International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(22s) 2026 pp. 412-424
https://doi.org/10.70917/ijcisim-2026-5440
Article

Hybrid Deep learning framework for automated Diabetic retinopathy classification using rential fundus images

Goka Tejaswini Nivedhitha[1], Dr.Naga Malleswari Dubba[2]

[1]Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Andhra Pradesh 522302, India. 2401050212cse@gmail.com
[2]Faculty of Dept. of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Andhra Pradesh 522302, India. nagamalleswary@kluniversity.in

Abstract: DR is a major cause of blindness. It is particularly crucial to make an early and accurate diagnosis of DR. The manual examination of retinal fundus images is time-consuming, subjective and prone to error and hence there is a motivation for automated computer aided solutions. Kaggle APTOS 2019 is used as the dataset. Images are scaled, normalised, augmented and feature extracted in order to setup parallel pipelines for classification and detection tasks. The range of models used include basic models (ResNet50, MobileNetV2 and DenseNet121), proposed models (SVM-RBF, Modified EfficientNetV2, EfficientNet–SVM hybrid) and additional hybrid or ensemble models (Xception, NASNet-Large and voting classifiers using several models). We use YOLOv5, YOLOv8, YOLOv9 and YOLOv11 as detectors in this work. Accuracy, precision, recall and F1-score are used for classification. DenseNet121 was 99.6 % for all of them. YOLOv5 had the highest precision (0.713) and mAP (0.788) for object detection and YOLOv9 had the highest recall for localising lesions. XAI approaches such as Grad-CAM can point out the important regions inside an image where the AI model based its conclusion on, thereby increasing the interpretability of the model. The web interface developed using flask provides the end to end inference and displays the real-time classified results to the user along with the annotated photographs which have the bounding boxes. It is a very accurate and interpretable hybrid system to detect and classify DR.
Keywords: Diabetic retinopathy, retinal fundus images, hybrid deep learning, EfficientNet, support vector machine, automated diagnosis”

1. INTRODUCTION

DR is a chronic complication of diabetes and it is a big threat to vision health particularly in developing and medium income countries [1]. It is a condition that impacts millions of people around the world. As a result of the damage to the blood vessels in the retina, the illness causes changes in the retina, including microaneurysms, blood leakage (haemorrhages) and swelling of the retina (macular oedema), and could ultimately result in permanent vision loss if not diagnosed early enough [2]. Ophthalmoscopy or retinal fundus imaging is a routinely used approach for the evaluation of DR, since it provides a complete picture of the retina and provides the physician with information about the pathological characteristics related to the progress of the illness [3]. New technology development, however, has made it possible for the intelligent systems to be introduced in the screening of eye illnesses so that fast and exact diagnosis can be established and the intelligent systems can help with clinical decisions [4].

The traditional diagnosis is mostly relied on ophthalmologists for manual examination of fundus images and appraisal of the severity of the disease. Manual grading is time consuming, costly and subject to inter-observer variability and therefore not scalable for large population screening programmes [5]. The variability in image quality and patient demographics and visibility of lesions would only further complicate the situation and would decrease the consistency and accuracy of the diagnosis [6]. The current automated systems suffer from the accuracy vs.

generalisation vs. interpretability trade-off and hence, novel approaches are required that can give accurate and interpretable diagnostic results to the diverse data sets [7].

The major objective of this study is to suggest enhanced intelligent framework for better detection of DR and rating of its severity in retinal fundus photos. The proposed system is expected to achieve improved reliability of the classification, better localisation skill of the lesion and better interpretability of the system in clinical aspect with the help of better visualisation technique. Furthermore, the system tries to evaluate performance for the different learning paradigms and robustness to changes of the retinal visual properties. This system is aimed at attaining diagnostic function with real time operation in clinical scenario with the help of advanced analytical technique and validation method [8].

An efficient automated DR identification system can be developed to considerably boost early detection, reduce ophthalmologist's diagnostic effort and allow for wide scale screening programs particularly in resource limited health care settings [9]. Further explainability approaches may be useful to elicit trust from medical professionals in automated systems. Furthermore, if screening is proven to be reliable and scalable, it can help lower global rates of diabetes-related blindness, enhance the outcomes of diabetes care and foster a preventive care strategy. Intelligent diagnostic technologies can revolutionise the paradigms of the eye screening approaches and make healthcare more accessible throughout the world.

2. LITERATURE REVIEW

In recent years, significant advancements have been made in the field of AI technology that has greatly influenced medical image processing, notably for independent detection of eye diseases. DL has been widely used in healthcare with many applications as well as quick development. Ahmed et al. [11] did a detailed review of DL modelling techniques. They pointed out that DL has the ability to automatically learn a hierarchical feature representation in complex medical data. A more uniform data representation is crucial for more interoperable and reliable data for AI-based medical applications that will improve diagnoses and model generalisation, Arora said [12]. The research show the promise of intelligent frameworks in healthcare analytics. But there are some problems that still need to be further explored, such as dataset variability, interpretability, and clinical dependability.

There is growing interest in hybrid modelling techniques for improving the classification accuracy for detection of DR. Taufiqurrahman et al. proposed a hybrid architecture of MobileNetV2 with SVM that gained better classification performance by using deep feature extraction and SVM based decision boundaries [13]. The model transparency and visualisation tools for the trust of clinicians and improvement of the decision-making process were pointed out as important in the review and assessment of XAI approaches in healthcare by Sadeghi et al. [14]. Shimpi and Shanmugam [15] suggested a hybrid approach with neural networks for early diagnosis of DR with an improvement in its DR categorisation capability. They studied the effectiveness of neural network hybrid system and showed that the hybrid system was able to classify DR with better accuracy. These approaches have produced promising results but there are difficulties about the adaptability of these methods to heterogeneous data and scalability in real-time clinical situations.

