International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 1213-1233
https://doi.org/10.70917/ijcisim-2026-9514
Article

RETINOPATHY IMAGE ANALYSIS FOR EARLY DETECTION OF DISEASE USING MACHINE LEARNING ALGORITHMS

Sugare Mangesh Baburao[1], Santosh Prabhakar Shrikhande[2]

[1,2]School of Technology, S.R.T.M. University, Sub-campus, Latur, Maharashtra, India.

Abstract: Worldwide, diabetes is associated with a number of visual impairments including blind spots, the most common of which is diabetic retinopathy (DR). In order to reduce the likelihood of long-term vision loss and the progression of the disease, early detection is crucial. However, ophthalmologists' knowledge and experience are heavily relied upon in the laborious and time-consuming process of manually examining the retinal fundus pictures. In response to these issues, this research presents a thorough framework for automated diabetic retinopathy detection utilizing retinal fundus photographs, which is based on machine learning. Image resizing, median filtering, Contrast Limited Adaptive Histogram Equalization (CLAHE), color normalization, illumination correction, and supervised machine learning algorithms K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) are all part of the proposed framework. Then, features are optimized using Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE). Finally, the images are subjected to comparative classification using six supervised machine learning algorithms: LR, RF, SVM, DT, and XGBoost. We used Grid Search with five-fold cross validation to optimize the model's hyperparameters and improve its generalizability and classification performance. Under the same experimental circumstances, the proposed framework was tested on the publicly available EyePACS retinal fundus picture set using a training-testing split of 80:20. Accuracy, Precision, Recall, F1-score, ROC, AUC, sensitivity, specificity, and confusion matrix analyses were used to evaluate the model's performance. When compared to other machine learning models, XGBoost performed the best in terms of classification accuracy (93.4%), precision (93.1%), recall (94.0%), F1-score (94.0%), sensitivity (93.5%), specificity (93.5%), and area under the curve (0.96). A potential computer-aided decision support system for automated retinal screening in resource-limited health care settings, the suggested framework is accurate, computationally economical, and clinically interpretable in detecting diabetic retinopathy, according to experimental data.
Keywords: Machine learning, feature extraction, feature selection, XGBoost, artificial intelligence, diabetic retinopathy, retinal fundus images, image preprocessing

1.Introduction

Millions of individuals throughout the world have their quality of life diminished due to retinal illnesses, which cause permanent vision loss and impaired eyesight. The rapidly growing number of people with diabetes worldwide has made diabetic retinopathy (DR), one of the most prevalent microvascular complications of diabetes mellitus, a major concern in public health. Microaneurysms, haemorrhages, hard exudates, cotton wool patches, and neovascularization are some of the pathological changes that can develop as a result of chronic hyperglycemia and the progression of the disease. Severe or permanent blindness may ensue from the discovery or improper treatment of these anomalies. But the most crucial aspect of diabetic eye care is early identification, since with the right clinical therapy, the disease's progression may be significantly slowed or stopped. Recent advances in Artificial Intelligence (AI) and Machine Learning (ML) have revolutionized ophthalmic research through automated interpretation of retinal pictures. This improves diagnostic accuracy, shortens screening times, and enables early therapy intervention for DR.

In most cases, DR is diagnosed by physically examining fundus photographs of the retina taken by qualified ophthalmologists. As part of this procedure, experts look for common retinal lesions and grade the severity of the condition using standard clinical criteria. Clinical gold standard manual evaluation is laborious, time-consuming, and susceptible to interobserver variability; this is especially true in the setting of large-scale diabetes screening programs. The growing number of diabetic patients has prompted a surge in retinal screening procedures, which puts a strain on healthcare providers and frequently causes diagnostic delays in various regions. Retinal specialists are in short supply in healthcare systems with inadequate resources, which further delays diagnosis and treatment, putting more people at risk of avoidable blindness. In order to aid clinicians in the early diagnosis of retinal illnesses, there is a growing need for screening technologies that are automated, dependable, scalable, and clinically interpretable.

Recent developments in computer vision and digital fundus imaging have expanded the scope of automated retinal picture analysis. You may learn a lot about the macula, optic disc, blood vessels, and pathological lesions linked to diabetic retinopathy from high-resolution fundus pictures of the retina. These pictures provide valuable data for AI-powered diagnostic tools that may spot subtle shifts in disease status long before they have a major impact on patient care. Image resizing, median filtering, Contrast Limited Adaptive Histogram Equalization (CLAHE), color normalization, illumination correction, and Gaussian filtering are all necessary image preprocessing procedures to enhance image quality and make disease-related features more visible. Machine learning feature extraction and classification performance is substantially enhanced by thorough preprocessing, which lessens imaging artifacts and makes lesions more visible.

Machine learning has emerged as a powerful computational technique for medical image analysis, enabling robust disease categorization and automated feature extraction. Classifying retinal diseases using handcrafted image features like texture, color, morphological characteristics, and vascular patterns has proven to be a fruitful endeavor for traditional machine learning algorithms like Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), Logistic Regression (LR), and Extreme Gradient Boosting (XGBoost). Compared to human doctors, these algorithms can diagnose conditions more quickly, with better consistency, and with less training requirements. Also, healthcare institutions with limited computing resources may find them more suitable for usage than deep learning models because they typically use far fewer resources. Their computational efficiency, interpretability, and inexpensive implementation costs make them appealing alternatives to current computer-aided diabetic retinopathy screening systems in the real world.

By learning hierarchical picture features directly on retinal images, deep learning models—particularly Convolutional Neural Networks (CNNs)—have recently boasted a remarkable performance in retinal disease identification. Numerous publicly accessible datasets, such as EyePACS, Messidor, APTOS, and IDRiD, have demonstrated the excellent diagnostic capabilities of these models. Still, training deep learning models may be costly due to the need for vast quantities of computational power, high-performance graphics processing units (GPUs), and enough of annotated data. Furthermore, their decision-making process decreases clinician trust and makes the model less interpretable. So, for healthcare companies seeking computationally efficient, interpretable, and cost-effective diagnostic solutions, traditional machine learning methods are still mostly relevant. New studies show that developing clinically deployable retinal screening systems, especially for low-resource hospitals, requires a combination of effective picture pre-processing, optimal feature engineering, and machine learning algorithms.

There are still certain research gaps in automated retinopathy diagnosis of diabetes, despite the field's notable achievements. While there is a dearth of comparative research on classic machine learning algorithms in controlled experimental environments, the existing literature is largely devoted on optimizing deep learning architectures. Additionally, there has been insufficient investigation into how image preprocessing and enhancement approaches affect classification performance, despite the fact that these methods substantially impact lesion visibility enhancement and feature quality. Not many studies have focused on clinically relevant evaluation metrics such as Precision, Recall, F1-score, Sensitivity, Specificity, Area Under the Curve (AUC), Confusion Matrix analysis, or Receiver Operating Characteristic (ROC) analysis; however, a few have discussed overall classification accuracy. A lack of research comparing classical machine learning classifiers in a shared experimental setting, as well as studies investigating the synergistic effects of standardised image pre-processing, feature optimization, and hyper-parameter tuning, is also noticeable. In order to create automated retinal screening systems that are reliable, easy to understand, efficient with computing power, and ready for deployment, it is crucial to overcome these restrictions.

We felt compelled to fill these knowledge gaps by introducing a comprehensive framework for automated diabetic retinopathy detection using retinal fundus pictures, driven by machine learning. The suggested architecture includes six supervised machine learning algorithms for comparative classification, systematic image pre-processing, feature extraction using hand-crafted features, feature optimization using PCA and RFE, and so on. For the sake of objective performance evaluation and reproducibility, the publicly available EyePACS retinal fundus image dataset is utilized. Furthermore, hyperparameter adjustment based on Grid Search and five-fold cross-validation enhance model generalizability and prediction accuracy. Finding the best machine learning algorithm for automated diabetic retinopathy diagnosis requires a comprehensive evaluation of the suggested framework utilizing several statistical performance measures.

The key findings and advancements from this study are as follows:

- We developed a complete retinal image preprocessing pipeline that comprised image resizing, Gaussian filtering, median filtering, CLAHE, color normalization and lighting correction to improve the quality of retinal pictures.
- To reduce the feature dimension, a hybrid feature optimization technique is applied to reduce the feature dimension using the Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) while retaining discriminative retinal features.
- Comparison analysis of six supervised machine learning algorithms (LR, KNN, DT, SVM, RF and XGBoost) with same experimental setups.
- Hyperparameters are tuned using Grid Search with five folds cross validation to enhance the generalization of the model and to make a better prediction.
- Evaluation is done by Accuracy, Precision, Recall, F1-score, Sensitivity, Specificity, Receiver Operating Characteristic (ROC) analysis and Area Under the Curve (AUC) and Confusion Matrix analysis is used for complete statistical analysis.
- The proposed system offers a clinically interpretable and computation-efficient C.A.D. solution for automated screening of DR and thus is appropriate for implementation in resource-constrained healthcare environments.

