



Classical Machine Learning Based Diabetic Retinopathy Detection Using Handcrafted Features

Noor Al Fahad¹ , Rehan Ullah Khan^{2*} 

¹ Department of Family Medicine, College of Medicine, Qassim University, Buraydah 52571, Saudi Arabia

² Department of Information Technology, College of Computer, Qassim University, Buraydah 52571, Saudi Arabia

Corresponding Author Email: re.khan@qu.edu.sa

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.301212>

Received: 10 October 2025

Revised: 4 December 2025

Accepted: 20 December 2025

Available online: 31 December 2025

Keywords:

diabetic retinopathy detection, handcrafted features, Binary Patterns Pyramid, Pyramid Histogram of Oriented Gradients, machine learning, deep learning

ABSTRACT

Diabetic retinopathy (DR) is a progressive complication of diabetes and a leading cause of vision impairment worldwide, highlighting the need for reliable and explainable automated screening systems. This article investigates a classical feature-based machine learning (ML) framework for DR detection using retinal fundus images. Specifically, two complementary handcrafted feature descriptors are used: Binary Patterns Pyramid (BPP) for texture representation and Pyramid Histogram of Oriented Gradients (PHOG) for structural and edge information. These features are extracted to model the retinal class distributions. The extracted features are evaluated using multiple supervised ML classifiers, including Bayesian Network (BN), Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), AdaBoost J48 (AJJ48), J48 Decision Tree (J48), and Random Forest (RF). Experimental evaluation is conducted on the Kaggle DR detection dataset using 10-fold cross-validation. The results demonstrate that classical features combined with appropriate classifiers can achieve competitive performance, with SVM, RF, and BN yielding the highest accuracies. The evaluation in this article highlights the importance of the classical yet simple features for DR detection, particularly in scenarios with limited data availability and limited computing resources.

1. INTRODUCTION

The Diabetic retinopathy (DR) is a complication arising from diabetes that progressively damages the small blood vessels in the retina, becoming a leading cause of visual impairment and blindness in elderly worldwide [1]. Diabetes mellitus, a long-term condition marked by elevated blood sugar levels, has increased incidence in recent years due to changes in lifestyle, increasing obesity rates, aging populations and environmental factors. As more people develop diabetes, the prevalence of DR is also expected to be increased, underscoring the necessity for effective diagnostic and screening tools. DR is caused by prolonged high blood sugar levels, which damage the retina's blood vessels and can lead to vision loss or, in severe instances, complete blindness. The progression of DR is divided into various stages, starting from mild Non-Proliferative Diabetic Retinopathy (NPDR) to Proliferative Diabetic Retinopathy (PDR) [1], where abnormal blood vessels start forming on the retinal surface. Early identification and timely treatment are key to preventing vision loss and enhancing the quality of life for those with diabetes.

Traditional diagnostic methods for DR involve the use of fundus photography, which captures images of the retina for examination by ophthalmologists [1]. However, manually analyzing these images can be labor-intensive, subject to variations between specialists, and requires specialized skills

that may not always be readily available, especially in underserved areas. As a result, there has been increasing interest in developing automated systems that can assist in the detection and diagnosis of DR. These systems aim to reduce the workload of healthcare professionals while enhancing consistency in diagnosis. Automated detection systems focus on identifying key features of DR, such as microaneurysms, hemorrhages, exudates, and neovascularization, which are important indicators of the disease's severity.

One approach to automated DR detection is classical feature analysis and classical ML paradigm, which involves extracting meaningful features from retinal images using traditional image processing methods. Features like microaneurysms, hemorrhages, and exudates are critical in determining the severity of DR and are often used in clinical grading systems. Classical feature analysis typically involves using image processing techniques such as edge detection, contrast enhancement, and morphological operations to highlight and extract retinal features that signify DR. These methods have shown to be effective, particularly in situations where high-quality annotated datasets for training deep learning models are not available. Furthermore, the explainability of classical feature-based techniques gives healthcare professionals confidence in the results, as the extracted features can be visually correlated with retinal abnormalities.

