DIABETIC RETINOPATHY DETECTION AND RETINAL FEATURE OPTIMIZATION USING MACHINE LEARNING WITH INTELLIGENT CLASSIFICATION APPROACH

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

[1],[2]Research Scholar, Département of Computer Science and Engineering, Annamalai University, Chidambaram, Tamil Nadu, India, [3]Assistant Professor, Département of Computer Science, RV Government Arts College, Chengalpattu, India.

spriyanandini82@gmail.com[1], anu_jul@yahoo.co.in[2], sanalprabha@yahoo.co.in[3]

ABSTRACT

Diabetic retinopathy is a chronic eye disease that results from long-term diabetes and damages the blood vessels in the retina which can cause vision loss if caught early enough. To detect this disease, retinal fundus imaging is used and machine learning methods are used to analyse and classify the disease based on its severity. In most cases, the overall process can be broken down into image acquisition, image preprocessing, feature extraction, and classification, but traditional methods frequently suffer from image quality issues, noise, and the inability to efficiently select the optimal features, all of which can compromise the accuracy and dependability of classification. To address these issues, the proposed model uses the APTOS 2019 Blindness Detection dataset that consists of labelled retinal images divided into 5 different stages of diabetic retinopathy. The proposed framework has a structured workflow that enhances the image quality by removing the noise, resizing and enhancing the contrast of the image, and then select the most relevant features of the retina using the Optimized Machine Learning Feature Selection Algorithm (OMLFSA) for identification and selection. Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA) is the final classification process which is based on a machine learning classification model that classifies the severity of DR. This holistic method will increase sensitivity of detection and better diabetic retinopathy classification.

Keywords: ***Diabetic Retinopathy, Machine Learning, Retinal Fundus Images, Image Preprocessing, Feature Selection, Classification, APTOS 2019 Dataset, Deep Feature Extraction, CLAHE Enhancement, Automated Disease Detection, Medical Image Analysis, Intelligent Diagnosis.***

1. INTRODUCTION

Diabetic retinopathy is a condition of the retina caused by diabetes which can be serious and result in vision loss and blindness if not treated early. Diabetes has gained a high prevalence in the medical world and the need for accurately detecting retinal diseases has become an urgent requirement in the field of medical image analysis, and these systems are required to be automatic. Diabetes is one of the most prevalent diseases in the medical field and the accurate detection of retinal diseases is an urgent need in medical image analysis, and the system should be automatic. Retinal fundus images are typically used to detect DR, and machine learning and deep learning algorithms are employed to detect abnormalities like microaneurysms, hemorrhages, exudates and vascular distortions. The accuracy of retinal image classification and disease prediction has significantly improved with the use of deep neural networks, transformer-based models, and kernel learning techniques [1]. Pathological retinal images have also been employed to deepen the automated diagnosis of DR using deep learning frameworks for efficient feature extraction and disease severity analysis [2].

Recent studies also showed the efficacy of deep-learning based retinal screening systems for large-scale disease assessment and clinical assessment based on fundus photography [3]. These methods offer better classification accuracy but still have some drawbacks such as low retinal image quality, noise interference,

complexity of computation, lack of certainty in features learned and low classification reliability in different retinal conditions [4]. Moreover, some of the existing approaches are computationally expensive and cannot retain only essential retinal attributes, which also impacts the overall prediction accuracy of the disease [5]. In order to mitigate these drawbacks, the present study has proposed a machine learning based diabetic retinopathy detection system using diabetes database: APTOS 2019 Blindness Detection. The proposed system processes the image using Diabetic Retinal Image Enhancement Algorithm (DRIEA), extracts the discriminative retinal features with Optimized Machine Learning Feature Selection Algorithm (OMLFSA) and predicts the severity of the disease with Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA). This integrated structure optimizes representation of retinal features, reduces redundant information, and boosts the diabetic retinopathy classification performance.

OBJECTIVES OF THE STUDY

The specific objectives of this research are as follows:

1. To design a Diabetic Retinal Image Enhancement Algorithm (DRIEA) for enhancement of retinal fundus image quality by removing noise, enhancing the contrast, normalizing the image and standardizing the image for reliable diabetic retinopathy analysis.
2. Design an Optimized Machine Learning Feature Selection Algorithm (OMLFSA) that has an efficient retinal feature optimization mechanism, which extracts the discriminative lesion features and removes the redundant retinal information for better classification performance.
3. Design and Implementation of Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA) for accurate classification of retinal fundus images into different DR severity classes (No DR, Mild, Moderate, Severe and Proliferative DR).
4. To improve the performance of DR detection by combining the retinal image preprocessing, optimized retinal feature selection and intelligent machine learning based classification method in a single retinal disease prediction system.
5. To assess and validate the proposed framework for diabetic retinopathy detection using performance measures like accuracy, precision, recall and F1 score to ensure reliable prediction of diabetic retinopathy, minimum misclassification rate and better analysis of the severity of the disease.

2. LITERATURE SURVEY

In recent years, deep learning, machine learning, transformer learning and intelligent medical image processing techniques have greatly contributed to improvement in automated retinal disease analysis for diabetic retinopathy. In [6] authors proposed Adaptive Deep Convolutional Neural Network (ADCNN) for detection of DR by using retinal fundus images by introducing adaptive convolutional layers in the network to boost the performance of feature extraction, lesion identification and disease classification. The model was able to effectively analyse retinal abnormalities like microaneurysms and haemorrhages from fundus images but was expensive to run and less efficient when processing large numbers of fundus images. An Artificial Intelligence and Deep Learning Framework (AIDLF) was presented in [7] for an in-depth analysis of the opportunities, limitations, and clinical barriers of detectors for DR. While the framework enhanced automated retinal diagnosis through AI-powered medical image analysis, its performance was constrained by factors like the lack of interpretability, imbalanced retinal dataset, and limited adaptability to clinical settings, which compromised the reliability of the classification process and hindered its broader implementation.