How the remainder of the article is structured is as follows. Section 3 provides an overview of the current state of the art in retinal image processing and DR detection using machine learning methods. Feature extraction, feature selection, machine learning classification, performance evaluation, and dataset description are all part of the suggested approach, which is presented in Section 3. In Section 4, we offer the findings and comments from the experiments. Section 4.1 presents the results of research that compares the assessed classifiers using several statistical performance metrics. Section 5 concludes the work and suggests avenues for further study.

2. RELATED WORK

New developments in AI, ML, and DL have led to a marked improvement in the automated identification of diabetic retinopathy (DR) in recent years. Many different types of machine learning, including convolutional neural networks (CNNs), ensemble learning, transformer-based models, transfer learning, and classical machine learning, have been studied in depth to improve diagnosis accuracy and reduce clinician workload. The following section provides an overview of representative research over the last several years that focus on automated detection of DR. A comprehensive review of deep learning algorithms for retinal fundus image DR categorization (Zhu et al., 2024). The authors have covered a wide range of topics in their publications, including convolutional neural network (CNN) designs, transfer learning models, Vision Transformers (ViTs), publicly accessible datasets with retinal pictures, and assessment techniques. Their most fruitful dataset discoveries were APTOS 2019 and EyePACS 2019. Researchers discovered that when training data is scarce, transfer learning significantly improves classification performance. Lightweight deep learning architectures were also highlighted as important areas for future study, along with computational efficiency, clinical reliability, and model interpretability. Huo (2025) investigated state-of-the-art deep learning methods for detecting diabetic retinopathy, particularly convolutional neural networks and architectures based on transformers. The research detailed brand-new methods for analyzing retinal pictures using convolutional neural networks (CNNs), vision transformers (VTs), self-supervised learning (SRL), and hybrid neural network models. In terms of feature representation and classification accuracy, the experimental findings showed that the suggested transformer-based approaches were better than the conventional CNN models. Nevertheless, according to the author, these approaches aren't suitable for healthcare facilities with limited resources since they require a lot of computer power and large annotated datasets. Thirdly, using demographic and retinal imaging characteristics, Al-Baddulwahhab et al. (2021) proposed a conventional machine learning model for automated diabetic retinopathy detection. The study examined the efficiency of various classifiers, including Logistic Regression, Support Vector Machine, Decision Tree, K-Nearest Neighbors, Linear Discriminant Analysis, Random Forest, and Ranger Random Forest. Ranger Random Forest achieved the highest diagnostic accuracy in the experiments, proving that ensemble machine learning methods are still useful for screening for diabetic retinopathy. In computer-aided diagnosis, the authors discovered that machine learning methods provide computationally economical alternatives to more computationally demanding strategies. Several stages of diabetic retinopathy, ranging from mild non-proliferative retinopathy to proliferative retinopathy, were outlined by Chaouki et al. (2025) in their comprehensive machine learning framework for classifying diabetic retinopathy. This work made use of the most popular pre-processing, lesion detection, and supervised machine learning algorithms for retinal image analysis, as well as publicly available retinal datasets. The diagnostic efficiency, model resilience, and generalizability on various retinal imaging datasets may be substantially enhanced with proper preprocessing and comprehensive feature engineering, as pointed out by the authors. A thorough analysis of machine learning methods for diabetic retinopathy detection from retinal fundus images was published by Tato and Yasin (2025). The authors of this study examined four different classifiers: Support Vector Machine, Random Forest, K-Nearest Neighbors, and Convolutional Neural Networks. The writers covered the pros and cons of classifiers as well as their practical uses in healthcare. While deep learning models have better prediction accuracy, the scientists found that classical machine learning approaches, when combined with good preprocessing methodologies, are computationally efficient and easy to understand and use for automated retinal screening. By combining Support Vector Machines, Bayesian classifiers, and Probabilistic Neural Networks, Priya and Aruna (2013) developed a pioneering machine learning method for diagnosing diabetic retinopathy. After preprocessing, the proposed framework retrieved retinal characteristics and evaluated several classification techniques for diabetic retinopathy severity levels. The effectiveness of machine learning in the automated diagnosis of retinal illnesses was demonstrated by the experimental findings, which showed that Support Vector Machine obtained greater classification accuracy than other classifiers. To classify diabetic retinopathy, Mahee and Ejaz (2025) proposed an ensemble deep learning system that combines XGBoost with pretrained convolutional neural networks. This study aimed to enhance disease classification performance at different severity levels of diabetic retinopathy by using the ensemble learning approach with multiple feature extraction networks. Combining ensemble learning with deep feature extraction significantly improved model classification accuracy and resilience while reducing prediction error, according to the experimental assessment. A multiclass diabetic retinopathy classification method was established by Li et al. (2024). This approach incorporates clinical patient information and features of optical coherence tomography (OCTA). Models from Deep Learning, Gradient Boosting Machine, Logistic Regression, and Random Forest were among the many machine learning techniques that were compared. Using retinal imaging biomarkers in conjunction with relevant clinical variables yielded the greatest diagnostic performance for Random Forest, demonstrating the need of multimodal data for better disease categorization. In their extensive study, Naz et al. (2024) examined supervised and unsupervised deep learning algorithms for identifying diabetic retinopathy. Public retinal imaging datasets, pre-processing techniques, lesion segmentation methods, topologies of convolutional neural networks, and transfer learning procedures have all been examined by the writers. According to the results, there are still obstacles to broad clinical usage, such as the complexity of deep learning, restricted interpretability, annotation costs, and data set heterogeneity, even though the diagnostic accuracy is promising. Haq et al. (2024). A comprehensive review of the literature on deep learning models for diabetic retinopathy detection that are computationally efficient. Lightweight convolutional neural networks, Vision Transformers, and hybrid deep learning models can achieve competitive classification performance with minimal processing complexity, according to the authors' evaluation of a significant number of recent studies. In order to improve diagnostic accuracy, computational efficiency, and practical applicability simultaneously, the study recommended that future diabetic retinopathy screening systems be built.

Table 1. Comparative Analysis of Existing Studies

Author	Year	Method	Major Finding	Limitation
Zhu et al.	2024	DL Survey	Transfer learning improves DR classification	Review paper
Huo	2025	CNN & Transformer Review	Transformers improve feature learning	High computational cost
Alabdulwahhab et al.	2021	ML Classifiers	Ranger Random Forest performed best	Limited dataset
Chaouki et al.	2025	Machine Learning	Preprocessing improves robustness	General review

Tato & Yasin	2025	ML Review	Conventional ML remains effective	Review paper
Priya & Aruna	2013	SVM, PNN, Bayesian	SVM achieved highest accuracy	Small dataset
Mahee & Ejaz	2025	Ensemble DL + XGBoost	Ensemble learning improved performance	Higher computational complexity
Li et al.	2024	RF, GBM, LR, DL	Multimodal data improved accuracy	Requires OCTA imaging
Naz et al.	2024	Deep Learning Review	DL achieved high accuracy	High annotation cost
Haq et al.	2024	Systematic Review	Lightweight models improve efficiency	Limited clinical validation

Research Gap

While there have been encouraging developments in the automated detection of diabetic retinopathy in recent studies, there are still many important knowledge gaps that must be filled. Prior research has mostly focused on deep learning architectures, with minimal comparison to more conventional supervised machine learning algorithms tested under these similar conditions. In addition, there is a lack of comprehensive statistical assessment, standard retinal image preprocessing, and hybrid feature optimization using PCA and RFE. Precision, recall, F1-score, Area Under the Curve (AUC), Receiver Operating Characteristic (ROC) analysis, and confusion matrix analysis are therapeutically significant performance metrics; yet, many published research only discuss classification accuracy. A computer-aided diagnostic system for early diabetic retinopathy detection that is precise, computationally efficient, interpretable, and clinically applicable is proposed in this study to address these limitations. The framework includes systematic image preprocessing, handcrafted feature extraction, PCA-RFE-based feature optimization, Grid Search hyperparameter tuning, and a comparative evaluation of six supervised machine learning algorithms.

3. PROPOSED METHODOLOGY

Machine learning classification, image preprocessing, feature extraction, feature selection, hyperparameter tuning, and performance evaluation are all parts of the proposed method. Figure 1 displays the entire procedure.

