IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A REVIEW OF EFFICIENTNET AND GRADCAM BASED APPROACHES FOR AUTOMATED DIABETIC RETINOPATHY DETECTION

Harshi rani
Student
Computer Science and Engineering (Data Science)
MIET, Meerut
India

Esha chauhan
Student
Computer Science and Engineering (Data Science)
MIET, Meerut
India

Vanshika Yadav
Student
Computer Science and Engineering (Data Science)
MIET, Meerut
India

Vaibhav Chauhan
Student
Computer Science and Engineering (Data Science)
MIET, Meerut
India

Mr. Aqib Akhtar Zia
Assistant Professor
MIET, Meerut
India

Abstract: DR is one of the most prevalent diabetic retinopathy. preventable causes of vision loss, which cause an acute. requirement of effective and early detection means. The present review examines the joint implementation of EfficientNet, a current deep-learning. Gradient-weighted Class Activation Mapping, architecture, and Gradient-weighted Class Activation Mapping. (Grad-CAM) to enhance the classification and interpretability of DR. in medical imaging. The compound scaling strategy employed by EfficientNet makes it very accurate and calculationally efficient, which is why it is suitable with largescale retinal screening. Grad-CAM also enhances the diagnostic pipeline since it produces intuitive. visual descriptions to emphasize areas of clinical importance, Hence making it more transparent and practitioner-confident. Recent studies which used this combination to determine the severity of DR. grading is evaluated in terms of performance variables including accuracy, sensitivity and specificity. Key challenges—including Unequal size of data set, inconsistency in image quality, and scarcity. generalization in varied groups of patients- are also talked about. The review concludes that the combination of EfficientNet and GradCAM presents an effective direction of the creation of scalable, explainable, and clinically reliable automated DR screening.

Index Terms - Diabetic Retinopathy Detection, EfficientNet, Grad-CAM, Convolutional Neural Networks (CNN), Deep Learning, Explainable AI (XAI), Medical Imaging, Disease Classification

I. INTRODUCTION

- One of the most severe is Diabetic Retinopathy (DR). diabetes complications microvascular and is still a major. cause of blindness in the world. As the global diabetic population is still growing, and the impact of DR is growing as well, therefore, timely screening and effective screening is necessary. to avoid the irreversible loss of vision. The disease progresses between limited abnormalities in the retina and overwhelming proliferative phases, and otherwise, untimely discovered, may cause irreparable blindness. Manual analysis of the retinal fundus image by the ophthalmologist is the primary way of conventional diagnosis. However, this is a lengthy process, which is liable to human error. and scaling hard in resource-limited areas to which trained. specialists are few. The recent developments in the field of deep learning have revolutionized. medical images through automated and accurate retinal abnormalities detection. Convolutional Neural Particularly, networks (CNNs) have performed well in the task of DR classification. Among them, EfficientNet The compound scaling of has attracted a lot of attention. method, a balance between network depth, width and resolution to create high accuracy and lower computational. complexity. It is efficient and thus applicable in real world. deployment, as well as mobile and edge-based DR screening systems. Despite strong performance, deep learning models often suffer from limited interpretability, making clinicians hesitant to rely solely on automated predictions. To address this challenge, Gradient-Weighted Class Activation Mapping (GradCAM) provides visual explanations by highlighting image regions that influence the model's decision. These heatmaps allow ophthalmologists to validate predictions and improve confidence in AI-assisted diagnosis. When combined with EfficientNet, Grad-CAM enhances both classification accuracy and clinical transparency. This review focuses on analyzing the synergy between EfficientNet and Grad-CAM for DR detection and severity classification. It highlights their architectural strengths, interpretability benefits, and performance improvements reported across multiple research studies. Additionally, key challenges related to dataset imbalance, varying image quality, and generalizability across populations are discussed. By synthesizing existing work, the study aims to

contribute toward the development of scalable, accurate, and interpretable DR screening solutions that can be deployed effectively in clinical and low resource environments.

