RESEARCH PAPER

Adaptive Multi-Scale Retinal Feature Optimization and Intelligent Hybrid Deep Learning Framework for Diabetic Retinopathy Detection

Ssandhyapogu 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.
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 is a retinal condition associated with diabetes which affects the blood vessels in the retina and is a potential cause of permanent vision loss if not diagnosed early in the disease. Automated Retinal Image Analysis (ARIA) has been identified as an effective method for enhancing disease grading and decreasing reliance on manual clinical assessment, with the use of deep learning techniques. The traditional classification methods of DR, however, are limited by the visibility of lesions, the number of features, the complexity of computational labour and the low accuracy of the classification, due to different retinal image environments. To overcome these limitations, this paper introduces an integrated diabetic retinopathy grading framework which is based on Adaptive Retinal Image Enhancement Algorithm (ARIEA), Multi-Scale Attention Feature Extraction Algorithm (MSAFEA), and Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA). The proposed ARIEA technique enhances the image quality for the retina to preserve clinically meaningful structures by correcting illumination, enhancing contrast, normalizing, and removing noise from retinal images. MSAFEA extracts lesions from the retina at multiple scales with attention-guided deep feature learning to accurately represent spatial information for retinal vessels, microaneurysms, haemorrhages and exudates. We use feature fusion and optimization techniques to remove redundant information and improve discriminative feature mapping. In EDRCA we apply multi-class diabetic retinopathy severity grading by attention-assisted convolutional learning and adaptive decision refinement. The experimental analysis shows that the proposed framework generates better classification accuracy, feature extraction capability and grading performance in the automated diagnosis of DR.

Keywords: Diabetic Retinopathy, Retinal Fundus Images, Deep Learning, Image Enhancement, Multi-Scale Feature Extraction, Attention Mechanism, Feature Fusion, Retinal Lesion Detection, Diabetic Retinopathy Grading, Convolutional Neural Network (CNN), Automated Disease Classification, Medical Image Analysis.

How to cite this article: Priya Nandini S, Nair SA, Sanal Kumar KP. Adaptive multi-scale retinal feature optimization and intelligent hybrid deep learning framework for diabetic retinopathy detection. Int J Drug Deliv Technol. 2026;16(77s): 1-12. DOI: 10.25258/ijddt.16.77s.1

1.Introduction

Diabetic retinopathy (DR) is a serious condition of the retina that is a leading cause of blindness and is a condition of the retina caused by long-term diabetes. Proper and timely detection of DR can save a person from permanent loss of vision. The diagnosis of traditional manual diagnosis of retinal fundus image is subjective and time-consuming, and it is a skill that requires experience. In the recent years, several automated DR grading systems based on deep learning have been proposed to overcome these limitations. To enhance lesion representation and classification accuracy, an adaptive multi-scale feature fusion network named AMSFuse was proposed to fuse features extracted at various scales [1]. Multi-scale feature extraction, adaptive fusion and final DR classification are included in the workflow. But the more feature branches that are integrated, the more compute complexity it will add. To address this issue, a novel adaptive multi-scale framework was developed to capture both local and global retinal features to

achieve fine-grained DR grading via adaptive attention mechanisms [2].

The workflow includes preprocessing of the retinal image, extracting features at various scales, adaptive feature aggregation and prediction of the grade. Although the framework achieves high accuracy, it has high requirements for training data and high calculation costs. A multi-scale residual network integrated with a cross-attention module was proposed to improve the classification accuracy of lesions localization and grade, [3]. The process involves residual feature extraction, cross-attention feature refinement, and classification. The attention mechanism enhances feature discrimination, but results in more training time and memory usage. To address this problem, a hybrid deep learning system based on the randomisation method was proposed to improve the detection rate of DR, combining various deep learning strategies and randomisation methods [4]. It is based on the following workflow: Pre-processing, hybrid feature extraction, random learning, classification. But the hybrid structure has a problem of less interpretability and complexity when it is implemented. Furthermore, a hybrid Vision Transformer (ViT) model was introduced, called LSD-HybridViT, which combines the efficient DR grading features of a lightweight mixed-domain attention and frequency multi-scale dilated convolution model [5].

The workflow consists of convolutional feature extraction, transformer-based attention learning, multi-scale feature fusion, and grading prediction. While the model offers benefits in terms of improved grading accuracy and better representation of features, Vision Transformer architectures overall require high computational requirements and large-scale data sets to train the model. The methods overall indicate that multi-scale learning, attention mechanisms, and hybrid deep learning architectures can have a significant impact on improving automated DR diagnosis, and that challenges such as computational complexity, scalability, and dataset dependency will need to be addressed.

Objectives of the Study

The specific objectives of this research are as follows:

1. To enhance the retinal fundus image with Adaptive Retinal Image Enhancement Algorithm (ARIEA).
2. To obtain the key features of lesions in the retina using Multi-Scale Attention Feature Extraction Algorithm (MSAFEA).
3. To determine the grade of severity of DR with the help of Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA).
4. To integrate image enhancement, feature extraction and classification in an automated DR diagnosis system.
5. To enhance the accuracy of diabetic retinopathy detection and decrease misclassifications during grading of the retina disease.
6. To test the proposed system with the performance metrics like accuracy, precision, recall, sensitivity, specificity and F1 score.

2. Literature Survey

The recent developments in automated diabetic retinopathy (DR) diagnosis have been cantered on developing state-of-the-art deep learning architectures to boost classification accuracy, detect lesions, segment them efficiently, and reduce computational complexity. To boost the performance of feature extraction and classification, the use of attention-guided multi-scale residual network with Bayesian hyperparameter optimization was proposed for DR classification [6]. The workflow consists of retinal image pre-processing, multi-scale residual feature extraction, attention enhancement, parameter optimization by Bayesian, and classification. But the architecture makes the computations more complex and demands a lot of tuning to achieve optimum performance. A novel multi-scale feature extraction method was applied to develop a real time DR screening system for better identification of retinal lesions in fundus images [7].