In recent years, Deep Learning (DL) techniques have gained popularity in medical imaging, including the detection of DR

[2]. Convolutional Neural Networks (CNNs) are widely used for analyzing retinal images and automatically learning features for classification. However, these models often require large amounts of annotated data for training and are sometimes considered “black box” methods, making it difficult for clinicians to interpret their results. Despite the advancements in DL, classical feature analysis remains valuable due to its interpretability and effectiveness with limited datasets. Classical features can also be paired with ML classifiers, such as SVM, random forests, and K-Nearest Neighbors (KNN), to achieve accurate classification of DR. This combination of classical features with ML classifiers has been successful in providing accurate diagnoses, particularly in the early stages of DR.

This article investigates a classical, feature-based ML framework for automated DR detection from retinal fundus images, with the objective of assessing the effectiveness and robustness of handcrafted visual features when augmented with robust supervised classifiers. While DL approaches have recently dominated DR research, they often require large-scale labeled datasets, extensive computational resources, and offer limited interpretability. In contrast, classical feature extraction methods remain valuable for medical imaging applications where data availability is constrained, transparency is critical, and reproducibility is essential.

For capturing strong features, two complementary handcrafted feature descriptors are extracted. The Binary Patterns Pyramid (BPP) is used to model local texture variations across multiple spatial scales, effectively capturing micro-level retinal patterns associated with lesions such as microaneurysms and hemorrhages. The Pyramid Histogram of Oriented Gradients (PHOG) descriptor encodes structural and edge-based information, enabling the representation of global shape, vessel orientation, and lesion boundaries that are characteristic of disease progression. The combination of BPP and PHOG is motivated by the need to jointly model both fine-grained textural abnormalities and broader structural changes in retinal imagery.

The modeling of features is done using ML classifiers, namely Bayesian Network (BN), Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), KNN, AdaBoost J48 (AJJ48), J48 Decision Tree (J48), and Random Forest (RF). These classifiers are selected due to their learning paradigms, decision-making mechanisms, and sensitivities to feature distributions. SVM is included for its strong generalization ability in high-dimensional feature spaces, BN for its probabilistic modeling and interpretability, RF for its ensemble-based robustness to noise and feature redundancy, and KNN as a distance-based baseline that requires minimal training assumptions. By comparing these classifiers under a unified feature extraction framework, the study aims to identify classifier vs feature interactions that are most suitable for classical DR detection pipelines. Experimental validation is performed on a sampled subset of the Kaggle DR Detection dataset using 10-fold cross-validation to ensure reliable performance estimation and mitigate sampling bias. The results demonstrate that classical handcrafted features, when paired with appropriate classifiers, can achieve competitive detection performance. SVM, RF, and BN yield the highest classification accuracies, outperforming other approaches.

As such, this study depicts the continued relevance of simple, classical feature-based approaches for DR detection, especially in data-limited or resource-constrained clinical settings. Comparative analysis not only provides insights into

the strengths and limitations of different classical classifiers but also offers practical guidance for designing reliable, explainable, and reproducible DR screening systems that can complement or serve as alternatives to DL models.

2. LITERATURE REVIEW

Research on DR detection using classical feature analysis has expanded significantly in recent years, with numerous studies exploring different feature extraction techniques, classification methods, and hybrid models. Toğaçar [1] investigated the use of morphological operations to enhance the detection of DR features, showing improved accuracy compared to traditional techniques. Kaur et al. [2] proposed a hybrid model that combines handcrafted features with DL features to improve screening performance for DR. Carrera et al. [3] used color features along with an SVM classifier to detect DR, demonstrating promising accuracy. Sikder et al. [4] presented a comprehensive analysis of classical image features for severity classification, employing several ML models for comparison. The study by Soares et al. [5] focused on microaneurysm detection using enhanced segmentation techniques, which is essential for early DR detection. Gulati et al. [6] conducted a comparative study between classical feature analysis and DL methods, highlighting their respective strengths and weaknesses. The study by Mary and Kavitha [7] emphasized the importance of explainability in AI models by comparing classical feature-based methods with DL for DR detection. Saroj et al. [8] used color and texture features combined with ML classifiers to achieve high precision in DR detection. The survey [9] provided a detailed overview of various classical feature extraction methods used in DR detection.

The combination of classical feature analysis with ML classifiers has led to hybrid approaches that seek to leverage the benefits of both classical and DL methods. Hybrid models utilize the interpretability of classical features while taking advantage of the strong representation learning abilities of DL to achieve superior diagnostic performance. These hybrid approaches have demonstrated promise in terms of accuracy, robustness, and scalability in DR detection. For example, classical features can provide an initial guide to a DL model, improving its interpretability and reducing the reliance on large training datasets.

