Advancement in Diabetic Retinopathy Prediction: Utilizing Voting Classifiers Techniques for Early Detection

Choon Kit Chan
Department of Mechanical Engineering, INTI International University, Malaysia
choonkit.chan@newinti.edu.my

P. Kavitha
Department of ISE, CMR Institute of Technology, Bengaluru, India
kavithapalani.n@cmrit.ac.in (corresponding author)

A. Kalaivani
Department of MCA, New Horizon College of Engineering, Bengaluru, India
ammu.chin24@gmail.com

Goutami Chenumalla
Department of CSE, BMSIT&M, Bengaluru, India
goutamich@bmsit.in

A. Vanathi
Department of CSE, Aditya University, Surampalem, Andhra Pradesh, India
vanathi.andiran@adityauniversity.in

K. H. Koushika
Department of CSE, Sai Vidya Institute of Technology, Bengaluru, India
koushika.kh@gmail.com

G. Likhitha
Department of BCA, Cambridge College, Bengaluru, India
likhitha.shdcfw@gmail.com

Received: 4 September 2025 | Revised: 14 November 2025 and 25 November 2025 | Accepted: 26 November 2025

 DOI: https://doi.org/10.48084/etasr.14526

ABSTRACT

Diabetic Retinopathy (DR), also known as diabetic eye disease, damages the retina and is linked to diabetes mellitus. According to previous studies, DR influences up to 80% of individuals and children who have had type I and II diabetes for more than 20 years. However, with proper care and cautious eye monitoring, severe blind forms of retinopathy and maculopathy can be avoided. This study used a voting classifier, combining the predictions of base models, such as Support Vector Machine (SVM), Random Forest (RF), and Gradient-Boosting (GB) classifiers, to improve the performance and accuracy of predictions.

Keywords-deep learning; ensemble techniques; voting classifier; random forest; Support Vector Machine (SVM); Gradient Boosting (GB)

I. INTRODUCTION

Diabetic Retinopathy (DR) is a widespread disease in working-age adults worldwide. Damage to the retinal blood vessels is a significant aspect of DR, which can cause partial vision loss or blindness [1, 2]. Microvascular abnormalities can be detected using retinal imaging modalities, such as fundus imaging and optical coherence. DR abnormalities, as shown in Figure 1, include microaneurysms, haemorrhage, soft and hard exudates, and neovascularization. Microaneurysms are small outpouchings in retinal vessels. Hemorrhages are the result of vessel damage, presented as blood spots. Soft exudates, lipid-rich deposits, and hard exudates, composed of lipoproteins, signal retinal leakage and lipid accumulation. Neovascularization refers to abnormal blood vessel growth, often fragile and prone to leakage, leading to vision-threatening complications [3]. Detection of these abnormalities in fundus images is crucial for timely assessment and therapeutic actions, helping to manage and reduce the risk of DR-related vision loss. By 2030, the number of people suffering from diabetes is expected to increase by approximately 12 %. Figure 2 compares the structure of a healthy and a diabetic retina.

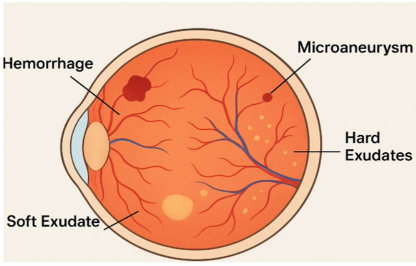

Fig. 1. Abnormalities in DR.

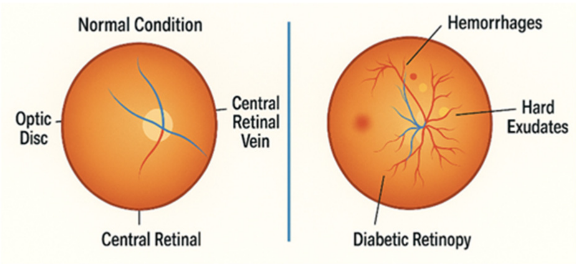

Fig. 2. Structures of a healthy and a diabetic eye.