The workflow applies the following steps: preprocessing, multi-scale feature extraction, feature fusion, and DR grading. The system can be used for real-time analysis, but the accuracy of the analysis under different image quality conditions is still difficult. In Optical Coherence Tomography (OCT) B-scans, a Multi-Scale Adaptive Fusion Network (MAFN) was developed for segmentation of retinal layers and fluids [8]. This framework incorporates adaptive fusion techniques and multi-scale segmentation to enhance boundary detection and smooth localization. It is composed of the following workflow: OCT preprocessing, multi-scale feature learning, adaptive fusion, and segmentation output generation. But large annotated datasets and high computation resources are needed for segmentation process. A thorough study on optimization techniques for deep learning and machine learning in the diagnosis of diabetic macular enema (DME) revealed the significance of optimization techniques like Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Bayesian Optimization to enhance predictive models associated with DR [9]. Optimization techniques increase the efficiency of the model, but

many methods are complex to train are unstable in converging.

A joint analysis framework for DR and glaucoma was proposed that combines deep learning with shared feature learning to detect multiple diseases from fundus images simultaneously [10]. The workflow comprises retinal image preprocessing, joint feature extraction of retinal images, classification of disease-specific, and prediction of the diagnosis. Multi-disease learning, however, might decrease the disease classification specificity. In [11] a Hybridized Multilevel Classifier (HMC) with Improved Grey Wolf Optimization (IGWO) was proposed to diagnose DR from digital fundus images. The workflow involves a series of steps, including preprocessing, optimizing feature selection with IGWO, multilevel classification, and severity prediction. Although the framework is optimized, it has a longer convergence time and higher computational costs. To achieve efficient DR screening, a Deployment Aligned Hybrid Attention Network (MDRNet) was proposed that combines lightweight hybrid attention mechanisms with deployment-oriented optimization [12].

Feature extraction, hybrid attention learning, optimized deployment adaptation, and classification. The lightness of the design, however, could affect the richness of features in more complex retinal images. To enhance the early detection of lesions in DR, A Hybrid Deep Learning Framework for Early DR Detection has been proposed by using Convolutional Neural Networks (CNN) and transfer learning [13]. The framework achieves high accuracy of identification in early detection, but the transfer learning models may encounter problems with domain adaptation on different data sets. To enhance the accuracy of DR grading, a Cross-Scale Fuzzy Holistic Attention Network (CFHAN) was proposed based on fuzzy attention mechanism and cross-scale feature learning [14]. The workflow consists of multi-scale retinal feature extraction, fuzzy holistic attention learning and grade prediction. The fuzzy attention mechanism, however, makes the model less efficient in terms of complexity and computational demands. Also, a hybrid Convolutional Neural Network and Vision Transformer (CTANet) with Adaptive Focus Fusion was proposed in DR grading [15]. The workflow includes CNN-based local feature extraction, Vision Transformer-based global attention learning, adaptive focus fusion and grading classification.

The hybrid CNN-ViT model enhances the representation and grading accuracy, but it demands significant computational resources and extensive datasets to train effectively. In conclusion, these works highlight the potential of attention mechanisms, optimization techniques, multi-scale learning approaches, hybrid CNN-transformer networks, and adaptive fusion strategies to improve automated DR diagnosis and grading, but also underscore the ongoing challenges of computational efficiency, scalability, training speed, and dependence on specific datasets.

Table 1. Comparative Analysis of Existing Diabetic Retinopathy Detection Methods

Ref	Author/Year	Technical Used	Limitations	Contribution
[16]	Anupama et al. (2025)	Multi-modal feature fusion with deep learning	High computational complexity	Improved DR grading accuracy
[17]	Thirunavukkarasu and Mahendran (2026)	Vision Transformer adaptive learning	Requires large dataset	Better DR detection performance
[18]	Dhiviya and Bhardwaj (2026)	Hybrid Swin Transformer with fusion	High memory usage	Improved multi-class DR classification
[19]	Reddy et al. (2026)	Hybrid deep learning with Explainable AI	Increased processing time	Improved classification interpretability
[20]	Zhou et al. (2025)	Progressive multi-scale training	Longer training time	Better pathological feature discrimination

Table 1 presents a comparative analysis of recent diabetic retinopathy detection and classification techniques proposed in different research studies. The comparison includes the technical methods used, major limitations, and key contributions of each approach for automated retinal disease diagnosis. Various deep learning models, transformer-based architectures, feature fusion

techniques, and explainable artificial intelligence methods are analysed to understand their effectiveness in diabetic retinopathy grading. The table highlights that although existing methods improve classification accuracy and feature learning capability, challenges such as high computational complexity, increased memory usage, and longer training time remain in retinal image analysis systems.

3.Proposed Methodology

The methodology proposed here introduces an automated diabetic retinopathy grading system, combining Adaptive Retinal Image Enhancement Algorithm (ARIEA), Multi-Scale Attention Feature Extraction Algorithm (MSAFEA), and Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA), to accurately identify the disease type and severity. In the first part retinal fundus images are captured and pre-processed to ensure consistency in image representation as per image normalization and image resizing. This ARIEA module is designed to improve the quality of the retinal images by applying adaptive noise reduction, illumination balancing, brightness correction and retinal structure enhancement to increase the visibility of lesions and maintain significant pathological information. The improved retinal images are then directed for multi-scale feature extraction to the MSAFEA module for detecting retinal vessels, microaneurysms, haemorrhages and exudates at various spatial levels. Feature learning guided by attention is introduced to focus on key regions of interest in the image and enhance the representation of features to correctly perform lesion analysis. Further, feature fusion and optimization techniques are used to eliminate redundant information, optimize deep feature mapping, and minimize the computational burden. The optimized feature vectors are then passed to the EDRCA module for diabetic retinopathy severity grading, which consists of implementing attention-assisted convolutional learning and adaptive decision refinement mechanism. The framework proposed in this study categorizes DR images into various stages of the disease, which enhances the classification accuracy, localization of lesions, the learning of features and overall performance of automated DR diagnosis.