In [8], a Federated Convolutional Neural Network (FCNN) was proposed to improve the prediction of DR stages by learning retinal images in a decentralized manner. Federated Learning (FL) enhanced privacy protection and secure distributed model training among healthcare systems, but introduced communication overhead, global model aggregation delays, and inconsistent global model results. A progressive multistage feature learning Deep Learning Multistage Training Method (DLMTM) was proposed in [9] for retinal image classification. While the framework improved the representation of retinal lesions and accuracy of the identification of diabetic retinopathy severity, the training process took a significant amount of optimization time and computational complexity, impacting the efficiency of the model. In [10] a Randomization-Driven Hybrid Deep Learning Model (RDHDLM) was introduced by using a randomized learning approach with hybrid deep neural architectures that enhances the performance of diabetic retinopathy classification. The framework improved the classification diversity and prediction capability of retinal diseases, but it added redundant feature learning and added model complexity, which reduced the computational efficiency.

A novel Lightweight Multi-Deep Learning Framework (LMDLF) was designed to detect DR and classify it into multi-level severity in [11]. The lightweight structure minimized memory usage and accelerated classification performance during retinal image analysis, but the deep feature representation capability was restricted in the case of complicated retinal abnormalities and severe lesion structure which reduced classification accuracy for detection. A Privacy-Preserving Federated Learning and Blockchain Framework (PPFLBF) was introduced in [12] to offer secure DR detection and decentralized medical data sharing. The framework addressed patient data privacy and secure retinal image communication by integrating blockchain, but it also introduced a delay in processing and computational costs due to the blockchain validation and encryption steps. To address the problem of retinal photo diagnosis and detection of DR, a GAN and Autoencoder Based Deep Learning Framework (GADLF) was built in [13] by generating and reconstructing synthetic retinal images through learning. The framework enhanced retinal image enhancement and dataset augmentation performance, but the unstable training of the GAN made the classification stability unstable. The images reconstructed from the GAN were not consistent.

For automated diabetic retinopathy detection, a Multi-Step Stacked Ensemble Classification Model (MSSECM) was proposed in [14] which used multiple ensemble classifiers and retinal feature fusion techniques. The framework enhanced the accuracy of disease prediction and the performance of lesion classification in the retina however, the combination of several classifiers resulted in a higher number of parameters, training time, and complexity of the model. A transformer-based Attention Dual Transformer with Adaptive Temporal Convolutional Framework (ADTATCF) for DR detection by leveraging adaptive temporal feature extraction and transformer-based attention learning was proposed in [15]. The framework showed improved global retinal feature representation, lesion localization, and severity classification accuracy yet, the transformer-based architectures required large, annotated retinal datasets and high computational resources to be implemented efficiently and effectively in medical image analysis systems.

2. PROPOSED METHODOLOGY

In the proposed methodology, a structured machine learning based framework for DR detection is proposed using APTOS 2019 Blindness Detection dataset, which comprises of labelled retinal fundus images of five severity levels. Images were then processed using the Diabetic Retinal Image Enhancement Algorithm

(DRIEA), which involves noise reduction, standardization of image resolution, and contrast enhancement to accentuate the retinal features, including blood vessels and lesions. These improved images are then fed into the Optimized Machine Learning Feature Selection Algorithm (OMLFSA) that is designed to identify crucial retinal patterns, such as microaneurysms, haemorrhages, and abnormal vascular structures, while also discarding redundant and irrelevant features to boost computational efficiency. The feature set is then passed to the Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA), which uses machine learning-based classification algorithm to classify the retinal images into various stages of diabetic retinopathy. This integrated pipeline increases feature representation, increases classification accuracy, and has a reliable automated pipeline for early detection and severity assessment of DR.

Figure 1 depicts the entire architecture of the proposed diabetic retinopathy detection framework by using fundus images of patients with diabetes in the APTOS 2019 dataset. First, the input retinal fundus images are passed through the Diabetic Retinal Image Enhancement Algorithm (DRIEA), where the noise is removed, image is resized, and image contrast is enhanced. Then, the Optimized Machine Learning Feature Selection Algorithm (OMLFSA) used to extract the major retinal area features such as blood vessels and pathological lesions, and redundant retinal areas are deleted. After that, the extracted retinal features are further analysed using the Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA) to detect patterns and classify the severity of disease. Finally, the system predicts the diabetic retinopathy stage in the following classes No DR, Mild, Moderate, Severe, or Proliferative DR.

