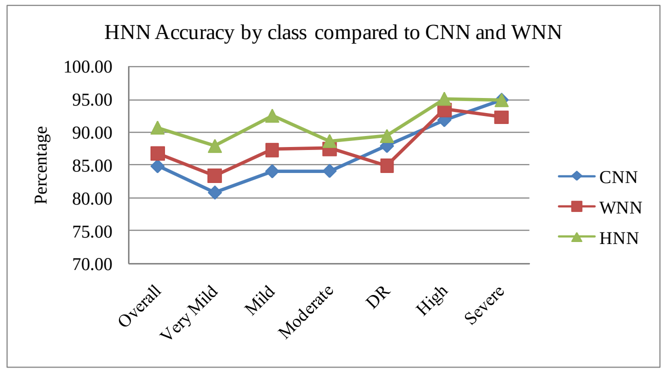


Figure 9. HNN Accuracy by class compared to CNN and WNN.

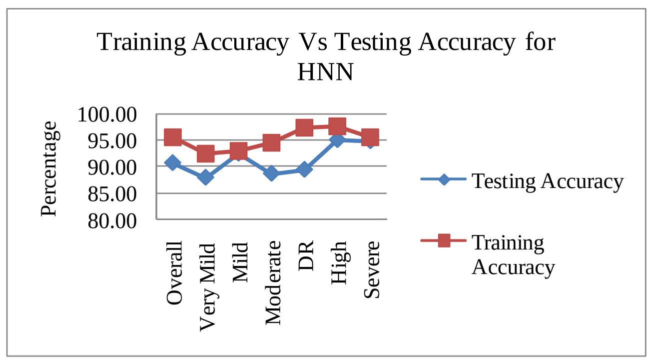


Figure 10. Training Accuracy Vs Testing Accuracy for Hybrid convolutional Neural Network

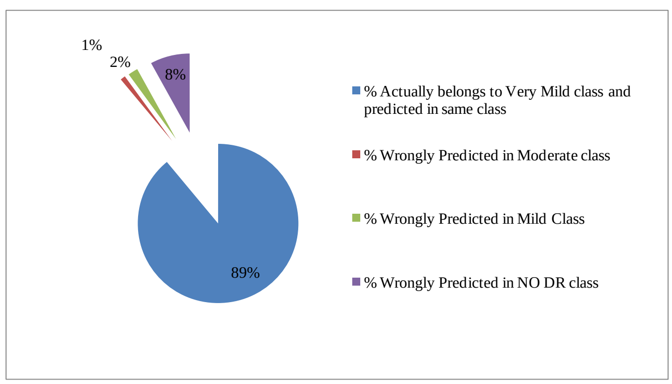


Figure 11. Percentage prediction in other classes for Hybrid

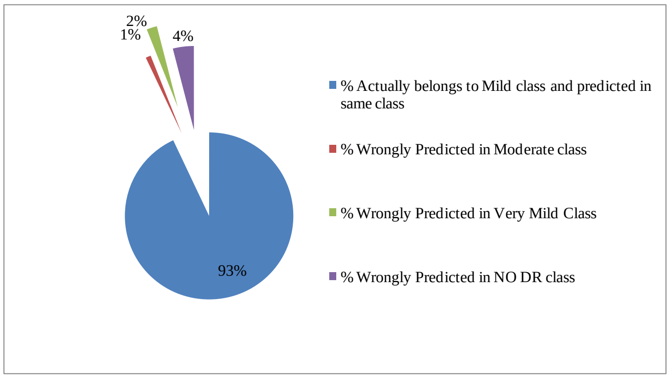


Figure 12. Actuallybelongs to the mild class, whereas Hybrid convolutional neural networks predict % in other classes.

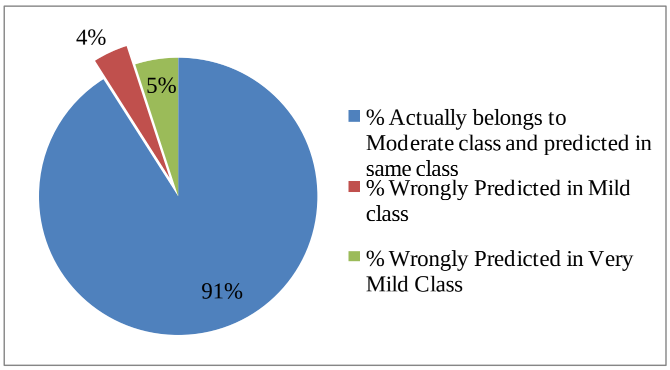


Figure 13. Actually, for Hybrid Convolutional Neural Network belongs to moderate class and% forecast in other classes.