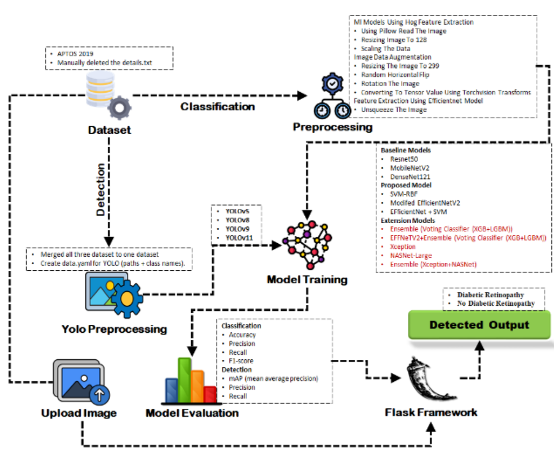
A few CNN-based models were proposed to improve the extraction and classification of the severity of an illness. Das et al. [16] proposed a CNN based method for feature extraction and classification for the detection of the lesions and the categorisation of DR severity level from fundus images which was found to be successful. Goel et al. [17] created a classification model based on DL for DR severity stages and enhanced the classification accuracy using colour fundus pictures. Moreover, Wiratama et al. have applied CNN model with MobileNetV2 for DR classification and demonstrated that it is more efficient in terms of computation and reduction of the size of the classifier and yet it has equal classification performance with other classifiers [18]. But generalisability of the CNN to different datasets and mappability of the diagnostic prediction have been a problem with previous CNN-based methods.

Recently, better DL architectures and ensemble learning methodologies have been developed to increase the diagnosis performance . Mohanty et al. have surveyed several DL architectures for DR detection/classification and recommended using complicated architecture for improving the prediction power and robustness of the architecture [19]. The diagnostic quality and the detection accuracy were further improved in a hybrid detection architecture, which was suggested by Taifa et al. [20] comprising a bespoke ML classifiers and various feature extractors. Although these studies have proven improvement for classification accuracy , the issue still remains to achieve high classification accuracy with a solid localisation of the lesions and high interpretable in a single diagnostic model.

Despite the reported success in the automated diagnosis of DR, there still exist multiple research gaps such as the generalisation of the model, interpretability and combining the categorisation and the detection of lesions into one and unified framework. Furthermore, one of the hurdles that still remains is to achieve high diagnostic accuracy in clinical applications in real-time. The aim of this work is to overcome these limitations by proposing a new intelligent diagnostic framework to achieve higher classification accuracy, better localisation capability and greater interpretability through new visualisation strategies, which can be towards more flexible and clinically applicable automated DR screening solutions.

3. MATERIALS AND METHODS

The proposed technique is aimed towards the accurate categorisation and localisation of Diabetic retinopathy from retinal fundus images of publicly available APTOS 2019 dataset. The system is structured around a process that involves ingesting data, upgrading the quality, segmenting data sets, training models and delivering trustworthy diagnostics. Architectures like EfficientNet can learn complex patterns of disease severity in the retina, as deep feature representations learned from them can. To improve the classification accuracy of the decision boundary and the robustness of classification, a hybrid classification system with SVM with radial basis function kernel and Modified EfficientNetV2 is utilised. To stimulate representation learning, we employ both structural and texture based descriptors such as Histogram of Orientated Gradients and tensor based feature transformations. The ensemble learning models (gradient boosting and complex convolutional network) improve the robustness and generalisation capability over the data and the XAI models (grad-cam) improve the Explainable ability of the models. The lesion localisation is carried out by using several YOLO based detection models and the XAI models based on Grad-CAM are utilised to enhance the interpretability of the models. We built an interface using Flask for real-time inference and scalable deployment, improved diagnostic accuracy, and increased clinical use.



“Fig.1 Proposed Architecture”

The system architecture for the categorisation of DR (diabetic retinopathy) utilising the retinal fundus pictures has three steps . The first approach is to directly use the EfficientNetV2 for end-to-end categorisation into DR and No-DR classes. The second method involves extracting features from the retinal fundus images using handcrafted features and using SVM with RBF kernel for classification. The third hybrid approach is hybrid system that consists of an EfficientNetV2 deep feature extractor and SVM (RBF kernel) as the classifier. The idea for the proposed technique is to utilize both deep and ML representations for enhancing the accuracy of the DR detection.

a) Dataset Collection:

The data set is taken from APTOS 2019 Blindness Detection from the Kaggle repository of retinal fundus images for automated DR screening. This data set consists of thousands of high-resolution colour fundus images from 5 severity groups from no DR to proliferative DR. It's really out of balance and has a large number of different image qualities, illumination and patient demographics. In real-world, the finer information regarding the grading and the clinical variability make it very attractive to develop robust, generalisable and scalable automated DR classification systems.

