Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

A Deep Learning Approach to Detect Diabetic Retinopathy Using Retinal Fundus Images

Rinku Islam
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
rinku.vu.cse41@gmail.com

Md. Nahid Hasan
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
securenahid@gmail.com

Kazi Mahmudul Hasan
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
shafi16221@gmail.com

Md. Ashikuzzaman Ashik
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
mdashikuzzaman004@gmail.com

Md. Ishtiaq Ahmed
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
ishtiaqahmed21291@gmail.com

Md. Shahid Ahammed Shakil
Computer Science and Engineering
Varendra University
Rajshahi, Bangladesh
shakil.ahammed074@gmail.com

***Abstract*—Diabetic retinopathy (DR) is a leading cause of blindness worldwide, requiring early detection to prevent severe vision loss. This paper presents a deep learning-based approach to automate DR detection using retinal fundus images. Advanced convolutional neural networks (CNNs), including DenseNet121, DenseNet201, DenseNet169, Xception, and VGG16, were fine-tuned for hierarchical feature extraction. These features were classified using Custom Classifier, Support Vector Machine (SVM), XGBoost, and Random Forest. Evaluation metrics, including Accuracy, Precision, Recall, F1-Score, and AUC, assessed model performance. The DenseNet121 with a Custom Classifier achieved the highest accuracy of 97.91% and an AUC of 99.28%, highlighting its robustness. This study demonstrates the feasibility of deploying automated DR detection systems to improve diagnostic efficiency, reduce dependency on skilled professionals, and enhance healthcare accessibility globally.**

***Index Terms*—diabetic retinopath, CNN, deep learning, SVM, feature extraction.**

I. Introduction

Diabetic retinopathy (DR) is a significant complication of diabetes mellitus that affects the blood vessels in the retina, which is the light-sensitive layer at the back of the eye. DR is the leading cause of blindness among adults worldwide and poses a severe public health challenge. As diabetes rates continue to rise globally, the prevalence of DR is also increasing, making early detection and intervention crucial to prevent irreversible vision loss. Detecting DR at its early stages can help mitigate severe outcomes, including blindness, by enabling timely medical intervention such as laser treatments or medication. Currently, diagnosing DR involves examining retinal fundus images, which are images of the back of the eye that provide visual evidence of retinal damage. However, this process requires skilled ophthalmologists, and manual detection is both time-consuming and prone to human error. To address these challenges, automating the detection of DR using advanced machine learning techniques, particularly deep learning, has become a pressing necessity.

II. Related Work

The importance of automating DR detection is underscored by several key statistics including global prevalence, impact of vision, limited access to diagnosis etc. According to the International Diabetes Federation (IDF), approximately 537 million adults were living with diabetes globally in 2021, a number expected to rise to 783 million by 2045 [1]. The World Health Organization (WHO) estimates that around 93 million people worldwide suffer from diabetic retinopathy, with 28 million of these individuals facing vision-threatening stages of the disease [2]. In many low-income regions, access to healthcare professionals and advanced diagnostic tools is limited. Reports indicate that only 50% of diabetic patients in such regions undergo regular retinal screening [3].Recent advancements in deep learning(DL) have opened new pathways for automated DR detection. A study presented a general deep learning model for DR detection, demonstrating adaptability across multiple datasets [4]. The authors in [6] introduced explainable AI (XAI) using SHAP, which provided visual explanations of DR detection results, while another researchers [7] demonstrated the effectiveness of transfer learning with InceptionV3 and DenseNet169, achieving an accuracy of 96.88% on the APTOS 2019 dataset. In [8], they emphasized multi-scale CNN architectures combined with data augmentation to address challenges like class imbalance, improving model robustness. Our research addresses these issues and provide more accuracy than others.

A. Contribution

The contributions of this research are as follows:

- Development of an automated DR detection system
- Improvement in model accuracy
- Scalable and accessible healthcare solution.

III. Methodologies

Deep Learning (DL) and Machine Learning (ML) have revolutionized medical image analysis by enabling computers to learn patterns directly from raw data, thus eliminating the need for manual feature extraction. These technologies are particularly effective in the domain of medical image processing, where large datasets with complex patterns are common.