Figure 3 shows the different stages of DR, where the Normal and Mild stages look visually similar. Hence, it is difficult to detect the Mild stage. In the initial stages of DR, patients are asymptomatic, but in advanced stages, it leads to floaters, blurred vision, distortions, and progressive loss of visual acuity. Hence, it is difficult but extremely important to detect DR using DL algorithms. Predicting a disease in its early stages can avoid the worst effects of later stages. The purpose is to provide a user-friendly interface to predict DR very accurately [4].

In clinical image analysis, automated AI-based solutions have emerged, particularly Convolutional Neural Networks (CNNs) with transfer learning. These strategies offer promising possibilities for advancing the prediction and handling of DR. Therefore, the applications of AI in DR detection focus mainly on the utilization of CNNs with transfer learning. Transfer learning is a popular technique in DR detection, offering a compelling way to address the challenges posed by insufficient labeling data and the intricacies of image classification. At its core, transfer learning involves repurposing knowledge gained from a neural network model pre-trained in a large-scale and diverse dataset, adapting it to a closely related task, such as detecting DR from fundus images. This approach is often used for the detection of DR with pre-trained deep learning models, as it can support complex and high-dimensional fundus images [5]. Xception is a well-known pre-trained model, often used in DR detection since it stands out for its use of depth-wise separable convolutions. This design allows the model to detect subtle patterns and structures in images for efficient hierarchical learning. Xception uses its pre-learned representations from various datasets to capture relevant features and serves as a highly effective feature extractor in DR detection [6]. The application of transfer learning using the Xception model in DR detection generally comprises two primary stages: feature extraction and fine-tuning. Other methods, such as U-Net, have also gained important recognition in the field of medical and biomedical image segmentation [7].

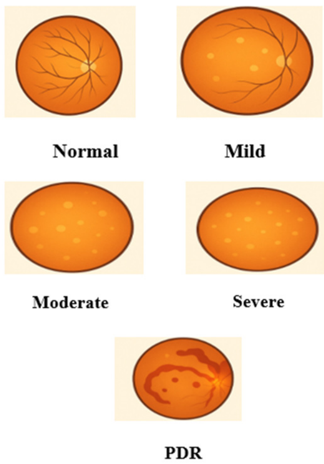

Fig. 3. Different stages of DR.

ML and DL methods have been used to examine fundus images for detecting DR. These approaches leverage segmented features from the fundus images, enhance the input images through k-means clustering, and extract features using pre-trained deep CNN models [8]. According to [9], DL frameworks based on transfer learning serve as instrumental tools for the diagnosis of DR along with other ocular diseases. The accuracy of Inception and ResNet models can be improved, identifying DR lesions better than advanced diagnostic methods [10]. However, detecting DR in fundus imaging using traditional computer vision-based methods is a time-consuming and resource-intensive process. Extracting complex features to improve model performance faces many challenges, mainly in accurately classifying the different stages of DR, detecting the disease in its early stages, and achieving high accuracy with existing transfer learning models.

II. PROPOSED METHOD

A. Experimental Dataset

This study offers a practical and convenient method for detecting DR, using a publicly available dataset that contains approximately 35,000 high-resolution retinal fundus images [11]. This dataset was labeled by ophthalmologists on a five-grade severity scale (0 = No DR, 1 = Mild, 2 = Moderate, 3 = Severe, 4 = Proliferative DR). This study aimed to accurately detect DR using algorithms such as Random Forest (RF), Support Vector Machines (SVM), and Gradient Boosting GB), improving accuracy and offering faster processing and improved decision-making capabilities. Combining algorithms with cross-validation helps examine the outcomes of different machine learning models and select the best-performing model for a given activity. Evaluating each model using the same cross-validation procedure offers a more accurate comparison and more reliable predictions. The dataset was divided into 80% training and 20% testing subsets.

B. Data Preprocessing

The dataset encompasses a diverse range of retinal affected images captured in various sizes and under various lighting conditions. To alleviate these variations, the images underwent several preprocessing steps for normalization. Figure 4 showcases the original state of fundus images before preprocessing.