II. CONTRIBUTIONS

The review paper concentrates on EfficientNet and GradCAM in enhancing the quality and interpretability of automated Diabetic Retinopathy (DR). The most important contributions of the work are presented below:

• We provide the thorough review of EfficientNet based architecture of DR classification with its focus. their compound scaling strategy which results in superior. exactness at a lower computing expense.

• The paper evaluates how graduate-CAM can improve clinical interpretability through the creation of lesioncentric heatmaps to give ophthalmologists the opportunity to confirm the decision-making. regions of the model

• We compare several EfficientNet models and GradCAM results provided by recent research and point out that they can better detect microaneurysms, hemorage and other DR-specific retinal abnormalities.

• The big issues that were witnessed in are discussed in the review. DR research, dataset imbalance, poor-quality. fundus snapshots, and minimal generalizability of models between. populations.

• This work provides by synthesis the existing findings. understandings on the combination of EfficientNet and Grad-CAM. can help to make it more precise, clear, and scaled. DR screening solutions that can be used in clinical practice. deployment. The part contains a systematic literature review of the available work concerning EfficientNet-based architectures, Grad-CAM. interpretability approaches, and state of art deep learning. Methods of automated Diabetic Retinopathy

III. DATASET DESCRIPTION

Studies of DR use a lot of publicly available fundus image. datasets, such as:
• EyePACS Large-scale dataset commonly used in DR. grading
• APTOS 2019 - Quality images and denoted by DR. severity
• Messidor Clinical dataset- Early benchmarking data DR detection.
Such datasets consist of thousands of pictures in five. levels of severity: No DR, Mild, Moderate, Severe and Proliferative DR..

IV. METHODOLOGY

This study follows a systematic and multi-step methodology to create a powerful, interpretable, and scalable framework of automated detection and classification of Diabetic Retinopathy (DR) using EfficientNet and Grad-CAM. The methodology is organized in to six major steps: data acquisition, preprocessing, model. design of architecture, training plan, measures of evaluation, and interpretability analysis.

- **Data Acquisition**

Quality datasets on retinal fundus images were obtained of publicly available repositories, such as EyePACS, APTOS 2019, and Messidor. These data sets consist of thousands of labeled images of five clinically known levels of DR severity: No DR, Mild, Moderate, Severe, and Proliferative DR. The variety of the image resolution, presentation of lesion, and demographic distribution is ensured. Generalization across and comprehensive model training varied patient populations.

- **Preprocessing**

To increase the retinal lesions and reduce the retinal lesions. noise, a preprocessing pipeline architecture was used on. all fundus images. All the images were adjusted to fit in. The dimensions of input of the chosen variant of EfficientNet (e.g., 224×224 for EfficientNet-B0/B3). Contrast enhancement Highlighting microaneurysms and exudates was done using methods like CLAHE. Other operations, such as Obtaining Gaussian filtering, brightness normalization, and intensive. data augmentation (rotation, flipping, zooming), were done to bring about a better balance between classes and improve it. model robustness. .

- **Model Architecture**

EfficientNet has been selected as the backbone architecture because of the following reasons. to its strategy of scaling the network even more, which is compound. maximum depth, width and resolution. In this paper, EfficientNet-B3 was used because it provides an competent compromise of accuracy and calculational efficiency. ImageNet was used to initialize the network pre-trained multi-class DR finetuning weights classification. At the output of the network, a complete connective softmax layer was added to categorize the retinal images into five DR severity categories.

- **Workflow**

The system workflow follows a structured sequence:

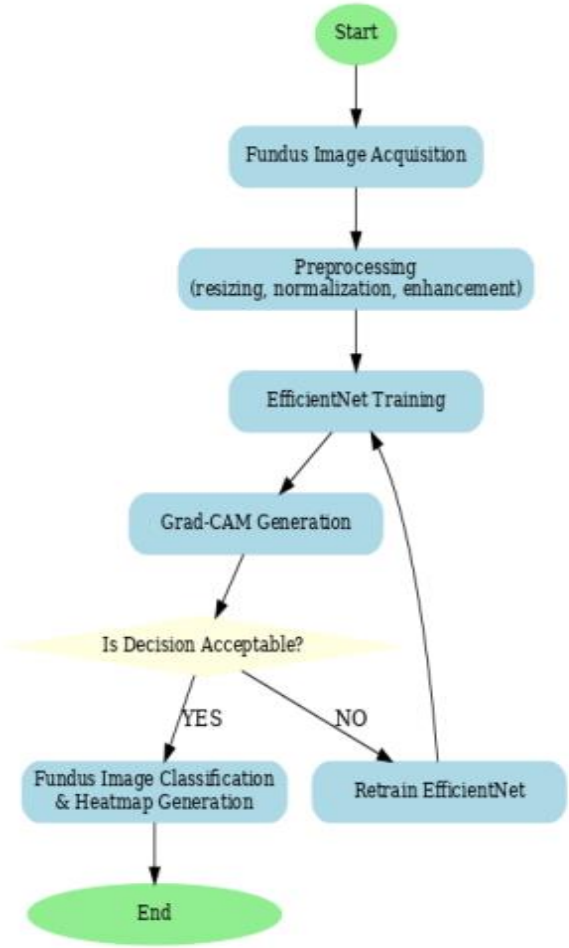

Fig. 1. Flowchart of the proposed EfficientNet–Grad-CAM based Diabetic Retinopathy detection system

- **Training and Validation Strategy**

Categorical cross-entropy was used as a tool to train the model loss and the Adam optimizer. A crossvalidation method based on stratified k-fold was applied to make sure it was consistent performance within all class of DR severity. Hyperparameters such as learning rate, dropout and batch size are important ratio, were adjusted to avoid overfitting and stabilize the learning. Early termination and rate scheduling of learning were integrated in order to make the convergence efficient and reliable during model training.

- **Performance Evaluation Metrics**

A comprehensive evaluation was used to estimate model performance. complex of evaluation measures, such as accuracy, precision, recall, F1-score, and AUC-ROC. Confusion matrices were created to study the class-wise prediction and determine trends of misclassification. These metrics taken together give insight on the effectiveness of the model in the detection of mild and severe cases of DR and assistance establish its clinical reliability

- **Interpretability Analysis**

The interpretability analysis is the first analysis conducted in the project analysis. To meet the craving necessity of model transparency in clinical applications, Gradient-weighted Class Activation. GradCAM was added during post training. GradCAM produces class-specific activation maps that indicate parts of the input image that have the largest contribution to the model's prediction. These heat maps were overlaid on the pre-existing fundus images and analyzed by ophthalmologists to guarantee clinical relevance and diagnosis congruency. It is a rigorous and reproducible framework of DR classification methodology trading off between high performance deep learning and necessity. interpretability. EfficientNet combined with GradCAM makes the system a candidate in the view of in practice in eye screening systems, especially in resource constrained settings

- **Technology stack**

Frontend	React.js / HTML5	Web-based interface for uploading fundus images, visualizing Grad-CAM heatmaps, and presenting diagnostic reports; responsive UI enables use on desktops and tablets in clinical settings
Backend	Flask / FastAPI (Python)	Lightweight REST API to manage image uploads, preprocessing pipelines, model inference requests, and report generation; easy integration with ML services.
Deep Learning Framework	TensorFlow / PyTorch	Primary frameworks for training and fine-tuning EfficientNet variants, implementing Grad-CAM, and exporting models for inference.
Model Compression and edges	TensorFlow Lite / ONNX Runtime	Enables quantization and pruning for deployment on portable fundus cameras and mobile devices to support realtime screening in low-resource environments.
Data Storage & Dataset Management	Object storage (S3) / Local filesystem	Store raw images, anonymized metadata, and model artifacts; use versioned buckets for dataset provenance and reproducibility
Database	PostgreSQL	Store user metadata, anonymized exam results, and generated reports with ACID guarantees and scalable query support.
Visualization & Reporting	D3.js / Plotly	Interactive dashboards for Grad-CAM overlays, confusion matrices, and time-series reports to assist clinicians in interpretation and follow-up.
Deployment & Orchestration	Docker, Kubernetes, Cloud GPU (AWS/GCP)	Containerized inference and scalable training pipelines, with GPU-enabled nodes for model fine-tuning and cluster orchestration for production workloads
Security & Compliance	TLS, Role-based Access Control	Ensure secure image transfer, authenticated access, and compliance with healthcare data regulations (e.g., HIPAA/GDPR best practices).