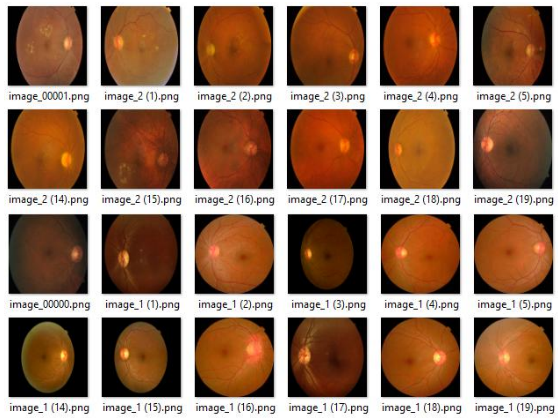


Fig.2 Dataset Collection

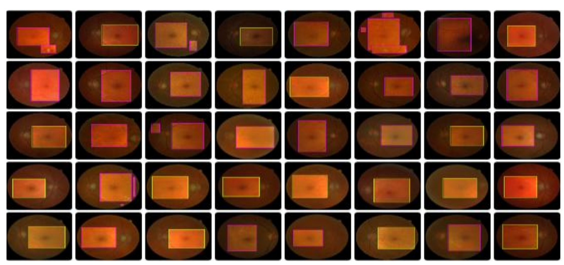


Fig.3

b) Pre-Processing:

To improve the quality of data, to represent it in a better way and to make the model reliable, it is important to preprocess the data. The proposed preprocessing approach performs the retinal images' homogeneity, introduces variation to the image set, and offers structured data for classification and detection tasks.

Image Standardization: The retina images were resized and data scaled for normalisation and to be used as inputs for the traditional ML classification. To reduce the size of the samples and the complexity of the computations, images were rescaled to a common spatial resolution. To minimise lighting and acquisition differences, the pixel intensity data were normalised by scaling. This step helps in making the features more consistent, helps the convergence of the model and ensures that the features for good DR categorisation are extracted uniformly.

Data Augmentation: Data Augmentation To improve the generalisation ability of the model, and to create diversity in the data sets, Data Augmentation techniques were applied. First, the pictures were scaled up to a higher resolution which is more appropriate for DL systems, followed by a slight rotation and flipping of the images in a controlled manner. These augmentations simulate variations in real retinal pictures and acquisition circumstances. The improved images can be converted to structured numerical representations and these representations can be processed well by the learning algorithms. This is to prevent over fitting and to enhance the performance on unseen data and classification accuracy.

Deep Feature Preparation: We prepared deep features with complex convolutional architectures to attempt to describe complex pathological patterns. The images were encoded in a multi-dimensional feature embedding structure for hierarchical learning frameworks. Other dimensional changes were made to allow the usage of deep feature extractors for effective representation learning . Such pre-processing approach will enable us to achieve fine-grained retinal texture and high-level features of lesions, which will increase the discriminative power and hence, improve the overall classification accuracy and model reliability.

Dataset Integration: A hurdle to lesion detection was overcome by creating a combined and diverse training set from multiple datasets of this retina. The annotations were also organised to provide the lesion category and spatial localisation information compatible with object detection frameworks. This stage guarantees the presence of the same limits and class of the lesion in all photos. The incorporation of many data sources boosts the diversity of the samples, improves the robustness of the detection model and allows for reliable localisation of abnormalities related with DR in various clinical imaging situations.

Parallel Dataset Structuring for Multi-Task Learning: Two parallel yet coordinated dataset pipelines were used for classification and detection, namely Parallel Dataset Structuring for Multi-Task Learning. The classification data set included image level diagnostic markers that could be used to predict disease severity. Each detection record in the detection dataset was annotated with spatial information needed to localise the lesions. The parallel data structures are persisted for the query to be optimised based on type, and for coordinating the model training. The architecture is flexible and allows the system to perform many image analysis and disease and lesion identification tasks concurrently, thus enabling a comprehensive automatic analysis of the diseases.

c) Algorithms:

Classification:

ResNet50: The deep convolutional architecture with residual learning (ResNet50) is used to enhance the classification performance. If we connect these, we call them skip connections, and they help to propagate the gradient and hence lead to stable training of deeper networks. The design is based on the concept of using learnt visual features in a hierarchical manner to enhance its representational capability and thus the exact sickness severity pattern can be recognized.

MobileNetV2: MobileNetV2 is a lightweight convolutional architecture which is computationally-efficient and faster in inference. It is based on depthwise separable convolution and inverted residuals to reduce the parameter complexity whilst attaining good feature extraction. The category classification is scalable, resource efficient without affecting the system performance significantly.

DenseNet121: The DenseNet121 is a dense connection for the powerful classification, where the feature information of each layer can be accessed by the subsequent layers. The strategy improves the usefulness of features, boosts the gradient flow and enhances the fine-grained retinal patterns. Architecture is more generic and stable, and the usage of parameters are less efficient.

Support Vector Machine (SVM–RBF): The RBF kernel function of SVM is implemented to create non-linear boundaries in a high dimensional feature space for classifications. It maximises margins of separations of classes. So, it increases stability and robustness. The kernel-based technique is good at handling intricate feature distributions and enhancing the classification consistency.

Modified EfficientNetV2: Modified EfficientNetV2, Optimised deep feature extraction architecture focusing on the balance of scalability of the network dimensions It can capture visual pattern of multi-scale. It has the benefits of enhancing the efficiency of learning and convergence stability. The design is able to improve classification reliability by obtaining discriminative and compact feature representations.

EfficientNet–SVM Hybrid: EfficientNet-SVM hybrid is a deep convolutional features learning and margin based classification. Good choices can be made by SVM and high level feature embedding can be done efficiently using EfficientNet. This improves the generalisation of the model, reduces the sensitivity to fluctuations in the data and thus increases the overall classification stability.

Ensemble Voting Classifier (XGBoost + LightGBM): To make the prediction robust, ensemble voting classifiers are used, where DT models with gradient boosting are used. Ensemble is composed of many learners whose probabilistic results are integrated together to make the ensemble more resilient and less volatile. The method is able to capture interaction of intricate features, and improve the generalisation for categorisation.