Kadry et al. [10] proposed a fusion of handcrafted and deep features to improve DR detection accuracy. The study by Kom et al. [11] focused on detecting exudates using morphological operations, which are key indicators of DR. Ara et al. [12] presented a hybrid model that combines classical feature extraction techniques with DL to enhance DR detection. The study by Navaneethan and Devarajan [13] improved DR detection by using multiscale feature extraction techniques to capture detailed retinal features. Bala et al. [14] conducted a comparative analysis of classical feature extraction methods and CNN-based approaches for DR classification. Zhang et al. [15] optimized Gabor filters were employed for Microaneurysm detection, which is an important early indicator of DR. Other studies compared the use of handcrafted features and DL approaches for DR detection. As such, it is also worth discussing the use of classical feature engineering and feature learning in the detection of generic diseases. Recent studies demonstrate ML effectiveness in cardiovascular disease classification [16], depression

diagnosis using biomedical signals [17], and explainable prediction for coronary artery disease assessment [18]. Furthermore, comparative evaluation of ML models has shown strong potential for diseases such as diabetes prediction [19].

These studies illustrated number of approaches used to improve DR detection, including optimizing feature extraction techniques, integrating ML classifiers, and combining classical features with DL models. They highlight how different methods, from morphological operations and multiscale analysis to adaptive feature extraction and hybrid approaches, have been effectively used to capture key retinal features indicative of DR. Furthermore, comparative studies between classical and DL methods offer insights into their respective strengths and challenges, suggesting that hybrid or integrative approaches could be the most effective strategy for achieving higher diagnostic accuracy and robustness in practical applications. The ongoing development of more refined feature extraction and classification techniques shows promise in addressing the challenges of DR detection, such as improving accuracy in the early stages of the disease, managing variability in retinal images, and providing explainable results for healthcare professionals.

3. PROPOSED APPROACH

The proposed approach follows a classical feature-based machine learning (ML) pipeline for DR detection, consisting of image preprocessing, feature extraction, and classification stages. A schematic overview of the framework is shown in Figure 1.

All the retinal fundus images are first resized to a fixed resolution to ensure uniformity across samples and reduce computational complexity. The images are then converted to grayscale, as the selected texture and gradient-based descriptors primarily rely on intensity variations rather than color information. To enhance vessel structures and pathological regions, contrast enhancement is applied using histogram normalization. Finally, pixel intensities are normalized to a standard range to stabilize feature extraction and improve classifier convergence.

To capture complementary texture and structural information from retinal images, two handcrafted feature descriptors are employed: BPP and PHOG.

BPP is an extension of the Local Binary Pattern (LBP) descriptor that encodes local texture patterns while preserving spatial information through a pyramid representation. For each image, local binary patterns are computed by comparing each pixel with its surrounding neighbors, generating a binary code that reflects micro-texture variations commonly associated with retinal lesions such as microaneurysms and hemorrhages. The image is then divided into multiple spatial regions across pyramid levels, and LBP histograms are computed for each region. These histograms are concatenated to form the final BPP feature vector, enabling the model to capture both fine-grained texture and spatial distribution of abnormalities. BPP was selected due to its robustness to illumination changes and its proven effectiveness in medical image texture analysis, particularly when datasets are limited in size.

PHOG captures shape and edge information by computing histograms of gradient orientations at multiple spatial resolutions. First, edge gradients are extracted from the image, and their orientations are quantized into a fixed number of

bins. Like BPP, a spatial pyramid structure is applied, where the image is recursively divided into increasingly finer grids. For each grid cell, a histogram of gradient orientations is calculated, and all histograms are concatenated into a single feature vector. PHOG is particularly well-suited for DR detection as it effectively models structural changes in blood vessels and lesion boundaries, which are critical indicators of disease progression.