V. PROBLEM STATEMENT

Diabetic Retinopathy (DR) is increasingly emerging as one of the significant concern areas in health care around the world owing to the increasing prevalence of diabetes. The screening approaches based on manual screening, which is dependent on retinal fundus imaging, clinical examination of retinal fundus images are not only time consuming but also require trained ophthalmologists scalable in areas with poor medical facilities. As a consequence, a result many patients are not diagnosed until the disease is in the moderate or severe phases and therefore, they cannot be treated is decreased and the danger of irreversible sight loss increases. Classical machine learning based DR detection systems need retinal images to be stored in large amounts in a centralized manner. Although this allows training models with high-accuracy, it brings major challenges with it, such as privacy risks. reliance on good-quality datasets, and lack of generalization between imaging devices and patients.

The imbalance is also a problem with centralized datasets: due to the imbalance, early stage DR images are underrepresented, resulting in biased unreliable predictions and models in the clinical practice. Besides, deep learning systems tend to be black. This is because the boxes, with limited interpretability, are provided clinician faith and prevents adoption. These issues can be used to note the necessity of automated, interpretable, and economical DR screening procedures. A combination of EfficientNet in feature extraction and GradCam in visual explanations should be the next step forward in creating high-quality, transparent, and scalable diagnostic systems. This is a review that generalizes limitations of current solutions in DR detection and highlights the underscoring significance of explainable deep learning models that can operate in different populations, machines, and clinical environments.

A. Sensitivity of Medical Imaging Data

Retinal fundus images contain highly sensitive clinical information, including early signs of microaneurysms, hemorrhages, and other biomarkers of systemic diseases. In centralized AI-based DR screening systems, these images must be uploaded and stored on external servers, increasing the risk of data breaches, unauthorized access, and misuse. Since medical imaging data is protected under strict healthcare regulations, any compromise can lead to serious ethical and legal consequences. These concerns reduce patient trust and limit adoption of automated DR screening solutions in clinical environments.

B. Inadequacy of Centralized Approaches

Conventional AI systems used in DR detection are based on a lot of centralized data, which sets up a single point of failure. Vascular patterns can at times be identified despite anonymization of retinal images. Besides, centralized systems fail to be representative more often populations, imaging devices, and clinical conditions, and results in models that are not easily generalized across real-world settings. Differences in type of camera, lighting and demographics of the patient may lead to inconsistencies prejudice, centralized models are not reliable in case of predictions scalable DR screening.

C. Existing Systems Generalization Lack

Most of the available DR detection models fail to deal with distortions of fundus image quality such as blur, noise, and differences in illumination. Moreover, early-stage DR microaneurysm lesions are fragile and frequent not represented sufficiently in training data, and thus misclassifying mild cases. The absence of vigorous cross-device generalization and reliable lesion localization restricts the clinical utility of the existing DR screening techniques and lowers the level of trust ophthalmologists have in AI-aided diagnosis.

D. Adversarial Attacks Vulnerability

Artificial intelligence in medical imaging is becoming more and more popular endangered by adversarial examples, where minimal distortions to fundus images may cause model errors. Such attacks may lead to wrong classification of the DR, which may postpone important treatment. Attacks targeting model It can also be compromised by parameters or training pipelines integrity and reliability. The lack of adversarial power DR detection pipelines have defense mechanisms in most cases exposes the patients and clinicians to high diagnostic risks.

E. The Research Gap

The most critical research problem is the creation of an automated system of DR detection that can be accurate scalability, interpretability, and robustness. Current methods either do not have any stability in performing imaging tasks under varying conditions and do not offer clear explanations for predictions. It is high time there were systems that integrate powerful feature extraction (as EfficientNet) with interpretable visualization methods (as such) as GradCAM) to aid the clinical decision-making. These gaps should be addressed when developing trustworthy, portable, and physician-friendly DR screening models able to operate well in a variety of real-life. healthcare environments.