EfficientNetV2 + Ensemble Voting: Hybrid approach of deep feature extraction and ensemble based classification approaches. The efficientNetV2 generates discriminative embeddings and the ensemble voting combines the results of various learners. The integrated technique is more resilient and less sensitive to the noise and is more robust in classification for complicated data distributions.

Xception: Xception, yet another very deep separable convolution, a depthwise and pointwise, to learn efficiently spatial and channel-wise feature representations. This is a design that is parameter efficient and preserves the expressive learning capacity. It can characterise the detailed textural differences and pattern of lesions for proper classification of illnesses.

NASNet-Large: We use NAS optimised CNN architectures for extracting multi-scale visual features, called NASNet-Large. The architecture automatically learns optimal network topologies, and enhances the representational power and the classification accuracy. It is aimed to increase generalisation and stability to more complicated modifications of pictures.

Xception–NASNet Ensemble: The Xception-NASNet ensemble combines many complimentary feature extraction algorithms that jointly enhance the classification resilience. To reduce the variance and bias of the prediction of each individual model, channel-wise separation and architecture optimised feature learning are used to create the ensemble, ensuring higher reliability of the classification.

Detection:

YOLOv5: YOLOv5 is an end-to-end detector, which means that it will predict the boxes and class probability at once. Optimised feature pyramids achieve a balance between detection speed and accuracy in the design. It is effective in anomaly detection and it is able to correctly identify the position and the boundary of the lesions.

YOLOv8: The anchor-free prediction of YOLOv8 and the improved feature aggregation lead to the better detection performance. The architecture can make the location design in the training process more precise and stable. It displays high learning capability of spatial and semantic lesion features, thus being helpful for efficient and scalable detection.

YOLOv9: YOLOv9 improves the ability to capture the context representation and multi-scale feature fusion to detect lesions. Little or subtle faults are remembered with the help of Architecture. It makes it easier to detect all the lesions and helps the design to cover all the lesions.

YOLOv11: YOLOv11 enhances single stage detection to achieve good spatial pattern modelling and learning efficiency. The design improvements enable fast and precise localization/categorization of the joints. The model is suitable and robust for the detection of different imaging situations.

d) Integration of XAI & Flask:

We use visualisation approaches to highlight the key aspects of the image that are relevant to categorisation judgements, and provide a more clear and interpretable prediction process. Using Grad-CAM for creating heatmaps for identifying diseased areas of retina related with diabetic symptoms. This interpretability mechanism assists with the clinical validation by providing the users a comprehension of the DL models decision making process thus improving the trust, reliability and diagnostic confidence in automated retinal image processing.

The interaction with trained models is realised in real time in an intuitive user interface developed using the web Flask framework. Users can upload retinal scans, launch predictive algorithms and get classification results with identified issue areas. This deployment architecture provides an efficient access to the model, smooth inference and scalable implementation. The relationship enhances the ease of use of the system and moves the analytical models used in the healthcare closer to the applications of the system.

4. EXPERIMENTAL RESULTS

Accuracy: The test accuracy is the ability to accurately discriminate cases with patients from cases without patients. Accuracy of a test is: the number of true positives and true negatives divided by the number of test instances. Mathematically that can be represented as:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} (1)$$

Precision: Accuracy = (number of samples correctly classified / number of samples classified as positive). Thus the formula to calculate precision will be given by:

$$Precision = \frac{True\ Positive}{True\ Positive\ + False\ Positive} (2)$$

Recall: Recall is an ML metric that shows its ability to retrieve all the instances of a given class. It is the ratio of number of correct observations to the total number of observations in actual class – provides an insight in how thorough the model is in capturing observations of the given class.

$$Recall = \frac{TP}{TP\ +\ FN} (3)$$

F1-Score: F1 Score is an ML measure that is used for evaluation of model accuracy. Combination of precision and recall of a model. Accuracy statistic: Total data / Correct prediction data of the model.

$$F1\ Score = 2 * \frac{Recall\ X\ Precision}{Recall + Precision} * 100(SEQ\ "equation"\ \backslash n\ \backslash *\ MERGEFORMAT\ 4)$$

mAP: One measure of quality ranking is MAP. Considers the number of relevant recommendations and its ranking position in the list of recommendations. The AP at K is averaged over a set of users or queries.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k \quad (5)$$

“Table.1 Performance Evaluation Table – Classification”

ML Model	Accuracy	Precision	Recall	F1-score
SVM-RBF	0.50215	0.51847	0.50215	0.50701
Ensemble	0.93974	0.93978	0.93974	0.93969
Modified EfficientNetV2-Small	0.97543	0.97540	0.97573	0.97543
EFFNeTV2 + SVM	0.49699	0.49663	0.49659	0.49656
EFFNeTV2 + Ensemble	0.93974	0.93978	0.93994	0.93461
MobileNetV2	0.99272	0.99268	0.99276	0.99272

ResNet50	0.99545	0.99543	0.95420	0.99545
DenseNet121	**0.99636**	**0.99636**	**0.99636**	**0.99636**
Xception	0.99545	0.99543	0.99543	0.99545
NASNetLarge	0.98180	0.98229	0.98150	0.98178
Ensemble	0.99545	0.99546	0.99544	0.99545