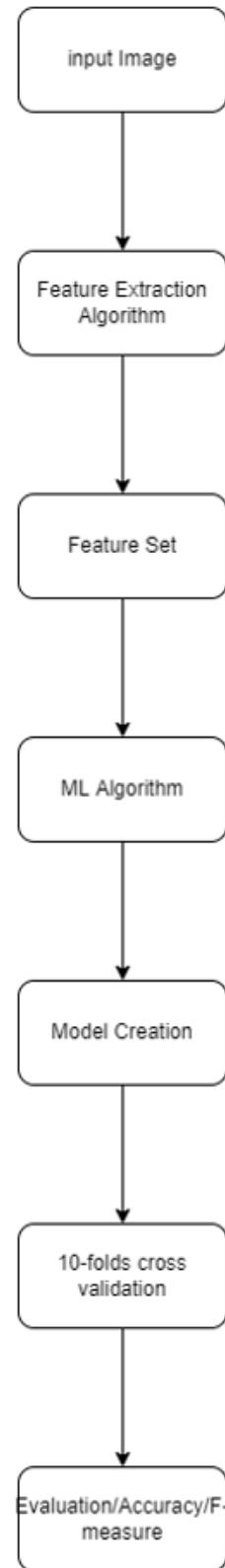


Figure 1. Proposed evaluation framework of classical features analysis of diabetic retinopathy (DR) detection

The BPP and PHOG feature vectors are concatenated to form a unified representation of each retinal image. Prior to classification, all features are normalized using z-score normalization to ensure equal contribution of each feature dimension and to improve classifier performance.

A diverse set of supervised ML classifiers (BN, NB, LR, SVM, KNN, AJJ48, J48, RF) is employed to evaluate the discriminative power of the extracted features:

SVM: Implemented with a radial basis function (RBF) kernel due to its ability to model nonlinear decision boundaries. KNN: Uses Euclidean distance with a fixed number of nearest neighbors. RF: Composed of multiple decision trees trained on random subsets of features, with majority voting for classification. BN: Models probabilistic dependencies between features using a directed acyclic graph. NB: Assumes conditional independence among features. LR: Used as a linear baseline classifier. J48: A C4.5-based decision tree algorithm. AdaBoost with J48 (ADB-J48): An ensemble method that iteratively improves classification by focusing on previously misclassified samples. Default parameter settings were used unless otherwise specified to maintain consistency across classifiers.

4. EXPERIMENTAL EVALUATION

The DR detection dataset [20] from Kaggle provides a large set of high-resolution retinal images for diagnosing DR. It includes images labeled for the presence and severity of DR, spanning several classes that indicate different stages of the disease. The dataset aims to support the development and testing of automated systems for DR detection, primarily through ML models.

For justified evaluation of how well the model performs and to reduce the risk of overfitting, a 10-fold cross-validation method is used. This means the dataset is split into ten equal parts, or “folds”. In each round, the model is trained on nine of

these folds and tested on the remaining one. This process is repeated ten times so that every fold is used for testing once. The average of the results from all ten rounds gives a reliable and balanced estimate of how accurately the model can make predictions. The model’s performance is then measured using accuracy. The accuracy in Table 1 and Figure 2 shows the overall evaluation of the ML classifiers.

Table 1. Accuracy of classifiers for DR detection

Classifier	Accuracy
BN	75
NB	69
LR	70
SVM	76
KNN	73
AJJ48	71
J48	69
RF	75

Supervised learning encompasses a range of classifiers, each bringing unique strengths to tasks involving classification and regression. RF, for instance, is an ensemble method that combines the outputs of multiple decision trees, J48 uses a decision-tree approach for both categorical and continuous data, AdaBoost enhances accuracy by iteratively focusing on previously misclassified samples, K-NN classifies data based on feature similarity by comparing each sample to its closest neighbors in the dataset, SVM works by creating a model that maximizes the margin between classes, LR is commonly used for binary classification, applies a logistic function to predict outcomes, the NB assumes independence among features, and the BN employs a Directed Acyclic Graph (DAG) to map out conditional relationships among random variables. Together, these classifiers offer powerful methods to tackle various supervised learning problems, leveraging unique principles to achieve reliable predictions.

