Natural Resources for Human Health

eISSN: 2583-1194

Original Research

Received 10 March 2026
Revised 20 April 2026
Accepted 24 May 2026

Natural Resources for Human Health 2026, Vol. 6 (1s): 345–359

KEYWORDS:
Diabetic retinopathy classification, Retinal fundus image analysis, Kolmogorov-Arnold Networks (KAN), Attention-enhanced deep learning, RegNet feature extraction, Grad-CAM++ explainability, Harris Hawks Optimization (HHO), Explainable artificial intelligence (XAI)

DiaKANet: An Explainable Attention-Enhanced Deep Feature Learning for Diabetic Retinopathy Classification from Retinal Fundus Images

S. Priya Nandini [1], S. Anu H Nair[2], K. P. Sanal Kumar[3]

[1] Research Scholar, Département of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India.
Assistant Professor, Département of Computer Engineering, Stanley Collège of Engineering and Technology, Hyderabad, Telangana.
E-mail : spriyanandini82@gmail.com
[2] Assistant Professor, Département of Computer Science and Engineering, Annamalai University, (Deputed To WPT, Chennai), India. E-mail : anu_jul@yahoo.co.in
[3]Assistant Professor, Département of Computer Science, RV Governement Arts College, Chengalpattu, India. E-mail : sanalprabha@yahoo.co.in

ABSTRACT: Diabetic retinopathy (DR) is a progressive complication of diabetes that significantly damages the blood vessels within the retina. However, it is difficult to detect in the early-stage owing to its asymptomatic nature; patients may experience intermittent or blurred vision as the condition develops. Retinal fundus images offer valuable information about the pathological changes in the retinal region, facilitating computer-aided detection and severity classification. Despite recent advances, existing techniques often struggle to deal with variations in retinal image quality and to extract subtle pathological patterns, thus limiting their reliability in multi-class classification. To tackle these issues, this study develops an Interpretable KAN-Based Framework for Multi-Class Diabetic Retinopathy Classification (IKAN-MDRC), aiming to develop an automated, robust, and explainable multi-class classification model by integrating advanced techniques using fundus images. To achieve this, the model uses a RegNet feature extraction strategy along with an efficient channel attention mechanism to effectively capture hierarchical retinal features and perform lightweight channel-wise attention to highlight informative retinal feature channels and eliminate less informative features. The extracted features are classified through the EfficientKAN classifier by effectively learning nonlinear relationships from the retinal features. In addition, the Harris Hawks optimization technique is used to select the most suitable hyperparameters to increase the classification performance and provide faster convergence. The model transparency is improved through incorporating the GradCAM++ explainability model by highlighting the significant retinal region that contributes more to the model decision. The experimental analysis is conducted using the Diabetic Retinopathy Detection dataset across various performance metrics. These findings demonstrate that the proposed IKAN-MDRC framework attains superior classification results with an accuracy of 97.61% compared to existing models and validates its potential in accurate categorization of multi-class DR with improved model interpretability to support clinical decision-making.

1. INTRODUCTION

Diabetic retinopathy (DR) is a frequent eye issue from high blood sugar that harms tiny blood vessels in the retina over time [1]. It shows up as various spots in eye scans, making regular spot detection vital. It is a top cause of preventable blindness in adults worldwide [2]. The earlier stages of DR often occur without symptoms, creating challenges in timely diagnosis. Conventional diagnosis methods mostly depend on ophthalmologists and manual evaluation of fundus images [3]. Moreover, these methods are time-intensive, computationally expensive, and relied on ophthalmologists. Such reliance on manual assessments leads to various challenges in areas with limited screening and also contributes to inconsistency in diagnostic accuracy [4]. Recently, progress in artificial intelligence (AI) and machine learning (ML) have developed as promising solutions for detecting and diagnosing DR [5]. ML-based methods trained using a large amount of fundus images have achieved high precision in automated DR classification [6]. Additionally, to enhance the detection performance, significant efforts have been directed toward developing automated approaches, which are practical and affordable, allowing them to be more useful for implementation in resource-constrained environments [7].

Deep learning (DL) is a subdivision of ML, which integrates convolutional neural networks (CNNs) for effective and precise diagnosis of DR [8]. CNNs play a major role in capturing visual features from fundus images to distinguish between normal and abnormal cases [9]. Compared to traditional ML methods, which require manual feature engineering, CNNs automatically capture hierarchical representations from fundus images, extracting low-level textural features as well as high-level pathological features such as hard exudates, neovascularization, and blot hemorrhages [10]. Visual representation is crucial in classifying DR by allowing the model to identify complex variations among various levels of DR. Therefore, well-trained DL-based methods can recognize clinically relevant areas, supporting understandability and decision-making for ophthalmologists [11, 12]. Although the DL-based approaches have advanced DR classification, they still contain some shortcomings. On one hand, the proposed models mainly perform visual feature extraction with traditional CNN-based approaches, which are insufficient in describing the channel-wise importance of discriminative retinal features for differentiating between severity stages of DR, especially when subtle visual differences exist between consecutive stages. On the other hand, traditional DR classification models generally implement conventional classifiers to perform classification using feature extraction-based representations, which fail to capture the complex nonlinear relationships among those retinal features effectively. Furthermore, few works focus on an integrated framework that includes efficient feature extraction, channel-wise feature refinement, and KAN-based classification with optimized parameters. These shortcomings motivate the development of an efficient framework for extracting more discriminative retinal representations and modeling the nonlinear relationships among features for multi-class DR classification. Therefore, develops an Interpretable KAN-Based Framework for Multi-Class Diabetic Retinopathy Classification (IKAN-MDRC). The primary contribution of this research is listed below:

- Integrates Regular network (RegNet) with Efficient channel attention (ECA) for feature extraction, in which RegNet captures hierarchical features related to DR patterns, while ECA learns channel-wise dependencies to enhance the discriminative quality of the extracted features.
- Employs the Efficient Kolmogorov-Arnold Network (Efficient-KAN) that learns complex nonlinear relationships from the extracted retinal features and effectively categorizes severity levels of diabetic retinopathy (DR)
- Incorporates Harris Hawks Optimization (HHO) to select an optimal hyperparameter for the EfficientKAN classifier, thus facilitating faster convergence, improved classification accuracy, and generalization capability.
- Integrates GradCAM++ to provide visualization of the most significant region to support model interpretability.
- Performs experimental evaluation using a diabetic retinopathy detection dataset across various evaluation measures to validate its robustness in accurately categorizing DR severity levels.

