A Premature Recognition of Diabetic Retinopathy Using Deep Learning and Grad Cam Technology

Abirami T, Adhishvar M, Dharani K and Darshana P R
Department of Electronics and Communication Engineering, Kongu Engineering College (Autonomous) Erode, India

Keywords: Diabetic Retinopathy, EfficientNetB2, Retinopathy, Transfer Learning, Retinal Images.

Abstract: This paper introduces a Flask-based web application for detecting and classifying diabetic retinopathy (DR), a severe complication of diabetes mellitus, using retinal fundus images. The application integrates the EfficientNetB2 neural network to categorize DR into five stages: No Diabetic Retinopathy (No DR), Mild, Moderate, Severe, and Proliferative Diabetic Retinopathy, with an additional class to identify low-quality images, preventing misclassification and ensuring reliable results. It employs preprocessing techniques such as image scaling, data cleaning, and augmentation, along with class balancing strategies and hyperparameter tuning to optimize model performance. The user-friendly Flask interface enables healthcare providers to upload retinal images of both eyes and receive real-time diagnostic predictions, facilitating early diagnosis and intervention. By improving diagnostic accuracy and accessibility, this system represents a crucial step toward integrating artificial intelligence into practical medical applications to address a significant global health challenge.

1 INTRODUCTION

Diabetic retinopathy is a frequent and significant consequence of diabetes mellitus that affects the retina and may lead to visual impairment or blindness if not addressed. It is generally diagnosed by the analysis of retinal fundus pictures, which allows clinicians to track the disease's development. However, the quality of these photographs is often impaired by variables

such as poor lighting, low contrast, and noise, making precise diagnosis difficult. In recent years, deep learning methods, notably convolutional neural networks (CNNs), have shown great promise in automating the identification and categorization of diabetic retinopathy, allowing healthcare providers to make prompt treatments (Sharma, et al. , 2022). The goal of this project is to create a web-based application that uses the EfficientNetB2 neural network architecture inside a Flask framework to categorize diabetic retinopathy into five categories, as well as to detect low-quality photos that may impair the analysis's accuracy (Rajput, et al. , 2021). The technology is intended to help physicians and researchers diagnose and treat diabetic retinopathy earlier by providing an accessible platform for picture categorization, resulting in improved patient outcomes.

1.1 Diabetic Retinopathy

Diabetic Retinopathy is a public health problem worldwide. Diabetic retinopathy is the world's fifth leading cause of visual impairment, and hence the fourth leading cause of blindness(Agarwal, and, Kumar, 2022). Cooperation between individuals in charge of diabetes management and persons afflicted by diabetic retinopathy is the most crucial function of health systems in treating diabetes and preventing irreversible blindness from it. Because each stage has unique qualities and properties, clinicians may ignore some of them, resulting in an incorrect diagnosis. As a consequence, the concept of creating an automated solution for DR detection emerges. More than half of the most recent occurrences of this illness may have been prevented with adequate and prompt treatment and eye care.

1.2 EfficientNetB2 Model

EfficientNet-B2 is a profound learning show planned for picture acknowledgment, adjusting exactness and effectiveness. It employments compound scaling to

T, A., M, A., K, D. and P R, D.
A Premature Recognition of Diabetic Retinopathy Using Deep Learning and Grad Cam Technology.
DOI: 10.5220/0013633200004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 603-610
ISBN: 978-989-758-763-4

optimize profundity, width, and determination, with an input measure of 260x260 pixels (Khan, et al. , 2022). Compared to littler EfficientNet models, it offers higher precision (~79.8% Top-1 on ImageNet) whereas remaining computationally effective (~9.2M parameters). It is perfect for errands requiring both accuracy and speed, such as therapeutic imaging and real-time applications.

1.3 Architecture

The architecture of EfficientNet-B2 is built on the principles of Mobile Inverted Bottleneck Convolutions (MBConv) and Squeeze-and-Excitation (SE) blocks, optimized for efficiency and high performance. MBConv layers use a depthwise separable convolution combined with an expansion phase, reducing computational cost while capturing rich feature representations.