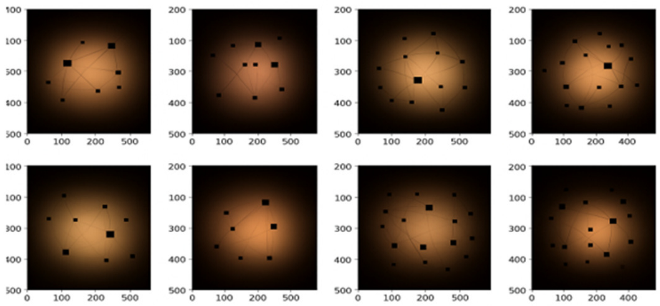

Fig. 4. Fundus images before preprocessing.

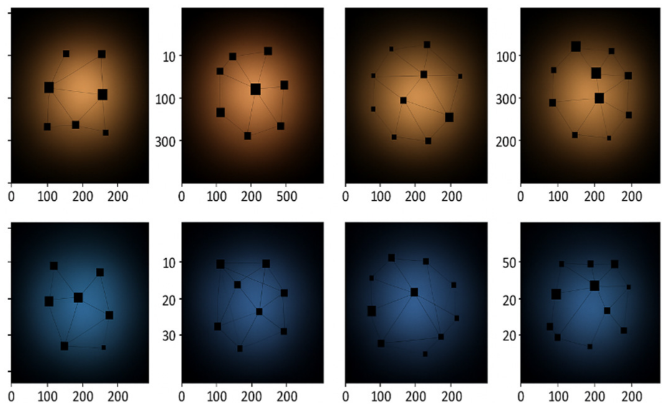

Fig. 5. Fundus image after resizing.

Resizing is a vital preprocessing step for DR fundus images. Bicubic interpolation is widely used for resizing, as it analyzes the average weight of neighboring pixels, favored for its ability to produce polished and graphically appealing outcomes in the resized images by minimizing artifacts and distortions, as seen in Figure 5 [12].

C. Data Augmentation

The effectiveness of the model is highly dependent on the size and diversity of the dataset. A large and diverse training dataset can help avoid overfitting to ensure generalization. Figure 6 illustrates the implementation of various techniques, such as flipping, rotating, cropping, and zooming, which were applied equally to all image classes. This augmentation strategy significantly enhances the comprehensiveness and representativeness of the training dataset, laying a solid foundation for the development of the model [13]. The augmented dataset was 3.6 times larger than the original.

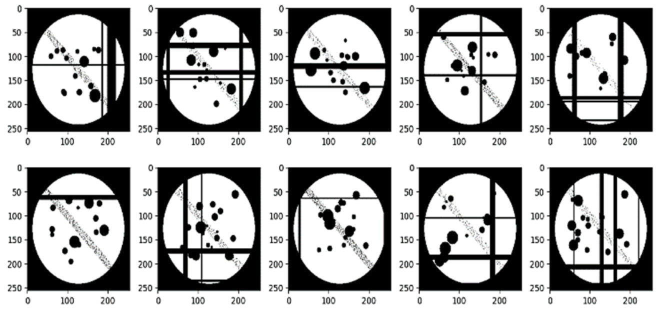

Fig. 6. Image samples after data augmentation.

D. Voting-Based Diabetic Retinopathy (DR) Detection

Ensemble algorithms are applied in detection and diagnosis to improve predictive performance by combining different individual models to strengthen the prediction. These methods leverage diverse sources of information and take advantage of the collective intelligence of several models to increase predictive precision, adaptability, and stability.

SVM is a powerful classification model. Leveraging extracted features such as texture patterns, vessel characteristics, and optical disk attributes, SVM learns to separate between diseased and healthy retinas. Through careful hyperparameter tuning and optimization, SVM can effectively describe a decision boundary in high-dimensional feature spaces, making the most of the margin separation between classes while minimizing classification errors.

RF is another technique that is widely used in the field of machine learning for both classification and regression tasks. During training, the user can specify how many decision trees will be used in an ensemble learning technique. Trees are combined by using random feature selection at each node, hence reducing the likelihood of overfitting. Each tree makes a separate prediction about the target variable, and the classification results are decided by majority vote and average regression. Each tree is trained on a resampled subset using bootstrap aggregation, increasing variety. Large, noisy datasets are no problem for RF, which is renowned for its resilience and scalability.

GB is an effective and powerful algorithm that is applied to both regression and classification problems, effectively handling unbalanced data, which is typical in datasets related to DR. The model is interpretable despite its complexity, which enables physicians to better comprehend key characteristics in the screening process and provide better-suited treatment plans.