Convolutional Neural Network and Actually belongs to relatively very mild class. Samples which truly belongs to classes Very mild, mild and moderate, but classified into different classes, for such samples analysis is carried out to get more focus on true negative samples and percentage false prediction in different classes as shown in figure 11,12 and 13 respectively.

True positive (TP) error and false positive (FP) error are two forms of prediction mistakes that might arise when evaluating an HNN (Hybrid Neural Network) model.

True Positive (TP) error:

A true positive mistake happens when the model predicts an inaccurately positive label for a falsely negative image. In other words, a feature or object that is not visible in the image is mistakenly identified by the model.[25]

False Positive (FP) error:

This error happens when the model predicts erroneously that a label that should be positive for a negative image. In other words, a feature or object that is not visible in the image is mistakenly identified by the model.[25]

When assessing an HNN model's performance, TP and FP errors are crucial metrics to take into account because they show where the model can be improved. High TP error means the model is too sensitive and picking up unimportant information, whereas high FP error means the model is too general and picking up features that aren't truly there.

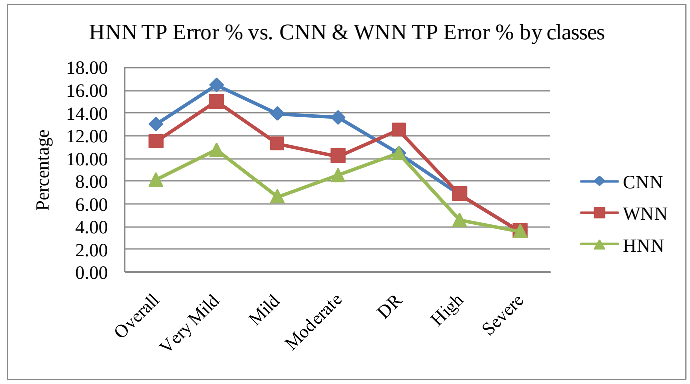


Figure 14. HNN TP Error % vs. CNN& WNN TP Error % by classes

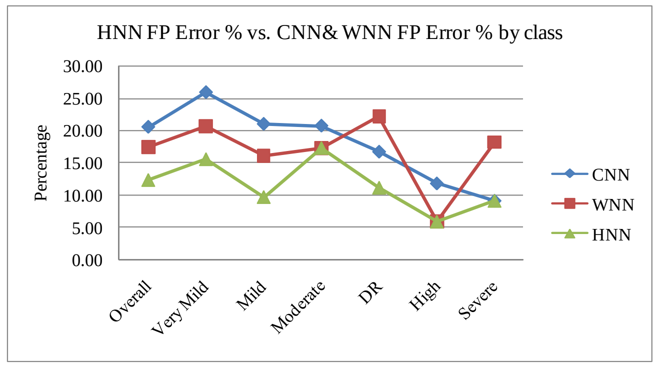


Figure 15. HNN FP Error % vs. CNN& WNN % FP Error by class

Techniques including changing the model architecture, maximizing hyperparameters, or expanding the size or diversity of the training dataset have all been used to reduce TP and FP errors. The results of HNN's TP Error & FP Error compared to CNN and WNN are presented in Figures 14 and 15, respectively. Another analysis employs Cohen's kappa, a statistical measure of inter-rater agreement, to evaluate the effectiveness of a Convolutional Neural Network (CNN) model. Kappa gauges the level of agreement between the model's predictions and the actual labels in real-world data, while adjusting for the agreement that could occur by chance.[25]

The range of kappa values is -1 to 1, with values closer to 1 suggesting a higher level of agreement between the model's predictions and the actual labels on the ground, and values closer to 0 or negative values indicating less or even no agreement.

By contrasting the predicted labels for a collection of images with the actual labels, the kappa value for a CNN model is determined.

In situations where the data is unbalanced or the classes are not equally represented, the kappa value offers further understanding of the model's properties. Figure 16 depicts the class-wise Kappa value for the hybrid convolutional neural network (HNN), which is contrasted with the CNN and WNN techniques in Figure 18. Finally, all the HNN parameters of the very mild and mild classes are compared with the CNN and WNN approaches utilized and found to be superior as shown in figures 17 and 18, respectively, with regard to the early diagnosis of diabetic retinopathy.