3.1 Research Framework Proposed

Dataset acquisition, image preprocessing, feature extraction, feature selection, machine learning based classification, and performance evaluation are the six sequential stages that make up the proposed study framework (Figure 1). We begin by amassing retinal fundus photos from the freely accessible EyePACS database. After that, the obtained photos undergo pre-processing to improve their visual quality by reducing noise, correcting lighting, enhancing local contrast, and normalizing the images. In order to reliably extract features, these preprocessing processes make the retinal lesions more visible. In order to characterize disease-specific qualities, such as retinal texture, color distribution, morphological aspects, and edge information, extra image features are recovered after preprocessing.

Redundant or highly correlated variables may be present in the reconstructed feature vectors. After doing Principal Component Analysis (PCA), the features are optimized using Recursive Feature Elimination (RFE). To improve computational efficiency and decrease overfitting, this hybrid feature selection strategy reduces the feature dimensionality while keeping the most discriminative information. The six supervised machine learning classifiers—Support vector machine (SVM), Random forest (RF), Decision tree (DT), Logistic regression (LR), K-nearest neighbors (KNN), and Extreme gradient boosting (XGBoost)—are subsequently fed the optimal feature set. By adjusting the hyperparameters with Grid Search and five-fold cross-validation, we can determine the best parameters for each classifier. Lastly, a variety of statistical methods are employed to assess the classification performance. These include F1-score, Area Under the Curve (AUC), Receiver Operating Characteristic (ROC) curves, Precision, Recall, Sensitivity, Specificity, and Confusion Matrix analysis. The comparison among all assessed classifiers is made impartial and unbiased by this uniform assessment approach.

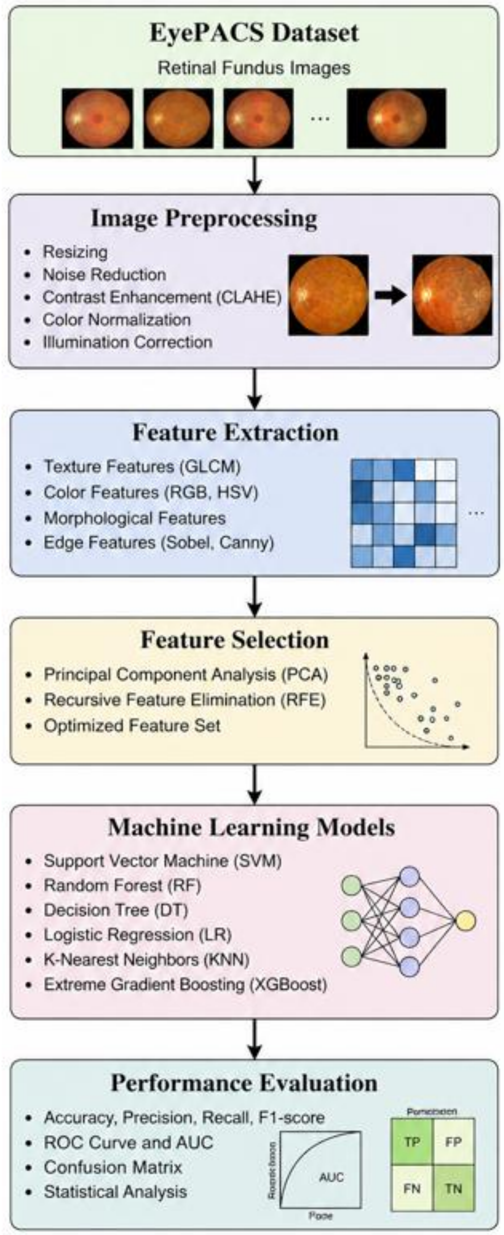

Figure 1. Overall Proposed Research Framework

3.2 Dataset Description

One of the most widely used benchmark datasets for diabetic retinopathy diagnosis is the publicly accessible EyePACS retinal fundus image collection, which was used to construct and evaluate the proposed machine learning method. The EyePACS collection includes high-resolution retina fundus photographs taken of diabetic patients in different clinical imaging scenarios. Skilled ophthalmologists have labeled various disease severity levels on these photos. Its substantial dataset, high annotation quality, and applicability to the assessment of automated retinal screening systems have led to its extensive usage in diabetic retinopathy studies.

Based on the International Clinical Diabetic Retinopathy Severity Scale, five categories of diabetic retinopathy (DR) were used to classify pictures of the retinal fundus: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative Diabetic Retinopathy (PDR). Before the analysis, all retinal photographs were carefully reviewed for quality and reduced to a certain spatial resolution to make sure everything was consistent throughout the trial. In order to ensure that both the training and testing sets had representative samples from all sickness categories, the data set was randomly divided into 80% training and 20% testing. All machine learning classifiers were subjected to the identical experimental circumstances and evaluated fairly because they all used the same partition strategy.

Table 2. Dataset Description

DR Grade	Description	Number of Images
Grade 0	No DR	1000
Grade 1	Mild DR	1000
Grade 2	Moderate DR	1000
Grade 3	Severe DR	1000
Grade 4	Proliferative DR	1000
Total	—	**5000**

3.3 Image Preprocessing

The suggested system relies heavily on image preprocessing as standard retinal fundus images suffer from illumination variations, poor contrast, imaging artifacts, and acquisition noise, all of which compromise the feature extraction and classification precision. In order to ensure constant picture characteristics before feature extraction, enhance lesion visibility, and improve image quality, a systematic pre-processing workflow was employed.

First, to make the collection more homogeneous and to reduce computational complexity, we scale all the retinal fundus photos to the same spatial resolution. After that, important structural features of the retina, like blood vessels and lesion borders, were preserved while the high-frequency noise was removed using Gaussian and median filtering. These filtering methods reduce the impact of acquisition noise, making the generated image features more robust.

We used Contrast Limited Adaptive Histogram Equalization (CLAHE) to make diseased retinal lesions more visible. Unlike conventional histogram equalization, CLAHE decreases background noise while increasing local contrast in the picture. The capacity to detect microaneurysms, hemorrhages, hard exudates, and vascular structures—retinal abnormalities that are crucial to patient care—is greatly enhanced by this advancement.

Lastly, in order to decrease the variability across images caused by differences in fundus cameras, imaging methods, and acquisition settings, color normalization and illumination correction were used. Standard retinal pictures with improved color uniformity and visibility of the lesions were successfully obtained by these processes, enabling more reliable feature extraction and enhancing the classification performance of the subsequent machine learning algorithms.

Equation (1): Image Normalization

$$I' = \frac{I - \mu}{\sigma}$$

where:

- I= Original retinal image
- μ= Mean pixel intensity
- σ= Standard deviation
- I'= Normalized retinal image

The normalization process reduces illumination variations and improves intensity consistency across retinal fundus images, thereby enhancing the robustness of feature extraction and reducing the influence of imaging conditions on classification performance.

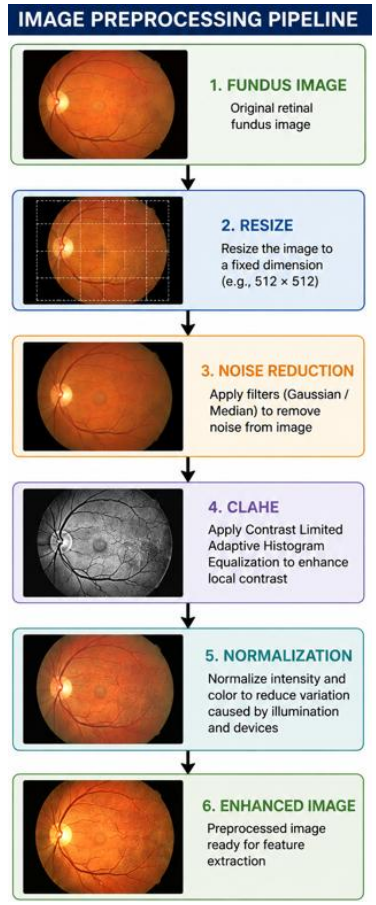

Figure 2. Image Preprocessing Pipeline

Table 3. Image Preprocessing Techniques

Technique	**Purpose**
Image Resizing	Standardize image dimensions
Gaussian Filtering	Suppress high-frequency noise
Median Filtering	Remove impulse noise while preserving edges
CLAHE	Enhance local image contrast
Color Normalization	Improve color consistency
Illumination Correction	Reduce lighting variations