2. LITERATURE REVIEW

Recent studies have explored numerous ML techniques for automated DR detection and multi-class severity classification from retinal fundus images. The authors [13] designed a new hybrid DL technique that incorporates recurrent neural networks (RNNs) and CNNs for timely DR recognition and progression surveillance through retinal fundus photographs. By using the successive features of disease progression, the study combines temporal data from many retinal scans to improve identification precision. Sathya and Valaramathi [14] proposed an Agentic-AI-based architecture for DR Analysis, which exploits intelligent agent-oriented learning systems to improve adaptability and contextual interpretability of retinal images. In [15], a hybrid strategy called EffNet-SVM is presented for the detection of DR and no-DR cases utilizing retinal fundus photographs. In EffNet-SVM, EfficientNetV2-Small is employed to learn discriminative representations from the input fundus images, followed by SVM-based classification using an RBF kernel. Abbasi et al. [16] incorporated quantile-driven histogram equalization to enhance the visibility of features in eye images. The article offers a significant advancement in fundus color images and is especially useful for lower-contrast image quality. Nayak and Arkachari [17] examined a new hybrid DL framework that incorporates ResNet50 for hierarchical feature extraction with Capsule Networks for accurate spatial association modeling. This approach aims to overcome these challenges of conventional CNNs, especially their inability to maintain spatial hierarchies and part-whole connections important for precise DR severity grading. The authors [18] A quantum denoising autoencoder was proposed as a hybrid quantum-classical approach that combines convolutional feature encoding with parameterized quantum circuits (PQCs) to improve retinal image denoising while preserving image quality. Sharma et al. [19] designed a DL approach for the automatic grading of disease severity named X-EyeNet. The methodology comprises a separate external-eye screening module for initial detection of abnormalities and a fundus-driven module for DR severity grading. Finally, the severity grade is determined by fundus analysis; for effective feature representation, the fundus is pre-processed, and gross examination is completed. The CORAL further determines the disease grade using a Consistent Rank Logits (CORAL) DL architecture.

Various ML and DL models have been developed to classify diabetic retinopathy (DR), but they still face difficulties in providing accurate and explainable solutions for severity classification. Most of the existing models use traditional feature extraction techniques, which may not adequately capture subtle retinal abnormalities, thereby reducing classification performance. Many previous studies focus mainly on binary classification and achieve better performance for detecting DR than for multi-class severity grading, highlighting the difficulty of distinguishing different DR stages. In addition, conventional classification models may have limited capability to learn complex nonlinear relationships among the extracted retinal features. The lack of effective attention mechanisms can also result in insufficient focus on important retinal regions and disease-related features. Thus, there is a need for a robust framework that can capture discriminative retinal features, emphasize informative feature channels, and effectively classify multiple DR severity levels, which motivates the development of the IKAN-MDRC model.

3. PROPOSED SYSTEM

Fig. 1 demonstrates the overall architecture of the IKAN-MDRC framework for multi-class DR classification. The proposed pipeline begins with circular retinal image cropping, illumination correction, contrast enhancement, resizing, and normalization operations to obtain high-quality and standardized retinal images. Then, the preprocessed images are fed into the RegNet-based feature extractor with the addition of ECA to enhance discriminative channel-wise features for selecting the most informative channels. Following this, the extracted discriminative features are passed into the EfficientKAN classifier to classify different grades of diabetic retinopathy. Finally, the HHO is utilized to optimize relevant parameters in the model to improve its performance in classification.

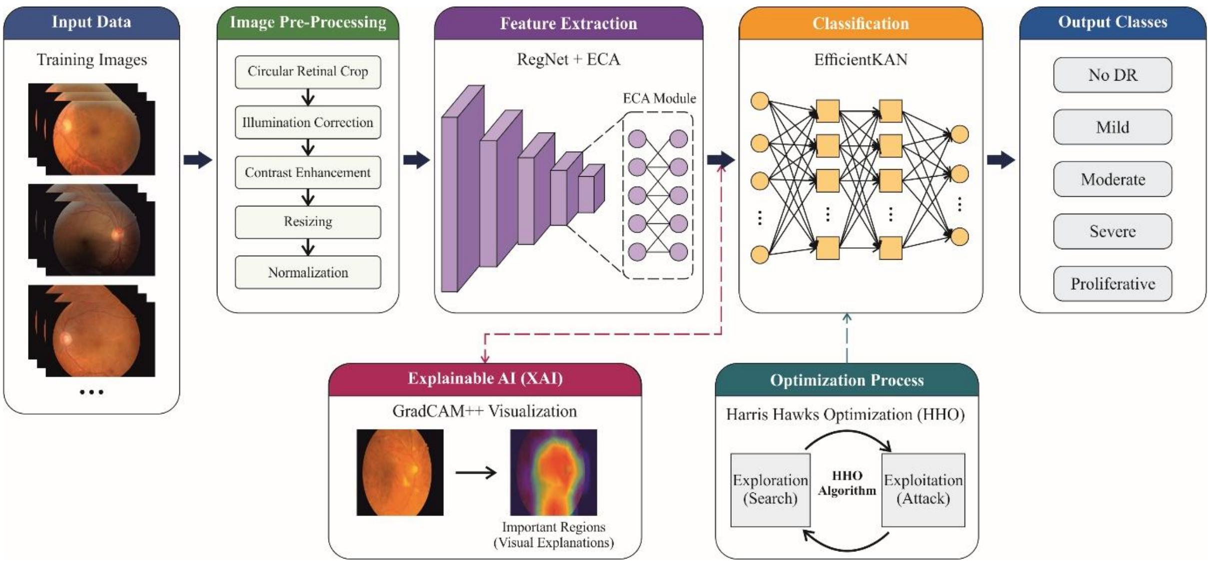

Fig. 1. Proposed IKAN-MDRC framework for multi-class DR classification

3.1. Image Preprocessing

Preprocessing of fundus images has been developed as useful in numerous scenarios. Various preprocessing methods can decrease blur and bias in images and make the lesion more visible, which can be more essential for an effective classification model. To accomplish accurate DR classification, this study uses five diverse preprocessing phases on the raw images as described below. Primarily, the circular retinal region is cropped to eliminate redundant black borders and background regions, thus permitting the model to focus on the significant retinal area. This stage also decreases irrelevant information and increases the visibility of significant lesions and retinal structures. Later, Illumination correction is utilized to decrease the background intensity differences due to unbalanced illumination of retinal images. As a result of this, fundus images may encompass non-uniform illumination throughout the retinal region. Each pixel intensity is computed as follows:

$$I_{corr} = \frac{I\,(u,v)}{B\,(u,v) + \epsilon} \tag{1}$$

From this expression, parameter $B\,(u,v)$ indicates the evaluated illumination component, $I\,(u,v)$ refers to the original image, and the symbol ϵ denotes a small constant applied for numerical stability. It helps to provide more consistent visibility of retinal structures within the images. Based on this, CLAHE is applied to enhance local contrast while reducing unnecessary noise amplification, thus increasing the visibility of subtle DR-related lesions. It enriches intensity variances in local retinal regions while retaining significant structural details. Following that, all the input images have been resized to a constant resolution of 224 x 224, confirming that all fundus images are presented in accurate dimensions that allow robust processing via learning methods. Lastly, min-max normalization is employed to scale pixel intensities to a range of [0,1]. The mathematical expression of min-max normalization is given by:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \tag{2}$$

Here, the term $I_{max} - I_{min}$ denotes its maximal and minimal intensity values correspondingly, and the parameter I represents the intensity of the original pixel. These preprocessing processes offer clean and standardized fundus images for efficient feature extraction and classification methods.

3.2. Feature Extraction Using RegNet-ECA

The pre-processed retinal fundus images include intricate visual patterns related to diverse stages of DR. Hence, an efficient feature extraction model is essential to learn both lower-level retinal regions and higher-level DR-associated features. RegNet integrated with ECA is used to capture distinctive and channel-refined feature representations for further classification.