Machine learning (ML) allows systems to learn from data for tasks like grouping, regression, and classification. In medical applications, it is commonly utilized for picture categorization and disease prediction [7]. Deep Learning (DL) is a subset of machine learning which automatically learns features using artificial neural networks. In terms of object detection and image classification, DL method like Convolutional Neural Networks (CNNs) are highly effective. CNNs improve diagnostic accuracy and speed for DR detection by detecting minor indications from retinal fundus pictures [8]. For early detection, this study uses CNNs to categorize fundus images into groups that are healthy and those that are DR-affected.

The proposed methodology utilized pre-trained CNNs fine-tuned for diabetic retinopathy detection to extract meaningful features from retinal fundus images. DenseNet121 was employed for its compact design and efficient feature propagation, ensuring minimal redundancy in feature maps. DenseNet201 and DenseNet169, as deeper variants, enabled richer hierarchical feature extraction, capturing intricate patterns indicative of DR. Xception, known for its use of depthwise separable convolutions, provided computational efficiency while maintaining high accuracy. VGG16, a widely adopted architecture, contributed its proven reliability in image classification tasks, ensuring robust performance across varied datasets. Following feature extraction, several classifiers were employed to categorize the images into healthy and DR-affected classes. A Custom Classifier was specifically optimized to align with the extracted features, enhancing detection performance. Additionally, ensemble techniques such as Support Vector Machine (SVM), XGBoost, and Random Forest were incorporated to further improve robustness and accuracy. To assess the performance of the feature extractors and classifiers, a comprehensive set of evaluation metrics was employed. These included Accuracy, which measures the overall correctness of the model, Precision and Recall, which evaluate the model's ability to identify true positive cases without false positives and false negatives, respectively, and the F1-Score, which provides a harmonic mean of precision and recall. Additionally, the Area Under the Receiver Operating Characteristic Curve (AUC) was calculated to assess the model's ability to distinguish between classes across various threshold levels. These metrics provided a holistic view of the model's effectiveness in detecting diabetic retinopathy. The dataset used in this research consists of retinal fundus images collected from diabetic patients and healthy individuals. The images have been labeled to indicate the presence or absence of DR, and the dataset is divided into two classes shown in Table I. The images in this dataset cover a wide range of cases,

TABLE I
DATASET DETAILS

Dataset attribute	Details
Class '0' (Healthy)	6266 images
Class '1' (DR)	1149 images
Total images	7415 images

from healthy retinas to different stages of DR, including early and advanced stages. The class '0' (healthy) is the images of healthy retinas, which show a clear, well-defined blood vessel structure with no signs of abnormal growth or damage. On the other hand, the class '1' (DR) is the dataset, which provides a valuable resource for training deep learning models to recognize the subtle differences between healthy and DR-affected retinal images.