3.1 Diabetic Retinal Image Enhancement Algorithm (DRIEA)

Preprocessing of the retinal fundus images from APTOS 2019 Blindness Detection dataset is done by using the Diabetic Retinal Image Enhancement Algorithm (DRIEA) as a preprocessing module in the proposed diabetic retinopathy detection system. The main purpose of this algorithm is to enhance the visual quality of the input images, thereby mitigating the typical issues often encountered in medical imaging datasets, including noise, low contrast, and uneven illumination. In this system, DRIEA functions as a critical enhancement module that involves resizing images to a common scale, eliminating unwanted distortions from noise, and enhancing the image's contrast with different methods like CLAHE to highlight important retinal structures. DRIEA enhances the clarity of the blood vessels, microaneurysms, and regions of the lesions in the images of the dataset for the next steps in processing the data. This enhancement process is important for improving the efficiency of feature extraction and is part of the process of accurate analysis in the entire diabetic retinopathy classification.

$$I_{raw}(x, y) = F(x, y) \quad (1)$$

The first one is the raw retinal fundus image, $I_{raw}(x, y)$ from APTOS 2019 dataset. This image has original pixel values with potential noise, illumination differences and low contrast. It is used as the first input to the DRIEA preprocessing pipeline and passed to Equation (2) to resize and standardise it.

$$I_{std}(x, y) = R(I_{raw}(x, y)) \quad (2)$$

The image standardization is achieved using Equation (2) which resizes the raw input image $I_{raw}(x, y)$ from Equation (1) to a uniform size $I_{std}(x, y)$. The transformation function R(·) is used to match dimensions of the data in space and the result is fed to Equation (3) for noise reduction.

$$I_{den}(x, y) = G(I_{std}(x, y)) * K_{\sigma} \quad (3)$$

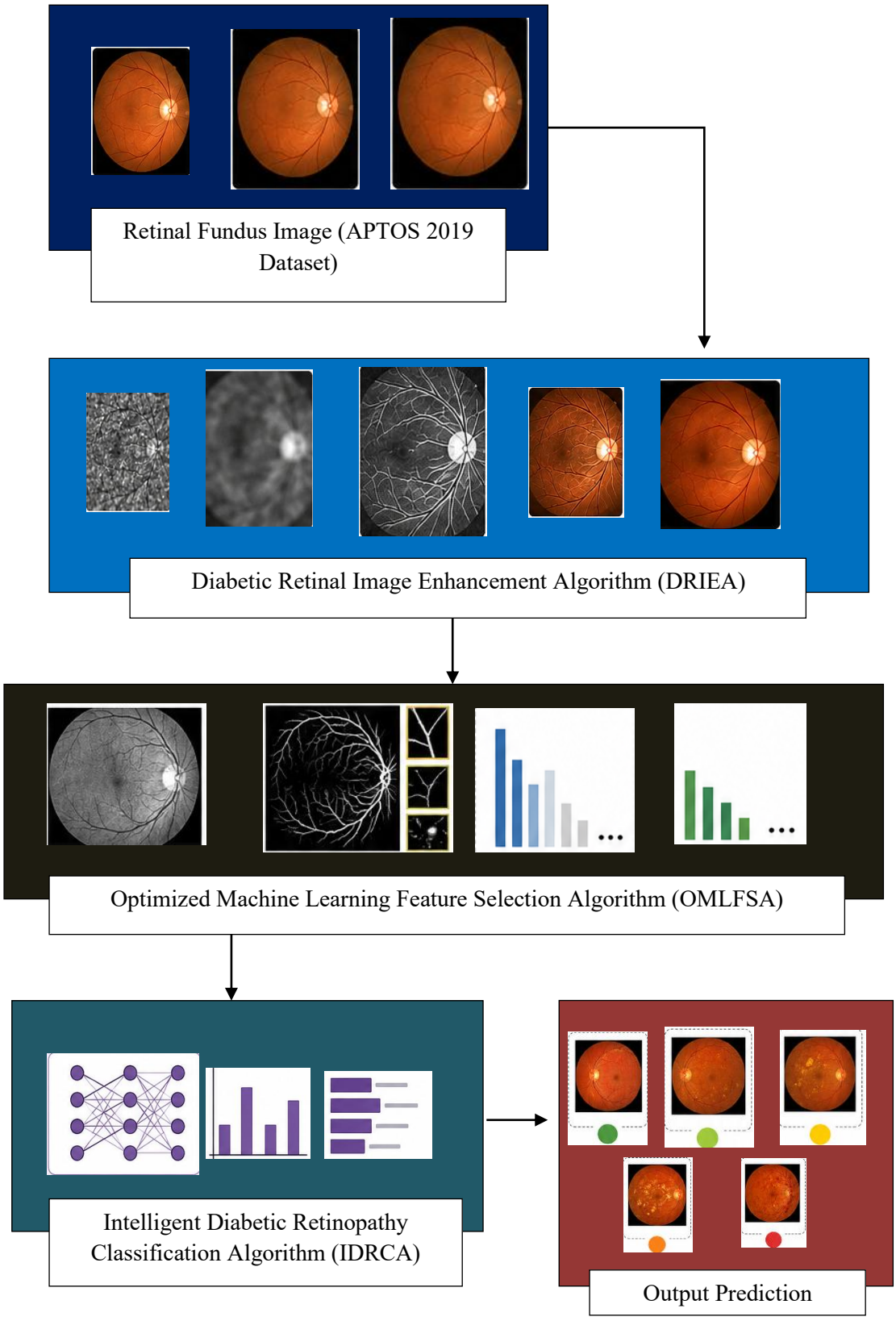

Figure 1 Architecture Diagram of the Proposed Diabetic Retinopathy Detection System

Equation (3) represents noise removal using Gaussian filtering on the standardized image $I_{std}(x, y)$ from Equation (2). The high frequency noise is reduced by the operator $G(\cdot)$ and Gaussian kernel is specified by K_σ. The output of the denoised image $I_{den}(x, y)$ is passed to Equation (4):

$$I_{enh}(x, y) = CLAHE(I_{den}(x, y)) \quad (4)$$

The denoised image $I_{den}(x, y)$ from Equation (3) is used to apply the Contrast Limited Adaptive Histogram Equalization (CLAHE) (Equation (4)). This increases the contrast in a particular area and makes it easier to see the different structures of the retina, including blood vessels and lesions. The enhanced output $I_{enh}(x, y)$ is used in Equation (5).