RegNet-based Feature Extraction

RegNet is efficiently utilized to transform the pre-processed fundus images into small and higher-level feature representations [20]. It is an efficient CNN model, which is generated to automatically process strong network topologies by changing important network parameters and using a group of consistent design rules. The process starts with a first convolutional layer that captures lower-level retinal features, including retinal textures, blood vessels, optic-disc structures, and lesion patterns. Based on a quantized linear rule, RegNet is organized into four major stages, with every stage gradually improving the channel counts:

$$w_i = w_0 + w_a\, i \tag{3}$$

Here, w_i denotes the width of the $i - th$ stage, w_0 implies the initial width, and w_a determines the channel-width expansion. The subsequent width can be quantized as:

$$W_q = q \cdot \left\lfloor \frac{w(i)}{q} \right\rceil \tag{4}$$

In this Eq. (4), q denotes the quantization factor. RegNet applies bottleneck blocks with group convolutions in every step to efficiently capture the distinctive retinal pattern. With increasing network depth, it learns more intricate features associated with exudates, microaneurysms, hemorrhages, and other DR features. Non-linear activation functions and batch normalization further enhance the robustness of the feature representation. The resulting feature mappings are combined utilizing Global Average Pooling (GAP):

$$F_{GAP} = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} F(i,j) \tag{5}$$

Let $F(i,j)$ denote the spatial feature activation. The final higher-level retinal characteristics are passed into the ECA mechanism for channel-wise refinement before EfficientKAN classification.

ECA

Although RegNet extracts hierarchical retinal features, different feature channels may contribute unequally to diabetic retinopathy recognition [21]. Therefore, ECA is incorporated to emphasize informative channels while maintaining low computational overhead. ECA is a lightweight channel-attention module that uses a 1D convolution to model local cross-channel dependencies without applying dimensionality reduction. Given the RegNet feature map, global average pooling first generates a channel descriptor:

$$Z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} F(i,j) \tag{6}$$

Here, $F_c(i,j)$ implies the activation of the c-th channel. The channel descriptor is further processed utilizing an adaptive 1D convolution and sigmoid function:

$$a = \sigma\big(Conv\ 1D_k(Z)\big) \tag{7}$$

Now, k denotes the adaptive kernel dimensions, and σ signifies the sigmoid function. The extracted attention weights are utilized to recalibrate the RegNet feature mappings:

$$F_c' = a_c F_c \tag{8}$$

Let, a_c denote the attention weight allocated to the c-th channel and F_c' represent the refined feature mapping. The adaptive ECA convolution highlights relevant retinal channels while reducing less important characteristics, enhancing the representation of DR-related structures. Consequently, RegNet-ECA integrates hierarchical feature representation with channel-wise attention to learn distinctive retinal characteristics that are passed into EfficientKAN for multi-class DR classification.

3.3. EfficientNET-based Diabetic Retinopathy Severity Classification

EfficientKAN then processes the learned feature representations to capture intricate non-linear interactions and achieve multi-class DR classification [22]. EfficientKAN is efficient for non-linear independence for these challenges since its trainable non-linear functions could model intricate relations while preserving low computational cost. Unlike CNNs that utilize fixed activation functions, KANs offer adaptive non-linear alterations that could enhance representation and classification ability. It is a memory-effective KAN variant developed to tackle the scalability bottlenecks of the original formulations. Every function $\phi_{ij}^{(\ell)}$ is parameterized as a weighted integration of a B-spline expansion and a fixed base activation:

$$\phi_{ij}^{(\ell)}(x) = w_{ij}^{(\ell)} b(x) + s_{ij}^{(\ell)} \sum_{k=1}^{g} \theta_{ijk}^{(\ell)} B_k^{(\ell)}(x) \tag{9}$$

Here, $b(x)$ is composed to enhance convergence and stabilize learning; $B_k^{(\ell)}(x)$ represents the basis functions, and $w_{ij}^{(\ell)}$, $s_{ij}^{(\ell)}$, and $\theta_{ijk}^{(\ell)}$ denote the trainable parameters. Even though this per-edge B-spline method flexibly captures non-liner relationship but improves memory and computational costs due to distinct transformations for all connections. In this method, the L1 and entropy regularization are estimated utilizing the average absolute spline, with each edge activity (i,j) in layer ℓ approximated as:

$$\bar{\theta}_{ij}^{(\ell)} = \frac{1}{g}\sum_{k=1}^{g} |\, \theta_{ijk}^{(\ell)}| \quad (10)$$

and the overall activity in the layer has been expressed by:

$$\bar{\Theta}^{(\ell)} = \sum_{i=1}^{n_\ell}\sum_{j=1}^{n_{\ell-1}} \bar{\theta} \quad (11)$$

The entropy penalty was further measured as:

$$S(\Phi^{(\ell)}) = -\sum_{i=1}^{n_\ell}\sum_{j=1}^{n_{f-1}} \frac{\bar{\theta}_{ij}^{(\ell)}}{\bar{\Theta}^{(\ell)}} \log\left(\frac{\bar{\theta}_{ij}^{(\ell)}}{\bar{\Theta}^{(\ell)}}\right) \quad (12)$$

The activity and entropy terms boost a compact and uniform set of trainable spline representations, helping efficient non-linear classification while preserving the computational complexity of EfficientKAN. The hyperparameters of EfficientKAN are further enhanced by employing HHO to attain a suitable configuration for effective multi-class DR classification. Therefore, the EfficientKAN classifier offers an effective non-linear decision module for mapping the RegNet-ECA feature extraction to the DR severity categories.

3.4. Harris Hawks Optimization Strategy

The selection of appropriate hyperparameters is vital for capturing optimal performance from the EfficientKAN classifier that enables consistent multi-class DR severity classification. Therefore, HHO is used to identify the optimal parameter setting by mimicking the cooperative hunting strategy of the Harris Hawk (HH) population employing diverse approaches for searching and feeding on prey, including exploration and exploitation stages, with transitions among the dual stages depending on variations in prey escape energy [23].

Exploration stage: In this stage, HH search diverse areas of the solution space to detect viable candidate values. The prey position and other random selection update every hawk's position. The position update rule is computed as:

$$P(t+1) = \begin{cases} P_{rand}(t) - r_1|P_{rand}(t) - 2r_2P(t)|, \; q \geq 0.5 \\ P_{rabbit}(t) - P_m(t) - r_3\big(LB + r_4(UB - LB)\big), \; q < 0.5 \end{cases} \quad (13)$$

Here, $P(t)$ implies the hawk's recent position, $P_{rabbit}(t)$ denotes the better solution, $P_{rand}(t)$ represents the hawk's random selection, $P_m(t)$ signifies the mean position, LB and UB represent the minimal and maximal boundaries of the search space, $r_1, r_2, r_3, r_4 \in [0,1]$ and are random solutions. In the presented system, the sizes of P denote the selected classifier hyperparameters, containing learning rate, regularization coefficient, dropout rate, and hidden-layer (HL) parameters.