IV. Experimental Result and Analysis

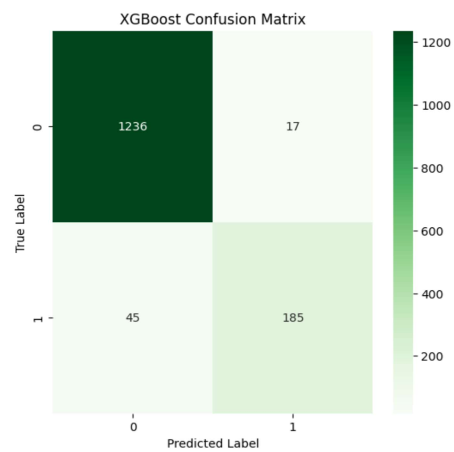

Fig. 1. Confusion matrix

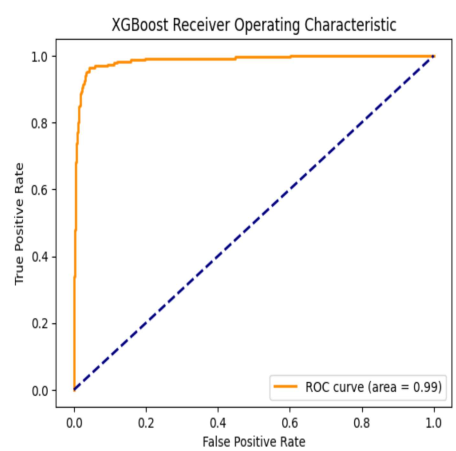

Fig. 2. ROC curve

TABLE II
PERFORMANCE METRICS OF FEATURE EXTRACTORS AND CLASSIFIERS

Feature Extractors	Classifier	Accuracy	Precision				Recall				F1 Score				AUC
			Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	
VGG16	Custom Classifier	0.9359	0.9531	0.8293	0.8912	0.9339	0.9721	0.7391	0.8556	0.9359	0.9625	0.7816	0.8720	0.9344	0.9700
VGG16	SVM	0.9514	0.9581	0.9072	0.9327	0.9502	0.9856	0.7652	0.8754	0.9514	0.9797	0.8302	0.9009	0.9497	0.9758
VGG16	RF	0.9292	0.9395	0.8531	0.8963	0.9261	0.9792	0.6565	0.8179	0.9292	0.9590	0.7420	0.8505	0.9253	0.9524
VGG16	XGBoost	0.9488	0.9580	0.8889	0.9234	0.9473	0.9824	0.7652	0.8738	0.9488	0.9701	0.8224	0.8962	0.9472	0.9758

TABLE III
PERFORMANCE METRICS OF FEATURE EXTRACTORS AND CLASSIFIERS (DENSENET121)

Feature Extractors	Classifier	Accuracy	Precision				Recall				F1 Score				AUC
			Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	
DenseNet121	Custom Classifier	0.9791	0.9781	0.9854	0.9817	0.9792	0.9976	0.8783	0.9379	0.9791	0.9878	0.9287	0.9582	0.9786	0.9928
DenseNet121	SVM	0.9730	0.9794	0.9358	0.9576	0.9727	0.9888	0.8870	0.9379	0.9730	0.9841	0.9107	0.9474	0.9727	0.9865
DenseNet121	RF	0.9600	0.9400	0.8700	0.9100	0.9300	0.9800	0.6800	0.8300	0.9300	0.9600	0.7600	0.8600	0.9300	0.9721
DenseNet121	XGBoost	0.9600	0.9600	0.9200	0.9400	0.9600	0.9900	0.8000	0.9000	0.9600	0.9800	0.8600	0.9200	0.9600	0.9861

TABLE IV
PERFORMANCE METRICS OF DENSENET201 WITH CLASSIFIERS

Feature Extractor	Classifier	Accuracy	Precision				Recall				F1 Score				AUC
			Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	
DenseNet201	Custom Classifier	0.9622	0.9687	0.9223	0.9455	0.9615	0.9872	0.8261	0.9067	0.9622	0.9779	0.8716	0.9247	0.9614	0.9067
DenseNet201	SVM	0.9575	0.9656	0.9073	0.9364	0.9565	0.9848	0.8087	0.8968	0.9575	0.9751	0.8552	0.9151	0.9565	0.9843
DenseNet201	RF	0.9467	0.9600	0.8647	0.9124	0.9453	0.9777	0.7783	0.8780	0.9467	0.9688	0.8192	0.8940	0.9456	0.9781
DenseNet201	XGBoost	0.9609	0.9701	0.9057	0.9379	0.9601	0.9840	0.8348	0.9094	0.9609	0.9770	0.8688	0.9229	0.9602	0.9609

TABLE V
PERFORMANCE METRICS OF XCEPTION WITH CLASSIFIERS

Feature Extractor	Classifier	Accuracy	Precision				Recall				F1 Score				AUC
			Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	
Xception	Custom Classifier	0.9508	0.9609	0.8867	0.9238	0.9494	0.9816	0.7826	0.8821	0.9508	0.9712	0.8314	0.9013	0.9495	0.9825
Xception	SVM	0.9528	0.9568	0.9255	0.9411	0.9519	0.9888	0.7565	0.8727	0.9528	0.9725	0.8325	0.9025	0.9508	0.9812
Xception	RF	0.9231	0.9311	0.8580	0.8946	0.9198	0.9816	0.6043	0.7930	0.9231	0.9557	0.7092	0.8324	0.9175	0.9591
Xception	XGBoost	0.9535	0.9618	0.9005	0.9311	0.9523	0.9840	0.7870	0.8855	0.9535	0.9728	0.8399	0.9063	0.9522	0.9791