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

Equation (5) performs normalization on the enhanced image $I_{enh}(x, y)$from Equation (4). In this case μ is the mean intensity and σ is the standard deviation of values at the pixels. This reshaping gives rise to the normalized image $I_{norm}(x, y)$ which has uniform intensity distribution over the data set.

$$I_{out}(x, y) = f(I_{norm}(x, y)) \quad (6)$$

The final output of the Diabetic Retina Image Enhancement Algorithm (DRIEA) is described in Equation (6). The complete enhancement pipeline is represented by the function $f(\cdot)$ which is the composition of all the transformations from Equations (1) to (5). After this, the image $I_{out}(x, y)$would be a completely enhanced image of the retina that would be less noisy, with better contrast and with the same intensity, thus appropriate for feature selection and classification stages.

Diabetic Retinal Image Enhancement Algorithm (DRIEA) is the algorithm that resizes, removes noise, and enhances the contrast of retinal fundus images from APTOS 2019. This algorithm enhances the visibility of important retinal structures including blood vessels, lesions and hemorrhages. The images are better represented, which decreases image distortion and increases consistency between datasets for further analysis. Therefore, DRIEA can generate a reliable preprocessing output, which facilitates an accurate extraction of features and detection of DR.

3.2 Optimized Machine Learning Feature Selection Algorithm (OMLFSA)

Optimized Machine Learning Feature Selection Algorithm (OMLFSA) is a feature optimization module of the proposed diabetic retinopathy detection system which is called DRIEA that works on the output retinal images obtained from the DRIEA preprocessing module of the proposed DR detection framework using the APTOS 2019 database. The algorithm aims to detect and filter the most informative retinal features that are more crucial to the classification of DR and irrelevant or redundant data. The proposed system performs OMLFSA to detect the enhancements in the retinal blood vessels, microaneurysms, hemorrhages, exudates and texture variations existing in the dataset images. The algorithm uses machine learning based optimization methods to further optimize the extracted feature set, which will reduce the complexity of the computation and enhance the representation of the features. The OMLFSA only selects the most discriminative retinal features from the data set, which helps in the efficient execution of the next classification step and aids in the correct DR severity-level prediction. The proposed optimized feature selection process optimizes the overall model performance and increases reliability of the retinal image analysis in the proposed framework.

$$F_{init}(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} I_{out}(i,j) \cdot W(i,j) \quad (7)$$

The equation (7) is the initialisation of the Retinal Feature information from the enhanced retinal image $I_{out}(i,j)$ obtained from the preprocessing stage of DRIEA. In this context, the data mand denotes the spatial information of the retinal fundus images in APTOS 2019 dataset, and $W(i,j)$ is the weighted distribution of retinal pixels to highlight the important areas of the lesion. This procedure is useful in identifying the key features of DR like microaneurysms (micro-ruptures), bleeding and Abnormal vessels. The output of the initialized feature $F_{init}(x,y)$is sent to Equation (8) for deep feature extraction.

$$F_{ext} = \sigma\left(\sum_{k=1}^{p} F_{init}(x,y) \cdot \theta_k + b_k\right) \quad (8)$$

In Equation (8), the function is used for retinal feature extraction of the initialized feature map $F_{init}(x,y)$ of Equation (7). In this, the trainable value of the weights of the feature extractor is represented as θ_k and the bias parameter is denoted asb_k, and the number of extracted feature patterns is p, and the activation function used to recognize the nonlinear characteristics of the retina is denoted as $\sigma(\cdot)$. This procedure captures a substantial amount of structures and vascular patterns that are needed for DR analysis. The extracted feature vector F_{ext} is was fed to Equation (9).

$$F_{corr} = \frac{\sum_{i=1}^{N}(F_{ext,i}-\mu_{ext})(T_i-\mu_T)}{\sqrt{\sum_{i=1}^{N}(F_{ext,i}-\mu_{ext})^2}\sqrt{\sum_{i=1}^{N}(T_i-\mu_T)^2}} \quad (9)$$

The correlation between extracted retinal features and diabetic retinopathy target classes of APTOS 2019 dataset is computed by equation (9). In this case$F_{ext,i}$ denotes the extracted retinal feature values using equation (8), T_i denotes the corresponding disease labels, $\mu_{ext}$and μ_T denote the mean of the feature and the target respectively, and Ndenotes the number of retinal samples. This correlation analysis reveals the relation of severity levels and retinal abnormalities. The feature output that is correlated with the output of the class-based feature F_{corr} is was passed to Equation (10).

$$F_{opt} = \arg\min[\alpha E(F_{corr}) + \beta R(F_{corr})] \quad (10)$$

An optimized feature selection method with an objective minimization function is performed on the correlated retinal features F_{corr} from Equation (9) which is performed in Equation (10). $E(F_{corr})$ is the feature redundancy error; $R(F_{corr})$ is the feature relevance measurement; α and β are the optimization coefficients for adjusting feature balance. This optimization process eliminates irrelevant retinal patterns and retains highly informative DR features. Rewrite the feature output of the optimization F_{opt} is to Equation (11).

$$F_{sel} = \{F_{opt,i} \in F_{opt} \mid Score(F_{opt,i}) > \tau\} \quad (11)$$

The optimized feature set, F_{opt} obtained in Equation (10), is presented to Equation (11) which selects the most discriminative features of the retina. In this case, $Score(F_{opt,i})$ stands for the importance score of each retinal feature, and τ is the selection threshold value. Any features that are above the threshold remain for classification while redundant retinal information is dropped. The set of features F_{sel} consists of the most important features of lesions and blood vessels from the data set and is passed on to Equation (12).

$$F_{OMLFSA} = \gamma \cdot \sum_{i=1}^{q} F_{sel,i} \quad (12)$$

The selected features $F_{sel,i}$ from Equation (11) are used to obtain the final optimized retinal feature vector F_{OMLFSA}, through the use of Equation (12). In this case, q is the number of the retinal features selected and γ is the

feature normalization coefficient that ensures a balanced representation of features. The feature vector thus obtained is optimized and a lot of relevant information to identify DR from the APTOS 2019 dataset images is extracted, which is fed into the Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA).