Transition stage from exploration to exploitation: The change from exploring an area to catching food relies on the prey's escape effort E, is expressed as:

$$E = 2E_0\left(1 - \frac{t}{T}\right) \quad (14)$$

Here, T is the maximum iteration count, E_0 and is the initial energy. The decreasing escape energy shifts the search from exploration to exploitation.

Exploitation stage: Let r denote the probability of prey escape, which is determined by the probability of the prey's escape energy; the HHs adapt 4 corresponding hunting behaviors.

Soft siege – For prey with high energy and limited escape capability, the hawk minimizes the distance to the prey. The new position is represented as:

$$P(t+1) = \Delta P(t) - E \mid JP_{rabbit}(t) - P(t) \mid \quad (15)$$

Hard siege- With a prey's energy and escape limited, the hawks make a direct attack.

$$P(t+1) = P_{rabbit}(t) - E \mid \Delta P(t) \mid \tag{16}$$

Soft siege with progressive fast swooping- For prey with higher energy and a greater chance to escape, HHO applies a soft encirclement approach in the later phases. The candidate value is set as:

$$Y = P_{rabbit}(t) - E \mid JP_{rabbit}(t) - P(t) \mid \tag{17}$$

The rapid-dive solution based on a Lévy-flight classification is expressed as:

$$Z = Y + S \times LF(D) \tag{18}$$

Let S denote a random vector, and $LF(D)$ denotes the Lévy-flight function.

Hard siege with progressive fast swooping: With prey's limited energy but high chance to escape, HHO uses tight encirclement with rapid diving movement. The final candidate solution is expressed as:

$$Y = P_{rabbit}(t) - E|JP_{rabbit}(t) - P_m(t)| \tag{19}$$

In this Eq. (19), $P_m(t)$ implies the mean hawk position. The final candidate solution is further refined by employing the gradual rapid-dive function.

Fitness Function: The quality of every candidate EfficientKAN hyperparameter method is assessed using a fitness function (FF) that depends on classification outcomes. The FF is expressed as:

$$F(P) = 1 - Accuracy_{val}(P) \tag{20}$$

In this Eq. (20), P implies a candidate EfficientKAN hyperparameter model, and $Accuracy_{val}$ denotes the validation accuracy. HHO reduces $F(P)$ to detect the pattern that provides a higher validation classification performance. The HHO-based optimization proceeds until the predefined termination criteria are met, after which the optimal hyperparameter method has been selected for EfficientKAN. An improved EfficientKAN classifier is then used for the multi-class DR classification tasks.

3.5. GradCAM++ Explainability

As DR severity is indicated by subtle retinal lesions that occupy only smaller areas of a fundus image, explaining the model's decision is vital to confirm whether the learned representation concentrates on diagnostically significant regions. Thus, Grad-CAM++ is integrated into the presented system, which is a modified version of the Grad-CAM model [24]. This enhancement helps visualize and interpret the decisions computed by DL models, emphasizing the retinal regions that are most significant for the classification decision. Particularly, Grad-CAM emphasizes a class-relevant areas by analyzing the gradients of the target class through convolutional layer's feature maps. It enhances this process by treating high-order gradients and the importance of specific spatial positions, facilitating more precise localization of fewer or more distinctive areas. It is beneficial for DR classification, while relevant retinal abnormalities could occupy small regions of the fundus image. Assume that y^c implies the output of the network for class c, and A^k represents the final convolution layer in the k- th feature map. A^k are then classified as A^k_{ij}; let (i, j) indicate the particular position on the feature maps. L^c is the heat map that will classify the most relevant areas for class c and is described for all points (i, j) as below:

$$L^c_{ij} = \sum_k w^c_k A^k_{ij} \tag{21}$$

where w^c_k denotes the weights for every k- th feature map. In the XAI model, these w^c_k weights are calculated utilizing the spatial contribution of the gradients and high-order derivatives as below:

$$w^c_k = \sum_i \sum_j a^{kc}_{ij} \cdot relu\left(\frac{\partial y^c}{\partial A^k_{ij}}\right) \tag{22}$$

Lastly, a^{kc}_{ij} denotes the significance weights related to the spatial position (i, j) of the $k - th$ feature mapping for class c. The weight is measured by utilizing the 2nd- and 3rd-order derivatives as:

$$a_{ij}^{kc} = \frac{\frac{\partial^2 y^c}{(\partial A_{ij}^k)^2}}{2 \cdot \frac{\partial^2 y^c}{(\partial A_{ij}^k)^2} + \sum_{q,b} A_{ab}^k \cdot \frac{\partial^3 y^c}{(\partial A_{ab}^k)^3}} \tag{23}$$

Now, c implies DR severity classes. The final Grad-CAM++ heatmap is produced from the convolutional feature mappings of the RegNet feature extraction and overlapped on the original retinal fundus image to emphasize the regions that contribute to the forecasted DR categories. Grad-CAM++ enhances the model transparency by visually explaining the classification decisions for the proposed system.

4. EXPERIMENTAL RESULTS AND DISCUSSION

This part offers the experimental results of the IKAN-MDRC system using the Diabetic Retinopathy Detection dataset [25], which comprises high-resolution retinal fundus images gathered under different imaging conditions. It is classified into 5 severity classes: No DR, Mild, Moderate, Severe, and Proliferative DR. The methodology should examine retinal images and precisely predict the corresponding DR grade while accounting for variations caused by distinct camera types, imaging settings, and anatomical orientations. Meanwhile, images may vary in appearance because of left/right eye views and inverted microscope lens representations utilized in medical analysis. As depicted in Table 1, the dataset consists of five severity classes containing a total of 35126 instances. Fig. 2 shows the sample retinal fundus images.

Table 1 Details of the dataset

Class	**Count**
No DR	25810
Mild	2443
Moderate	5292
Severe	873
Proliferative DR	708
Total count	**35126**

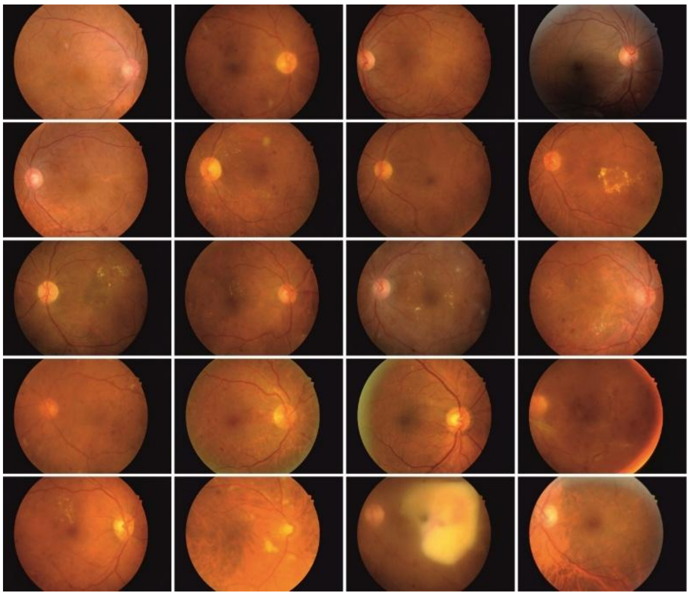

Fig. 2. Sample retinal fundus images from the DR dataset

Fig. 3 offers the confusion matrices of the IKAN-MDRC method with an 80:20 train-test split. Both confusion matrices show higher values along the diagonal, signifying the effective classification of DR severity levels for each image sample.