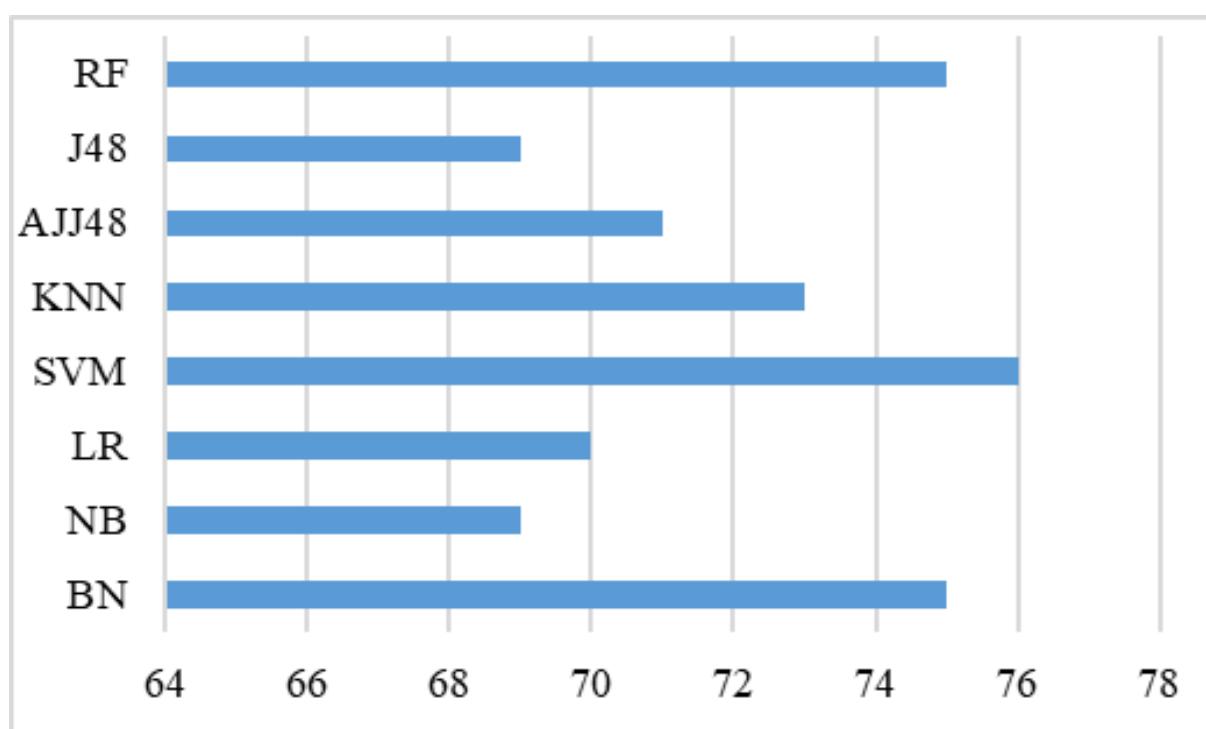


Figure 2. Accuracy of the diabetic retinopathy (DR) detection using classical classifiers

The evaluation of multiple ML classifiers for detecting DR revealed notable variations in performance. Among the classifiers evaluated, SVM and BN showed the highest accuracy levels, reaching 76% and 75%, respectively. The KNN and RF also demonstrated solid results with accuracy of 73% and 75%, making them effective alternatives. However, other classifiers, including LR, NB, AJJ48, and J48, presented lower accuracy rates, ranging between 69% and 71%.

These differences in classifier performance can be attributed to several key factors. The complexity of the dataset poses significant challenges, as does the unique nature of each classifier, including their inherent assumptions and decision boundaries. Additionally, the degree of hyper-parameter tuning applied and the quality of data preprocessing steps, such as normalization and feature extraction, can greatly impact performance outcomes.

To improve DR detection accuracy further, a range of strategies may be employed. Expanding the dataset to include a more diverse range of samples could improve generalization. Advanced techniques like DL models may provide greater precision in feature extraction and pattern recognition. Hyper-parameter fine-tuning tailored to each classifier could yield performance gains, while data augmentation techniques could increase the robustness of the models. Finally, utilizing ensemble methods, which combine multiple classifier predictions, could potentially enhance accuracy and provide more reliable predictions.