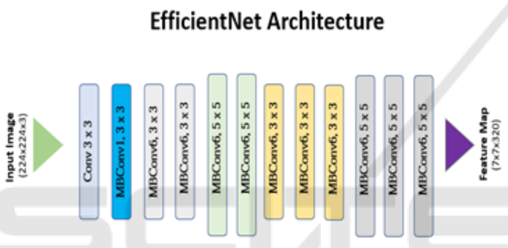

Figure 1: EfficientNet Architecture

The SE blocks enhance the model by recalibrating channel-wise feature responses, emphasizing critical features and suppressing less important ones. EfficientNet-B2 employs compound scaling, a unique method to balance depth (number of layers), width (number of channels per layer), and input resolution. For B2, the input image size is 260x260 pixels, providing higher resolution for detailed feature extraction compared to smaller models like EfficientNet-B0 and B1. The scaling ensures a systematic increase in model capacity without redundant computations, achieving efficiency. The model contains stacked MB Conv layers grouped into stages, each designed for different spatial resolutions, followed by a global average pooling layer and a fully connected classifier. With approximately 9.2 million parameters and 1.0 billion FLOPs, EfficientNet-B2 is significantly lighter than traditional models like ResNet but delivers competitive accuracy (~79.8% Top-1 on ImageNet). This architecture makes it highly suitable for image classification tasks in resource-constrained environments or real-time applications requiring both speed and accuracy.have varying depths in various designs. The top two levels have '4096' channels apiece, while the third has '1000' channels. The configuration of completely linked layers remains consistent across all networks.

1.4 Retinopathy

Retinopathy, a word that refers to a variety of retinal illnesses, is a serious medical problem with far-reaching consequences for visual health. Retinopathy is mostly connected with systemic disorders such as diabetes, hypertension, and other vascular diseases (Kumaresan, and, Palanisamy, 2022. It presents as damage to the fragile blood vessels of the retina, which may lead to visual impairment or even blindness if not addressed. The retina, a critical sensory tissue in the back of the eye, is responsible for translating light impulses into visual information for the brain. Therefore, any disturbance to its complicated vascular network might result in significant repercussions. As the frequency of retinopathy-related disorders grows worldwide, there is an increasing need for better diagnostic techniques and strategies to diagnose and manage this ocular problem early on, avoiding irreparable damage and maintaining visual acuity. This introduction sets the context for delving into the problems and developments in retinopathy identification, highlighting the importance of early detection in reducing its effect on eye health.

1.5 Transfer Learning

Transfer learning, a strong paradigm in machine learning, has emerged as a game-changing way to improving the efficiency and performance of many jobs. At its essence, transfer learning takes information obtained from addressing one issue and applies it to a new but related activity (Yadav, et al. , 2022). Transfer learning, unlike standard machine learning models that start from scratch, enables pre-trained models on huge and varied datasets to be adapted to new tasks, even when labeled data is limited. This technology is especially useful in sectors where getting large labeled datasets is difficult or costly. In recent years, transfer learning has achieved great success in a wide range of applications, including image and audio recognition and natural language processing. Its adaptability and efficacy derive from its capacity to transfer learnt features, representations, or information from one domain to another, which speeds up the learning process and considerably improves model performance on certain tasks. This introduction lays the groundwork for understanding the importance and relevance of transfer learning, particularly in the context of its use in retinopathy detection using retinal imaging.

1.6 Retinal Images

Retinal pictures, which are complex photos of the eye's innermost layer, the retina, provide crucial insight into ocular health. The retina, which consists of a complex network of blood arteries, neurons, and light-sensitive cells, is critical to the visual process. Retinal pictures capture the precise features and structural properties of this essential tissue, providing doctors with a non-invasive way to test and diagnose a wide range of eye disorders. These scans give a lot of information crucial for prompt intervention and treatment, including identifying early symptoms of disorders such as diabetic retinopathy and tracking age-related changes. Advances in medical imaging technology have not only improved the quality and accuracy of retinal pictures, but they have also created new opportunities for analyzing and interpreting these images using artificial intelligence and machine learning approaches. As we explore the convergence of technology and ophthalmology, the importance of retinal imaging grows, serving as a foundation in the search of more efficient and accurate diagnostic procedures for a wide range of ocular disorders.