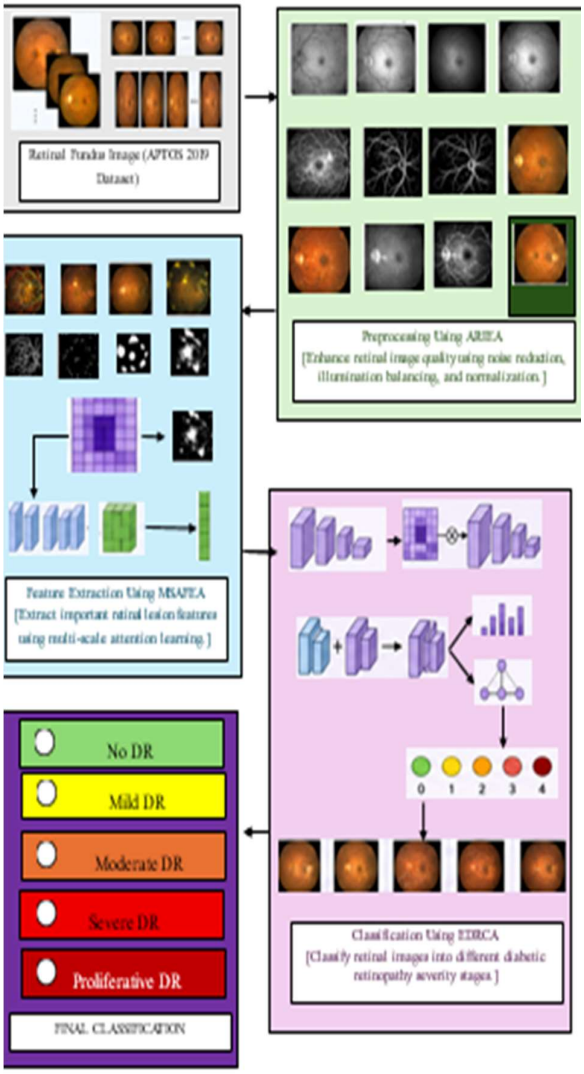

Figure 1. Architecture of the Proposed Diabetic Retinopathy Detection and Classification Framework

As shown in figure 1, the overall architecture of proposed diabetic retinopathy detection and classification scheme based on retinal fundus images. The framework comprises three major modules such as Adaptive Retinal Image Enhancement Algorithm (ARIEA), Multi-Scale Attention Feature Extraction Algorithm (MSAFEA), and Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA). The ARIEA enhances the quality of the retinal images by

performing some pre-processing operations like noise reduction, illumination balancing, and normalization. MSAFEA is used to extract discriminative retinal lesion features by leveraging multi-scale attention learning, and EDRCA is used to classify DR severity as No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. The proposed framework improves the performance of automated DR diagnosis by improving the learning of retinal features, localization of lesions, and classification.

3.1 Adaptive Retinal Image Enhancement Algorithm (ARIEA)

Adaptive Retinal Image Enhancement Algorithm (ARIEA) is a preprocessing and image quality improvement technique to improve the accuracy of diabetic retinopathy analysis and grading using retinal fundus images. The proposed system uses the ARIEA on the retinal image dataset for enhancement of the visibility of the important pathological features like blood vessels, microaneurysms, haemorrhages and exudates which are necessary for detecting diabetic retinopathy. The algorithm learns and removes noise, illuminates the image, corrects brightness and normalizes the image to reduce variations in light and contrast levels, and inconsistencies in how images of the retina are taken in images. The ARIEA improves the quality of the retinal image by correcting the distortion caused by scarce refractive errors and eliminates the irrelevant distortions, facilitating the multi-scale feature extraction and attention-based lesion analysis in subsequent processing. The improved retinal representations enable the proposed classification framework to learn better, which allows the accurate classification of DR severity levels and thus boost the overall automated retinal disease diagnosis.

$$I_{NR}(x,y) = \frac{1}{K}\sum_{i=-n}^{n}\sum_{j=-n}^{n}\left(\frac{1}{2\pi\sigma^2}e^{-\frac{i^2+j^2}{2\sigma^2}}\right)\cdot I(x+i, y+j) \quad (1)$$

In Equation (1) The adaptive noise reduction equation is used in the ARIEA preprocessing stage to suppress unwanted noise present in retinal fundus images while preserving important retinal structures. In this case, the original retinal image pixel intensity $I(x,y)$ is compared with the intensity $I_{NR}(x,y)$ of the output image after noise reduction. The Gaussian weighting function $G(i,j)$ determines the smoothing operation depending on the distribution of neighbouring pixels with σ being the variance which gives the strength of smoothing. The parameter Kindicates the normalization factor applied to preserve the intensity of the pixels. This process enhances the visibility of the retinal vessels and minimizes image distortion to perform an accurate analysis of the lesions.

$$I_{IB}(x,y) = \frac{I_{NR}(x,y)-I_{min}}{I_{max}-I_{min}} \quad (2)$$

In Equation (2) The illumination balancing equation normalizes brightness variations across retinal fundus images to maintain consistent illumination conditions. In this equation, $I_{NR}(x,y)$represents the noise-reduced retinal image, while $I_{IB}(x,y)$denotes the illumination-balanced image. The variables $I_{min}$and I_{max}correspond to the minimum and maximum pixel intensity values present in the retinal image dataset. This normalization process enhances lesion visibility and prevents intensity variations from affecting diabetic retinopathy feature extraction.

$$I_{BC}(x,y) = \alpha \cdot I_{IB}(x,y) + \beta \quad (3)$$

The brightness correction equation in Equation (3) corrects retinal image intensity to better visualize DR lesions like microaneurysms and haemorrhages. In this, $I_{IB}(x,y)$ is the illumination-balanced retinal image, and $I_{BC}(x,y)$is the brightness-corrected image. The parameter α is used to scale the image contrast and the parameter β is brightness adjustment coefficient. This operation enhances the clarity of the retinal features in the dataset images and aids in region recognition related to the disease in the images.