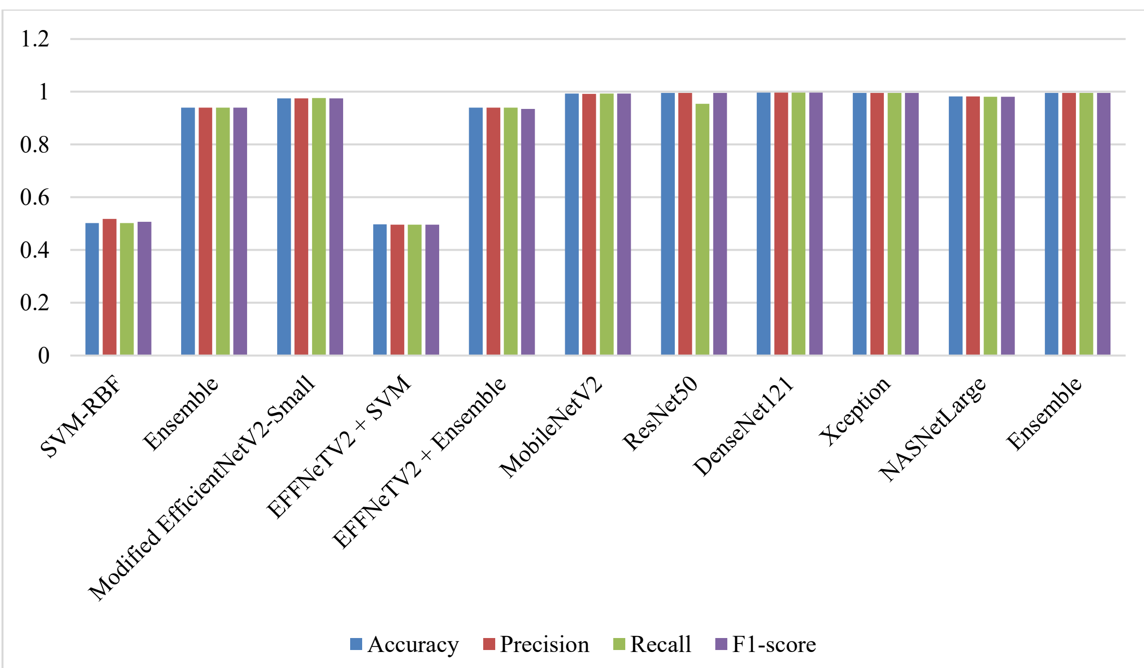
The comparative performance analysis of different ML and DL models for DR classification is shown in Fig.1. The DenseNet121, Xception, ResNet50 and Ensemble models can be seen to be performing well in terms of accuracy, precision, recall and F1-scores.

*"**Table.2*** Performance Evaluation Table – Detection"

ML Model	Precision	Recall	mAP
YOLOv5	**0.713**	**0.796**	**0.788**
YOLOv8	0.680	0.788	0.763
YOLOv9	0.549	0.877	0.707
YOLOv11	0.509	0.863	0.601

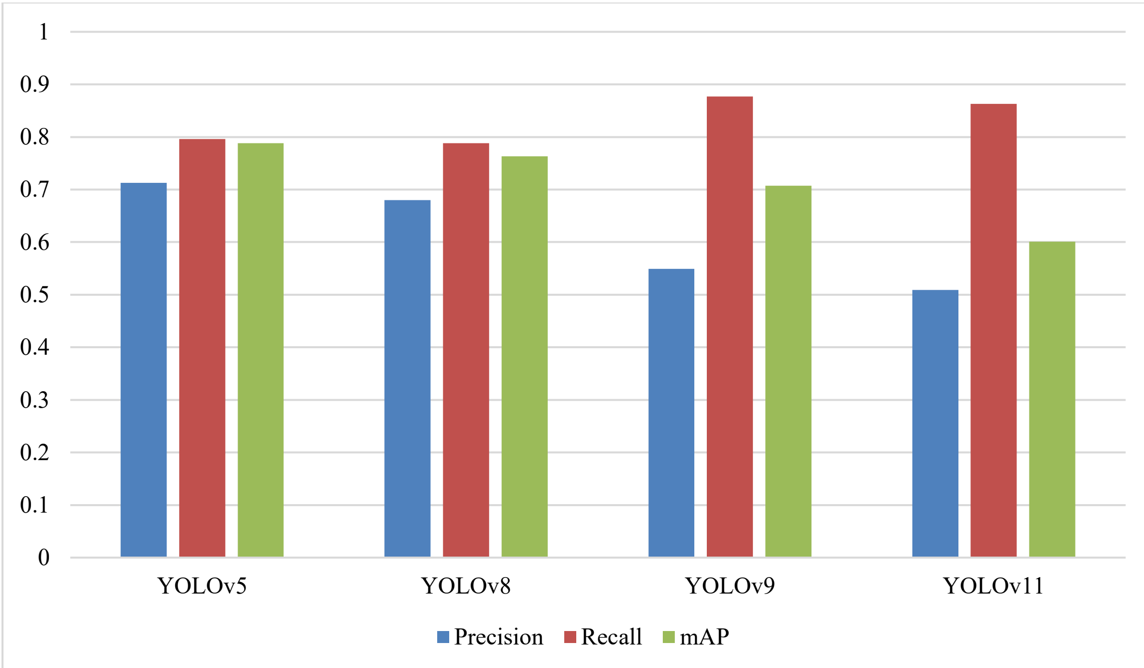
The accuracy and sensitivity of the object identification models were compared and the results show that the YOLOv5 is more accurate with more precision and mAP. In contrast, YOLOv9 and YOLOv11 had higher recall rates, which implies a trade-off between accuracy and sensitivity.

*"**Graph.1*** Comparison Graph – Classification"



In Graph.1, some ML and DL models are compared to show their performances. DenseNet121, Xception, ResNet50 and Ensemble attains the best accuracy, precision, recall and F1 score that outperforms SVM based approaches.