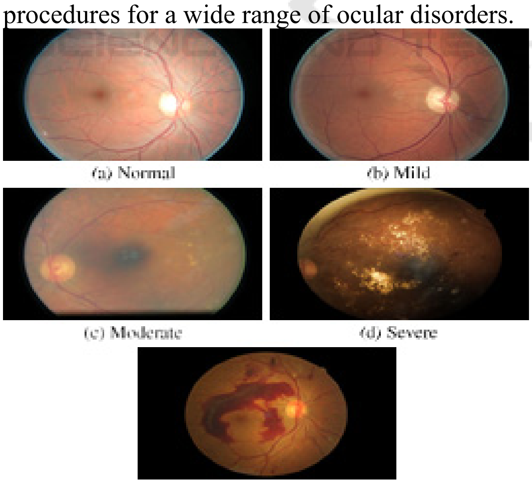

Figure 1: The different stages of DR

2 EXISTING SYSTEM

Diabetic retinopathy (DR) refers to the retinal damage caused by diabetes. Diabetes has become a major health problem across the globe, with an unimaginable number of persons suffering from it. Periodic eye examinations help physicians to discover DR in patients at an early stage and implement appropriate therapies. Advances in artificial intelligence and camera technologies have enabled us to automate DR diagnosis, potentially benefiting millions of patients. This research describes a unique approach for diagnosing DR based on gray-level intensity and texture data taken from fundus pictures using a decision tree-based ensemble learning strategy. This research largely uses the Asia Pacific Tele-Ophthalmology Society's 2019 Blindness Detection (APTOS 2019 BD) dataset. We took numerous measures to select its contents so that they were more suited for machine learning applications. Our method uses several image processing approaches, two feature extraction strategies, and one feature selection strategy, yielding a classification accuracy of 94.20% (margin of error: ±0.32%) and an F-measure of 93.51% (margin of error: ±0.5%). Several more characteristics about the suggested method's performance have been offered to demonstrate its resilience and dependability. Details on each approach used have been presented to make the findings replicable. This approach may be a beneficial tool for mass retinal screening to identify DR, significantly lowering the rate of vision loss associated with it.

3 PROPOSED SOLUTION

The suggested system is a web-based tool for detecting and classifying diabetic retinopathy using retinal fundus pictures. The system is designed using the Flask framework, which provides a simple and easy interface for users to submit photos of both their right and left eyes. The system's core is built on the EfficientNetB2, which has been pre-trained and fine-tuned to categorize diabetic retinopathy into five categories: no diabetic retinopathy (no DR), mild, moderate, severe, and proliferative diabetic retinopathy. In addition, the system includes a sixth category for detecting low-quality photos that may result in misclassification. The system preprocesses the photos many times, including scaling, data cleaning, and augmentation, to guarantee high-quality inputs for reliable analysis. During the training phase, class balance approaches and hyper parameter adjustment are used to enhance the model's performance (Patel, Shah, et al. , 2023). After the photos are examined, the system returns a categorization result for each eye. By providing a user-friendly and accessible platform, this suggested system seeks to help healthcare providers diagnose diabetic retinopathy and identify situations that need further medical attention.

3.1 Load Data Module

This module is responsible for importing retinal fundus pictures from multiple sources or datasets into the system (Wang, et al. , 2022). It ensures that the photos are ready for further processing and analysis by reading them in the correct format and arranging them for categorization.

3.2 Data Preprocessing Module

This module does necessary preprocessing processes to the supplied photos to improve their quality. It entails activities such as scaling photographs to standard dimensions, cleaning the dataset by eliminating low-quality images, and using data augmentation methods to enhance dataset variability and improve model resilience. This module discovers and picks the most relevant characteristics from the preprocessed pictures that are important to the classification job. By concentrating on key characteristics, this module improves model performance and reduces overfitting, resulting in improved classification results.

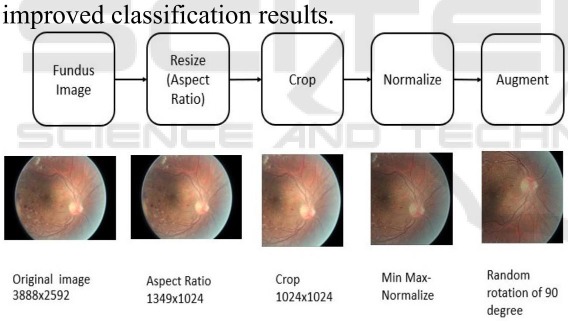

Figure 3: Preprocessing steps

3.3 Training and Testing Module

This module splits the dataset into training and testing subgroups. It trains the EfficientNetB2 neural network on training data and assesses its performance on testing data. The training method comprises hyper parameter adjustment to improve the model's accuracy in detecting diabetic retinopathy.