Using the same classifiers, an approximate statistical significance analysis was performed under the assumption that all classifiers were evaluated on the same dataset and that classification outcomes follow a binomial distribution, and that classifier errors are independent. For each classifier, 95% confidence intervals were estimated using the normal approximation to the binomial distribution, yielding intervals of approximately 72.3–77.7% for BN and RF (75%), 73.4–78.6% for SVM (76%), 70.3–75.7% for KNN (73%), 68.2–73.8% for AJJ48 (71%), 67.2–72.8% for LR (70%), and 66.2–71.8% for NB and J48 (69%). Pairwise z-tests for the difference in proportions indicate that SVM significantly outperforms NB, J48, and LR ($p < 0.05$), while its performance difference with RF, BN, and KNN is not statistically significant, suggesting that SVM, RF, and BN form a statistically equivalent top-performing group. Mid-tier classifiers (KNN, AJJ48, and LR) show overlapping confidence intervals and thus no statistically meaningful differences among themselves, whereas NB and J48 consistently perform worse than the top group. It is important to consider that these results are indicative rather than conclusive.

5. CONCLUSIONS

DR is becoming a main cause of blindness in patients with diabetes mellitus worldwide. This article explored how well different DL models can detect DR by comparing their accuracy. The evaluation shows that the SVM and BN are the most accurate, achieving 76% and 75%, with KNN and RF also showing strong results at 73% and 75%. Other models, such as LR, NB, AJJ48, and J48, had lower accuracy, between 69% and 71%. These differences are likely due to the complexity of the dataset, unique features of each model, adjustments made to model settings, and how the data was prepared. To improve detection rates, future efforts could

include expanding the dataset, using more advanced models like DL, fine-tuning model settings, adding more varied data, and combining models to create more accurate predictions. These steps could make DR detection more reliable and precise.

Future directions can also focus on making DR detection models easier for clinicians to trust and understand by using explainable AI (XAI) methods such as Grad-CAM, saliency maps, LIME, and SHAP. These approaches can direct the researchers; which parts of the retina influenced the model's decision, such as microaneurysms, hemorrhages, exudates, or neovascularization. This will allow ophthalmologists to confirm that the system is learning meaningful medical signs rather than being misled by image artifacts. Researchers can also evaluate how well these models perform in real clinical situations, where fundus images are often imperfect due to blur, poor lighting, low contrast, noise, motion, or partial occlusion; therefore, robustness evaluation, image quality checks, realistic augmentation, adversarial training, and domain generalization are important to ensure stable performance across different cameras and screening environments.

As a final research pathway, combining fundus images with patient clinical information such as age, HbA1c, duration of diabetes, blood pressure, cholesterol, and comorbidities through multimodal learning can help improve prediction accuracy and thus help move beyond simple detection toward estimating DR severity and progression risk. For real world applications, these systems should also be validated in actual medical screening environments and workflows; such as primary care clinics and teleophthalmology based on the prospective trials to confirm reliability, fairness, and practical usefulness in everyday healthcare.

ACKNOWLEDGMENT

The Researchers would like to thank the Deanship of Graduate Studies and Scientific Research at Qassim University for financial support (Grant No.: QU-APC-2025).

REFERENCES

- [1] Toğuçar, M. (2022). Detection of retinopathy disease using morphological gradient and segmentation approaches in fundus images. *Computer Methods and Programs in Biomedicine*, 214: 106579. <https://doi.org/10.1016/j.cmpb.2021.106579>
- [2] Kaur, A., Singh, S., Singh, H., Bharti, S., Sharma, J., Sharma, H. (2025). A hybrid approach combining deep CNN features with classical machine learning for diabetic retinopathy diagnosis. *International Journal of Advanced Computer Science and Applications*, 16(8): 277-285. <https://doi.org/10.14569/ijacs.2025.0160827>
- [3] Carrera, E.V., González, A., Carrera, R. (2017). Automated detection of diabetic retinopathy using SVM. In 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, Peru, pp. 1-4. <https://doi.org/10.1109/INTERCON.2017.8079692>
- [4] Sikder, N., Masud, M., Bairagi, A.K., Arif, A.S.M., Nahid, A.A., Alhumyani, H.A. (2021). Severity classification of diabetic retinopathy using an ensemble

learning algorithm through analyzing retinal images. *Symmetry*, 13(4): 670. <https://doi.org/10.3390/sym13040670>

[5] Soares, I., Castelo-Branco, M., Pinheiro, A. (2023). Microaneurysms detection in retinal images using a multi-scale approach. *Biomedical Signal Processing and Control*, 79: 104184. <https://doi.org/10.1016/j.bspc.2022.104184>