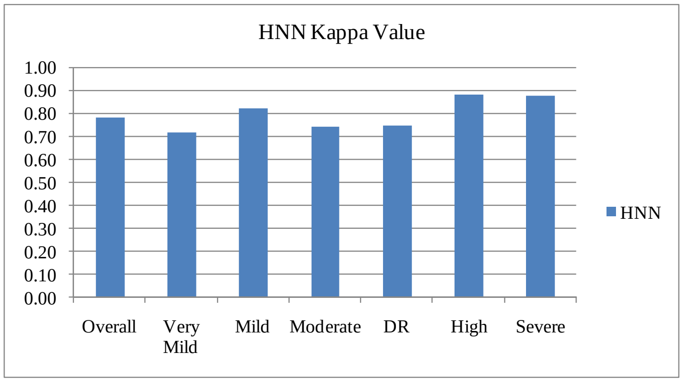


Figure 16. Class wise Kappa Value for Hybrid convolutional Neural Network

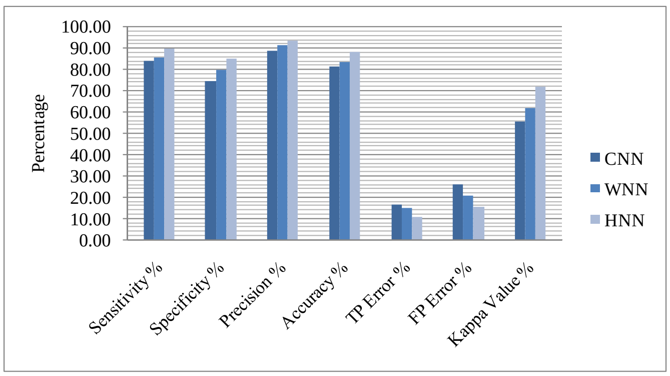


Figure 17. HNN parameter comparison with very mild grading class compared to CNN and WNN.

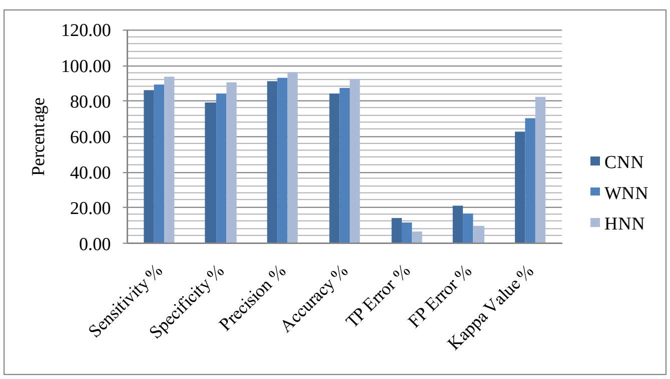


Figure 18. HNN parameter comparison with Mild grading class compared to CNN and WNN.

The proposed model synergistically integrates Hybrid convolutional neural network with a Convolutional Neural Network (CNN). This innovative combination harnesses the advantages of both methodologies, culminating in superior performance across all three datasets. The model demonstrates remarkable classification accuracies, achieving 91.06% on the IDRiD dataset, 95.34% on the DDR dataset, and 92.69% on the EyePACS dataset. These results significantly surpass the performance metrics of existing models, highlighting the efficacy of this integrated approach.

LIMITATIONS

The research study delves into the practical implementation of a deep learning model designed to detect diabetic retinopathy (DR). Throughout this process, several noteworthy limitations and challenges emerged, shedding light on areas that warrant further investigation to enhance the model's performance and real-world applicability.

The first significant challenge lies in cross-dataset testing and generalization. When evaluating the model's performance across different datasets, a notable drop in accuracy was observed during cross-dataset testing compared to within-dataset testing. Above results suggest that the model's ability to generalize across diverse datasets is limited, potentially due to overfitting to specific training data characteristics.

The second challenge pertains to computational efficiency and variability. The time required for processing and training varied significantly across datasets. Specifically, the processing time per image was 5.13 seconds for IDRiD, 6.7 seconds for DDR, and 5.4 seconds for EyePACS. Training times per image were even longer: 3.97 seconds for IDRiD, 8.56 seconds for DDR, and 4.89 seconds for EyePACS. Managing this variability poses challenges when scaling the model to larger datasets or deploying it in resource-constrained environments.