III. PERFORMANCE ANALYSIS

The models were trained five times on different folds, and the final measurements are combined into one average. The performance of the model is evaluated using accuracy, sensitivity, and specificity metrics [14]. Performance in differentiating DR from non-DR was assessed in this context. Precision determines the ability to confirm correct positive cases from all cases, which is essential in medical diagnosis to prevent unnecessary treatment errors [15]. Recall or sensitivity determines how well a model detects its actual cases, so that health professionals can prevent missing true DR instances. Specificity determines the proportion of accurate negative case identifications to reduce false alarms. F1-score offers balanced accuracy by minimizing incorrect positive and negative classifications in the results. The ROC curve evaluates the discrimination ability of the model between positive and negative cases, and the AUC value indicates performance quality. Such metrics enable the evaluation of both the beneficial characteristics and areas for improvement within a single analytical framework [16].

A. Model Performance Comparison

The results in Table I show that the Voting Classifier excelled with its highest accuracy and AUC-ROC score, establishing it as the top model for DR detection. The Voting Classifier significantly outperformed individual techniques by combining their strengths, effectively reducing overfitting and enhancing predictive performance. GB delivered strong results, especially in recall, effectively reducing missed DR cases [17]. RF, while slightly lower in accuracy compared to GB, showed robust generalization. SVM, despite its acceptable accuracy, exhibited lower recall, signaling the risk of false negatives in DR detection.

TABLE I. PERFORMANCE RESULTS

Models	Accuracy	Precision	Recall	F1-score
SVM	91.50%	92.00%	90.50%	91.20%
RF	94.20%	93.80%	94.50%	94.10%
GB	95.30%	95.00%	95.60%	95.30%
Voting Classifier (Ensemble)	96.80%	97.00%	96.50%	96.70%

B. Performance Visualization

Figure 7 visualizes the accuracy, precision, recall, and F1-score of the models, demonstrating that the proposed Voting Classifier provides better predictive accuracy than individual ones.

C. ROC Curve Analysis

To compare the capability of each model, the ROC curve was plotted to distinguish non-DR and DR cases. The highest AUC of 0.98 was achieved using the Voting Classifier, offering the best balance between sensitivity and specificity. As shown in Figure 8, SVM had a lower AUC compared to GB and RF.

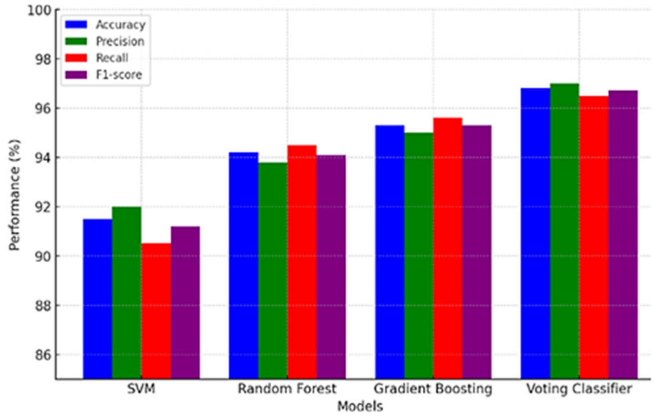

Fig. 7. Performance comparison of the models for DR detection.

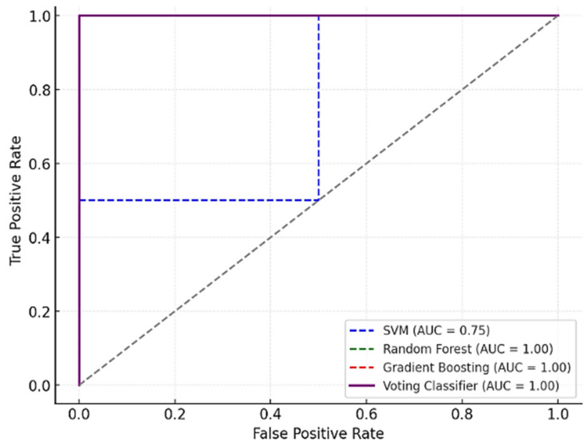

Fig. 8. ROC curve for DR detection models.

D. Confusion Matrix Analysis

Table II and Figure 9 provide the predictions of the Voting Ensemble classifier.