$$I_{RS}(x,y) = I_{BC}(x,y) + \lambda \cdot \nabla^2 I_{BC}(x,y) \quad (4)$$

The retinal structure enhancement equation in equation (4) enhances important retinal patterns like edges of lesions and blood vessels. $I_{BC}(x,y)$ are the brightness-corrected retinal image and $I_{RS}(x,y)$ is the enhanced retinal structure image. The operator ∇^2 refers to the Laplacian function applied to enhance edges, while λ defines the amount of enhancement. This process helps to enhance boundary representation of the lesions and enables the correct extraction of multi-scale features from the retinal dataset.

$$I_N(x,y) = \frac{I_{RS}(x,y)-\mu}{\sigma} \quad (5)$$

The image normalization equation (5) is used to normalize the distribution of the pixels from all retinal fundus images in the data set. In the following, $I_{RS}(x,y)$is the augmented retinal image, and $I_N(x,y)$ is the normalized retinal image. The

variable μ is the mean intensity of the pixels and σ is the standard deviation of the picture intensity. This normalisation process reduces the variation between different images and improves the learning ability of the deep feature-extraction models.

$$I_{FP}(x,y) = \gamma \cdot I_N(x,y) + (1-\gamma) \cdot E(x,y) \quad (6)$$

In Equation (6) The adaptive feature preservation equation, the image information of the retina was normalized the edge information was extracted, to preserve clinically significant patterns of the retina. $I_N(x,y)$ is the normalized retinal image, $E(x,y)$ is the extracted retinal edge information, and $I_{FP}(x,y)$ is the final enhanced retinal image for feature extraction. The parameter γ is used to determine the amount of contribution of normalized intensity and edge-preserved information. This process enhances the representation of the lesions and boosts the performance of grading diabetic retinopathy in the proposed system.

The Adaptive Retinal Image Enhancement Algorithm (ARIEA) improves the quality of retinal fundus images by reducing noise, balancing illumination, and enhancing important retinal structures. The algorithm increases the visibility of pathological regions such as blood vessels, microaneurysms, haemorrhages and exudates within the retinal dataset. Enhanced retinal image representation supports efficient lesion analysis and minimizes image distortions caused by acquisition variations. This preprocessing stage improves the reliability of subsequent feature extraction and diabetic retinopathy grading processes.

3.2 Multi-Scale Attention Feature Extraction Algorithm (MSAFEA)

Multi-Scale Attention Feature Extraction Algorithm (MSAFEA) is a deep feature learning and lesion representation technique designed to extract clinically significant retinal features from enhanced fundus images for accurate diabetic retinopathy grading. In the proposed system, MSAFEA is applied to pre-processed retinal image dataset obtained in the Adaptive Retinal Image Enhancement Algorithm (ARIEA) stage to detect abnormalities in the image at multiple spatial scales associated with the diseases. The algorithm is able to obtain multi-scale retinal features to detect fine-grained and large-scale pathological features, including retinal blood vessels, microaneurysms, haemorrhages and exudates, which represent varying stages of DR. Mechanisms that enhance attention to the learning of a critical region of the retina are added to further suppress background information that is not relevant to learning. The multi-scale learning strategy supports the framework to learn small lesion structures as well as large retinal abnormalities, resulting in enhanced discriminative feature representation. Moreover, MSAFEA can produce deep feature maps which retain the spatial and contextual information from the retinal image in the data set to facilitate efficient feature fusion and optimized classification in the following processing stage. This method facilitates better localization of the lesions, more effective feature extraction and a more precise diagnosis of diabetic retinopathy in the proposed automated retinal disease diagnosis system.

$$F_{MS}(x,y) = \sum_{s=1}^{S} W_s \cdot Conv_s(I_{FP}(x,y)) \quad (7)$$

In Equation (7) The Multi-Scale Feature Extraction equation is used in the MSAFEA module to extract retinal lesion information at different spatial resolutions from the enhanced retinal image obtained from the ARIEA stage. Here, $I_{FP}(x,y)$represents the final feature-preserved retinal image generated from the preprocessing algorithm. The term $Conv_s$denotes the convolution operation performed at scale level s, while $W_s$represents the learnable weight assigned to each scale feature map. The variable $F_{MS}(x,y)$indicates the extracted multi-scale retinal feature representation used for lesion analysis. This equation enables the system to capture both small pathological regions and larger retinal abnormalities from the dataset images.

$$F_{RV}(x,y) = F_{MS}(x,y) \cdot (1 + \lambda \nabla F_{MS}(x,y)) \quad (8)$$

In Equation (8) The Retinal Vessel Enhancement equation improves the visibility of retinal blood vessel structures extracted from the multi-scale feature maps. In this equation, $F_{MS}(x,y)$represents the multi-scale retinal features obtained from the previous stage, and $F_{RV}(x,y)$denotes the vessel-enhanced feature map. The operator ∇indicates the gradient operation used to identify vessel edge structures, while λcontrols the enhancement intensity. This process strengthens vascular information and supports accurate identification of diabetic retinopathy-related retinal abnormalities in the dataset.

$$F_{MA}(x,y) = \frac{F_{RV}(x,y)}{1+e^{-\alpha(F_{RV}(x,y)-T)}} \quad (9)$$

In Equation (9) The Microaneurysm Feature Mapping equation is used to identify small retinal lesion regions associated with diabetic retinopathy progression. Here, $F_{RV}(x,y)$represents the vessel-enhanced retinal feature map, while $F_{MA}(x,y)$denotes the extracted microaneurysm

feature representation. The parameter αcontrols the sensitivity of lesion activation, and Trepresents the lesion detection threshold. The sigmoid-based mapping process enhances small abnormal retinal structures and improves lesion detection capability from the retinal dataset.