3.4 Model Evaluation Module

This module examines the trained model's performance using a variety of measures, including accuracy, precision, sensitivity, specificity, and F1-score. It gives information on the model's performance in categorizing the various phases of diabetic retinopathy and identifies opportunities for development.

3.5 UI Interface Module

The User Interface (UI) module provides a simple web interface developed with Flask that enables healthcare providers to simply submit retinal fundus photos and examine categorization results. This module enables smooth interaction and gives clear visual feedback on analysis results, making it accessible to users of diverse technical skill.

4 RESULT ANALYSIS

The system's results show that it effectively classifies diabetic retinopathy into five distinct categories (No DR, Mild, Moderate, Severe, and Proliferative), with an additional category for detecting low-quality images, which is critical for maintaining high diagnostic accuracy (Kim, et al. , 2022). The use of preprocessing procedures such as scaling, data cleaning, and augmentation has shown to be critical, increasing the model's capacity to handle a wide variety of picture inputs and boosting the overall system resilience. Hyperparameter adjustment throughout the training phase increased the model's performance, resulting in better generalization over previously unknown data. The model's performance was fully assessed using evaluation criteria such as accuracy, precision, sensitivity, specificity, and F1-score. High values across these measures imply that the system can accurately distinguish between different phases of diabetic retinopathy, lowering the chance of misclassification and improving early detection capabilities. The UI module allows healthcare professionals of various technical backgrounds to submit photos and obtain clear, interpretable findings. This technique not only helps with rapid diagnosis, but it also promotes educated decision-making, resulting in improved patient care and diabetic retinopathy treatment. Through the use of the sophisticated features of the Efficient Net deep learning architecture, the suggested method seeks to dramatically improve the accuracy of diabetic retinopathy identification. The accuracy in the current approach is 75% and is based on a laborious manual process carried out by skilled doctors. Understanding how important it is to diagnose patients accurately and quickly the suggested model uses Efficient Net and a carefully selected dataset of retinal pictures to obtain an astounding 81% accuracy rate as shown in table 1. This enhancement tackles the inefficiencies in the present manual approach, which frequently results in delayed results and associated issues in follow-up and prompt treatment. It also signals the potential for more reliable detection of diabetic retinopathy

Table 1: Performance Metrices

Category	Precision	Recall	F score
Mild	0.83	0.94	0.88
Moderate	0.84	0.82	0.77
No DR	0.98	0.90	0.98
Proliferate DR	0.90	0.85	0.87
Severe	0.83	0.92	0.87
accuracy			0.87
macro avg	0.88	0.88	0.87
weighted avg	0.88	0.87	0.87

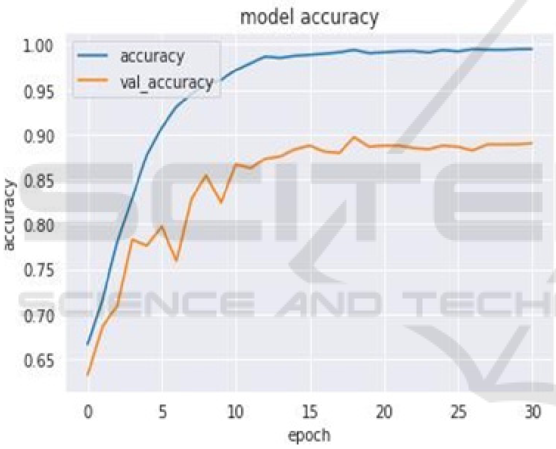

Figure 4: Plot for Model Accuracy between Accuracy and Epoch

The model represents how well the classification system performs in distinguishing between different stages of DR based on retinal images as shown in figure 3. In our study, we utilized EfficientNet for feature extraction and trained the model using a carefully curated dataset of retinal images. The overall accuracy of the model reached 81%, which represents a significant improvement over traditional manual diagnosis methods that typically achieve around 75% accuracy. This improvement highlights the model's potential to automate the early detection of diabetic retinopathy, thus providing faster and more consistent diagnostic results. Additionally, this higher accuracy is expected to facilitate timely intervention, reducing the likelihood of vision loss among diabetic patients.