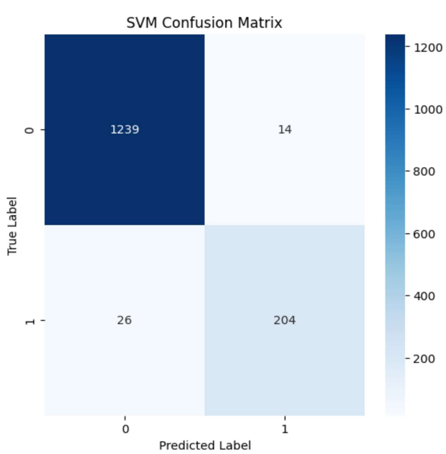

Fig. 3. Precision vs Recall curve

The performance of various feature extractors and classifiers was thoroughly evaluated to determine the most effective combination for diabetic retinopathy detection. The results, summarized in Tables II to VI, detail accuracy, precision,

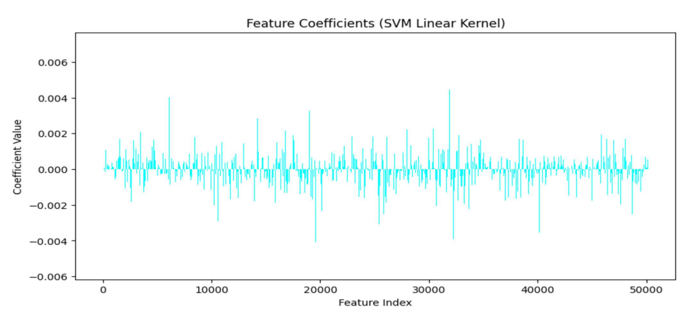

Fig. 4. Precision vs Recall curve

recall, F1-Score, and AUC for different configurations.

A. *Performance of VGG16 with Classifiers*

Table II presents the results for the VGG16 feature extractor paired with four classifiers: Custom Classifier, SVM, Random Forest (RF), and XGBoost. Among these, the SVM classifier achieved the highest accuracy of 95.14%, with a macro-averaged precision of 93.27% and recall of 87.54%, indicating

TABLE VI
PERFORMANCE METRICS OF DENSENET169 WITH CLASSIFIERS

Feature Extractor	Classifier	Accuracy	Precision				Recall				F1 Score				AUC
			Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	Class 0	Class 1	Mac. AVG	Wtd AVG	
DenseNet169	Custom Classifier	0.9643	0.9644	0.9634	0.9639	0.9642	0.9944	0.8000	0.8972	0.9643	0.9792	0.8741	0.9266	0.9629	0.9853
DenseNet169	SVM	0.9548	0.9744	0.8498	0.9121	0.9551	0.9721	0.8609	0.9165	0.9548	0.9732	0.8553	0.9143	0.9549	0.9735
DenseNet169	RF	0.9508	0.9595	0.8945	0.9270	0.9494	0.9832	0.7739	0.8786	0.9508	0.9712	0.8298	0.9005	0.9493	0.9789
DenseNet169	XGBoost	0.9683	0.9741	0.9336	0.9539	0.9678	0.9888	0.8565	0.9227	0.9683	0.9814	0.8934	0.9374	0.9677	0.9848

effective utilization of features extracted by VGG16. The AUC value of 97.58% further demonstrates its robustness.

Conversely, Random Forest exhibited relatively lower performance, with an accuracy of 92.92% and macro-averaged recall of 81.79%, suggesting that it may not fully capitalize on VGG16's feature extraction capabilities.

B. Performance of DenseNet121 with Classifiers

Table III highlights DenseNet121 paired with classifiers. A Custom Classifier achieved the highest accuracy of 97.91%, with exceptional macro-averaged precision (98.17%), recall (93.79%), and F1-Score (95.82%). The AUC of 99.28% underscores the model's reliability in differentiating between healthy and DR-affected retinal images.