Graph.2 Comparison Graph – Detection



The performance of YOLOv5, YOLOv8, YOLOv9 and YOLOv11 is compared in Fig.2, which shows that YOLOv5 has higher precision and mAP while YOLOv9 and YOLOv11 have higher recall, indicating a tradeoff between accuracy and sensitivity.

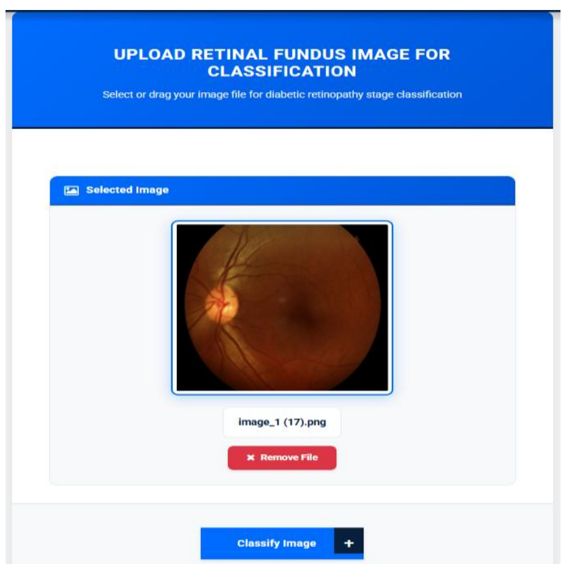


Fig.4 Input a Flie

The retinal fundus image provided via the online web interface is used to classify the stage of DR as shown in Fig.4. The image(s) to be analysed automatically are obtained, displayed and uploaded by the Input module of the system.

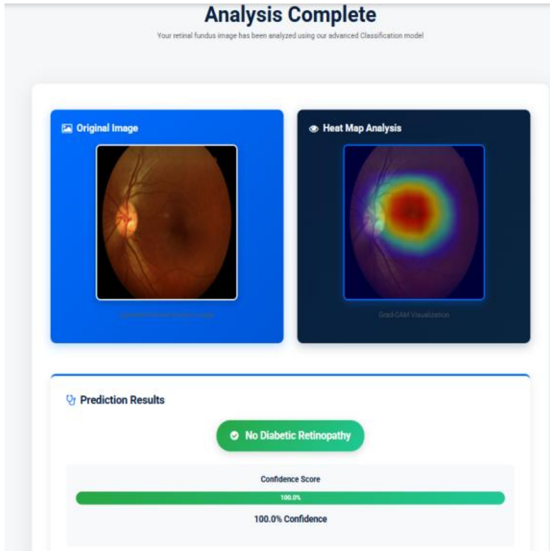


Fig.5 Predicted Result

Figure 5 represents the study result that estimates "No Diabetic Retinopathy" with 100.0% confidence. Original retinal fundus image with Grad-CAM heatmap which shows the nonpathological areas the model utilised to make the decision.

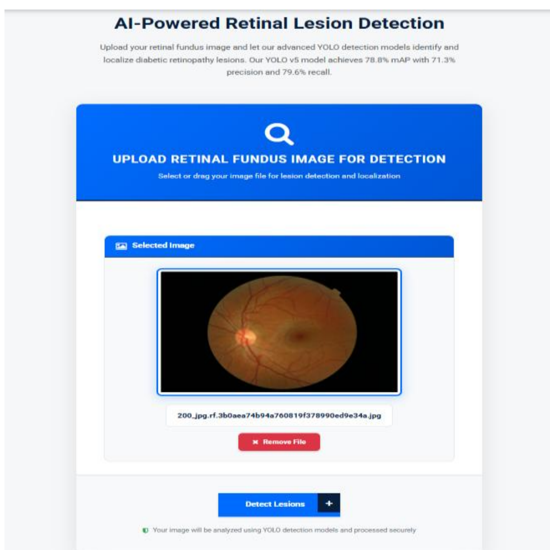


Fig.6 Input a Flie

Fig. 6. Retinal Lesion Identification using AI with a preview of the uploaded fundus image The user can delete the selected file or see the "lesion detection" button to automatically detect DR screening and analysis.

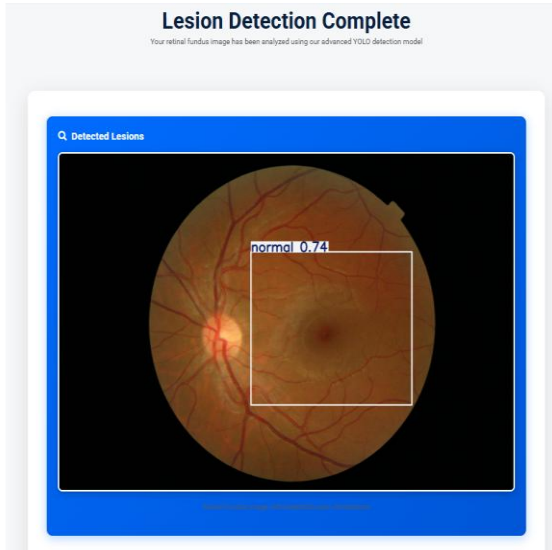


Fig.7 Predicted Result

The results of the lesion detection are displayed in Fig.7. The confidence score of the analysed retinal fundus image for the bounding box highlighted as “normal” is 0.74, which shows that there is no aberrant lesion in the area.