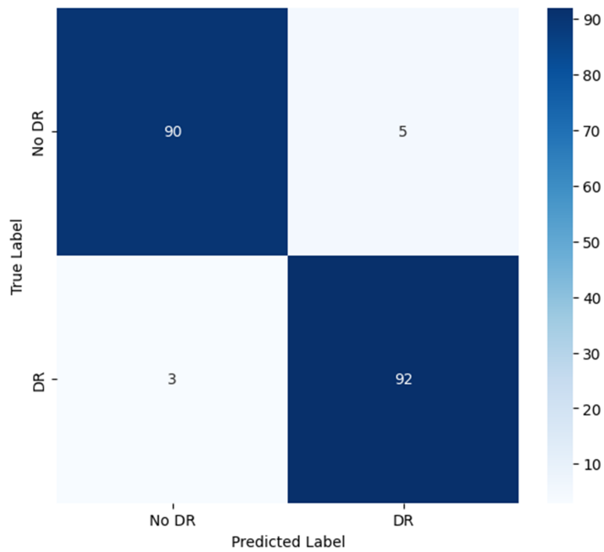

Fig. 9. Confusion matrix of the Voting Classifier.

TABLE II. CONFUSION MATRIX OF THE VOTING ENSEMBLE MODEL

Actual/Predicted	No DR	DR
No DR	90	5
DR	3	92

E. Comparative Analysis with Existing Works

A comparison was made with existing works on DR detection. However, it should be noted that the datasets may vary in resolution, balance, and evaluation protocols. The study in [8] used the Kaggle EyePACS dataset of ~35,000 images, achieving 92.6% accuracy with SVM. In [9], DiaCNN with transfer learning achieved 94-95% accuracy. In [14], a patch-based CNN achieved ~95% accuracy on APTOS 2019. Single-model approaches risk overfitting and poor detection in imbalanced DR categories. The proposed Voting Classifier, using the SVM, RF, and GB base models, achieved 96.8% accuracy and 0.98 AUC on [11], offering superior generalization of real-world DR screening, which consolidates the strengths of individual learners by reducing overfitting [18, 19].

IV. CONCLUSIONS

This study presented a robust ensemble-based framework for early detection of DR, combining machine learning and deep learning models, achieving a high accuracy of 96.8% with an AUC of 0.98. The reported accuracies for earlier methods were in the range of 92-95%. Fewer false positives and false negatives improve the consistency. These results highlight the diagnostic importance of using a variety of models in retinal image analysis. Future work will focus on improving accuracy and explaining the decisions made by the model, exploring hyperparameter tuning, larger training datasets, and explainable AI techniques. Forthcoming work will also examine the role of clinical metadata and multimodal imaging methods to enhance future early screening development. In summary, this framework provides a reproducible, clinically relevant, and scalable diagnostic tool that can help healthcare professionals improve DR grading and prevent detrimental effects, such as loss of vision, in patients.

REFERENCES

[1] A. E. Ilesanmi, T. Ilesanmi, and G. A. Gbotoso, "A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks," *Healthcare Analytics*, vol. 4, Dec. 2023, Art. no. 100261, https://doi.org/10.1016/j.health.2023.100261.

[2] G. Bushi *et al.*, "Creatinine as a predictor of proliferative diabetic retinopathy among patients with type 2 diabetes mellitus: a systematic review and meta-analysis," *International Urology and Nephrology*, vol. 57, no. 12, pp. 4119–4129, June 2025, https://doi.org/10.1007/s11255-025-04590-3.

[3] M. Shorfuzzaman, M. S. Hossain, and A. El Saddik, "An Explainable Deep Learning Ensemble Model for Robust Diagnosis of Diabetic Retinopathy Grading," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 17, no. 3s, pp. 1–24, Oct. 2021, https://doi.org/10.1145/3469841.

[4] S. Lakshmi and B. Sargunam, "An Ensemble Deep Learning Model for Detecting Diabetic Retinopathy Stages," *Bharati Vidyapeeth Medical Journal*, vol. 5, no. 2, pp. 117–121, June 2025, https://doi.org/10.4103/BVMJ.BVMJ_17_25.