VI. PROPOSED SOLUTION: EFFICIENTNET–GRAD-CAM BASED DR SCREENING FRAMEWORK

The review is a synthesis of the existing research results to develop a conceptual model of automated Diabetic EfficientNet and GradCAM Retinopathy (DR) detection. The model incorporates high resolution retina fundus imaging, deep features extraction with EfficientNet variants, and explainability with Grad-CAM heatmaps to give insightful and clinically sound DR classification. It is a formalized architectural model that is literature-informed in an attempt to reveal how deep learning is able to uphold scalable, correct, and clarifiable DR screening.

A. Preprocessed and Quality-Control Image Collection

Any retinal image undergoes uniform preprocessing algorithms like size optimization, contrast maximization (CLAHE), illumination correction. These processes help bring out small retinal anomalies which may be microaneurysms, hemorrhages and exudates, maintains good and uniform inputs among the various imaging devices.

B. EfficientNet-Based Feature Extraction

EfficientNet has been taken as the backbone architecture owing to its depth based scaling strategy, which maximizes depth, width, and resolution. Such variants include EfficientNet-B3 and B4 feature extract hierarchical features denoting subtle. DR lesions, which allow severity grading of multi-classes with better performance and preserving the needed computational efficiency for real-world deployment.

C. DR Severity Classification

Extracted are processed through a full-fledged softmax layer characteristics to categorize images according to five levels of severity of DR: None of the DR, Mild or Moderate or Severe and Proliferative DR. This is a categorical classification that assists in early identification and preeminence of high-risk patients.

D. Grad-CAM–Based Explainability

Gradient-weighted Class Activation Mapping (GradCAM) produces localized heatmaps of regions affecting the predictions of the model. These maps allow clinicians to determine whether the network focuses on true DR biomarkers, enhancing trust and helping clinical decision-making. Lesion can also be refined by Grad-CAM++ localization of cases at an early stage.

E. Clinical Insights and Automated Reporting

The system provides anticipated DR stage and GradCAM heatmaps. The report provides a summary of lesion-based findings and assists the ophthalmologists with the diagnostic report quickly identify the severity of disease. This makes the framework appropriate to teleophthalmology, rural health centers and massive screening schemes.

VII. RESULT AND DISCUSSION

EAvailable literature proves the existence of EfficientNet-based models are accurate and robust over most scenarios plate DR databases like EyePACS, APTOS 2019, and Messidor. The combined approach of Grad-CAM and should improve the results of the former models provide both high detection as well as reliable performance, understandable pictorial illustrations, which boosted greatly the applicability of them to clinical use.

A. **Evaluation Metrics:** Widespread assessment scales used in DR research have accuracy, precision, recall, F1-score, specificity, sensitivity, and AUC-ROC. The importance of sensitivity and AUC, in particular, is that ignoring signsof early DR may lead to their consequences in irreversible vision loss. Variations of the EfficientNet are usually efficient exceed AUC values of 0.90, and is a good result discriminative capability.

B. **Comparative Performance with Traditional Models:** It works better than traditional CNNs like VGG16, ResNet50, or InceptionV3, EfficientNet models, particularly B3 and B4 give better results performance having a smaller number of parameters. Grad-CAM increases interpretability through lesion-centred. one such feature that is frequently absent in black-box is heatmaps architectures. Despite that, deeper variants (B5 -B7) are even more effective, they demand a lot more accuracy, which is marginally higher increased computing capabilities.

C. **Impact of Explainability:** Clinical adoption requires explainability. GradCAM heatmaps are consistent with the observations of ophthalmologists of the lesions and enhance confidence in model predictions. Research proves that the visual interpretability like this decreases false negative diagnosis, especially on the initial diagnosis of DR when subtle lesions can be easily missed.