Lastly, the impact of preprocessing steps cannot be overlooked. While essential for enhancing model accuracy, preprocessing steps (such as normalization, augmentation, and sampling) introduce complexity and variability into the preprocessing pipeline. Balancing the need for accuracy improvement with maintaining consistency and robustness during training and evaluation is essential.

FUTURE SCOPE

The future of diabetic retinopathy grading is quite bright, and developments in science and technology are probably going to have a big influence on how this condition is identified and treated.

With the help of cutting-edge imaging techniques like optical coherence tomography (OCT) and adaptive optics imaging, diabetic retinopathy can be detected earlier and graded with more accuracy.There is also ongoing research to create new treatments for diabetic retinopathy, such as gene therapy, stem cell therapy, and novel medication delivery methods that can address the pathophysiology of the condition.

The creation of more potent treatments for diabetic retinopathy is another area that needs progress. The outcome of ongoing clinical studies for novel drugs, gene therapies, and cell therapies could have a substantial impact on lowering the prevalence and severity of diabetic retinopathy.

The grading of diabetic retinopathy may also be improved with improved healthcare policies and access to care. This entails raising awareness among the general population regarding the value of routine eye exams, enhancing accessibility to screening and diagnostic equipment, and ensuring that patients have access to prompt and efficient treatments. A multidisciplinary strategy that incorporates improvements in technology, research, and healthcare policies is required to improve diabetic retinopathy grading. Working together to address these issues will help to increase the precision and efficacy of diabetic retinopathy grading and, in the end, lessen the toll that this condition has on both people and society as a whole.[25]

The proposed strategy for the early diagnosis and management of diabetic retinopathy signifies a notable advancement in addressing this prevalent complication of diabetes. Through the integration of innovative technologies such as telemedicine-enabled retinal imaging and AI-driven diagnostic algorithms into healthcare infrastructures, the approach aims to expedite detection, augment accessibility to screening modalities, optimize clinical throughput, and enhance patient outcomes. Early recognition of diabetic retinopathy facilitates expeditious intervention, thereby attenuating the incidence of vision impairment and mitigating the socioeconomic ramifications on affected individuals and societal sectors. The collaborative engagement of multidisciplinary healthcare cohorts is imperative for the seamless assimilation of these methodologies into extant frameworks, thereby ensuring comprehensive and coordinated care provision for diabetic retinopathy patients.

Conclusion

In summary, the Hybrid Neural Network (HNN) model exhibits enhanced efficacy in the detection and grading of diabetic retinopathy compared to conventional weighted neural networks (WNN) and convolutional neural networks (CNN). The HNN's elevated sensitivity, specificity, precision, and accuracy render it a robust tool for the early detection of diabetic retinopathy indicators, such as microaneurysms and exudates. This early detection is critical for preventing or delaying vision loss in diabetic patients. The superior performance of the HNN highlights its potential as an advanced diagnostic tool for evaluating the severity of diabetic retinopathy and informing appropriate therapeutic interventions.

Conflict of Interest Statement

The authors confirm that they have no conflicts of interest to disclose regarding this research. There are no financial or personal relationships with individuals or organizations that could be perceived as influencing the outcomes or interpretation of this study. The results and conclusions presented are solely those of the authors and are free from external influence.