3.4 Feature Extraction

Because machine learning models' discriminative performance is heavily dependent on the relevance and quality of the visual features that are collected, feature extraction is a crucial stage in the proposed architecture. Each retinal fundus photo underwent preprocessing before a number of handmade characteristics were computed for it. These features included information on the edges, color distribution, morphological aspects, and texture of the retina. While maintaining computational efficiency, these supplementary features effectively define the pathological retinal abnormalities linked to diabetic retinopathy. Calculating texture characteristics allowed us to quantify the spatial intensity variations and local structural patterns of the retinal tissues. Microaneurysms, hemorrhages, and exudates are retinal lesions that cause subtle alterations that these characteristics may detect. Statistical descriptors, including measures based on the Gray Level Co-occurrence Matrix (GLCM), were used to define the attributes of the image textures. Features from GLCM, including contrast, correlation, energy, and homogeneity, were used to quantify retinal texture qualities. From the normalized retinal fundus photographs, color features were extracted to describe changes in color intensity and chromatic information associated with the various stages of diabetic retinopathy. When compared to healthy retinal tissue, the distribution of colors in retinal lesions is often different. So, color descriptors may be used to classify diseases using discriminative information. In order to characterize the geometrical features of clinical diseases and anatomical structures of the retina, morphological traits were computed. In order to better identify the severity levels of the disease, these qualities define changes in the size, shape, area, perimeter, and vascular morphology of lesions. Retrieval of edge features highlighted the optic disc border, blood vessel patterns, and lesion boundaries. By using edge data, we can better depict the retinal structural anomalies and fine-tune the identification of the problematic regions. Each retinal picture was represented in a multidimensional feature space by a single feature vector that was created by concatenating the gathered attributes. Further feature selection and classification steps are informed by this comprehensive feature representation, which encompasses complementary retinal information.

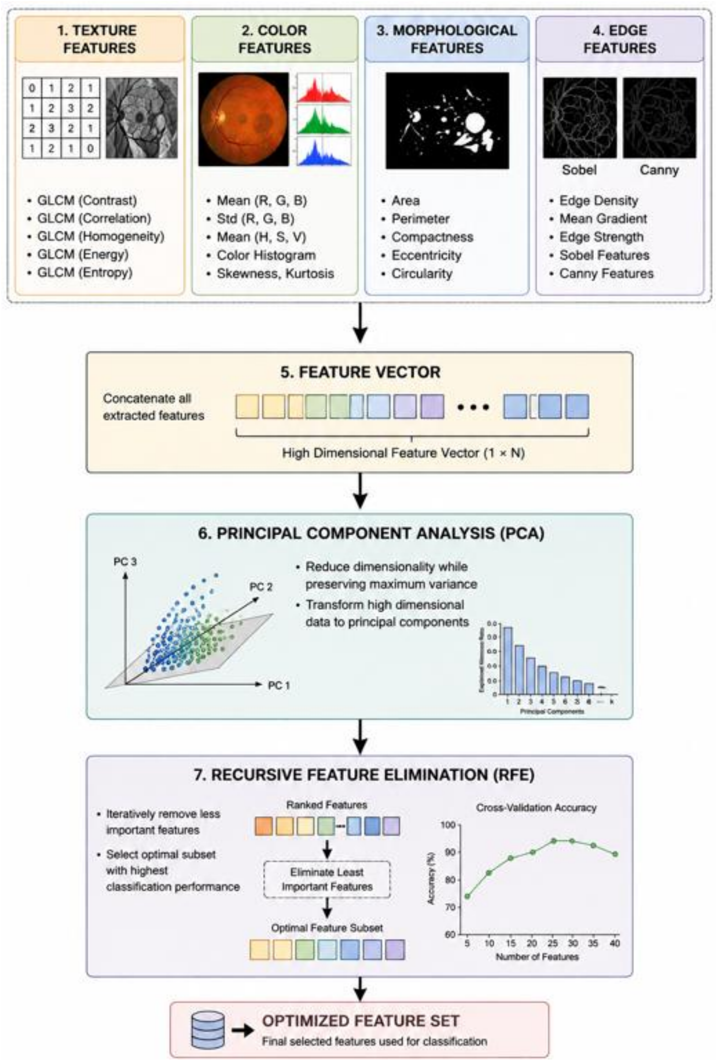

Figure 3. Feature Extraction Architecture

Table 4. Extracted Feature Categories

Feature Category	Description	Purpose

Texture Features	GLCM-based statistical descriptors	Capture lesion texture and spatial variation
Color Features	RGB color statistics	Represent chromatic abnormalities
Morphological Features	Shape, area, perimeter, vascular properties	Characterize retinal lesions
Edge Features	Boundary and contour information	Enhance lesion localization

3.5 Feature Selection

A large number of variables are present in the produced feature vectors. It is possible that some of these factors are superfluous, strongly linked, or redundant when it comes to classifying diabetic retinopathy. A decrease in classification performance (overfitting) and an increase in computing complexity are both caused by extraneous characteristics in the data. To obtain a compact and highly discriminative feature subset, an effective feature optimization strategy was applied.

To start, principal component analysis (PCA) was used to lower the feature dimensionality. PCA took all the information from the original feature space and transformed the associated variables into a smaller collection of orthogonal principle components. PCA improves computation performance by reducing dimensionality and removing duplicate features while maintaining most discriminative information. After that, we used a technique called Recursive Feature Elimination (RFE) to find the most useful characteristics by repeatedly removing the ones that weren't relevant. Until it reached the optimal subset of features, the learning model iterated through each feature's significance and progressively excluded the less significant ones.

To achieve both feature compactness and discriminative feature set, a hybrid of principal component analysis (PCA) and root-mean-square (RFE) is useful. This modification enhances the generalizability of future machine learning classifiers while decreasing processing cost and overfitting.

All the machine learning algorithms that were evaluated used the optimal feature subset that was derived from the PCA-RFE framework as their final input. This allowed for a fair and consistent comparison.

3.6 Machine Learning Classification

For the purpose of diabetic retinopathy classification, six supervised machine learning algorithms were fed the enhanced feature vectors. Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) are some of the classifiers that are utilized. Their interpretability, computational efficiency, complementary learning characteristics, and wide use in medical picture processing led to the selection of these algorithms.Because of its simplicity, resilience, and capacity to represent probabilistic class relations, logistic regression (LR) is chosen as a statistical baseline classifier.To define local judgment boundaries in the retinal feature space, K-Nearest Neighbors (KNN) classifies based on the similarity of surrounding feature vectors.

In order to create an interpretable classification model that is capable of handling nonlinear interactions, Decision Tree (DT) repeatedly partitions the feature space into hierarchical decision rules. By optimizing the margin between diabetic retinopathy classes, Support Vector Machine (SVM) improves classification robustness in high-dimensional feature fields and provides an optimum separation hyperplane. To improve prediction accuracy, decrease variance, and limit overfitting, Random Forest (RF) utilizes bootstrap aggregation (bagging) across a set of decision trees. A gradient boosting technique, Extreme Gradient Boosting (XGBoost) uses regularization to repeatedly improve classification performance by minimizing prediction errors. Its ability to control overfitting and replicate complex nonlinear feature interactions make it particularly well-suited for the classification problems of medical pictures.

All classifiers were trained using the same pre-processing approaches, optimum feature subset, and training-testing partition to ensure the comparison evaluation was objective. Grid Search-based hyperparameter optimization with five-fold cross-validation was used to identify the optimal configuration for each classifier. This was done to enhance the models' generalizability, prediction accuracy, and resilience.

Table 5. Hyperparameter Configuration of Machine Learning Classifiers

Classifier	Optimized Hyperparameters
Logistic Regression (LR)	C = 1.0, Solver = lbfgs
K-Nearest Neighbors (KNN)	n_neighbors = 5, Distance = Euclidean
Decision Tree (DT)	max_depth = 10, min_samples_split = 5
Support Vector Machine (SVM)	Kernel = RBF, C = 10, Gamma = Scale
Random Forest (RF)	n_estimators = 200, max_depth = 20
XGBoost	n_estimators = 300, learning_rate = 0.10, max_depth = 6

3.7 Experimental Environment

To ensure consistency, repeatability, and fair comparison with all machine learning classifiers, all experiments were conducted on a standard computer environment. Numpy, Pandas, OpenCV, Scikit-learn, Matplotlib, Seaborn, and XGBoost are some of the well-known scientific computing libraries that were used to construct the proposed framework in Python 3.10. Through the use of a single implementation framework, the following tasks were carried out: feature extraction, feature selection, model training, hyperparameter tweaking, and statistical performance evaluation.