[5] S. Mounika and V. RaviSankar, "Diabetic Retinopathy Detection using the Genetic Algorithm and a Channel Attention Module on Hybrid VGG16 and EfficientNetB0," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21319–21325, Apr. 2025, https://doi.org/10.48084/etasr.9720.

[6] A. Sebastian, O. Elharrouss, S. Al-Maadeed, and N. Almaadeed, "A Survey on Deep-Learning-Based Diabetic Retinopathy Classification," *Diagnostics*, vol. 13, no. 3, Jan. 2023, Art. no. 345, https://doi.org/10.3390/diagnostics13030345.

[7] M. Sushith, A. Sathiya, V. Kalaipoonguzhali, and V. Sathya, "A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images," *Scientific Reports*, vol. 15, no. 1, Apr. 2025, Art. no. 15166, https://doi.org/10.1038/s41598-025-99309-w.

[8] A. M. Fayyaz, M. I. Sharif, S. Azam, A. Karim, and J. El-Den, "Analysis of Diabetic Retinopathy (DR) Based on the Deep Learning," *Information*, vol. 14, no. 1, Jan. 2023, Art. no. 30, https://doi.org/10.3390/info14010030.

[9] M. R. Shoaib *et al.*, "Deep learning innovations in diagnosing diabetic retinopathy: The potential of transfer learning and the DiaCNN model," *Computers in Biology and Medicine*, vol. 169, Feb. 2024, Art. no. 107834, https://doi.org/10.1016/j.compbiomed.2023.107834.

[10] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey," *IEEE Access*, vol. 10, pp. 28642–28655, 2022, https://doi.org/10.1109/ACCESS.2022.3157632.

[11] "Diabetic Retinopathy Detection." Kaggle, [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection/data.

[12] M. S. Patil, S. Chickerur, C. Abhimalya, A. Naik, N. Kumari, and S. Maurya, "Effective Deep Learning Data Augmentation Techniques for Diabetic Retinopathy Classification," *Procedia Computer Science*, vol. 218, pp. 1156–1165, 2023, https://doi.org/10.1016/j.procs.2023.01.094.

[13] S. Qummar *et al.*, "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019, https://doi.org/10.1109/ACCESS.2019.2947484.

[14] V. Deepa, C. S. Kumar, and T. Cherian, "Ensemble of multi-stage deep convolutional neural networks for automated grading of diabetic retinopathy using image patches," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 6255–6265, Sept. 2022, https://doi.org/10.1016/j.jksuci.2021.05.009.

[15] S. Kumari and P. Singh, "Deep learning for unsupervised domain adaptation in medical imaging: Recent advancements and future perspectives," *Computers in Biology and Medicine*, vol. 170, Mar. 2024, Art. no. 107912, https://doi.org/10.1016/j.compbiomed.2023.107912.

[16] N. N. Harun, S. Jamaludin, and N. H. Mohtar, "Diabetic Retinopathy Detection with Deep Learning: A ResNet-CNN Model Enhanced by Attention Mechanism and Ensemble Learning," *Open Journal of Applied Sciences*, vol. 15, no. 03, pp. 688–699, 2025, https://doi.org/10.4236/ojapps.2025.153044.

[17] N. Almakayeel *et al.*, "Hyperparameter tuned deep learning-driven medical image analysis for intracranial hemorrhage detection," *PLOS ONE*, vol. 20, no. 7, 2025, Art. no. e0326255, https://doi.org/10.1371/journal.pone.0326255.

[18] S. Madhu, Dr. K. N. Bhargavi, M. V. S. Ramprasad, S. Gautam, and S. Bhavana, "Accurate diabetic retinopathy segmentation and classification model using gated recurrent unit with residual attention network," *Biomedical Signal Processing and Control*, vol. 102, Apr. 2025, Art. no. 107348, https://doi.org/10.1016/j.bspc.2024.107348.

[19] A. S. Narasimha Raju, S. J. Hussain, M. Rajababu, R. K. Gatla, K. Venkatesh, and V. Guntreddi, "CADxPolydetect: a clinically explainable hybrid deep learning system for multi-class colorectal lesion detection using augmented colonoscopy images," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, Sep. 2025, Art. no. 335, https://doi.org/10.1186/s12911-025-03176-4.