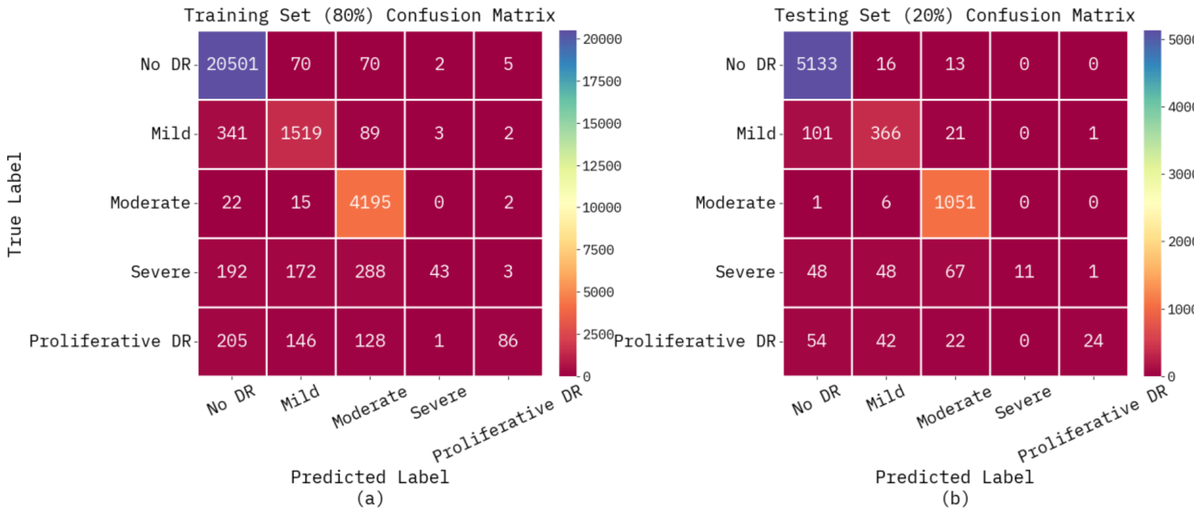

Fig. 3. Confusion matrices of the IKAN-MDRC approach with (a-b) an 80:20 train-test split

The lower off-diagonal values display misclassifications between diverse severity categories, indicating the effectiveness and robustness of the IKAN-MDRC method for the DR severity classification task.

Table 2 shows the DR severity classification outcomes of the IKAN-MDRC method with an 80:20 training/testing split. The proposed IKAN-MDRC approach achieves average performance with accuracy of 97.50% and 97.49% for the training and testing sets, demonstrating its high consistency and ability for accurate DR severity level classification.

Table 2 DR severity classification results of the IKAN-MDRC method with an 80:20 split

Class	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (80%)					
No DR	96.77	96.43	99.29	97.84	91.65
Mild	97.02	79.03	77.74	78.38	76.78
Moderate	97.81	87.95	99.08	93.18	92.12
Severe	97.65	87.76	6.16	11.51	22.90
Proliferative DR	98.25	87.76	15.19	25.90	36.11
Average	**97.50**	**87.78**	**59.49**	**61.36**	**63.91**
Testing Set (20%)					
No DR	96.68	96.18	99.44	97.78	91.43
Mild	96.66	76.57	74.85	75.70	73.91
Moderate	98.15	89.52	99.34	94.18	93.26
Severe	97.67	100.00	6.29	11.83	24.78
Proliferative DR	98.29	92.31	16.90	28.57	39.10
Average	**97.49**	**90.92**	**59.36**	**61.61**	**64.50**

Fig. 4 exhibits the classification performance of the IKAN-MDRC methodology with ROC and PR curves. Fig. 4a implies a robust discriminative outcome across all five DR severity categories, with AUC scores ranging from 0.902 to 0.996. The Adequate DR class obtains the maximum AUC of 0.996, signifying superior class discrimination. Moreover, Fig. 4b illustrates the PR results, with AP scores ranging from 0.563 to 0.987. The No DR class got the highest AP of 0.987, whereas the lower AP values for Severe and Proliferative DR indicate similarly problematic classification.

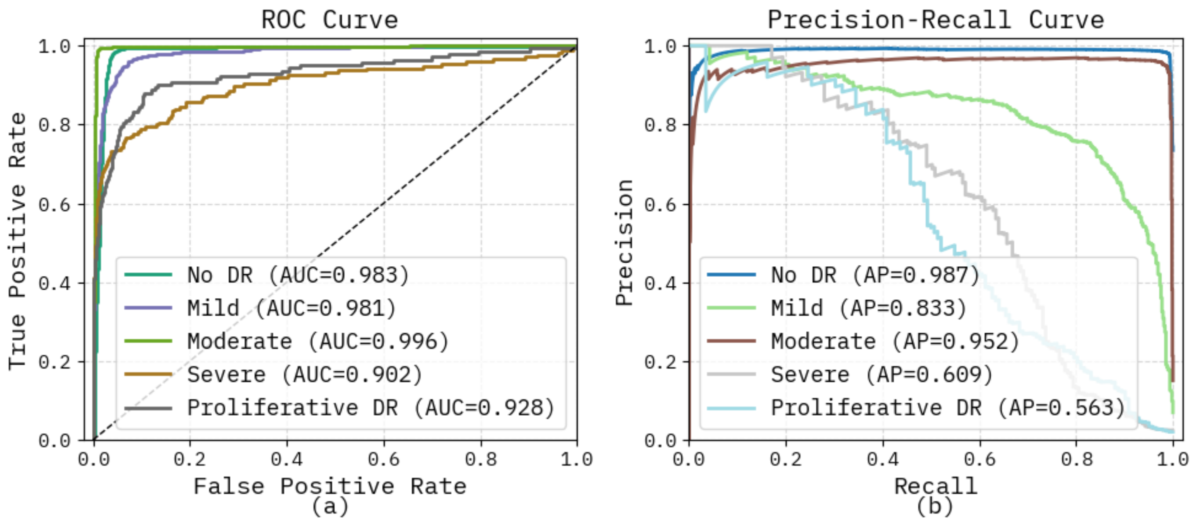

Fig. 4. Classification result of the IKAN-MDRC approach with (a-b) ROC/PR curves

Fig. 5 illustrates the confusion matrices of the IKAN-MDRC system on 70:30 train-test split. Both confusion matrices indicate higher diagonal values, implying the effective detection of DR severity levels, with few samples misclassified among distinct severity classes, demonstrating the efficiency and robustness of the IKAN-MDRC algorithm for the DR severity detection task.