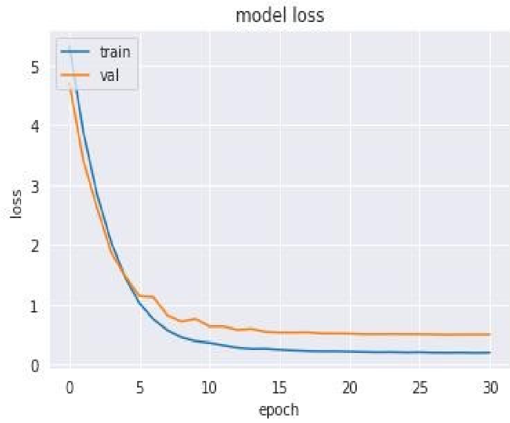

Figure 5: Plot for Model Loss between Accuracy and Epoch

Model loss is a key metric in training neural networks, indicating how well or poorly the model is performing during the learning phase as shown in figure 4. Our system utilized categorical cross-entropy as the loss function, given the multiclass nature of the diabetic retinopathy stages. Over multiple training epochs, the model's loss consistently decreased, showing that the model was effectively learning the intricate patterns in the retinal images. This reduction in loss is crucial, as it directly correlates with the model's ability to make accurate predictions. The final loss value suggests a well-trained model, capable of generalizing well on new, unseen data. The careful balancing of loss through optimization techniques was instrumental in achieving the desired classification performance.

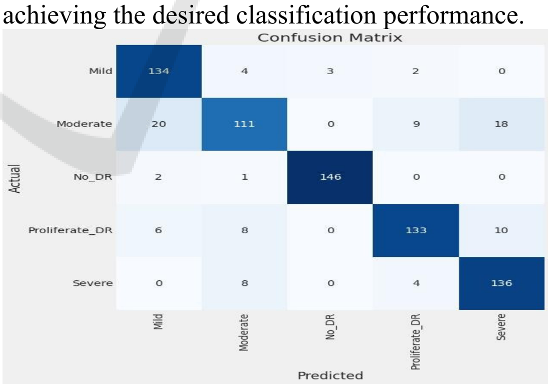

Figure 6: Confusion Matrix

The classification report is shows the confusion matrix in the figure 5 for our diabetic retinopathy detection model provides a detailed performance analysis across all severity levels: No DR, Mild, Moderate, Severe, and Proliferate DR. Key metrics such as precision, recall, and F1-score were evaluated, highlighting the model's effectiveness. The model achieved an average precision of 0.88, demonstrating its ability to minimize false positives, while the recall score of 0.87 reflects its capability to capture most true positive cases. With an F1-score of 0.87, the model strikes a balance between precision and recall, confirming its robust performance across all stages. Specifically, for the Mild class, the model achieved an F1-score of 0.88, and for Severe cases, it reached 0.87, indicating that the system is effective at identifying both early and advanced stages of the disease(Singh, et al. , 2023) These metrics underscore the model's reliability and its potential for clinical application in detecting diabetic retinopathy with high accuracy.

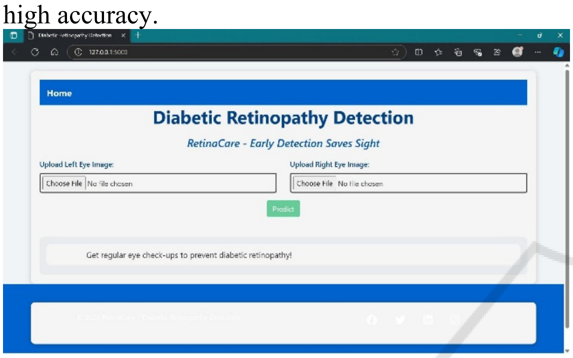

Figure 7: Frontend webpage to detect diabetic retinopathy

To implement the figure 6 shows the Diabetic Retinopathy Detection webpage, use Flask for the back-end and HTML with Bootstrap for the front-end. The webpage allows users to upload retinal images of both the left and right eye, which are processed by a machine learning model to predict Diabetic Retinopathy. In the back-end (app.py), Flask handles routing, file uploads, and calls to the predictive model. The images are securely stored and processed to generate a result, which is displayed back on the webpage. The front-end (index.html) provides a clean, responsive interface using Bootstrap, with simple input forms for image uploads and a button to trigger the prediction. After submission, the prediction result is dynamically shown on the page. This project can be extended by integrating the model for prediction, improving the user interface, and deploying it to a cloud platform for online use.