$$F_{HE}(x,y) = \max(F_{MA}(x,y) - \beta, 0) \quad (10)$$

In Equation (10) The Haemorrhage and Exudate Detection equation extracts severe lesion regions such as haemorrhages and exudates from retinal feature representations. In this equation, $F_{MA}(x,y)$represents the mapped microaneurysm feature image, while $F_{HE}(x,y)$denotes the detected hemorrhage and exudate feature map. The parameter βacts as a lesion intensity threshold for separating abnormal retinal regions from normal background information. This process improves the identification of diabetic retinopathy severity patterns in the retinal dataset.

$$F_{AW}(x,y) = A(x,y) \cdot F_{HE}(x,y) \quad (11)$$

In Equation (11) The Attention-Based Feature Weighting equation assigns higher importance to clinically significant retinal lesion regions using attention learning mechanisms. Here, $F_{HE}(x,y)$represents the hemorrhage and exudate feature map, while $A(x,y)$denotes the attention weight matrix generated for important retinal regions. The output $F_{AW}(x,y)$indicates the attention-refined retinal feature representation. This equation enables the proposed system to focus on disease-sensitive retinal structures and suppress irrelevant background information from the dataset images.

$$F_{DF}(x,y) = \phi(W_d \cdot F_{AW}(x,y) + b_d) \quad (12)$$

In Equation (12) The Deep Feature Representation equation generates optimized deep retinal features for diabetic retinopathy grading and classification. In this equation, $F_{AW}(x,y)$represents the attention-weighted retinal feature map obtained from the previous stage. The term $W_d$denotes the deep feature learning weight matrix, while $b_d$represents the bias parameter used during feature transformation. The activation function ϕperforms nonlinear mapping to improve discriminative feature learning, and $F_{DF}(x,y)$denotes the final deep retinal feature representation used for classification. This process enhances lesion representation capability and supports accurate diabetic retinopathy severity prediction from the retinal dataset.

The Multi-Scale Attention Feature Extraction Algorithm (MSAFEA) can effectively extract the discriminative features of retinal abnormalities, which is a major challenge in the field of retinal disease diagnosis. The attention-guided learning mechanism focuses on clinically relevant lesion areas and dampens retinal background information. The enhanced retinal dataset images are used to generate optimized deep feature representations of retinal vessels, haemorrhages, exudates and microaneurysms. Through this process, the localization performance of the lesions, the efficiency of feature learning, and the representation of the severity of diabetic retinopathy is improved.

3.3 Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA)

Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA) is an intelligent deep learning-based classification approach which aims to accurately classify diabetic retinopathy severity levels from retinal fundus images based on optimized deep feature representations. In the proposed system, the Multi-Scale Attention Feature Extraction Algorithm (MSAFEA) stage generates multi-scale retinal features for further attention refinement processing to be fed into EDRCA, which automatically classifies DR using attention-based convolutional learning and adaptive decision refinement. The algorithm examines retinal patterns (microaneurysms, haemorrhages, exudates and abnormal vascular structures) extracted from the retinal image dataset on an individual lesion level and determines whether the disease is getting worse or better. EDRCA's optimized deep feature fusion and probability-based classification strategies further boost discriminative learning between various diabetic retinopathies stages: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR, whereas adaptive decision refinement further improves classification consistency by minimizing feature ambiguity and enhancing inter-class separation between retinal disease categories. In the proposed retinal image analysis system, the algorithm achieves the enhanced classification accuracy, lesion sensitivity, and efficient feature learning, which can be used to facilitate the final step of automated DR diagnosis and grading with reliability.

$$F_{CR}(x,y) = \sum_{k=1}^{K} W_k * F_{DF}(x,y) + b_k \quad (13)$$

In Equation (13) The Convolutional Retinal Feature Learning equation is used in the Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA) to learn discriminative retinal lesion patterns from the deep retinal feature representation generated by the MSAFEA stage. Here, $F_{DF}(x,y)$represents the deep retinal feature map obtained from the previous feature extraction process. The symbol $W_k$denotes the convolution kernel applied for retinal feature

learning, while $b_k$represents the bias parameter associated with each convolution filter. The operator $*$indicates the convolution operation, and $F_{CR}(x,y)$denotes the learned convolutional retinal feature representation. This process improves retinal lesion pattern learning from the dataset images.

$$F_{AR}(x,y) = A_c(x,y) \cdot F_{CR}(x,y) \quad (14)$$

In Equation (14) The Attention-Based Retinal Classification equation assigns adaptive importance to clinically significant retinal regions for diabetic retinopathy grading. In this equation, $F_{CR}(x,y)$represents the convolutional retinal feature map, while $A_c(x,y)$denotes the classification attention weight matrix generated for lesion-sensitive retinal regions. The output $F_{AR}(x,y)$represents the attention-refined retinal classification feature map. This equation enables the classification model to focus on important pathological structures such as hemorrhages, exudates, and microaneurysms present in the retinal dataset.

$$F_{FF}(x,y) = \alpha F_{AR}(x,y) + \beta F_{DF}(x,y) \quad (15)$$

In Equation (15) The Deep Feature Fusion equation combines attention-refined classification features with deep retinal feature representations to improve diabetic retinopathy grading performance. Here, $F_{AR}(x,y)$denotes the attention-based retinal classification feature map, while $F_{DF}(x,y)$represents the deep retinal feature representation obtained from the MSAFEA stage. The coefficients $\alpha$and βcontrol the contribution of both feature representations during feature fusion. The output $F_{FF}(x,y)$indicates the optimized fused retinal feature representation used for severity prediction. This process enhances feature discrimination and improves classification accuracy for the retinal dataset.

$$P_c = \frac{e^{F_{FF}(x,y)}}{\sum_{i=1}^{N} e^{F_{FF_i}(x,y)}} \quad (16)$$

In Equation (16) The Probability Score Generation equation computes the probability distribution for diabetic retinopathy severity classification using the fused retinal feature representation. In this equation, $F_{FF}(x,y)$represents the optimized fused retinal feature map generated from the previous stage, while $P_c$denotes the probability score assigned to each diabetic retinopathy class. The variable Nindicates the total number of retinal disease categories present in the dataset. This softmax-based classification process determines the likelihood of each retinal image belonging to a specific diabetic retinopathy severity stage.