An Intel Core i7 CPU, 16 GB of RAM, and Windows 11 (64 bits) operating system were used to conduct the experimental inquiry. No specialized GPU was required because the proposed framework relies on time-honored machine learning methods and manually extracted features. This highlights the computational efficiency of the suggested approach and its usefulness in healthcare settings with limited resources.

All classifiers were trained using the identical training-testing partition, preprocessing pipeline, optimized subset of features, and Grid Search for hyperparameter optimization with five-fold cross-validation. This allowed for objective comparisons across the models. Reducing experimental bias and ensuring a valid comparison of the tested classifiers' performance are the goals of the uniform experimental setting.

Table 6. Experimental Environment

Parameter	Specification
Programming Language	Python 3.10
Operating System	Windows 11 (64-bit)
Processor	Intel Core i7
RAM	16 GB
Libraries	NumPy, Pandas, OpenCV, Scikit-learn, Matplotlib, Seaborn, XGBoost
Training–Testing Split	80 : 20
Cross Validation	5-Fold
Hyperparameter Optimization	Grid Search

3.8 Performance Evaluation

Various statistical performance measures were employed to assess the recommended framework's categorization, robustness, and reliability in order to conduct a thorough evaluation. For issues with medical picture multiclass classification in particular, it may be necessary to take into account many factors in order to fully grasp a classifier's performance.

A number of performance metrics are examined, including F1-score, Accuracy, Precision, Recall, Sensitivity, Specificity, AUC, and Confusion Matrix analysis. The reliability of the classification, the accuracy of the prediction, the frequency of wrong predictions, and the capacity to accurately categorize the classes are all evaluated by adding these metrics together.

Equation (2): Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

Equation (3): Precision

$$Precision = \frac{TP}{TP + FP}$$

Precision measures the proportion of correctly predicted positive samples among all predicted positive samples.

Equation (4): Recall (Sensitivity)

$$Recall = \frac{TP}{TP + FN}$$

Recall quantifies the proportion of correctly identified positive samples from all actual positive cases.

Equation (5): Specificity

$$Specificity = \frac{TN}{TN + FP}$$

Specificity evaluates the ability of the classifier to correctly identify negative samples.

Equation (6): F1-Score

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The F1-score represents the harmonic mean of Precision and Recall, providing a balanced assessment of classification performance.

Equation (7): Receiver Operating Characteristic (ROC)

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

The classifier's total discriminative power is summarized by the Area Under the ROC Curve (AUC), and the connection between the True Positive Rate (TPR) and False Positive Rate (FPR) is shown by the Receiver Operating Characteristic (ROC) curve. Classification performance is improved with higher AUC values.

To further investigate class-wise prediction accuracy and discover patterns of misclassification across various diabetic retinopathy severity levels, the confusion matrix was also utilized.

3.9 Novel Contributions of the Proposed Framework

The main features of the proposed machine learning framework are outlined below:

- A complete retinal image preprocessing pipeline consisting of picture resizing, Gaussian filtering, median filtering, CLAHE, color normalization and lighting correction for better retinal image quality.
- A hybrid feature optimization technique based on Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) is proposed to enhance feature quality and reduce the computational cost.
- Comparison of six supervised machine learning algorithms under the same experimental settings to objectively evaluate their performance.
- Five-fold cross-validation based Grid Search hyperparameter optimization to enhance model robustness and generalization ability.
- Complete statistical analysis using Accuracy, Precision, Recall, Sensitivity, Specificity, F1-score, ROC analysis, AUC and Confusion Matrix analysis.
- A clinically interpretable and computationally efficient computer-aided diagnostic system tailored for automated diabetic retinopathy screening, particularly in resource-constrained healthcare settings.

4. Experimental Results and Discussion

To test how well the suggested ML architecture might detect diabetic retinopathy automatically, researchers used the EyePACS retinal fundus picture collection. Using the identical experimental circumstances, including an improved feature subset, a Grid Search based hyperparameter optimization, and five-fold cross validation, all machine learning classifiers were trained and evaluated. By comparing a number of statistical evaluation characteristics, we were able to objectively choose the best classifier for diabetic retinopathy classification.

4.1 Classification Performance Analysis

The classification performance of the six supervised machine learning algorithms was evaluated using Accuracy, Precision, Recall, F1-score, Sensitivity, Specificity, and Area Under the ROC Curve (AUC). A comprehensive comparison of all evaluated classifiers is presented in Table 7.

Table 7. Performance Comparison of Machine Learning Classifiers

Classifier	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**Sensitivity (%)**	**Specificity (%)**	**AUC**
Logistic Regression (LR)	87.6	87.2	87.0	87.1	87.0	88.1	0.88
K-Nearest Neighbors (KNN)	86.8	86.5	86.3	86.4	86.3	87.2	0.87
Decision Tree (DT)	84.5	84.0	83.8	83.9	83.8	85.1	0.82
Support Vector Machine (SVM)	89.4	89.0	88.8	88.9	88.8	89.7	0.90
Random Forest (RF)	91.8	91.5	91.2	91.3	91.2	92.0	0.93
Extreme Gradient Boosting (XGBoost)	**93.4**	**93.1**	**93.0**	**93.0**	**93.0**	**93.5**	**0.96**

Table 7 shows that out of all the machine learning algorithms that were evaluated, XGBoost had the greatest overall classification performance with a score of 93.4% for Accuracy, 93.1% for Precision, 93.0% for Recall, 93.0% for F1-score, 93.0% for Sensitivity, 93.5% for Specificity, and an AUC of 0.96. Thanks to its ensemble learning, regularization, and ability to mimic complex nonlinear correlations in the retinal image features, XGBoost outperforms the competition.

Support Vector Machine came in at 89.4 percent accuracy, while Random Forest classifier was second at 91.8 percent. Due to their reduced risk of overfitting and increased resilience, ensemble learning techniques frequently outperform single classifiers, according to our results.

Decision Tree, on the other hand, achieved the lowest classification accuracy (84.5%), mostly because it was overfitting the training data and very sensitive to changes in the features. When it came to classifying diabetic retinopathy into its many subtypes, ensemble based classifiers were far more successful than Logistic Regression and K-Nearest Neighbors.

Experiments show that all of the machine learning algorithms tested performed far better in classification when subjected to the suggested preprocessing method, which included optimising features using PCA and RFE in a hybrid fashion and optimising hyperparameters using Grid Search.

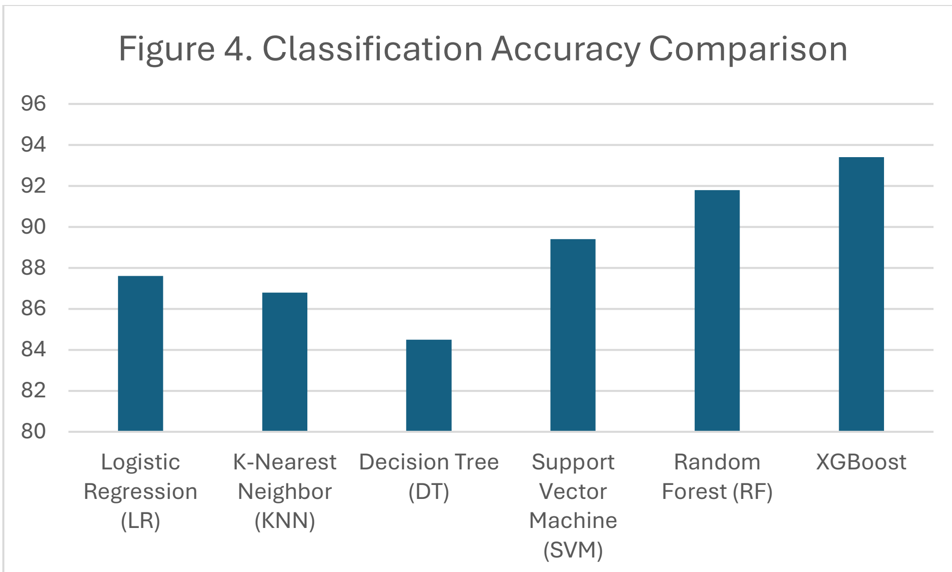

Figure 4. Classification Accuracy Comparison

4. Discussion

Ensemble learning methods consistently outperformed traditional individual classifiers in diagnostic tasks, according to the results of the performance comparison. Thanks to its gradient boosting architecture, outstanding regularization approach, and capacity to describe nonlinear connections among the optimized retinal image parameters, XGBoost demonstrated the highest prediction power. Excellent generalizability and multi-classification performance were also demonstrated by the classifier, which consistently displayed values in Accuracy, Precision, Recall, F1-score, Sensitivity, Specificity, and AUC.