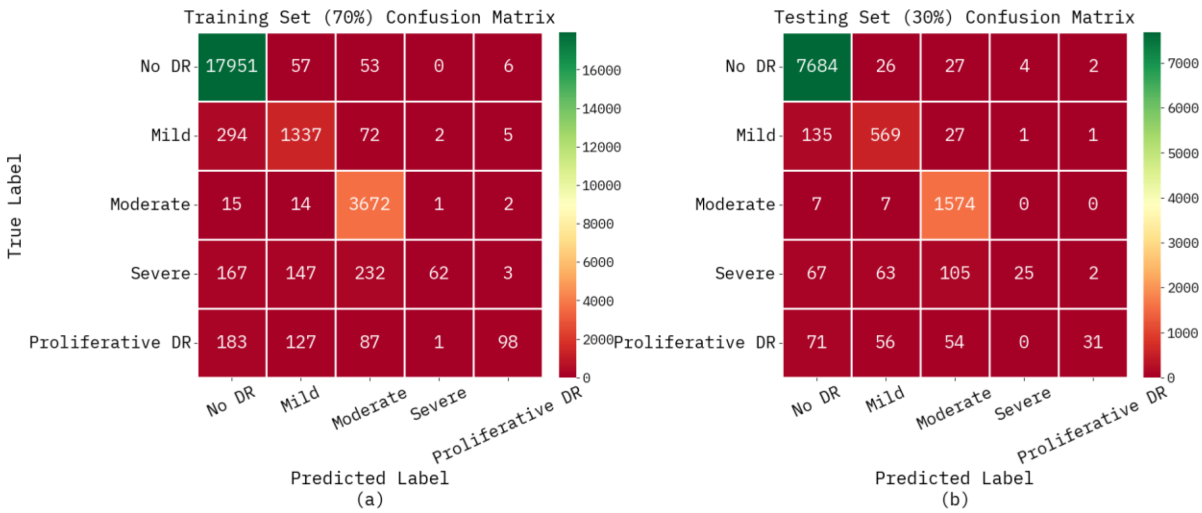

Fig. 5. Confusion matrices of the IKAN-MDRC approach with (a-b) 70:30 train-test split

Table 3 depicts the DR severity recognition results of the IKAN-MDRC system with a 70:30 training/testing split. The proposed IKAN-MDRC technique obtained an average accuracy of 97.61% and 97.51% for the training and testing sets, signifying its higher reliability and capability for precise DR severity level detection.

Natural Resources for Human Health
eISSN: 2583-1194

Table 3 DR severity classification results of the IKAN-MDRC method with a 70:30 split

Class	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (70%)					
No DR	96.85	96.46	99.36	97.89	91.85
Mild	97.08	79.49	78.19	78.83	77.27
Moderate	98.06	89.21	99.14	93.91	92.95
Severe	97.75	93.94	10.15	18.32	30.48
Proliferative DR	98.32	85.96	19.76	32.13	40.75
Average	**97.61**	**89.01**	**61.32**	**64.22**	**66.66**
Testing Set (30%)					
No DR	96.78	96.48	99.24	97.84	91.67
Mild	97.00	78.92	77.63	78.27	76.66
Moderate	97.85	88.08	99.12	93.27	92.22
Severe	97.70	83.33	9.54	17.12	27.74
Proliferative DR	98.23	86.11	14.62	25.00	35.07
Average	**97.51**	**86.59**	**60.03**	**62.30**	**64.67**

Fig. 6 reveals the classification outcomes of the IKAN-MDRC model with ROC and PR curves. Fig. 6a presents effective discriminative performance across each of the 5 DR severity categories, with AUC values between 0.929 and 0.994. The Moderate DR class reaches a maximal AUC of 0.994, representing exceptional class distinction. Moreover, Fig. 6b exhibits the PR outcomes, with AP values between 0.544 and 0.992. The No DR class attains a better AP of 0.992, while the minimal AP values for Severe and Proliferative DR show a moderate classification performance.

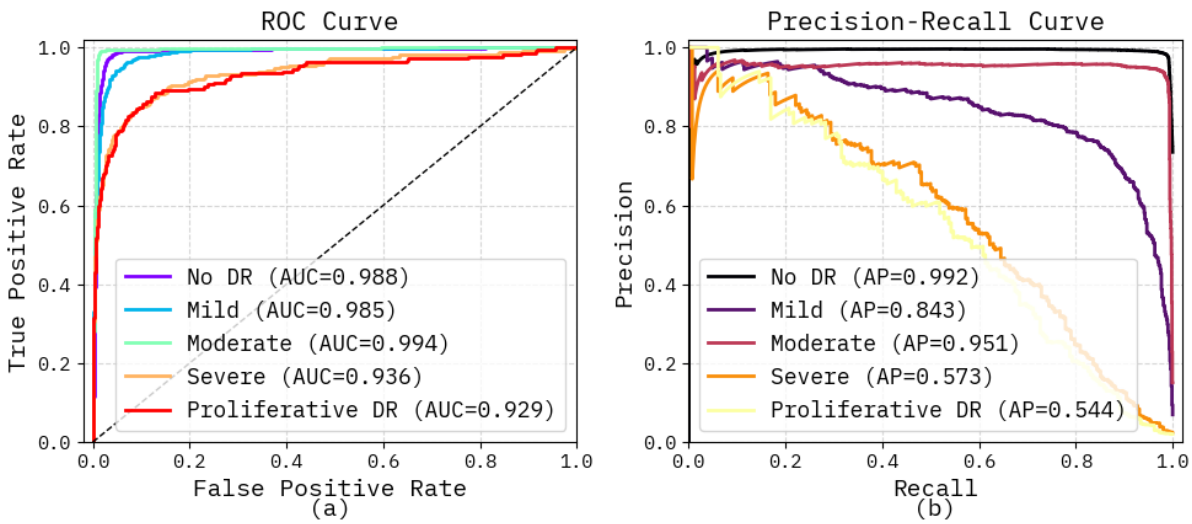

Fig. 6. Classifier results of the IKAN-MDRC approach with (a-b) ROC/PR curves

Table 4 and Fig. 7 depict the comparative results of the IKAN-MDRC framework with existing methods using evaluation metrics [26, 27]. The findings imply that the ViT CNN and MPDCNN methods attain lower performance, while the Deep

Residual CNN and MVDRNet models gain slightly enhanced results. Likewise, the DLGP-DR, ViT-MAE, and Xception approaches have reached moderate and reasonable performance. On the other hand, the proposed XCV2-GTN method has achieved superior performance with an accuracy of 97.61%, precision of 89.01%, recall of 61.32%, and an F1-Score of 64.22%, providing its efficiency for DR severity level detection.