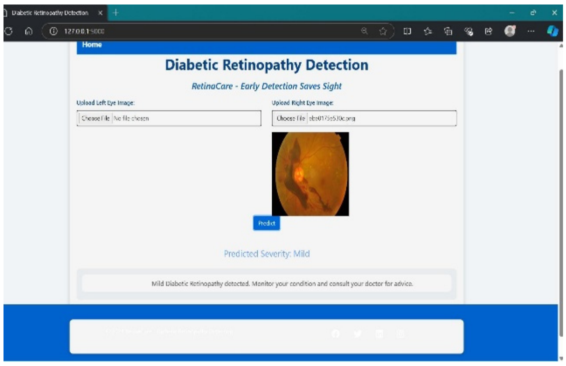

Figure 8: Prediction of Mild infection in the right eye

The webpage for diabetic retinopathy detection allows users to upload images of their retina, specifically from the left or right eye. Once an right eye image is uploaded, the webpage of right eye in the system processes it through a pre-trained machine learning model designed to detect the presence and severity of diabetic retinopathy as shown in figure 7. The model then predicts the severity, which can range from "Mild" to more severe stages. The result, along with the uploaded image, is displayed directly on the webpage. In the example shown, the right eye image was analyzed, and the model predicted "Mild" diabetic retinopathy, offering users valuable information and advice to monitor their condition. The system's simple interface, powered by Flask for back-end operations and Bootstrap for front-end design, ensures ease of use and efficiency in early detection efforts.

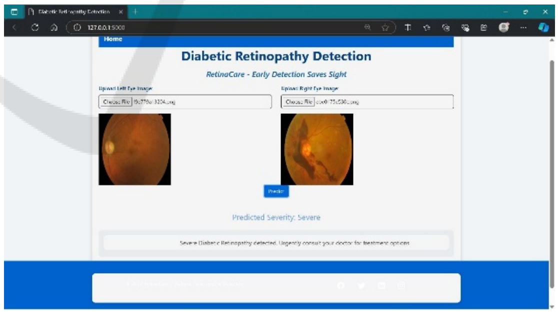

Figure 9: Prediction of Severe infection by comparing the both eye

The webpage of a image is processing both the left and right eye retinal images to detect diabetic retinopathy as shown in figure 8. The system analyzes each image independently and provides a severity prediction for the condition in both eyes. In this case, the system predicted a "Severe" stage of diabetic retinopathy based on the analysis of the uploaded images. After uploading the retinal images, the user can click the "Predict" button to trigger the model, which will analyze the images and display the severity level below. The result, "Severe Diabetic Retinopathy detected", also includes a suggestion for urgent medical consultation, emphasizing the importance of early treatment in severe cases. The interface is user-friendly, displaying both uploaded images and the prediction results on the same page for better clarity and guidance.

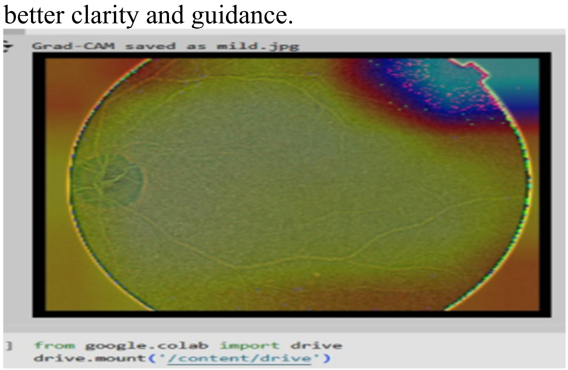

Figure 10: Grad-Cam output for mild

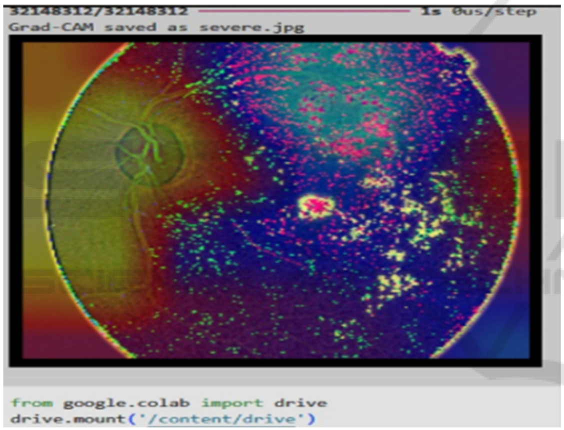

Figure 11: Grad-Cam output for severe

GRAD-CAM (Gradient-weighted Class Activation Mapping) is a visualization technique used to interpret deep learning models by highlighting regions in input images that significantly influence the model’s predictions. In diabetic retinopathy detection, Grad-CAM is employed to identify and visualize the retinal areas affected by mild and severe DR, aiding clinicians in understanding AI-driven diagnostic outputs