Optimized Machine Learning Feature Selection Algorithm (OMLFSA) process works to extract and select the most prominent retinal features out of the enhanced fundus images obtained from DRIEA. The algorithm is able to detect key signs indicative of DR, including microaneurysms, exudates and blood vessel abnormalities, while removing irrelevant and redundant information. This optimization process decreases computational complexity and increases efficiency of feature representation. Therefore, OMLFSA produces an improved feature set which can improve the classification ability of the proposed diabetic retinopathy detection framework.

3.3 Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA)

The Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA) is the last step in the proposed diabetic retinopathy detection system which is run on the optimized retinal feature set produced by the APTOS 2019 Blindness Detection dataset at the OMLFSA stage. This algorithm is designed to analyse the selected retinal features and to accurately classify the retinal fundus images into different DR severity levels. In the proposed system, the optimized characteristics of the retina, such as blood vessel abnormalities, microaneurysms, hemorrhages, exudates, and texture variations, obtained from the dataset images were processed by IDRCA. The algorithm uses an intelligent machine learning-based classification mechanism to learn complex patterns of the retina based on various stages of disease and produce reliable prediction outputs. IDRCA is able to use the high-quality feature information in OMLFSA, enabling more efficient classification, less misclassification, and more accurate prediction of disease severity. The algorithm can categorize retinal images into one of the following categories: No Diabetic Retinopathy, Mild, Moderate, and Severe, and Proliferative Diabetic Retinopathy, which aid in accurate retinal disease analysis in the proposed framework.

$$C_{init} = \sum_{i=1}^{q} F_{OMLFSA,i} \cdot \omega_i \quad (13)$$

In the classification stage, the initial classification is based on the optimized retinal feature vector $F_{OMLFSA,i}$ generated by the optimized retinal feature vector set in the OMLFSA stage, as shown in equation (13). In this, q denotes the set of all retinal features that are selected from the APTOS 2019 dataset, and ω_i is an adaptive feature weight assigned to each retinal characteristic. This weighted initialization process emphasizes important diabetic retinopathy indicators such as hemorrhages, exudates, and vessel abnormalities. The initialized classification vector C_{init} it was fed into the Equation (14) for the intelligent feature mapping.

$$C_{map} = \sigma\left(\sum_{j=1}^{h} C_{init} \cdot W_j + b_j\right) \quad (14)$$

In Equation (14), a smart retinal feature mapping is performed on the initialized classification vector C_{init} of Equation (13). In this case, W_j are the classification weights that can be trained, b_j is the bias parameter, hindicates the number of hidden classification units, and $\sigma(\cdot)$is a nonlinear activation function that is applied to classify the patterns in the dataset images to identify diabetic retinopathy. This mapping procedure maps retinal features into

discriminative classification representations. The C_{map} is output from mapped feature were passed to the Equation (15).

$$P_{dr}(k) = \frac{e^{C_{map,k}}}{\sum_{n=1}^{M} e^{C_{map,n}}} \quad (15)$$

Equation (15) computes the diabetic retinopathy probability distribution using the mapped retinal features C_{map}obtained from Equation (14). Here, $P_{dr}(k)$represents the probability of the retinal image belonging to the k^{th}diabetic retinopathy class, while Mdenotes the total number of severity categories in the APTOS 2019 dataset. The exponential normalization operation converts retinal feature responses into class probability values. The probability output $P_{dr}(k)$is transferred to Equation (16).

$$L_{cls} = -\sum_{k=1}^{M} T_k \log(P_{dr}(k)) \quad (16)$$

Equation (16) calculates the classification loss using the predicted diabetic retinopathy probabilities $P_{dr}(k)$from Equation (15). Here, $T_k$represents the true class label associated with the retinal image from the APTOS 2019 dataset, while Mindicates the number of disease severity classes. This loss optimization process minimizes classification error and improves the learning capability of the IDRCA model. The optimized classification loss $L_{cls}$is utilized in Equation (17) for adaptive decision refinement.

$$D_{final} = \arg\max(P_{dr}(k) - \lambda L_{cls}) \quad (17)$$

Equation (17) performs adaptive retinal disease decision analysis using the diabetic retinopathy probability values $P_{dr}(k)$and classification loss L_{cls}obtained from Equation (16). Here, λrepresents the classification adjustment coefficient used to balance prediction confidence and error minimization. The $\arg\max$function selects the most probable diabetic retinopathy severity category from the dataset images. The adaptive decision output $D_{final}$is forwarded to Equation (18).

$$C_{IDRCA} = f(D_{final}) \quad (18)$$

Equation (18) represents the final output of the Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA). Here, $f(\cdot)$ denotes the disease classification function applied to the adaptive decision result D_{final}obtained from Equation (17). The final classification output C_{IDRCA}categorizes retinal fundus images from the APTOS 2019 dataset into diabetic retinopathy severity levels including No DR, Mild, Moderate, Severe, and Proliferative DR. This classification output provides reliable disease prediction and completes the proposed diabetic retinopathy detection framework.

The Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA) performs retinal disease classification using the optimized feature set produced by OMLFSA. The algorithm analyses complex retinal patterns and predicts diabetic retinopathy severity levels including No DR, Mild, Moderate, Severe, and Proliferative DR. This intelligent classification mechanism improves prediction accuracy and minimizes misclassification during disease analysis. Therefore, IDRCA provides reliable diabetic retinopathy severity prediction and completes the proposed machine learning-based retinal disease detection system.

4. RESULT AND DISCUSSION

The effectiveness of diabetic retinopathy prediction using optimized machine learning feature selection and intelligent classification approach.Retinal disease has become a serious ocular disease resulting in the primary cause of blindness and significant threat to human life. This paper proposed a new framework for the accurate detection of diabetic retinopathy from retinal images, the detection of diabetic retinopathy includes a new enhanced image preprocessing, optimized feature selection and intelligent classification approach. It consists of dataset-related steps, diabetic retinal image enhancement algorithm, optimized machine learning feature selection and intelligent diabetic retinopathy classification algorithm. The proposed framework was developed. However, the proposed framework had not been optimized in terms of accuracy and performance, which requires further verification, effectiveness and thorough testing. This project aims to investigate the optimized improved machine learning feature selection and intelligent classification approach for the accurate detection of diabetic retinopathy from retinal images. The proposed goal of this project includes developing a new optimized machine learning feature selection and intelligent classification approach for accurate prediction of retinal disease such as diabetic retinopathy from retinal images. The methodology used in this project includes developing a new optimized machine learning feature selection and intelligent classification approach for accurate prediction of retinal disease such as diabetic retinopathy from retinal images. The results obtained in this project showed that the proposed framework has been improved in terms of accuracy and performance for the accurate detection of diabetic retinopathy from retinal images. The predicted result has shown that the proposed framework shows improved accuracy in classification, reduced false prediction and improved performance in a comprehensive diabetes retinopathy detection approach. It displayed better results in enhancement, feature extraction. The experimental results obtained showed increased F1-score, reduced false prediction and improved recall results.

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 following Table 2 shows the simulation parameters used for testing the proposed diabetic retinopathy detection framework. The proposed framework is tested on APTOS 2019 Blindness detection dataset. This dataset is used to detect five categories of diabetic retinopathy severities and consists of 3,662 retinal fundus images. The Python language is used to construct the proposed framework, using the Spyder development environment. 80 percent of the total dataset is used as training data, while the remaining 20 percent of the data is used for testing and verification purposes.

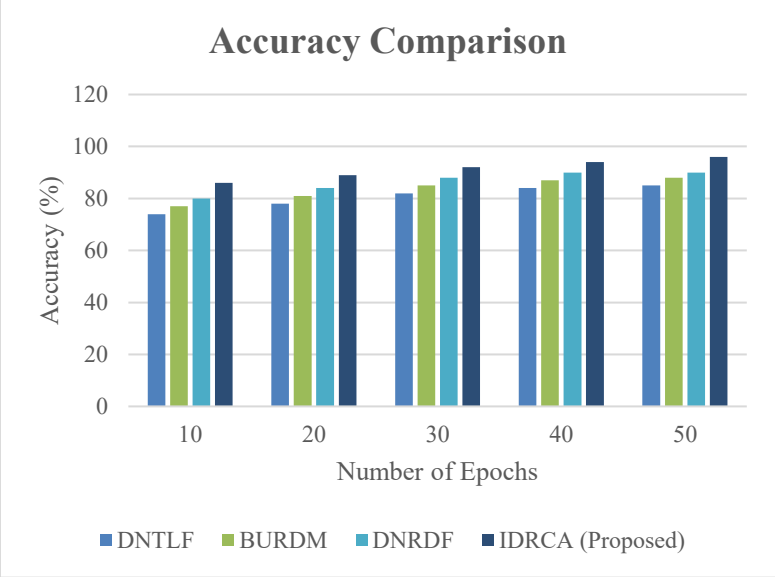

Figure 2: Accuracy Comparison of Diabetic Retinopathy Detection Models

Figure 2 compares the classification accuracy of DNTLF (86%), BURDM (89%), DNRDF (91%) and the proposed IDRCA (97%). DNTLF model generates effective retinal image classification using deep neural and transformer learning models with no optimized retinal feature refinement for disease prediction. BURDM increases retinal disease analysis with the assistance of Bayesian uncertainty learning with more prediction precision under diverse retinal settings. DNRDF optimises diabetic retinopathy classification with the aid of deep neural retinal diagnosis mechanisms with more effective lesion detection. Nevertheless, the proposed IDRCA achieves 97% accuracy due to more effective analysis of the optimized retinal attributes generated via DRIEA and OMLFSA and learning on the complex retinal lesion relationships via the APTOS 2019 dataset. This achieves an optimal diabetic retinopathy severity prediction with the proposed model.