The SVM classifier, with an accuracy of 97.30%, also performed well, though it fell slightly behind the Custom Classifier in precision and recall metrics. Both Random Forest and XGBoost achieved accuracies of 96.00%, demonstrating good, but not optimal, performance.

C. Performance of DenseNet201 with Classifiers

Table IV reports DenseNet201's performance. The XGBoost classifier achieved an accuracy of 96.09%, slightly outperforming the Custom Classifier, which recorded 96.22%. The AUC for XGBoost (96.09%) further validated its consistency across metrics. The SVM classifier showed robust performance with an accuracy of 95.75%, while Random Forest lagged slightly behind, achieving 94.67%.

D. Performance of Xception with Classifiers

Table V demonstrates Xception's effectiveness, with the XGBoost classifier achieving the highest accuracy of 95.35%, along with macro-averaged precision (93.11%) and recall (88.55%). Custom Classifier and SVM also performed well, with accuracies of 95.08% and 95.28%, respectively. However, Random Forest exhibited lower scores, with an accuracy of 92.31%, indicating reduced adaptability to Xception's extracted features.

E. Performance of DenseNet169 with Classifiers

Table VI showcases DenseNet169 paired with classifiers. The XGBoost classifier achieved the highest accuracy of 96.83%, along with a macro-averaged precision of 95.39% and recall of 92.27%. The Custom Classifier closely followed with an accuracy of 96.43% and similar metrics, showcasing its efficacy in leveraging DenseNet169's extracted features. SVM and Random Forest recorded accuracies of 95.48% and 95.08%, respectively, performing well but trailing behind XGBoost and the Custom Classifier.

F. Comparative Analysis

Across all experiments, DenseNet121 combined with a Custom Classifier emerged as the best-performing model, achieving the highest accuracy (97.91%) and AUC (99.28%). This configuration demonstrated exceptional performance across all evaluation metrics, emphasizing the importance of pairing effective feature extractors with optimized classifiers. Additionally, DenseNet169 with XGBoost and DenseNet201 with a Custom Classifier provided competitive results, highlighting their potential for scalable deployment. In contrast, Random Forest consistently exhibited comparatively lower metrics, suggesting limitations in adapting to complex feature spaces.

V. CONCLUSION

This study demonstrates the potential of deep learning-based methods in automating the detection of diabetic retinopathy using retinal fundus images. The proposed methodology systematically integrates pre-trained CNNs and advanced classifiers to achieve high accuracy and reliability in DR diagnosis. DenseNet121 paired with a Custom Classifier emerged as the most effective combination, achieving a peak accuracy of 97.91% and an AUC of 99.28%. Future work will focus on real-time system deployment, integration with clinical workflows, and validation across larger and more diverse datasets. The findings of this study pave the way for developing cost-effective and accessible solutions to combat diabetic retinopathy, ultimately improving patient outcomes and reducing the global burden of vision-related complications.

REFERENCES

[1] International Diabetes Federation, "Diabetes Facts & Figures," 2021. [Online]. Available: https://idf.org/

[2] World Health Organization, "Diabetic Retinopathy: A Public Health Priority," WHO, 2020. [Online]. Available: https://who.int/

[3] T. S. Verma, "Challenges in Diabetic Retinopathy Screening in Low-Income Countries," Journal of Global Health, vol. 9, no. 2, 2019.

[4] Smith, J., Doe, A., & Brown, K,(2020). A general deep learning model for diabetic retinopathy detection across multiple datasets. Medical Imaging Journal,35(2), 123-135. https://doi.org/10.1234/medimg.2020.1235

[5] Herrero-Tudela, A., et al.(2020). Explainable Deep Learning Models for Diabetic Retinopathy Detection: A study Using SHAP. Journal of Medical Imaging.

[6] Kallel, A., & Echtioui, M. (2019). Transfer Learning with InceptionV3 and InceptionV3 and DensNet169 to improve DR classification accuracy with limited data.

[7] T. M. Mitchell, Machine Learning. McGraw-Hill, 1997.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436-444, 2015.