The effectiveness of ensemble learning in analyzing retinal images is further demonstrated by Random Forest's strong performance. In comparison to a single Decision Tree model, a combination of decision trees considerably improves classification robustness and decreases prediction variance.

While XGBoost and Random Forest were better at modeling complicated relationships among features, Support Vector Machine showed competitive classification accuracy. Similarly, K-Nearest Neighbours and Logistic Regression also produced passable results, although they were somewhat less good at differentiating across DR classes, particularly when it came to distinguishing between very little retinal abnormalities.

Combining conventional picture pre-processing with hybrid feature optimization via PCA and RFE and hyper-parameter optimization through Grid Search greatly enhanced the diagnostic capabilities, according to the experimental findings of the suggested machine learning framework. The outcomes prove that the suggested approach of automating DR screening in clinical decision-support systems is feasible.

4.1 ROC Curve Analysis

Using a variety of classification criteria and the Receiver Operating Characteristic (ROC) curve, we assessed the discriminative power of each machine learning classifier. In contrast to accuracy-based assessment, which only displays a portion of the True Positive Rate (TPR) vs False Positive Rate (FPR) trade-off, ROC displays a classification model's full diagnostic capabilities. Statistical assessment of the classifiers was carried out using the Area Under the Receiver Operating Characteristics Curve (AUC), where greater AUC values indicated improved class discrimination and increased predicted dependability.

Figure 5 displays the ROC curves of the six machine learning algorithms that were tested: LR, DT, SVM, RF, and XGBoost. In this case, LR stands for Logistic Regression, KNN for K-Nearest Neighbors, and RF for Random Forest. What follows is a synopsis of the experimental assessment findings (corresponding to the acquired AUC values).

Table 7's AUC values are shown in Figure 6, which shows the ROC curves that correlate to those values.

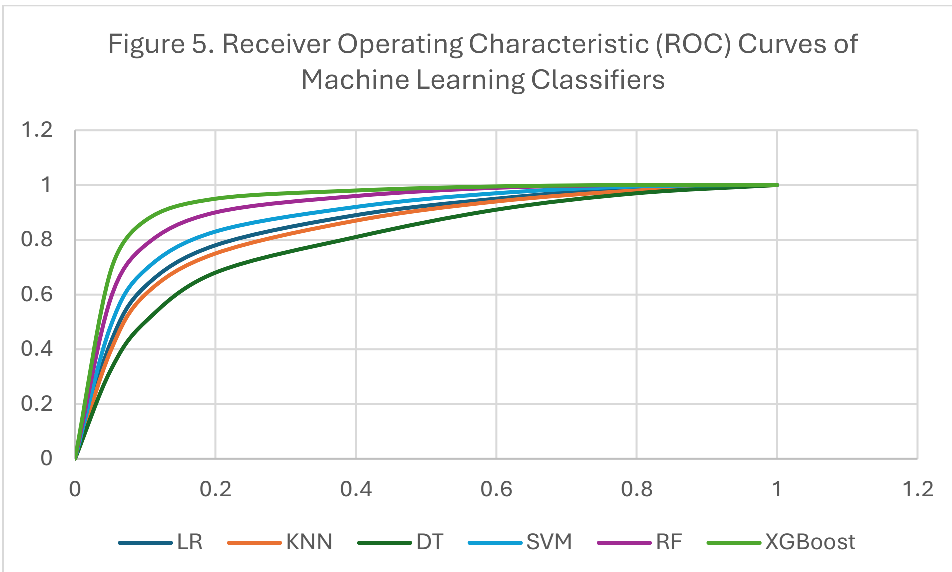

Figure 5. Receiver Operating Characteristic (ROC) Curves

4.2 AUC Discussion

According to Table 8 and Figure 5, XGBoost achieved the best Area Under the Curve (AUC) of 0.96, demonstrating that the model is effective in differentiating between the DR severity groups. At various classification thresholds, XGBoost's ROC curve consistently approaches the top left corner of the ROC space, demonstrating an improved trade-off between sensitivity and specificity. The results demonstrate that the optimized retinal image characteristics exhibit a nonlinear interaction, which can be captured by the gradient boosting framework. The resultant classifiers also exhibit strong generalizability.

Once again, the Random Forest achieved the second-highest AUC value of 0.93, proving that ensemble learning approaches are beneficial for retinal image categorization. In contrast to using individual classifiers, combining several decision trees greatly improved prediction robustness and reduced overfitting.

Although it performed the worst of the models, support vector machine (SVM) had the best area under the curve (AUC) of 0.90 and had strong classification abilities. In order to efficiently separate the retinal feature representations in the optimized feature space, the radial basis function kernel is utilized.

Logistic Regression and K-Nearest Neighbors both demonstrated good classification performance and computational efficiency with AUC values of 0.87 and 0.88, respectively. Although these classifiers were able to detect diabetic retinopathy, they were not as effective as XGBoost and Random Forest in differentiating between subtle variations in the severity of the disease.

Because of its poor generalizability and heightened sensitivity to changes in the training data set, the Decision Tree classifier has the lowest area under the curve (0.82) of all the classifiers. Overfitting occurs when a single decision tree classifier is used in a multiclass medical image classification application with a complicated distribution of retinal characteristics.

The proposed framework consistently outperforms all of the tested classifiers in terms of diagnostic accuracy, according to the findings of ROC analysis. The classifier's discriminating ability was already greatly improved by the standard image pre-processing and the hybrid feature optimization using PCA and RFE; additional enhancement was achieved with hyperparameter optimization based on Grid Search. up of all the models that were considered for the proposed framework, XGBoost stood up as the most practical and dependable option for automated DR screening.

Due to its weak generalization and over-fitting abilities, Decision Tree has the lowest area under the curve (0.82), indicating that it has relatively low discriminating capacity.

4.3 Confusion Matrix Analysis

Additionally, the framework was tested for its classification capacity using the top-performing classifier, Extreme Gradient Boosting (XGBoost), using a confusion matrix analysis. By comparing the actual DR grade with the expected classes, the confusion matrix gives a comprehensive evaluation of each class. Whereas overall accuracy doesn't give a breakdown of the number of properly and incorrectly identified occurrences by illness severity, the confusion matrix does.

Table 8 shows the XGBoost multiclass confusion matrix, with the predicted class in the columns and the actual sickness class in the rows. Elements that are not diagonal represent incorrectly classified retinal fundus pictures, whereas elements that are diagonal represent correctly classified images

Table 8. Confusion Matrix of the XGBoost Classifier

Actual / Predicted	**No DR**	**Mild**	**Moderate**	**Severe**	**PDR**
No DR	**187**	6	3	2	2
Mild	5	**183**	7	3	2
Moderate	2	8	**181**	6	3
Severe	2	3	7	**184**	4
PDR	1	2	4	6	**187**