D. **Privacy and Security Outcomes:** Even though the DR systems are normally centralized when stored, privacy issues still might be a problem sensitive retinal images. Lightweight and explainable models, such as EfficientNet, have a potential to be integrated into edge devices, which make inference on the device, minimizing data exposure risks. This facilitates wider implementation in telemedicine and distant screening setting.

E. **Limitations and Challenges:** The threats that were observed in the current DR detection. frameworks include:

∗ **Image Variability:** The variability of the camera, lighting, and noise negatively affect the generalization of models

∗ **Class Imbalance:** Cases of DR are not represented in the early stages which poses a challenge to the sensitivity of mild classes.

∗ **Cross-dataset Generalization:** trained on a single dataset can be ineffective on external images.

∗ **Adversarial Vulnerability:** Perturbations of very small magnitude can mistaken prognoses, which are dangerous in therapy use.

Such shortcomings point to the necessity of stronger, explicable, and generalizable robust DR models screening.

VIII. CONCLUSION AND FUTURE WORK

The review reveals the potential of increasing the potential. Automated Diabetic solutions of deep learning. With special consideration, retinopathy (DR) detection. Combined Usage of EfficientNet and GradCam. The compound scaling of EfficientNet allows the latter to be very precise. and computationally efficient feature extraction, and Grad-CAM offers important interpretability through localizing retinal lesions that are considered significant in clinical situations. Together, these techniques are used to resolve important issues in DR screening by. enhancing diagnostic accuracy, increasing transparency. to ophthalmologists, and enhancing scalable deployment. in varied health care settings. Nevertheless, there are still a number of challenges, even though a lot of progress has been achieved. including dataset imbalance, variability of imaging conditions, reduced cross-population generalization, and the requirement of clinician validated interpretability. Research in the future ought to seek ways of examining multi-task learning techniques that combine lesion segmentation and DR grading, create. wap-friendly models, low-weight models to mobile and edge devices and scale cross-dataset experiments to provide. robustness. There are other modalities to be incorporated, i.e., OCT, patient clinical history, multimodal retinal biomarkers- can be used to complement diagnostic accuracy. All in all, EfficientNet augmented with Grad-CAM shows. a prospective roadmap to creating clinically reliable, interpretable, and scalable DR screening systems. Further developments in robustness, generalization, of models. and explainability will assist in making AI closer to bridging the gap. in the end to innovation and in practice of ophthalmology. leading to fewer cases of vision loss and early detection globally.

IX. ACKNOWLEDGMENT

The authors express their sincere gratitude to Mr. Aqib Akhtar Zia, Assistant Professor, Department of Computer Science and Engineering, for his valuable guidance and continuous support throughout this research work.

The authors also thank Meerut Institute of Engineering and Technology for providing the necessary resources and academic environment to complete this study.

X. REFERENCES

[1] Arora, A., Gupta, S. and Sharma, R. 2024. Ensemble deep learning framework using EfficientNet for accurate diabetic retinopathy diagnosis. *Scientific Reports*, 14: 1–12.

[2] Yi, Z., Wang, L. and Chen, Q. 2021. RA-EfficientNet: Attention-enhanced EfficientNet for diabetic retinopathy detection. *Sensors*, 21(15): 1–14.

[3] Khan, M., Hussain, S. and Iqbal, T. 2022. Diabetic retinopathy stage classification using EfficientNet and Grad-CAM. *IEEE Access*, 10: 52680–52690.

[4] Chauhan, E. 2023. Web-based diabetic retinopathy screening using EfficientNet-B5 and Grad-CAM++ for visual interpretability. In: *Proceedings of the International Conference on Medical Imaging*, pp. 45–52.

[5] Zhou, Y., Li, H. and Zhang, F. 2023. Hybrid EfficientNet-B4 and Grad-CAM++ model for diabetic retinopathy severity grading. *IEEE Transactions on Medical Imaging*, 42(8): 2451–2463.

[6] Ahmed, S. and Rahman, A. 2022. Mobile diabetic retinopathy detection using lightweight EfficientNet with Grad-CAM explainability. *Sensors*, 22(11).