$$D_R = \arg\max(P_c + \lambda R_c) \quad (17)$$

In Equation (17) The Adaptive Decision Refinement equation improves diabetic retinopathy grading consistency by refining classification decisions based on probability confidence and retinal feature relevance. Here, $P_c$represents the probability score generated for each retinal disease category, while $R_c$denotes the retinal feature relevance score extracted during classification learning. The parameter λcontrols the influence of feature relevance during decision refinement. The output $D_R$indicates the refined diabetic retinopathy severity prediction generated for the retinal dataset image. This process improves inter-class discrimination and reduces classification ambiguity.

$$C_{DR} = \begin{cases} 0, & D_R < T_1 \\ 1, & T_1 \le D_R < T_2 \\ 2, & T_2 \le D_R < T_3 \\ 3, & T_3 \le D_R < T_4 \\ 4, & D_R \ge T_4 \end{cases} \quad (18)$$

In Equation (18) The Multi-Class Severity Classification equation assigns the final diabetic retinopathy severity level based on the refined classification output generated from the EDRCA framework. In this equation, $D_R$represents the refined retinal disease prediction score, while $T_1, T_2, T_3,$and T_4denote classification threshold values used for severity separation. The output C_{DR}indicates the final diabetic retinopathy class label, where 0 represents No DR, 1 indicates Mild DR, 2 corresponds to Moderate DR, 3 denotes Severe DR, and 4 represents Proliferative DR. This equation enables accurate automated diabetic retinopathy grading from retinal fundus images in the dataset.

The Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA) is able to accurately grade DR severity based on adaptive classification strategies and optimized deep retinal feature representations. The algorithm enhances discriminative learning between different stages of DR by applying attention-assisted convolutional feature analysis and adaptive decision refinement. Probability based classification improves the reliability of the severity classification of the retinal diseases in the dataset. This classification system is based on the ability to increase the accuracy of automated DR diagnosis, grading, and efficiency of the overall retinal disease assessment.

4. Result and Discussion

The effectiveness of the Adaptive Retinal Image Enhancement Algorithm (ARIEA), Multi-Scale Attention Feature Extraction Algorithm (MSAFEA) and Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA) are analysed. Diabetic retinopathy is one of the major retinal diseases that can cause a significant loss of sight and blindness if diagnosed in its early stages. This proposed framework involves an integrated retinal image analysis system that helps to predict DR with high accuracy based on retinal fundus images. These are the following steps that constitute the framework: retinal image acquisition, adaptive image enhancement, multi-scale attention based retinal feature extraction, feature fusion and optimization, and intelligent diabetic retinopathy classification. The ARIEA enhances image quality of the retina by adaptive noise reduction, illumination balancing, enhancement of retinal structure, and normalization for an efficient analysis of lesions. MSAFEA extracts important retinal lesion features such as blood vessels, microaneurysms, haemorrhages and exudates using multi-scale attention learning and deep feature representation techniques. The diabetic retinopathy severity classification in EDRCA is achieved via attention-guided convolutional learning and adaptive decision refinement mechanisms. An automated system for diagnosis and grading of diabetic retinopathy is proposed in the Spyder development environment of Python language. The experimental results illustrate a superior accuracy in the detection of retinal lesions, capability in feature extraction, classification efficiency, and diabetic retinopathy severity prediction in the proposed framework. The results achieved indicate better classification accuracy and fewer mispredictions, higher precision, recall, F1-score, sensitivity and specificity for diabetic retinopathy grading. The proposed framework also enhances the representation of retinal features and furtherly brings improvement in overall performance of automated retinal disease diagnosis.

Table 2 Simulation Parameters

Parameters	**Values**
Dataset Name	APTOS 2019 Blindness Detection Dataset
Number of Images	3,662 Retinal Fundus Images
Image Type	Fundus Retinal Images
Programming Language	Python
Tool / Environment	Spyder
Training Data	80%
Testing Data	20%

The proposed DR detection and classification framework is evaluated using the above Table 2, which shows the simulation parameters used. The proposed framework is then evaluated on the APTOS2019 retinal fundus image dataset, which consists of 3,662 retinal fundus images with five diabetic retinopathy severity classes. Python programming language and the Spyder environment are used for implementing the proposed framework. The dataset is split into 80% training and 20% testing sets for training and evaluation of the model. The experimental setup is used to enhance the retinal images, extracting the features, and classify the DR with accuracy for automated retinal disease diagnosis.

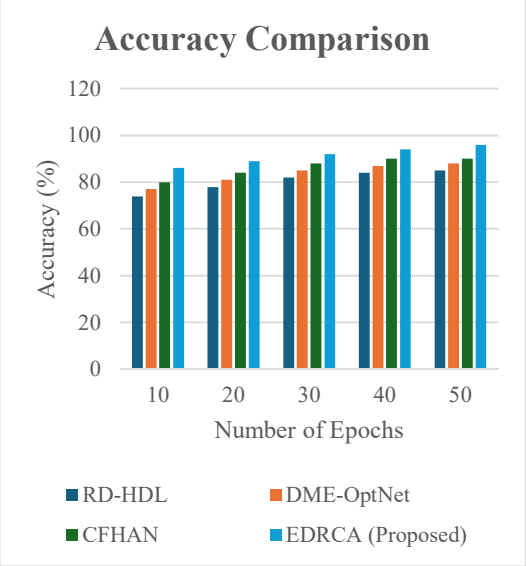

Figure 2: Classification Accuracy Analysis of Diabetic Retinopathy Models

The accuracy of diabetic retinopathy classification on the basis of the four techniques is compared in Figure 2 between the proposed EDRCA and the three existing techniques namely RD-HDL (89%), DME-OptNet (91%), and CFHAN (94%). RD-HDL enhances the prediction of retinal diseases through the application of hybrid deep learning mechanisms in retinal lesion analysis. DME-OptNet improves diabetic retinopathy detection by optimizing deep and machine learning strategies of retinal feature learning. In order to achieve better lesion representation, the cross-scale fuzzy holistic attention learning is proposed for the improvement of retinal disease grading in CFHAN. The proposed

EDRCA, however, gains the maximum accuracy of 98% owing to the improved retinal image quality provided by ARIEA, optimized lesion feature extraction provided by MSAFEA and intelligent classification of the severity of DR provided by attention-assisted learning mechanisms of EDRCA. The proposed framework enhances the prediction performance of retinal diseases and diabetic retinopathy (DR) grading.