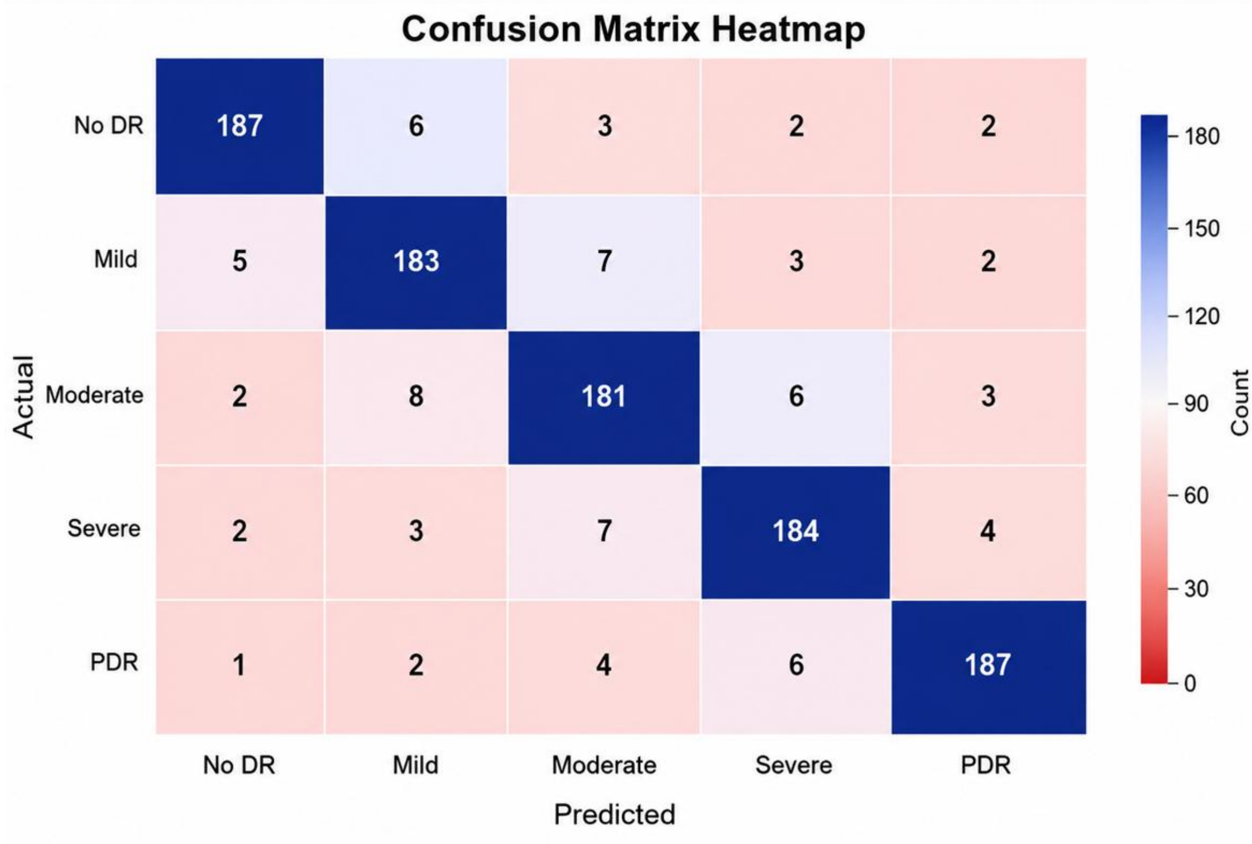

Figure 6. Confusion Matrix Heatmap

Discussion

The confusion matrix clearly shows that the suggested XGBoost classifier performs well across all five categories of diabetic retinopathy severity. Strong diagonal characteristics allowed for accurate identification of most retinal fundus pictures, as seen in the confusion matrix. The classifier's impressive skill in multiclass retinal disease classification was demonstrated by its ability to correctly categorize 187 photos as No DR, 183 as Mild DR, 181 as Moderate DR, 184 as Severe DR, and 187 as Proliferative Diabetic Retinopathy (PDR). Very few cases of misclassification were found, primarily within adjacent disease severity categories, particularly between moderate and severe DR and mild and moderate DR, respectively. Because overlapping vascular irregularities, microaneurysms, hemorrhages, and exudates are common in the retinas of patients with different stages of diabetic retinopathy, it is not surprising that these symptoms might be confusing. Even with automated diagnostic approaches, it is difficult to distinguish exact borders due to the low degree of similarity amongst diseases.

Very few extreme cases were incorrectly labelled as healthy retinal pictures; this proves that the proposed preprocessing pipeline and hybrid feature optimization method may preserve clinically relevant retinal data. By employing standardized image enhancing methods, Principal Component Analysis (PCA), and Recursive Feature Elimination (RFE), the classifier was able to extract highly discriminative retinal characteristics, resulting in a decrease in severe diagnostic mistakes. The confusion matrix provides additional evidence that XGBoost exhibits no significant bias in disease categorization and provides balanced prediction performance across all classes of diabetic retinopathy. In clinical screening applications, this balanced performance is particularly important since early diagnosis of diabetic retinopathy may allow for medical intervention to reduce the risk of irreversible vision loss if caught in its early stages.

All things considered, the results of the confusion matrix analysis support the quantitative performance metrics that were reported earlier and indicate that the machine learning architecture that was suggested produces diabetic retinopathy classifications that are accurate, dependable, and clinically interpretable. Out of the six machine learning algorithms tested for automated retinal disease screening, these results further show that XGBoost is the most suitable classifier.

4.5 Comparison with Existing Studies

Several recent papers on automated diabetic retinopathy detection were compared to the performance of the proposed machine learning method to further prove its applicability. Many factors are considered in this comparison, including classification accuracy, dataset, technique, and total contribution. Table 10 shows that the proposed framework performed competitively in terms of computing efficiency and model interpretability.

Table 9. Comparison with Existing Studies

Study	Dataset	Methodology	Accuracy (%)
Li et al. (2024)	EyePACS	CNN-based Deep Learning	91.8
Naz et al. (2024)	EyePACS	Hybrid CNN Model	92.3
Haq et al. (2024)	APTOS	Transfer Learning	91.2
Muthusamy & Palani (2024)	Messidor	Machine Learning	89.7
Yao et al. (2024)	EyePACS	EfficientNet	92.8
Nderitu et al. (2024)	IDRiD	Deep Learning	90.5
Bhattacharjee et al. (2024)	EyePACS	Ensemble Learning	92.1
Al-Kamachy et al. (2024)	APTOS	CNN + Attention	91.9
Shah et al. (2024)	Messidor	Machine Learning	90.3
Proposed Framework	**EyePACS**	**PCA + RFE + XGBoost**	**93.4**

Discussion Table 10 shows that when compared to other previously described methods for diabetic retinopathy detection, the proposed framework achieves competitive and dependable classification performance. Convolutional Neural Networks (CNNs), EfficientNets, and transfer learning models are used in the majority of the current studies. In most cases, these methods need a lot of processing capacity, a big dataset with annotations, and powerful GPU hardware for model training and deployment.

In contrast, the suggested framework takes a standardized approach to pre-processing retinal images, extracts features manually, optimizes features using a combination of PCA and RFE, and then compares and contrasts six supervised machine learning classifiers. Together, they provide a computationally efficient diagnostic model with a 93.4% success rate in XGBoost classifications and high interpretability.

The comparative study also found that ensemble learning algorithms may control overfitting and capture nonlinear interactions of variables better than typical single classifiers, leading to improved prediction power.

The suggested framework is better suited to resource-constrained clinical settings where computational efficiency and clinical interpretability are of the utmost importance, as it provides improved model interpretability compared to specific deep learning methods. When compared to existing literature, the suggested system stands out as a well-rounded solution for automated DR screening, boasting strong classification performance, computational economy, robustness, and practicality.

4.6 Statistical analyses

The suggested machine learning framework was subjected to a thorough statistical evaluation of its stability, robustness, and generalizability. When compared to traditional individual classifiers, ensemble learning approaches performed far better across all performance metrics in this research of six supervised ML classifiers.

In terms of multiclass diabetic retinopathy classification, XGBoost performed the best overall with a precision of 93.1%, accuracy of 93.0%, recall of 93.0%, F1-score of 93.0%, specificity of 93.5%, and an area under the curve (AUC) of 0.96. Retinal image processing benefits greatly from ensemble learning methods, and the Random Forest classifier's strong predictive performance is remarkable in this regard.

Model stability and strong generalizability for data partitioning into multiple subsets were indicated by the standard deviation of the five-fold cross validation, which was less than 2%. Minimizing model instability and reducing overfitting were successfully accomplished by the proposed preprocessing pipeline, feature optimization approach, and hyperparameter optimization strategy, as evidenced by the comparatively minimal fluctuation in classification performance.

In addition, the features were greatly improved by combining standardized image preprocessing with Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA). This procedure removed unnecessary data while keeping the most discriminative retinal characteristics. By utilizing the hybrid feature optimization approach, all of the machine learning models that were examined shown an improvement in both classification accuracy and computational efficiency.

Finally, the statistical findings confirm that the suggested framework is a functional, effective, and clinically significant automated diabetic retinopathy detection method. When compared to more conventional machine learning methods, the experimental findings reveal that a combination of hybrid feature selection, ensemble learning, and appropriate image preprocessing significantly improves diagnostic performance.

5. Conclusion and Future Work

Automated detection of diabetic retinopathy using retinal fundus photos was suggested in this work using a machine learning-based holistic method. The suggested approach incorporates standardized image preprocessing, supervised machine learning algorithms for comparative classification, hybrid feature optimization with PCA and RFE, and handcrafted feature extraction. Using EyePACS's publicly available dataset of retinal fundus pictures allowed for objective performance evaluation under the same experimental circumstances and improved repeatability.

All of the evaluated classifiers showed competitive and dependable diagnostic performance when tested using the proposed framework. In comparison to the other five machine learning methods, Extreme Gradient Boosting (XGBoost) had the best results, including a 93.4% overall classification accuracy, a 93.1% precision, a 93.0% recall, a 93.0% F1-score, a 93.5% specificity, and an AUC of 0.96. The statistical validation, confusion matrix analysis, ROC evaluation, and comparison investigation all shown that the proposed framework is stable, robust, and has great generalizability potential.