5. CONCLUSION

The proposed method is able to fulfill the most crucial requirement to achieve correct and automated DR diagnosis and to address the drawbacks of manual fundus picture interpretation which is sometime time consuming, subjective and prone to error. Image categorisation and detection are done using the APTOS 2019 Kaggle dataset, including parallel pipelines for processing and analysis. Models such as ResNet50, MobileNetV2, DenseNet121, SVM-RBF, Modified EfficientNetV2, EfficientNet–SVM hybrid, hybrid and ensemble models such as Xception, NASNet-Large and voting classifiers of several models are being utilized. For localization and detection of abnormalities, lesion detection, YOLOv5, YOLOv8, YOLOv9 and YOLOv11 are utilized. We use XAI techniques (Grad-CAM) to visualize the vision part of the model prediction and a Flask based web-interface for user input and for running the model in real-time. The suggested method offers a better accuracy, precision, recall and F1-score of 99.6% in all the accuracy measurement metrics with DenseNet121. YOLOv5 has the highest detection precision of 0.713 and mean average precision of 0.788 whilst YOLOv9 obtains the best recall in localisation. To summarise, this hybrid approach provides a solution to enable early diagnosis, better clinical decision making and has scalable use in healthcare settings in a thorough and interpretable way.

This technique can be extended further by adding more multi-modal retinal imaging modalities e.g. OCT with fluorescein angiography, to improve the diagnosis accuracy in different DR phases. Further optimisation of feature representation and classification performance can be achieved by adopting more complicated DL architectures and attention approaches. Access can be further broadened by deploying the model in cloud or mobile platforms, in remote or resource limited clinical environments. Patient information may be longitudinal and permit prediction and early intervention. Additionally, other sophisticated XAI techniques can further boost transparency, further improving clinicians' trust in and interpretability of complex decision-making of automated DR screening.

REFERENCES

1. Gupta, I. K., Patil, S., Mahadevkar, S., Kotecha, K., Mishra, A. K., & Rodrigues, J. J. (2025). Retinal fundus imaging-based diabetic retinopathy classification using transfer learning and fennec fox optimization. MethodsX, 14, 103232.
2. Sahu, P., Kumar, A., Jain, R., Upreti, K., Yadav, D. K., & Radhakrishnan, G. V. (2025). Explainable Hybrid Deep Learning Framework with Multimodal Inputs for Diabetic Retinopathy Detection. An-Najah University Journal for Research-A (Natural Sciences), 9999(9999), None-None.