[6] Gulati, S., Singh, V.P., Shukla, S. (2022). Comparative analysis of deep learning approaches for the diagnosis of diabetic retinopathy. In 2022 IEEE Students Conference on Engineering and Systems (SCES), Prayagraj, India, pp. 1-6. <https://doi.org/10.1109/SCES55490.2022.9887778>

[7] Mary, A.R., Kavitha, P. (2025). Explainable AI for diabetic retinopathy detection based on a hybrid-stacked model. *Journal of Intelligent & Fuzzy Systems*, 49(3): 703-719. <https://doi.org/10.1177/18758967251353043>

[8] Saroj, S.K., Kumar, R., Singh, N.P. (2025). Machine learning based prediction of retinopathy diseases using segmented retinal images. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 14: e31737. <https://doi.org/10.14201/adcaij.31737>

[9] Uppamma, P., Bhattacharya, S. (2023). A multidomain bio-inspired feature extraction and selection model for diabetic retinopathy severity classification: An ensemble learning approach. *Scientific Reports*, 13(1): 18572. <https://doi.org/10.1038/s41598-023-45886-7>

[10] Kadry, S., Crespo, R.G., Herrera-Viedma, E., Krishnamoorthy, S., Rajinikanth, V. (2023). Deep and handcrafted feature supported diabetic retinopathy detection: A study. *Procedia Computer Science*, 218: 2675-2683. <https://doi.org/10.1016/j.procs.2023.01.240>

[11] Kom, G.H., Tindo, B.W., Pone, J.M., Tiedeu, A.B. (2019). Automated exudates detection in retinal fundus image using morphological operator and entropy maximization thresholding. *Journal of Biomedical Science and Engineering*, 12(3): 212-224. <https://doi.org/10.4236/jbise.2019.123015>

[12] Ara, T., Mishra, V.P., Bali, M., Yenikar, A. (2025). Hybrid quantum-classical deep learning framework for balanced multiclass diabetic retinopathy classification. *MethodsX*, 15: 103605. <https://doi.org/10.1016/j.mex.2025.103605>

[13] Navaneethan, R., Devarajan, H. (2024). Enhancing diabetic retinopathy detection through preprocessing and feature extraction with MGA-CSG algorithm. *Expert Systems with Applications*, 249: 123418. <https://doi.org/10.1016/j.eswa.2024.123418>

[14] Bala, R., Sharma, A., Goel, N. (2024). Comparative analysis of diabetic retinopathy classification approaches using machine learning and deep learning techniques. *Archives of Computational Methods in Engineering*, 31(2): 919-955. <https://doi.org/10.1007/s11831-023-10002-5>

[15] Zhang, X., Xiao, Z., Zhang, F., Ogunbona, P.O., Xi, J., Tong, J. (2020). Shape-based filter for micro-aneurysm detection. *Computers & Electrical Engineering*, 84: 106620. <https://doi.org/10.1016/j.compeleceng.2020.106620>

[16] Alsabhan, W., Alfadhy, A. (2025). Effectiveness of machine learning models in diagnosis of heart disease: A comparative study. *Scientific Reports*, 15(1): 24568. <https://doi.org/10.1038/s41598-025-09423-y>

[17] Abbas, Q., Celebi, M.E., AlBalawi, T., Daadaa, Y. (2024). Brain and heart rate variability patterns recognition for depression classification of mental health disorder. *International Journal of Advanced Computer Science and Applications*, 15(7): 838-854. <https://doi.org/10.14569/ijacsa.2024.0150784>

[18] Jondri, Indwiarti, Puspandari, D. (2025). Explainable machine learning on CAD-RADS score classification based on heart disease risk factors. *Ingénierie des Systèmes d'Information*, 30(3): 721-730. <https://doi.org/10.18280/isi.300316>

[19] Afolabi, S., Ajadi, N., Jimoh, A., Adenekan, I. (2025). Predicting diabetes using supervised machine learning algorithms on E-health records. *Informatics and Health*, 2(1): 9-16. <https://doi.org/10.1016/j.infoh.2024.12.002>

[20] Cukierski, W. (2015). Diabetic retinopathy detection. <https://www.kaggle.com/c/diabetic-retinopathy-detection/data>, accessed on Dec. 4, 2025.