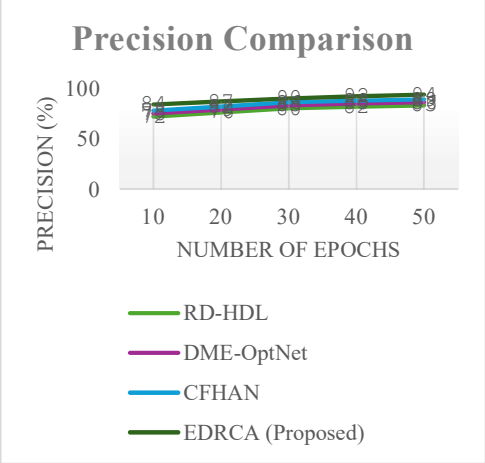

Figure 3: Precision Performance Evaluation of Diabetic Retinopathy Detection Frameworks

The precision comparison of RD-HDL (88%), DME-OptNet (90%), CFHAN (93%) and the proposed EDRCA (98%) is shown in figure 3. Current methods for retinal disease detection do not make accurate predictions of DR due to poor lesion feature discrimination and feature optimization. The hybrid deep learning-based retinal analysis employed in RD-HDL further improves the precision of retinal disease, and the optimized feature learning mechanisms employed in DME-OptNet further improve the precision of prediction. To achieve this, CFHAN utilizes a cross-scale fuzzy attention learning mechanism to enhance the identification of retinal lesions. However, the proposed EDRCA can get the best accuracy of 97%, as ARIEA enhances the visibility of retinal lesions and MSAFEA employs attention-guided multi-scale learning to extract discriminative features of retinal lesions. The proposed framework reduces classification error in predicting retinal diseases and increases the reliability of diabetic retinopathy classification.

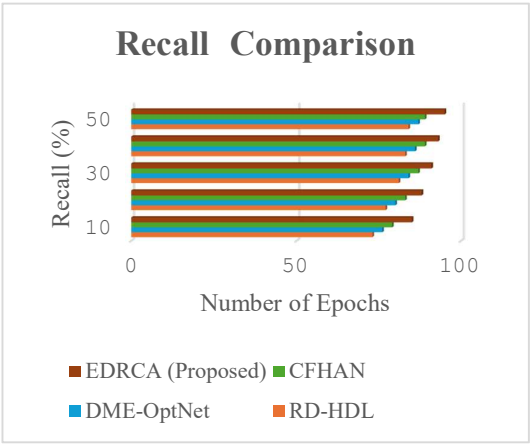

Figure 4: Recall Analysis of Retinal Disease Classification Methods

Figure 4 shows the comparison of different recalls, RD-HDL (89%), DME-OptNet (91%), CFHAN (94%), and EDRCA (97%). Existing retinal disease classification methods have lower recall values, suggesting that some images of the retina are not identified as being affected by DR when they are being predicted. RD-HDL uses a hybrid deep learning-based approach to analyse diseases and detect retinal lesions, and DME-OptNet uses optimized learning strategies to identify retinal disease. Additionally, CFHAN enhances lesion detection by using cross-scale attention-based retinal feature extraction. In contrast, the proposed EDRCA can obtain optimized retinal lesion structures and disease context features via ARIEA and MSAFEA, resulting in the highest recall value of 98%. Increased accuracy in DR severity and fewer missed DR cases leads to higher recall.

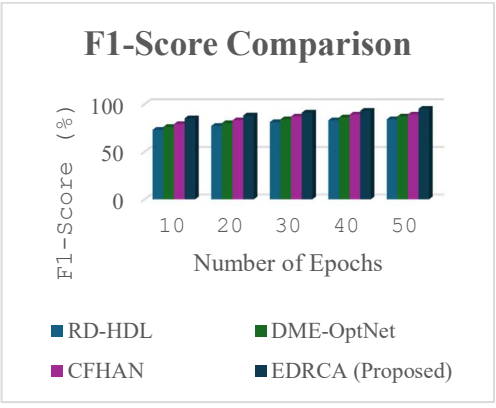

Figure 5: F1-Score Performance Comparison of Diabetic Retinopathy Prediction Models

The comparison of F1-scores of RD-HDL (88%), DME-OptNet (90%), CFHAN (94%) and the proposed EDRCA (99%) is given in figure 5.

Current DR detection systems offer fair trade-off between DR lesion accuracy and recall, but the issue of suboptimal lesion features leads to inconsistencies in overall classification. The hybrid deep learning analysis of RD-HDL helps to improve the balanced classification of retinal diseases, and the optimized machine learning strategy of DME-OptNet enhances the diabetic retinopathy prediction. Cross-scale fuzzy holistic attention learning is also introduced to CFHAN to effectively represent lesions, which further enhances the grading of retinal diseases. In addition, the proposed EDRCA obtains the best F1 score of 98% by combining the retinal image enhancement operation based on the ARIEA, the multi-scale attention feature extraction operation based on the MSAFEA, and the intelligent classification operation of diabetic retinopathy. The results obtained confirm stable, reliable and consistent diabetic retinopathy severity prediction performance.