Table 4 Comparison outcomes of the IKAN-MDRC method with other approaches

Methods	Accuracy	Precision	Recall	F1-Score
MVDRNet [26]	89.71	78.42	52.18	61.24
Deep Residual CNN [26]	87.20	75.63	48.76	58.01
MPDCNN [26]	86.81	73.85	47.32	56.82
DLGP-DR [27]	93.23	84.76	57.84	59.21
Xception [27]	94.00	86.32	59.18	60.12
ViT CNN [27]	82.50	70.84	43.27	53.74
ViT-MAE [27]	93.42	87.16	59.47	61.08
IKAN-MDRC	**97.61**	**89.01**	**61.32**	**64.22**

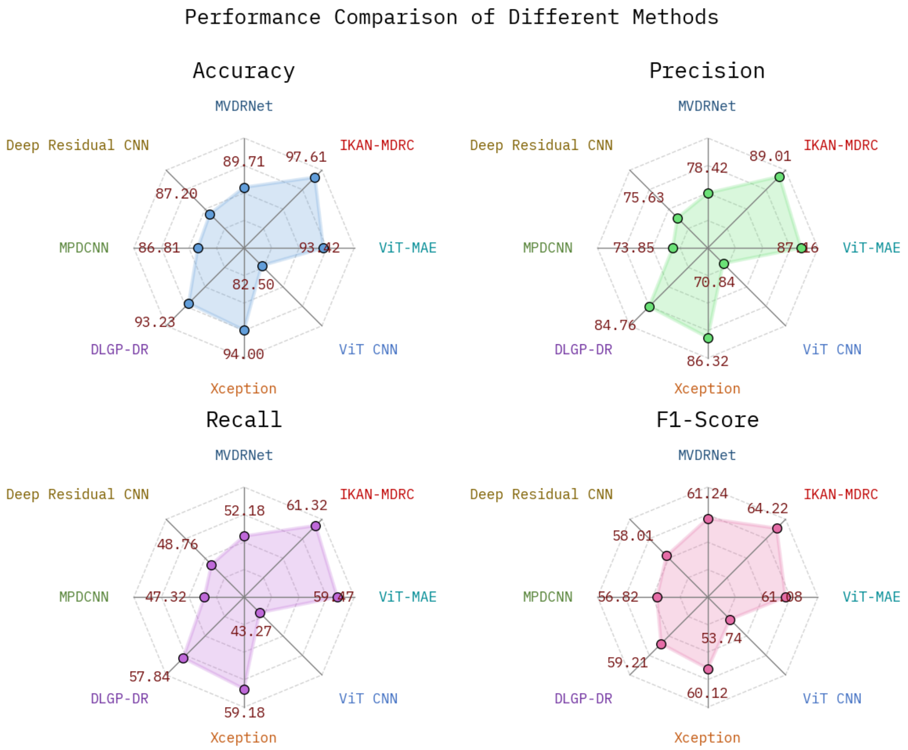

Fig. 7. Performance comparison of the IKAN-MDRC method with metrics

Table 5 shows the ablation analysis of the IKAN-MDRC approach with several metrics. The outcome implies that the baseline EfficientKAN approach obtains fewer solutions among all configurations. The integration of RegNet and EfficientKAN algorithms progressively improves feature representation and classification capability. Moreover, incorporating the RegNet + ECA + EfficientKAN improves the extraction of significant channel-wise characteristics and enhances classification performance. The complete IKAN-MDRC (RegNet + ECA + EfficientKAN + HHO) approach achieves higher performance with an accuracy of 97.61%, precision of 89.01%, recall of 61.32%, and F1-Score of 64.22%, confirming the efficiency of every incorporated component for DR severity level detection.

Table 5 Ablation study of the XCV2-GTN method with measures

Methods	Accuracy	Precision	Recall	F1-Score
EfficientKAN	91.42	78.16	48.73	58.82
RegNet + EfficientKAN	93.68	82.47	53.84	64.01
RegNet + ECA + EfficientKAN	95.24	86.15	57.92	69.18
IKAN-MDRC (RegNet + ECA + EfficientKAN + HHO)	**97.61**	**89.01**	**61.32**	**64.22**

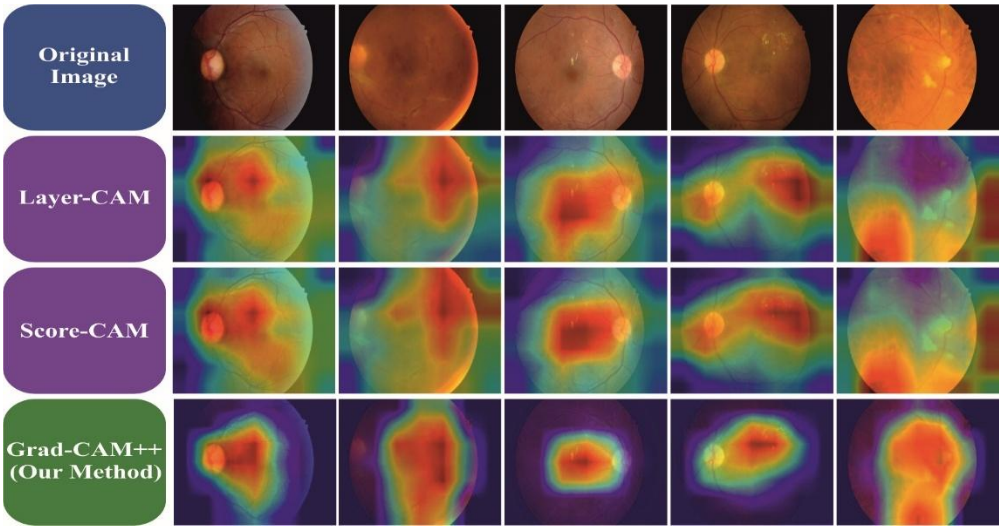

Fig. 8. Visual comparison of different class activation mapping methods for retinal images

Fig. 8 represents the original retinal images are depicted along the highlighted regions generated by Layer-CAM, Score-CAM, and the proposed Grad-CAM++ methodologies. Layer-CAM and Score-CAM recognizes relevant regions but expose wider and more scattered activation patterns under few images. When compared to, Grad-CAM++ generates more focused heatmaps that are strongly associated with the visually significant retinal areas. Hence the proposed system offers clearer and more localized visual explanations, which make the emphasized regions easily interpreted.

The experimental results show the robustness of the IKAN-MDRC system for the classification of DR severity. The preprocessing tasks improve the quality of fundus images, hence allowing the RegNet with ECA to learn more discriminative retinal features, which facilitates distinguishing between different levels of severity of DR. Meanwhile, the EfficientKAN classifier is designed to transform the selected features into a higher dimension so that the distance between features that correspond to different diabetic retinopathy severity levels can be expanded. Furthermore, the HHO is designed to optimize the parameters in the model to ensure better performance and stable classification. Thus, the metrics performance, as well as the confusion matrix and per-class performance analysis, indicate that the proposed model IKAN-MDRC can effectively differentiate between severity levels of DR. In brief, it is reasonable to integrate efficient feature extraction, a channel attention mechanism, a KAN-based classifier, and an optimizer in a hybrid model for automated DR grade classification.

5. CONCLUSION AND FUTURE WORK

In this paper, an IKAN-MDRC is proposed for the automated grading of DR from fundus retinal images. The IKAN-MDRC framework comprises multiple stages: circular retinal cropping, illumination correction, contrast enhancement, resizing, and normalization to improve the quality of the input images; RegNet with an ECA network to extract discriminative retinal representations; EfficientKAN classifier to enhance nonlinear feature transformation for multi-class DR classification; and HHO for further optimization of model parameters. Experiments are carried out on the dataset to demonstrate the performance of the proposed method compared with other state-of-the-art works. The overall results show the efficiency of the IKAN-MDRC approach in assisting in distinguishing and grading DR images at different severity levels. Furthermore, a synergistic combination of an efficient feature extraction network, channel attention, KAN-based classification, and an optimizer is employed in this method, creating a high-performing automatic DR grade classification approach. In future work, this method could be extended to incorporate larger multi-center datasets and lesion-based diabetic retinopathy classification. Additionally, more advanced XAI techniques could be included to help clinicians better understand and interpret IKAN-MDRC's performance

Data Availability Statement: The data that support the findings of this study are openly available in Kaggle repository at https://www.kaggle.com/competitions/diabetic-retinopathy-detection/data, reference number [25].