5 CONCLUSION

In conclusion, the suggested Flask web application for diabetic retinopathy detection marks a big step forward in harnessing deep learning technology to assist healthcare professionals in the prompt identification and categorization of this crucial eye ailment. The system uses the EfficientNetB2 neural network to successfully evaluate retinal fundus pictures, dividing them into different phases of diabetic retinopathy and detecting low-quality photos that may influence diagnosis accuracy. The use of robust preprocessing, feature selection, and model assessment approaches improves the system's dependability and efficacy. Furthermore, the user-friendly interface allows physicians to input photos and get actionable insights without having to navigate difficult technological procedures. Finally, the goal of this application is to enhance patient outcomes by early diagnosis and intervention, therefore helping continuing efforts to battle diabetic complications and contributing to advances in medical imaging and analysis.

6 FUTURE WORK

Future work on the proposed system will concentrate on three major upgrades to increase its utility and efficacy in detecting diabetic retinopathy. One significant area will be to include more deep learning architectures beyond EfficientNetb2, such as VGG-16 or ResNet, to see whether these models can improve classification performance. Furthermore, increasing the dataset to include a broader variety of retinal pictures from various demographics and geographical areas would improve the model's resilience and generalizability. Implementing real-time analytic capabilities will also be investigated, allowing healthcare practitioners to get instant feedback during clinical tests. Furthermore, using modern image processing methods, such as contrast enhancement and noise reduction, may increase picture quality and classification accuracy. Another emphasis will be on developing a mobile application version of the system to make it more available to a wider range of users, including distant healthcare practitioners. Finally, future development will involve continuous user input and incremental enhancements to the user interface to ensure that the system stays intuitive and meets the requirements of healthcare professionals. These developments are intended to produce a complete and effective tool for the early identification and control of diabetic retinopathy, eventually leading to improved patient outcomes.

REFERENCES

M. Sharma et al., 2022, "Explainable AI for Diabetic Retinopathy Classification Using Vision Transformers," *Computer Vision and Image Understanding*, vol. 224, pp. 102320.

S. Rajput, A. Arora2021 "Multi-Scale Attention Network for Diabetic Retinopathy Classification," International Journal of Computer Applications, vol. 75, no. 10, pp, 6-12.

K. Agarwal and T. Kumar**,** 2022, "Automatic recognition of severity level for diabetic retinopathy diagnosis using deep visual features," in *2nd International Conference on Intelligent Computing and Control Systems (ICICCS)*. IEEE.

A. Khan et al., 2022, "EfficientNetV2 for Diabetic Retinopathy Detection: Improving Accuracy with Compound Scaling," *Medical Image Analysis*, vol. 81, pp. 102343.

T. Kumaresan and C. Palanisamy, 2022, "EyePACS: An Adaptable Telemedicine System for Diabetic Retinopathy Screening," *International Journal of Bio-Inspired Computation*, vol. 10, no. 1, pp. 142-156.

P. Kumar and R. Yadav, 2022, "Grad-CAM Based Visualization for Enhancing Explainability in Diabetic Retinopathy Detection," *Artificial Intelligence in Medicine*, vol. 126, pp. 102190.

V. Patel and K. Shah, 2023,"Diabetic Retinopathy Grading Using Hybrid CNN-LSTM Model," *Expert Systems with Applications*, vol. 202, pp. 117319.

L. Wang et al.,2022, "Multimodal Deep Learning for Diabetic Retinopathy Classification Using Fundus and Optical Coherence Tomography Images," *IEEE Transactions on Medical Imaging*, vol. 41, no. 3, pp. 789-799.

J.Kim et al.,2022, "Federated Learning for Diabetic Retinopathy Detection: Addressing Privacy Concerns in AI Healthcare," *Nature Machine Intelligence*, vol. 4, no. 7, pp. 570-578.

R. Singh et al.,2023 "Improved Robustness in Diabetic Retinopathy Detection Using Adversarial Training," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2781-2789.