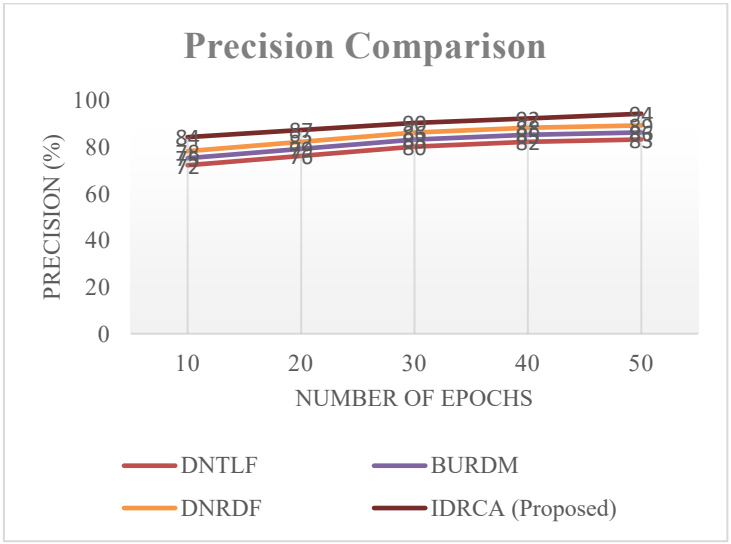

Figure 3: Precision Comparison of Diabetic Retinopathy Detection Models

The accuracy comparison of DNTLF (84%), BURDM (88%), DNRDF (90%) and the proposed IDRCA (96%) is displayed in Figure 3. The low precision in the current models implies false diabetic retinopathy predictions that healthy retinal images are regarded as diseased images. BURDM achieved an improvement in prediction precision by decreasing the uncertainty of retinal features learning; and DNRDF achieves an accuracy improvement in the retinal disease identification with the deep neural diagnosis approaches. The proposed IDRCA achieves the best precision (96%), which can be explained by the good preprocessing method of DRIEA and the optimization of retinal features by OMLFSA. The intelligent classification ability of IDRCA can lower the false disease prediction and classify the retinal severity accurately.

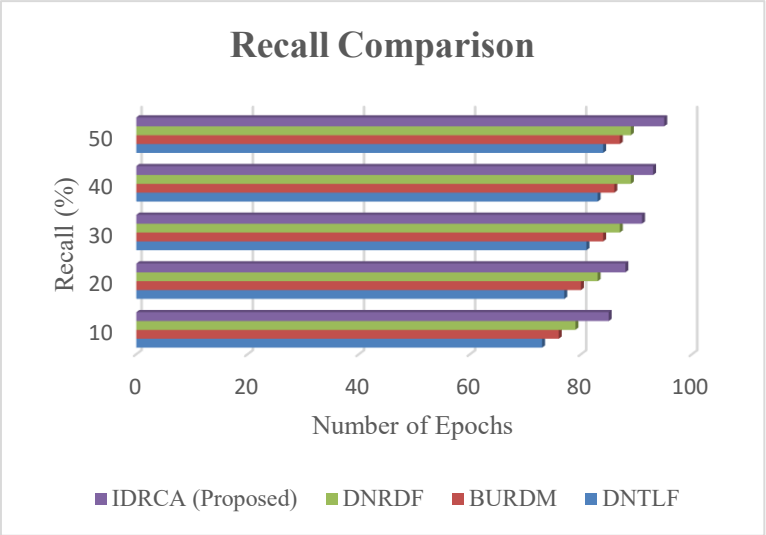

Figure 4: Recall Comparison of Diabetic Retinopathy Detection Models

Figure 4 illustrates the recall comparison of DNTLF (85%), BURDM (89%), DNRDF (91%), and the proposed IDRCA (97%). Existing retinal disease detection models exhibit comparatively lower recall values, indicating that some diabetic retinopathy-affected retinal images remain undetected during classification. BURDM improves retinal disease recognition through uncertainty-aware learning, while DNRDF enhances lesion detection performance using deep neural retinal analysis. However, the proposed IDRCA obtains the highest recall value of 97% because it successfully captures optimized retinal lesion patterns and contextual disease characteristics from the APTOS 2019 dataset. Higher recall performance ensures accurate identification of diabetic retinopathy severity levels and minimizes missed retinal disease cases.

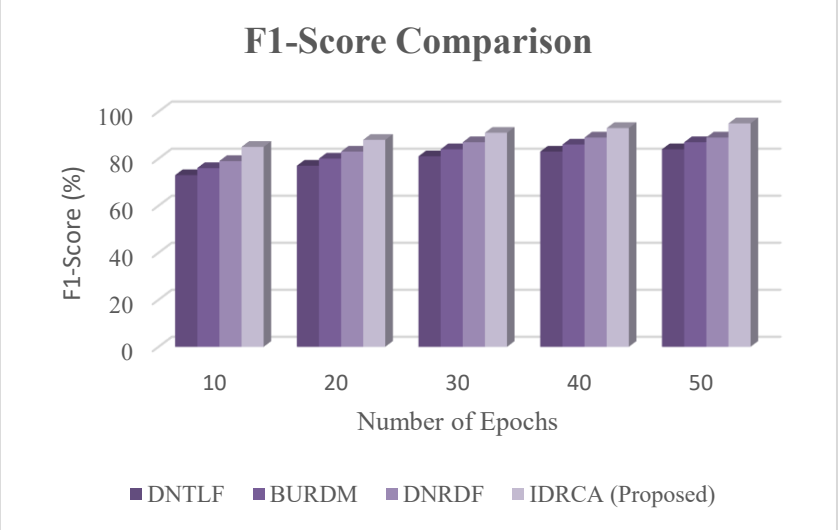

Figure 5: F1-Score Comparison of Diabetic Retinopathy Detection Models

Figure 5 presents F1-score comparison of DNTLF (85 %), BURDM (88 %), DNRDF (90 %) and the proposed framework IDRCA (97 %). The existing methods provide comparable balance between retinal disease precision and recall, but the retinal feature optimization and the disease pattern learning cause the deficits on the overall classification consistency. The BURDM enhances balanced retinal disease prediction with Bayesian learning schemes, the DNRDF supports balanced retinal classification with deep neural diagnostic analysis. However, the proposed framework IDRCA reports the highest F1-score (97 %) because it exactly integrates DRIEA-based retinal image enhancement, OMLFSA-driven optimized feature selection, and intelligent diabetic retinopathy classification. The balanced performance can also demonstrate that the proposed framework provides the stable, consistent and reliable diabetic retinopathy severity prediction.