3. Zhu, S., Xiong, C., Zhong, Q., & Yao, Y. (2024). Diabetic retinopathy classification with deep learning via fundus images: A short survey. IEEE Access, 12, 20540-20558.
4. Hai, Z., Zou, B., Xiao, X., Peng, Q., Yan, J., Zhang, W., & Yue, K. (2024). A novel approach for intelligent diagnosis and grading of diabetic retinopathy. Computers in Biology and Medicine, 172, 108246.
5. Sadek, N. A., Al-Dahan, Z. T., & Rattan, S. A. (2024). Comprehensive Survey of the State-of-the-Art Deep Learning Models for Diabetic Retinopathy Detection and Grading Using Retinal Fundus Photography. Al-Nahrain Journal for Engineering Sciences, 27(2), 155-163.
6. S. S. Senjam, "Diabetes and diabetic retinopathy: The growing public health concerns in India," *Lancet Global Health*, vol. 12, no. 5, pp. e727–e728, May 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214109X24000755
7. M. M. Butt, D. N. F. A. Iskandar, S. E. Abdelhamid, G. Latif, and R. Alghazo, "Diabetic retinopathy detection from fundus images of the eye using hybrid deep learning features," *Diagnostics*, vol. 12, no. 7, p. 1607, Jul. 2022. [Online]. Available: https://www.mdpi.com/2075-4418/12/7/1607
8. S. Madathil and S. Kutti Padannayil, "MC-DMD: A data-driven method for blood vessel enhancement in retinal images using morphological closing and dynamic mode decomposition," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 8, pp. 5223–5239, Sep. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1319157822001720
9. A. Jabbar, S. Naseem, J. Li, T. Mahmood, M. K. Jabbar, A. Rehman, and T. Saba, "Deep transfer learning-based automated diabetic retinopathy detection using retinal fundus images in remote areas," *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, p. 135, May 2024. [Online]. Available: https://link.springer.com/10.1007/s44196-024-00520-w
10. R. Sayres, A. Taly, E. Rahimy, K. Blumer, D. Coz, N. Hammel, J. Krause, A. Narayanaswamy, Z. Rastegar, D. Wu, S. Xu, S. Barb, A. Joseph, M. Shumski, J. Smith, A. B. Sood, G. S. Corrado, L. Peng, and D. R. Webster, "Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy," *Ophthalmology*, vol. 126, no. 4, pp. 552–564, Apr. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0161642018315756
11. S. F. Ahmed, M. S. B. Alam, M. Hassan, M. R. Rozbu, T. Ishtiak, N. Rafa, M. Mofijur, A. B. M. Shawkat Ali, and A. H. Gandomi, "Deep learning modelling techniques: Current progress, applications, advantages, and challenges," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13521–13617, Nov. 2023. [Online]. Available: https://link.springer.com/10.1007/s10462-023-10466-8
12. A. Arora, "The value of standards for health datasets in artificial intelligence-based applications," *Nature Med.*, vol. 29, no. 11, pp. 2929–2938, Oct. 2023. [Online]. Available: https://www.nature.com/articles/s41591-023-02608-w
13. S. Taufiqurrahman, A. Handayani, B. R. Hermanto, and T. L. E. R. Mengko, "Diabetic retinopathy classification using a hybrid and efficient MobileNetV2-SVM model," in Proc. IEEE REGION 10 Conf. (TENCON), Nov. 2020, pp. 235–240.
14. Z. Sadeghi, R. Alizadehsani, M. A. Çifçi, S. Kausar, R. Rehman, P. Mahanta, P. K. Bora, A. Almasri, R. S. Alkhawaldeh, S. Hussain, B. Alatas, A. Shoeibi, H. Moosaei, M. Hladík, S. Nahavandi, and P. M. Pardalos, "A review of explainable artificial intelligence in healthcare," *Comput. Electr. Eng.*, vol. 118, Jun. 2024, Art. no. 109370. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790624002982
15. J. K. Shimpi and P. Shanmugam, "A hybrid diabetic retinopathy neural network model for early diabetic retinopathy detection and classification of fundus images," *Traitement du Signal*, vol. 40, no. 6, pp. 2711–2722, Dec. 2023. [Online]. Available: https://iieta.org/journals/ts/paper/1018280/ts.400631
16. D. Das, S. K. Biswas, and S. Bandyopadhyay, "Detection of diabetic retinopathy using convolutional neural networks for feature extraction and classification (DRFEC)," *Multimedia Tools Appl.*, vol. 82, no. 19, pp. 29943–30001, Aug. 2023. [Online]. Available: https://link.springer.com/10.1007/s11042-022-14165-4
17. S. Goel, S. Gupta, A. Panwar, S. Kumar, M. Verma, S. Bourouis, and M. A. Ullah, "Deep learning approach for stages of severity classification in diabetic retinopathy using color fundus retinal images," *Math. Problems Eng.*, vol. 2021, pp. 1–8, Nov. 2021. [Online]. Available: https://www.hindawi.com/journals/mpe/2021/7627566/
18. A. B. Wiratama, Y. N. Fuadah, S. Saidah, R. Magdalena, I. D. S. Ubaidah, and R. B. J. Simanjuntak, "Diabetic retinopathy classification based on fundus image using convolutional neural network (CNN) with MobilenetV2," in *Proc. 3rd Int. Conf. Electron., Biomed. Eng., Health Inform.*, vol. 1008, Jan. 2023, pp. 89–102.
19. C. Mohanty, S. Mahapatra, B. Acharya, F. Kokkoras, V. C. Gerogiannis, I. Karamitsos, and A. Kanavos, "Using deep learning architectures for detection and classification of diabetic retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, Jun. 2023. [Online]. Available: https://www.mdpi.com/1424-8220/23/12/5726
20. I. A. Taifa, D. M. Setu, T. Islam, S. K. Dey, and T. Rahman, "A hybrid approach with customized machine learning classifiers and multiple feature extractors for enhancing diabetic retinopathy detection," *Healthcare Analytics*, vol. 5, Jun. 2024, Art. no. 100346. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772442524000480
21. APTOS. APTOS, 2019 Blindness Detection. Accessed: Jan. 28, 2025. [Online]. Available: https://www.kaggle.com/competitions/aptos2019-blindness-detection/data
22. S. S. T. Mandiga, S. P. Mallavarapu, J. Nayani, and R. Mathi, "Retinal blindness detection due to diabetes using MobileNetV2 and SVM," in *Proc. Int. Conf. Advancements Smart, Secure Intell. Comput. (ASSIC)*, Nov. 2022, pp. 1–6.
23. S. G. Kobat, N. Baygin, E. Yusufoglu, M. Baygin, P. D. Barua, S. Dogan, O. Yaman, U. Celiker, H. Yildirim, R.-S. Tan, T. Tuncer, N. Islam, and U. R. Acharya, "Automated diabetic retinopathy detection using horizontal and vertical patch division-based pre-trained DenseNET with digital fundus images," *Diagnostics*, vol. 12, no. 8, p. 1975, Aug. 2022.

24. Y. Liu, "Automatic grading diabetic retinopathy in color fundus image: Cascaded hybrid attention network," *J. Radiat. Res. Appl. Sci.*, vol. 17, no. 4, Dec. 2024, Art. no. 101160.
25. R. Romero-Oraá, M. Herrero-Tudela, M. I. López, R. Hornero, and M. García, "Attention-based deep learning framework for automatic fundus image processing to aid in diabetic retinopathy grading," *Comput. Methods Programs Biomed.*, vol. 249, Jun. 2024, Art. no. 108160.
26. J. Kim, S. Lee, E. Hwang, K. S. Ryu, H. Jeong, J. W. Lee, Y. Hwangbo, K. S. Choi, and H. S. Cha, "Limitations of deep learning attention mechanisms in clinical research: Empirical case study based on the Korean diabetic disease setting," *J. Med. Internet Res.*, vol. 22, no. 12, Dec. 2020, Art. no. e18418.
27. B. Bai, J. Liang, G. Zhang, H. Li, K. Bai, and F. Wang, "Why attentions may not be interpretable?," in *Proc. 27th ACM SIGKDD Conf. Knowl. Discovery Data Mining*, Aug. 2021, pp. 25–34.
28. V. Venugopal, N. I. Raj, M. K. Nath, and N. Stephen, "A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images," *Decis. Anal. J.*, vol. 8, Sep. 2023, Art. no. 100278.
29. D. Keerthana, V. Venugopal, M. K. Nath, and M. Mishra, "Hybrid convolutional neural networks with SVM classifier for classification of skin cancer," *Biomed. Eng. Adv.*, vol. 5, Jun. 2023, Art. no. 100069.
30. M. Z. Sajid, M. F. Hamid, A. Youssef, J. Yasmin, G. Perumal, I. Qureshi, S. M. Naqi, and Q. Abbas, "DR-NASNet: Automated system to detect and classify diabetic retinopathy severity using improved pretrained NASNet model," *Diagnostics*, vol. 13, no. 16, p. 2645, Aug. 2023.