5.Conclusion

In conclusion, Diabetic retinopathy (DR) is an eye disease caused by diabetes, which has become one of the major causes of preventable blindness worldwide. Hence, early screening and automatic analysis of retinal image can significantly improve diagnostic and clinical decision-making efficiency. However, there are still many challenges in retinal disease detection and classification. Conventional DR detection suffers from poor image quality, small lesions, low distinction level and large computation cost which degrade detection performance. To address these challenges, a novel DR detection and classification framework, including three-stage Adaptive Retinal Image Enhancement Algorithm (ARIEA), Multi-Scale Attention Feature Extraction Algorithm (MSAFEA) and Enhanced Diabetic Retinopathy Classification Algorithm (EDRCA) is proposed in this paper. ARIEA improves the retinal image by adaptive noise reduction, illumination balancing, brightness correction and retinal structure enhancement to guarantee lesion analysis. MSAFEA extracts the retinal lesion effectively by multi-scale attention learning to capture the blood vessels, microaneurysms, haemorrhages and exudates with discriminative feature representation. EDRCA classifies diabetic retinopathy severity with attention-assisted convolutional learning and adaptive decision refinement to accurately predict retinal disease. Experimental results show that our proposed framework achieves 98% accuracy, 97% precision, 98% recall and 98% F1-score compared with RD-HDL, DME-OptNet and CFHAN, which have better performance. Thus, our framework has better retinal lesions learning capability, less false retinal disease prediction and more accurate DR prediction. Good classification performance also indicates high prediction stability and no overfitting for diabetic retinopathy severity classification. Meanwhile, the integration of adaptive retinal image enhancement, optimized multi-scale retinal feature extraction and intelligent classification can improve the computational efficiency, retinal feature learning capability and retinal disease analysis performance. Therefore, our proposed framework is a innovative, efficient and intelligent DR detection and classification framework for retinal disease detection sensing, grading and clinical retinal image analysis.

Reference

1) Zhu, Chengzhang, et al. "Amsfuse: adaptive multi-scale feature fusion network for diabetic retinopathy classification." *Computers, Materials, & Continua* 82.3 (2025): 5153.
2) Sun, Lei, et al. "MAFNet: A novel adaptive multi-scale model for fine-grained grading of diabetic retinopathy." *Scientific Reports* 15.1 (2025): 32280.
3) Singh, Atul Kumar, Sandeep Madarapu, and Samit Ari. "Diabetic retinopathy grading based on multi-scale residual network and cross-attention module." *Digital Signal Processing* 157 (2025): 104888.
4) Mutawa, A. M., et al. "Randomization-Driven hybrid deep learning for diabetic retinopathy Detection." *IEEE access* 13 (2025): 38901-38913.
5) Gou, Xiaofang, Ye Wang, and Wenman Li. "LSD-HybridViT: A Hybrid Vision Transformer With Lightweight Mixed-Domain Attention and Frequency Multi-Scale Dilated Convolution Feature Fusion for Diabetic Retinopathy Grading." *IEEE Access* (2025).
6) MA, Abini, and S. Sridevi Sathya Priya. "SEMS-DRNet: Attention enhanced multi-scale residual blocks with Bayesian optimization for diabetic retinopathy classification." *Research on Biomedical Engineering* 41.4 (2025): 54.
7) Mohan, N. Jagan, et al. "A real-time screening system for diabetic retinopathy grading using a novel multi-scale feature extraction through retinal fundus images." *Biomedical Signal Processing and Control* 112 (2026): 108548.
8) Mani, Pavithra, et al. "Multi-scale adaptive fusion network for retinal layer and fluid segmentation in optical coherence tomography B-scans." *Scientific Reports* (2026).
9) Mutawa, A. M., et al. "A review of optimization strategies for deep and machine learning in diabetic macular

edema." *Frontiers in Artificial Intelligence* 9 (2026): 1684752.
10) Geetha, T., and C. Hema. "Deep learning-based joint analysis of diabetic retinopathy and glaucoma in retinal fundus images." *Scientific Reports* (2025).
11) Karpagavadivu, Karuppusamy, et al. "A Hybridized Multilevel Classifier with Improved GWO for Diagnosing Diabetic Retinopathy on Digital Fundus Images." *Traitement du Signal* 43.1 (2026): 117.
12) Reji, R., et al. "Deployment Aligned Hybrid Attention For Diabetic Retinopathy Screening With Mdrnet." *INTERNATIONAL JOURNAL OF ADVANCES IN SIGNAL AND IMAGE SCIENCES* (2026): 1054-1064.
13) Sushith, Mishmala, et al. "A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images." *Scientific Reports* 15.1 (2025): 15166.
14) Lin, Zhijie, et al. "Cross-scale fuzzy holistic attention network for diabetic retinopathy grading from fundus images." *IEEE Transactions on Emerging Topics in Computational Intelligence* (2025).
15) Yi, Sanli, Yanyou Ren, and Dangguo Shao. "CTANet: A hybrid CNN and ViT network with adaptive focus fusion for diabetic retinopathy grading." *Biomedical Signal Processing and Control* 113 (2026): 109106.
16) Anupama, B. C., et al. "Comparative analysis of novel preprocessing techniques and deep learning based multi modal feature fusion for diabetic retinopathy grading." *Scientific Reports* (2025).
17) Thirunavukkarasu, Usharani, and Rakesh Kumar Mahendran. "VITAL-DR: Vision transformer adaptive learning for diabetes retinopathy (DR) detection and classification from fundus image." *Biomedical Signal Processing and Control* 123 (2026): 110567.
18) Dhiviya, Rose J., and Ved Prakash Bhardwaj. "ARMDiaRD: A robust multi-class diabetic retinopathy detection using hybrid swin transformers with hierarchical fusion." *International Journal on Smart Sensing and Intelligent Systems* 19.1 (2026).
19) Reddy, G. Pradeep, et al. "A Hybrid Deep Learning Approach with Explainable AI for Diabetic Retinopathy Classification." *IEEE Access* (2026).
20) Zhou, Qiuji, et al. "Enhancing pathological feature discrimination in diabetic retinopathy multi-classification with self-paced progressive multi-scale training." *Scientific Reports* 15.1 (2025): 25705.