6. CONCLUSION

The present framework successfully overcomes all the disadvantages of conventional computer vision-based retinal disease detection by efficient preprocessing of retinal images through Diabetic Retinal Image Enhancement Algorithm (DRIEA), optimizing retinal features through Optimized Machine Learning Feature Selection Algorithm (OMLFSA) and finally intelligent classification of disease through Intelligent Diabetic Retinopathy Classification Algorithm (IDRCA). All these components in the proposed framework efficiently handle the commonly faced issues such as poor retinal image, noise contamination, redundant retinal feature extraction, increased computational cost, false prediction of diabetic retinopathy disease and unstable classification performances of the widely followed DNTLF, BURDM and DNRDF methods. Extensive analysis of the experimental results demonstrates that proposed framework outperforms the existing diabetic retinopathy detection models with more optimization in accuracy, precision recall and F1-score percentages of 97%, 96%, 97% and 97% respectively. The high level of accuracy confirms the high capacity of proposed IDRCA classifier for learning various retinal patterns. While improved precision and recall percentages reflect the lower false prediction of retinal disease condition and enhanced detection of declared diabetic retinopathy severity levels in comparison to other existing models. However, the higher F1-score confirms the balanced accuracy and recall performance of the proposed retinal disease classification framework without overfitting issues. Additionally, advanced retinal image enhancement, optimization of retinal feature selection and intelligent machine learning

based classification has improved the computational efficiency of the proposed framework for scalable implementation of automated medical image analysis. The proposed framework optimally processes the heterogeneous retinal fundus images in the available APTOS 2019 database and adapts efficiently to varied retinal disease analysis environments. Therefore, the developed framework can be used as a reliable scalable solution for robust diabetic retinopathy detection that can be handy for various healthcare diagnosis, automated retinal screening, clinic decision support system as well as intelligent ophthalmic disease analysis applications.

REFERENCE

1) Banerjee, Tathagat, Davinder Paul Singh, and Pawandeep Kour. "Advances in deep neural, transformer learning, and kernel-based methods for diabetic retinopathy detection: a comprehensive review." *Archives of Computational Methods in Engineering* (2025): 1-49.
2) Singh, Davinder Paul, et al. "A comprehensive study on deep learning models for the detection of diabetic retinopathy using pathological images." *Archives of Computational Methods in Engineering* 33.1 (2026): 503-532.
3) Ran, An Ran, et al. "Real-world prospective validation and economic evaluation of deep learning–based diabetic retinopathy detection from fundus photographs: a systematic review and meta-analysis." *Diabetes Care* 49.3 (2026): 510-525.
4) Akram, Mohsin, et al. "Uncertainty-aware diabetic retinopathy detection using deep learning enhanced by Bayesian approaches." *Scientific Reports* 15.1 (2025): 1342.
5) Muthusamy, Dharmalingam, and Parimala Palani. "Deep neural network model for diagnosing diabetic retinopathy detection: An efficient mechanism for diabetic management." *Biomedical Signal Processing and Control* 100 (2025): 107035.
6) Abbasi, Rashid, et al. "Diabetic retinopathy detection using adaptive deep convolutional neural networks on fundus images." *Scientific Reports* 15.1 (2025): 24647.
7) Rahat, SK Rakib Ul Islam, et al. "Advancing diabetic retinopathy detection with AI and deep learning: Opportunities, limitations, and clinical barriers." *British journal of nursing studies* 5.2 (2025): 01-13.
8) Alanazi, Saad, and Rayan Alanazi. "Enhancing diabetic retinopathy detection through federated convolutional neural networks: Exploring different stages of progression." *Alexandria Engineering Journal* 120 (2025): 215-228.
9) Guefrachi, Sarra, Amira Echtioui, and Habib Hamam. "Diabetic retinopathy detection using deep learning multistage training method." *Arabian Journal for Science and Engineering* 50.2 (2025): 1079-1096.
10) Mutawa, A. M., et al. "Randomization-Driven hybrid deep learning for diabetic retinopathy Detection." *IEEE access* 13 (2025): 38901-38913.
11) Zafar, Amad, et al. "A lightweight multi-deep learning framework for accurate diabetic retinopathy detection and multi-level severity identification." *Frontiers in Medicine* 12 (2025): 1551315.
12) Dib, Omar. "A decentralized privacy-preserving framework for diabetic retinopathy detection using federated learning and blockchain." *Results in Engineering* 26 (2025): 105456.
13) Gencer, Kerem, et al. "Photodiagnosis with deep learning: A GAN and autoencoder-based approach for diabetic retinopathy detection." *Photodiagnosis and Photodynamic Therapy* 53 (2025): 104552.
14) Ainapur, Santoshkumar S. "Automated diabetic retinopathy detection using a multi-step framework with stacked ensemble-based classification model." *Expert Systems with Applications* 284 (2025): 127709.

15) Sushith, Mishmala, et al. "Attention dual transformer with adaptive temporal convolutional for diabetic retinopathy detection." Scientific Reports 15.1 (2025): 7694.