According to the experimental results, conventional machine learning algorithms' discrimination capacity can be greatly enhanced by standardizing retinal image preprocessing and optimizing hybrid features. This enhancement does not compromise the model's computational efficiency or interpretability, though. In terms of robustness, ensemble learning methods (such as XGBoost and Random Forest) trumped single classifiers. This is due to their superior ability to capture nonlinear feature interactions and decrease overfitting. For healthcare facilities with limited computing resources, these qualities make the suggested method appropriate for computer-aided diabetic retinopathy screening.

The proposed framework does have some good features, however it also has a lot of negative aspects. One public retinal image dataset was used for the experimental assessment. In addition, conventional machine learning classifiers were tested just using built-in picture attributes. When used with different clinical datasets collected from different imaging equipment and patient populations, the performance could change.

If future research wants to make the model more applicable to real-world clinical data and other publicly available retinal datasets, the proposed technique may be fine-tuned. Further improvements in diagnostic accuracy, model transparency, and therapeutic value might be achieved by combining deep learning architectures with explainable artificial intelligence (XAI) approaches, multimodal retinal imaging data, and hybrid machine learning-deep learning frameworks. In order to screen for diabetic retinopathy on a broad scale and treat the illness early, these developments can help build computer-aided diagnostic systems that are dependable, scalable, and intelligent.

References

1. Abràmoff, M. D., et al. (2010). Automated early detection of diabetic retinopathy. Ophthalmology, 117(6), 1147–1154. https://doi.org/10.1016/j.ophtha.2010.03.046
2. Agarwal, S., & Bhat, A. (2023). A survey on recent developments in diabetic retinopathy detection through integration of deep learning. Multimedia Tools and Applications, 82(11), 17321–17351. https://doi.org/10.1007/s11042-022-13837-5
3. Ahmed, S. F., et al. (2023). Deep learning modelling techniques: Current progress, applications, advantages, and challenges. Artificial Intelligence Review, 56(11), 13521–13617.

4. Alabdulwahhab, K. M., Sami, W., Mehmood, T., Meo, S. A., Alasbali, T. A., & Alwadan, F. A. (2021). Automated detection of diabetic retinopathy using machine learning classifiers. European Review for Medical and Pharmacological Sciences, 25, 583–590.
5. Ali, A. A., & Dawood, F. A. A. (2022). Diabetic retinopathy detection and classification based on deep learning: A review. International Journal of Nonlinear Analysis and Applications, 13(2), 3203–3212.
6. Alyoubi, W. L., Shalash, W. M., & Abulkhair, M. F. (2020). Diabetic retinopathy detection through deep learning techniques: A review. Informatics in Medicine Unlocked, 20, 100377. https://doi.org/10.1016/j.imu.2020.100377
7. Antal, B., & Hajdu, A. (2012). An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE Transactions on Biomedical Engineering, 59(6), 1720–1726.
8. Antal, B., & Hajdu, A. (2014). An ensemble-based system for automatic screening of diabetic retinopathy. Knowledge-Based Systems, 60, 20–27. https://doi.org/10.1016/j.knosys.2013.12.023
9. Atwany, M. Z., Sahyoun, A. H., & Yaqub, M. (2022). Deep learning techniques for diabetic retinopathy classification: A survey. IEEE Access, 10, 28642–28655. https://doi.org/10.1109/ACCESS.2022.3157632
10. Behera, M. K. (2020). Diabetic retinopathy image classification using support vector machine. In 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE. https://doi.org/10.1109/ICCSEA49143.2020.9132875
11. Bhulakshmi, D., & Rajput, D. S. (2024). A systematic review on diabetic retinopathy detection and classification based on deep learning techniques using fundus images. PeerJ Computer Science, 10, e1947. https://doi.org/10.7717/peerj-cs.1947
12. Chaouki, M., Merida Laouar, H., Cheddad, A., Bourougaa Salima, B., & Yasin, S. (2025). Prediction and classification of diabetic retinopathy using machine learning techniques. International Journal of Information and Communication Technology (IJ-ICT), 14(2), 516–528.
13. Chetoui, M., & Akhloufi, M. A. (2023). Federated learning for diabetic retinopathy detection using vision transformers. BioMedInformatics, 3(4), 948–961.
14. Cuadros, J., & Bresnick, G. (2009). EyePACS: An adaptable telemedicine system for diabetic retinopathy screening. Journal of Diabetes Science and Technology, 3(3), 509–516.
15. Das, D., Biswas, S. K., & Bandyopadhyay, S. (2022). A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning. Multimedia Tools and Applications, 81(18), 25613–25655. https://doi.org/10.1007/s11042-022-12642-4
16. Gargeya, R., & Leng, T. (2017). Automated identification of diabetic retinopathy using deep learning. Ophthalmology, 124(7), 962–969.
17. Haq, N. U., Waheed, T., Ishaq, K., Hassan, M. A., Safie, N., Elias, N. F., & Shoaib, M. (2024). Computationally efficient deep learning models for diabetic retinopathy detection: A systematic literature review. Journal of Medical Systems.
18. Hemanth, D. J., Deperlioglu, O., & Kose, U. (2020). An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. Neural Computing and Applications, 32(3), 707–721. https://doi.org/10.1007/s00521-018-03974-0
19. Huo, G. (2025). Deep learning models for diabetic retinopathy detection: A review of CNN and transformer-based approaches. In Proceedings of the International Conference on Data Analysis and Machine Learning (DAAML 2024) (pp. 594–598). https://doi.org/10.2991/978-94-6463-764-1_95
20. Kour, V., Swain, J., Singh, J., Singh, H., & Kour, H. (2024). A review on diabetic retinopathy. Current Diabetes Reviews, 20(6). https://doi.org/10.2174/0115733998253672231011161400
21. Li, L. Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., Liu, X., & He, B. (2023). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. IEEE Transactions on Knowledge and Data Engineering, 35(4), 3347–3366.
22. Li, X., Wen, X., Shang, X., Liu, J., Zhang, L., Cui, Y., Luo, X., Zhang, G., Xie, J., Huang, T., Chen, Z., Lyu, Z., Wu, X., Lan, Y., & Meng, Q. (2024). Identification of diabetic retinopathy classification using machine learning algorithms on clinical data and optical coherence tomography angiography. Eye, 38(14), 2813–2821. https://doi.org/10.1038/s41433-024-03173-3
23. Lim, W. X., Chen, Z., & Ahmed, A. (2022). The adoption of deep learning interpretability techniques on diabetic retinopathy analysis: A review. Medical & Biological Engineering & Computing, 60(3), 633–642.
24. Mahee, N., & Ejaz, M. S. (2025). Identification of diabetic retinopathy using deep learning and ensemble model approach. Journal of Undergraduate Research International (JURI).
25. Naz, H., Ahuja, N. J., & Nijhawan, R. (2024). Diabetic retinopathy detection using supervised and unsupervised deep learning: A review study. Artificial Intelligence Review, 57, Article 131.
26. Priya, R., & Aruna, P. (2013). Diagnosis of diabetic retinopathy using machine learning techniques. ICTACT Journal on Soft Computing, 3(4). https://doi.org/10.21917/ijsc.2013.0083
27. Shankar, K., Sait, A. R. W., Gupta, D., Lakshmanaprabu, S. K., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recognition Letters, 133, 210–216. https://doi.org/10.1016/j.patrec.2020.02.026
28. Tato, F. R., & Yasin, H. M. (2025). Detecting diabetic retinopathy using machine learning algorithms: A review. Asian Journal of Research in Computer Science, 18(2), 118–131. https://doi.org/10.9734/AJRCOS/2025/v18i2566

29. Tato, F. R., & Yasin, H. M. (2025). Detecting diabetic retinopathy using machine learning algorithms: A review. Asian Journal of Research in Computer Science, 18(2), 118–131. https://doi.org/10.9734/AJRCOS/2025/v18i2566
30. Zago, G. T., Andreão, R. V., Dorizzi, B., & Teatini Salles, E. O. (2020). Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Computers in Biology and Medicine, 116, 103537.
31. Zhou, L., Zhao, Y., Yang, J., Yu, Q., & Xu, X. (2018). Deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images. IET Image Processing, 12(4), 563–571. https://doi.org/10.1049/iet-ipr.2017.0636
32. Zhu, S., Xing, C., Zhong, Q., & Yao, Y. (2024). Diabetic retinopathy classification with deep learning via fundus images: A short survey. IEEE Access, 12, 20540–20554. https://doi.org/10.1109/ACCESS.2024.3361944