REFERENCES

[1] Datta, H. and Bhandari, A.K., 2026. Neural search guided CNN–transformer for diabetic retinopathy prediction from fundus images. Optics & Laser Technology, 198, p.114944.

[2] Akhtar, S., Aftab, S., Ali, O., Ahmad, M., Khan, M.A., Abbas, S. and Ghazal, T.M., 2025. A deep learning based model for diabetic retinopathy grading. Scientific reports, 15(1), p.3763.

[3] Hariobulesu, P. and Shaik, F., 2025. Enhanced multi-grade diabetic retinopathy detection and classification via ensembled deep learning model from retinal fundus images. Expert Systems with Applications, 285, p.128116.

[4] Aftab, S. and Akhtar, S., 2025. Diabetic retinopathy severity classification using data fusion and ensemble transfer learning. Journal of Software Engineering and Applications, 18(1), pp.1-23.

[5] Ikram, A. and Imran, A., 2025. ResViT FusionNet Model: An explainable AI-driven approach for automated grading of diabetic retinopathy in retinal images. Computers in biology and medicine, 186, p.109656.

[6] Kuruba, C. and Gopalan, N.P., 2025. 3DECNN: a novel method for segmentation of the diabetic retinopathy in retinal fundus images using 3D-edge CNN. Neural Computing and Applications, 37(21), pp.16187-16201.

[7] Gupta, I.K., Patil, S., Kotecha, K., Srivastava, S., Farouk, A. and Rodrigues, J.J., 2026. An ensemble of deep learning models with falcon optimization assisted diabetic retinopathy diagnosis on retinal fundus images. Scientific Reports.

[8] Sheejakumari, V., Sundravadivelu, K., Pushparani, S., Senthilvel, P.G., Balasubramaniam, S., Shah, M.A., Alqahtani, M.S. and Abbas, M., 2025. Diabetic retinopathy detection based on mobile maxout network and weber local descriptor feature selection using retinal fundus image. Scientific Reports, 15(1), p.13858.

[9] Minarno, A.E., Bagaskara, A.D., Bimantoro, F. and Suharso, W., 2025. Classification of diabetic retinopathy based on fundus image using inceptionv3. JOIV: International Journal on Informatics Visualization, 9(1), pp.23-28.

[10] Nandanapu, R., Krishna, G.S.V.V. and Reddy, V.S.S., 2025, February. Predicting Sevearity of Type 2 Diabetic Retinopathy Using Retinal Fundus Images with new DeepLearning Architectures from the Inception and Resnet Families. In 2025 International Conference on Artificial Intelligence and Data Engineering (AIDE) (pp. 550-555). IEEE.

[11] Pamungkas, Y., Triandini, E., Yunanto, W. and Thwe, Y., 2025. Enhancing diabetic retinopathy classification in fundus images using CNN architectures and oversampling technique. Journal of Robotics and Control (JRC), 6(1), pp.413-425.

[12] Zacharowicz, R., Men, Y., Nakayama, L. and Behar, J.A., 2026. Multisource domain training with retinal expert-in-the-loop for accurate diabetic retinopathy staging from fundus images. Machine Learning: Health, 2(1), p.015010.

[13] Sushith, M., Sathiya, A., Kalaipoonguzhali, V. and Sathya, V., 2025. A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images. Scientific Reports, 15(1), p.15166.

[14] Sathya, R. and Valaramathi, A., 2025. Detection and diagnosis of diabetic retinopathy in retinal fundus images using agentic AI approaches. Scientific Reports, 16(1), p.3780.

[15] Naveen, K.V., Anoop, B.N., Siju, K.S., Kar, M.K. and Venugopal, V., 2025. EffNet-SVM: A hybrid model for diabetic retinopathy classification using retinal fundus images. IEEE Access, 13, pp.79793-79804.

[16] Abbasi, R., Amin, F., Alabrah, A., Choi, G.S., Khan, S., Bin Heyat, M.B., Iqbal, M.S. and Chen, H., 2025. Diabetic retinopathy detection using adaptive deep convolutional neural networks on fundus images. Scientific Reports, 15(1), p.24647.

[17] Nayak, S. and Arkachari, M., 2025, November. Deep Learning for Early Detection of Diabetic Retinopathy Using Retinal Fundus Images. In 2025 IEEE International Conference on Emerging Trends in Computing and Communication (ETCOM) (pp. 1-6). IEEE.
[18] Chilukuri, R., P, P., Gatla, R.K. and Almenweer, R.A., 2026. Quantum denoising autoencoder improves retinal fundus image quality for early diabetic retinopathy screening. Scientific Reports, 16(1), p.5970.
[19] Sharma, A., Sai, S.S. and Vidhya, S., 2026, April. X-EyeNet: An AI-based System for Diabetic Retinopathy Severity Grading using Retinal Fundus Images. In 2026 9th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 2083-2089). IEEE.
[20] Sathiamoorthy, J. and Usha, M., 2025. Strawberry Leaf Disease Detection using Deep Learning based RegNet. International Journal of Data Science and Artificial Intelligence, 3(01), pp.24-32.
[21] Yutra, A.Z., Zheng, J., Li, X. and Endris, A., 2023, December. Skinaacn: An efficient skin lesion classification based on attention augmented convnext with hybrid loss function. In Proceedings of the 2023 7th International Conference on Computer Science and Artificial Intelligence (pp. 295-300).
[22] Villagómez, E.L. and Mahalec, V., 2025. Comparative Evaluation of Kolmogorov-Arnold Autoencoders and Orthogonal Autoencoders for Fault Detection with Varying Training Set Sizes. arXiv preprint arXiv:2508.02860.
[23] Zhang, Q., Sheng, J., Zhang, Q., Yang, Z., Xin, Y., Wang, B., Zhang, R. and Alzheimer's Disease Neuroimaging Initiative, 2025. Optimizing deep learning with improved Harris Hawks optimization for Alzheimer's disease detection. Artificial Intelligence Review, 58(10), p.301.
[24] Dadoulis, G.I., Manolis, G.D. and Katakalos, K., BRIDGING ANALYTICAL AND EXPERIMENTAL DOMAINS WITH EXPLAINABLE AI: GRAD-CAM++ FOR DAMAGE DETECTION IN LIGHTWEIGHT STRUCTURES UNDER MOVING MASSES.
[25] https://www.kaggle.com/competitions/diabetic-retinopathy-detection/data
[26] Muthusamy, D. and Palani, P., 2024. Deep learning model using classification for diabetic retinopathy detection: an overview: Muthusamy and Palani. *Artificial Intelligence Review*, *57*(7), p.185.
[27] Touati, M., Touati, R., Nana, L., Benzarti, F. and Ben Yahia, S., 2025. DRCCT: enhancing diabetic retinopathy classification with a compact convolutional transformer. *Big Data and Cognitive Computing*, *9*(1), p.9.