References and notes

1. P. Porwal, S. Pachade, M. Kokare, et al. IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge. *Med. Image Anal.* **2020**, 59.
2. L. Dai, L. Wu, H. Li, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat. Commun.* **2021**, 12 (1).
3. M.M. Islam, H.C. Yang, T.N. Poly, W.S. Jian, Y.C. (Jack) Li. Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis. *Comput. Methods Programs Biomed.* **2020**, 191.
4. G.T. Zago, R.V. Andreão, B. Dorizzi, E.O. Teatini Salles. Diabetic retinopathy detection using red lesion localization and convolutional neural networks. *Comput. Biol. Med.* **2020**, 116.
5. R.E. Hacisoftaoglu, M. Karakaya, A.B. Sallam. Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems. *Pattern Recognit. Lett.* **2020**, 135, 409–417.
6. G. Quellec, K. Charrière, Y. Boudi, B. Cochener, M. Lamard. Deep image mining for diabetic retinopathy screening. *Med. Image Anal.* **2017**, 39, 178–193.
7. W.L. Alyoubi, W.M. Shalash, M.F. Abulkhair. Diabetic retinopathy detection through deep learning techniques: A review. In *Informatics in Medicine Unlocked*; Elsevier Ltd, **2020**; Vol. 20.
8. J. de la Torre, A. Valls, D. Puig. A deep learning interpretable classifier for diabetic retinopathy disease grading. *Neurocomputing* **2020**, 396, 465–476.
9. P. Nagaraj, P. Deepalakshmi, R.F. Mansour, A. Almazroa. Artificial flora algorithm-based feature selection with gradient boosted tree model for diabetes classification. *Diabetes, Metab. Syndr. Obes.* **2021**, 14, 2789–2806.
10. L. Cheng, X.H. Wu, Y. Wang. Artificial flora (AF) optimization algorithm. *Appl. Sci.* **2018**, 8 (3).
11. S. Qummar, F.G. Khan, S. Shah, et al. A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. *IEEE Access* **2019**, 7, 150530–150539.
12. M. Alkhodari, M. Rashid, M.A. Mukit, et al. Screening Cardiovascular Autonomic Neuropathy in Diabetic Patients with Microvascular Complications Using Machine Learning: A 24-Hour Heart Rate Variability Study. *IEEE Access* **2021**, 9, 119171–119187.
13. A.K. Jaiswal, P. Tiwari, S. Kumar, et al. Deep Learning-Based Smart IoT Health System for Blindness Detection Using Retina Images. *IEEE Access* **2021**, 9, 70606–70615.
14. O.M. Al-hazaimeh, A.A. Abu-Ein, N.M. Tahat, M.A. Al-Smadi, M.M. Al-Nawashi. Combining Artificial Intelligence and Image Processing for Diagnosing Diabetic Retinopathy in Retinal Fundus Images. *Int. J. online Biomed. Eng.* **2022**, 18 (13), 131–151.
15. Y. Li, H. Al Hajj, P.-H. Conze, et al. Multimodal information fusion for the diagnosis of diabetic retinopathy. *Investigative ophthalmology & visual science*. 2022, p 2988-F0258.
16. S. Basu, S. Mukherjee, A. Bhattacharya, A. Sen. Segmentation of Blood Vessels, Optic Disc Localization, Detection of Exudates, and Diabetic Retinopathy Diagnosis from Digital Fundus Images. In *Proceedings of Research and Applications in Artificial Intelligence: RAAI 2020*; Springer Singapore, Singapore, **2021**; pp 173–184.
17. E. Bhatti, P. Kaur. DRAODM: Diabetic Retinopathy Analysis Through Optimized Deep Learning with Multi Support Vector Machine for Classification. *Commun. Comput. Inf. Sci.* **2019**, 1036, 174–188.
18. X. Zeng, H. Chen, Y. Luo, W. Ye. Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. *IEEE Access* **2019**, 7, 30744–30753.
19. L. Dai, L. Wu, H. Li, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat. Commun.* **2021**, 12 (1), 12.
20. V. Bansal, A. Jain, N. Kaur Walia. Diabetic retinopathy detection through generative AI techniques: A review. *Results Opt.* **2024**, 16, 1–11.
21. S.I. Pao, H.Z. Lin, K.H. Chien, et al. Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network. *J. Ophthalmol.* **2020**, 2020.
22. H. Seetah, N. Singh, P. Meel, T. Dhudi. A Convolutional Neural Network Approach to Diabetic Retinopathy Detection and its Automated Classification. In *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*; ICACCS, IEEE, **2021**; Vol. 1, pp 1000–1006.
23. M.M. Farag, M. Fouad, A.T. Abdel-Hamid. Automatic Severity Classification of Diabetic Retinopathy Based on DenseNet and Convolutional Block Attention Module. *IEEE Access* **2022**, 10, 38299–38308.
24. W.L. Alyoubi, M.F. Abulkhair, W.M. Shalash. Diabetic retinopathy fundus image classification and lesions localization system using deep learning. *Sensors* **2021**, 21 (11), 3704.
25. S.H. Abbood, H.N.A. Hamed, M.S.M. Rahim, et al. Hybrid Retinal Image Enhancement Algorithm for Diabetic Retinopathy Diagnostic Using Deep Learning Model. *IEEE Access* **2022**, 10, 73079–73086.